

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO RIO GRANDE
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM COMPUTACIONAL

**UMA ABORDAGEM METODOLÓGICA PARA A CONSTRUÇÃO DE CRITÉRIOS DE
INFORMAÇÃO A PARTIR DE SUPERFÍCIES QUÁDRICAS**

por

Celso Nobre da Fonseca

Tese de Doutorado apresentada para a obtenção do título de Doutor em Modelagem Computacional

Orientadora: Profa. Dra. Viviane Leite Dias de Mattos
Coorientadora: Profa. Dra. Andrea Cristina Konrath

Rio Grande - RS

2024

Celso Nobre da Fonseca

“Uma Abordagem Metodológica Para a Construção de Critérios de Informação a partir de Superfícies Quádricas”

Tese apresentada ao Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande - FURG, como requisito parcial para obtenção do Grau de Doutor. Área de concentração: Modelagem Computacional.

Aprovado em: 10/12/2024

BANCA EXAMINADORA

Documento assinado digitalmente
 **VIVIANE LEITE DIAS DE MATTOS**
Data: 11/12/2024 15:34:10-0300
Verifique em <https://validar.iti.gov.br>

Profa. Viviane Leite Dias de Mattos (Orientadora - FURG)

Documento assinado digitalmente
 **ANDREA CRISTINA KONRATH**
Data: 11/12/2024 10:40:55-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Andréa Konrath (Co-orientadora - UFSC)

Documento assinado digitalmente
 **ADRIANO MENDONCA SOUZA**
Data: 11/12/2024 10:30:03-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Adriano Mendonça de Souza (UFSM)

Documento assinado digitalmente
 **Luiz Ricardo Nakamura**
Data: 10/12/2024 16:22:51-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Luiz Ricardo Nakamura (UFLA)

Documento assinado digitalmente
 **TIAGO DA CRUZ ASMUS**
Data: 10/12/2024 16:54:41-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Tiago da Cruz Asmus (FURG)

Documento assinado digitalmente
 **ANDRE ANDRADE LONGARAY**
Data: 10/12/2024 16:59:02-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. André Andrade Longaray (FURG)

Rio Grande - RS
2024

Ficha Catalográfica

F676a Fonseca, Celso Nobre da.

Uma abordagem metodológica para a construção de critérios de informação a partir de superfícies quádricas / Celso Nobre da Fonseca . – 2024.
205 f.

Tese (doutorado) – Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Modelagem Computacional, Rio Grande/RS, 2024.

Orientadora: Dra. Viviane Leite Dias de Mattos.

Coorientadora: Dra. Andréa Cristina Konrath

1. Critérios de Informação 2. Modelagem 3. Simulação 4. Superfícies Quádricas I. Mattos, Viviane Leite Dias de II. Konrath, Andréa Cristina III. Título.

CDU 004.414.23

Catálogo na Fonte: Bibliotecário Simone Tarouco Przybylski CRB 10/1166

Agradecimentos

A conclusão desta tese representa o resultado de um esforço coletivo que envolveu o apoio e a colaboração de muitas pessoas. A todos que de alguma forma contribuíram para a realização deste trabalho, meu sincero agradecimento.

Primeiramente, agradeço às minhas orientadoras, Prof^a. Dra. Viviane e a Prof^a. Dra. Andrea, pelas orientações, paciência e dedicação ao longo de toda a pesquisa. Suas sugestões e críticas construtivas foram fundamentais para o desenvolvimento deste trabalho.

Agradeço também à minha esposa e filho, que estiveram ao meu lado em todos os momentos deste processo.

Agradeço ainda ao meu amigo e colega, Arlindo, pela amizade e pelas revisões minuciosas de diversas equações ao longo do desenvolvimento da pesquisa. Sua contribuição foi de grande importância para o aperfeiçoamento do conteúdo técnico deste trabalho.

Agradeço também aos membros da banca examinadora, que com seus conhecimentos e contribuições enriqueceram este trabalho, o meu sincero agradecimento. Suas observações e sugestões foram fundamentais para aprimorar a qualidade final da pesquisa.

Gostaria de expressar minha eterna gratidão aos meus pais, cuja essência continua a me guiar com amor, inspiração e força. Este trabalho é, em grande parte, fruto dos valores, ensinamentos e apoio incondicional que recebi de ambos.

Meu progresso é reflexo do caminho trilhado por aqueles que, com sua sabedoria e dedicação, abriram as portas para novas possibilidades.

RESUMO

A escolha de critérios de informação é crucial na seleção de modelos, impactando diretamente a precisão, eficiência e robustez das previsões. Em contextos onde os dados apresentam características complexas, como variações não lineares, a utilização de critérios de informação eficazes se torna parte essencial para garantir a adequação do modelo e a qualidade das previsões. Este estudo foi desenvolvido com o objetivo de avaliar a possibilidade de, a partir de uma equação Quádrica, formular um critério de informação abrangente que seja adaptável ao tipo de modelo escolhido e para ser utilizado em modelagens que envolvam séries temporais. O estudo iniciou com a análise do desempenho dos Critérios de Informação tradicionais, AIC, AICc e BIC, o que foi feito em um estudo de caso, seguido de uma análise do comportamento da penalidade destes critérios em função da quantidade de parâmetros e do comprimento da série. Após, foi estudada a relação das penalidades de Critérios de Informação com curvas Cônicas e superfícies Quádricas. Em sequência, foi proposto um Critério de Informação com uma penalidade abrangente e adaptável. Finalizando, foi elaborado um estudo de simulação que envolveu planejamento e análise dos resultados de um experimento para verificar a eficácia e adaptabilidade das formulações propostas. Neste experimento foram geradas séries temporais tendo modelos autoregressivos e de médias móveis como mecanismo gerador de dados, com e sem constante, com e sem tendência, e diferentes tamanhos. As séries foram geradas com diferentes intensidades de coeficientes autoregressivos e de médias móveis. Para cada condição experimental, geraram-se 3000 séries, a partir das quais foram calculadas as proporções de identificações corretas. Após uma análise exploratória dos dados, foi realizada uma Análise de Variância não paramétrica (teste de Friedman), seguida de testes de Wilcoxon para comparações múltiplas, e uma Regressão Beta para avaliar o impacto de cada fator, bem como suas interações, nas proporções obtidas. A análise realizada permitiu comprovar que o desempenho dos critérios de informação na modelagem de uma série temporal depende do tipo de mecanismo gerador dos dados. Três versões do critério proposto superaram os critérios tradicionais AIC, AICc e BIC. A versão T3, a melhor delas, é uma alternativa promissora para análise e seleção de modelos, especialmente em séries temporais com dados gerados por modelos autoregressivos ou de médias móveis.

Palavras-chave: Critérios de Informação, Modelagem, Simulação, Superfícies Quádricas

ABSTRACT

The choice of appropriate information criteria plays a crucial role in model selection, as these models directly affect the accuracy, efficiency, and robustness of predictions. In contexts where data exhibit complex characteristics, such as nonlinear variations, the use of effective information criteria becomes essential to ensure model adequacy and the quality of predictions. This study was developed with the aim of evaluating the possibility of formulating a comprehensive information criterion based on a quadratic equation, suitable for use in time series modeling. The research began with an analysis of the performance of traditional information criteria AIC, AICc, and BIC using a case study, followed by an analysis of the penalty behavior of these criteria in relation to the number of parameters and the length of the series. Next, the relationship between the penalties of information criteria and conic curves and quadratic surfaces was explored. Subsequently, a new information criterion with a more comprehensive penalty was proposed. Finally, a simulation study was conducted to verify the effectiveness and applicability of the proposed formulations, involving the generation of time series using autoregressive and moving average models as data-generating mechanisms, with and without constants, trends, and varying sizes. For each experimental condition, 3000 series were generated, and the proportions of correct model identification were determined. After an exploratory data analysis, non-parametric analysis of variance (Friedman test) was performed, followed by Wilcoxon tests for multiple comparisons and Beta regression to evaluate the impact of each factor and their interactions on these proportions. The analysis confirmed that the performance of information criteria in time series modeling depends on the type of data-generating mechanism. Furthermore, three adaptations of the proposed criterion outperformed traditional criteria AIC, AICc, and BIC. The best-performing adaptation T3 shows promise as an alternative for model selection, particularly for time series with autoregressive or moving average data-generating mechanisms.

keywords: Information Criteria, Modeling, Simulation, Quadratic Surfaces

Lista de Figuras

Figura 1:	Comportamento FAC (à esquerda) e FACP (à direita) para um modelo AR(1).	31
Figura 2:	Comportamento FAC (à esquerda) e FACP (à direita) para um modelo MA(1).	31
Figura 3:	Comportamento FAC (à esquerda) e FACP (à direita) para um modelo ARMA(1,1).	32
Figura 4:	Comportamento Gráfico das Superfícies Quádricas em termos do eixo z (Fonte: Stewart (2013)).	50
Figura 5:	Representação gráfica da função $w_{AIC} = w_{AIC}(n,k) = 2k$ utilizada na determinação do Critério de Informação AIC .	51
Figura 6:	Representação gráfica da função $w_{AICc} = w_{AICc}(n,k) = 2k + \frac{2k(k+1)}{n-k-1}$ utilizada na determinação do Critério de Informação AICc	51
Figura 7:	Representação gráfica da função $w = w_{BIC}(n,k) = k \ln(n)$ utilizada na determinação do Critério de Informação BIC .	52
Figura 8:	Representação de um ponto particular de um critério $Z=Z(k,n)$	53
Figura 9:	Comportamento gráfico da penalidade quando $n = 10$. . .	54
Figura 10:	Reta formada pela interseção do plano $n = 2$ (destacado em preto) com o gráfico da penalidade de AIC (em vermelho) .	55
Figura 11:	A Hipérbole destacada na interseção do plano $n = 2$ (em preto) com o gráfico da penalidade de AICc (em vermelho) .	55
Figura 12:	Reta gerada pela interseção do plano $n = 2$ (em preto) com o gráfico da penalidade de BIC (em vermelho)	56
Figura 13:	Comportamento da penalidade de BIC para $n = 10$, $n = 20$ e $n = 30$	56
Figura 14:	Comportamento da penalidade de AICc para $n = 10$, $n = 20$ e $n = 30$	57
Figura 15:	Comportamento da penalidade de AICc para $n = 30$, $n = 40$, $n = 50$, $n = 60$ e $n = 70$	57
Figura 16:	Comportamento da penalidade de AIC, AICc e BIC com $k = 1$	59
Figura 17:	Resultados da penalidade encontrada nos Critérios de Informação AIC, AICc e BIC para os dados analisados	61
Figura 18:	Comportamento da razão dos valores de AIC entre os modelos e o mais parcimonioso.	63

Figura 19:	Comportamento da razão entre o modelo mais parcimonioso e os demais, ordenado pelo AICc para os dados analisados. .	64
Figura 20:	Comportamento da razão entre o modelo mais parcimonioso e os demais, ordenado pelo BIC para os dados analisados. .	65
Figura 21:	Comportamento de $-2*LH$, AIC, AICc e BIC para os 25 modelos candidatos	66
Figura 22:	Comportamento da interseção da penalidade de AIC com a penalidade de AICc	68
Figura 23:	Comportamento da interseção da penalidade de AIC com a penalidade de BIC	69
Figura 24:	Comportamento da interseção da penalidade de AICc com a penalidade de BIC	70
Figura 25:	Gráfico da função descrita pela Equação (63)	71
Figura 26:	Gráfico da função descrita pela Equação (68)	73
Figura 27:	Comportamento da Hipérbole sobre o menor valor de AICc (melhor modelo)	74
Figura 28:	Comportamento da Hipérbole sobre o maior valor de AICc (pior modelo)	75
Figura 29:	Gráfico da Hiperboloide obtida a partir da penalidade do Critério de Informação AICc	78
Figura 30:	Triângulo Retângulo para ψ	81
Figura 31:	Comportamento gráfico para $\frac{d\alpha}{dn}$ em termos de n	84
Figura 32:	Comportamento gráfico para $\frac{d\gamma}{dn}$ em termos de n	85
Figura 33:	Comportamento gráfico para α (em verde) e γ (em vermelho) em termos de n	85
Figura 34:	Comportamento gráfico para γ , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma,E,F}$ com a penalidade do BIC para diferentes valores de n	88
Figura 35:	Comportamento gráfico para γ_1 , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma,E,F}$ com a penalidade do AIC para diferentes valores de n	89
Figura 36:	Comportamento gráfico para γ_1 , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma,E,F}$ com a penalidade do AICc para diferentes valores de n	89
Figura 37:	Fluxograma para a estratégia proposta	92
Figura 38:	Metodologia usada modelar uma série temporal com $n = 50$	94

Figura 39:	Saída dos resultados armazenados em uma planilha eletrônica para modelos de tamanho 50	96
Figura 40:	Boxplot das proporções de acertos para cada critério de informação	98
Figura 41:	Boxplot das proporções de acertos por modelo ARMA	98
Figura 42:	Gráfico Boxplot com os modelos considerados para o critério AIC	100
Figura 43:	Gráfico Boxplot com os modelos considerados para o critério AICc	100
Figura 44:	Gráfico Boxplot com os modelos considerados para o critério BIC	101
Figura 45:	Gráfico Boxplot com os modelos considerados para o critério T1	101
Figura 46:	Gráfico Boxplot com os modelos considerados para o critério T2	101
Figura 47:	Gráfico Boxplot com os modelos considerados para o critério T3	101
Figura 48:	Gráfico Boxplot com os modelos considerados para o critério T4	102
Figura 49:	Gráfico Boxplot com os modelos considerados para o critério T5	102
Figura 50:	Gráfico Boxplot com os modelos considerados para o critério T6	102
Figura 51:	Gráficos envelope para modelos Beta de AIC de cima para baixo ARMA(0,1),ARMA(1,0) e ARMA(1,1)	106
Figura 52:	Interação entre ϕ e θ para ARMA(1,1) para AIC	107
Figura 53:	Interação entre ϕ e $B1$ para ARMA(1,1) para AIC	107
Figura 54:	Interação entre ϕ e $B0$ para ARMA(1,1) para AIC	107
Figura 55:	Interação entre ϕ e n para ARMA(1,1) para AIC	107
Figura 56:	Interação entre θ e $B1$ para ARMA(1,1) para AIC	108
Figura 57:	Interação entre θ e $B0$ para ARMA(1,1) para AIC	108
Figura 58:	Interação entre θ e n para ARMA(1,1) para AIC	108
Figura 59:	Interação entre $B1$ e $B0$ para ARMA(1,1) para AIC	108
Figura 60:	Interação entre $B1$ e n para ARMA(1,1) para AIC	109
Figura 61:	Interação entre ϕ e n para ARMA(1,0) para AIC	110
Figura 62:	Interação entre $B1$ e $B0$ para ARMA(1,0) para AIC	110
Figura 63:	Interação entre $B1$ e n para ARMA(1,0) para AIC	110

Figura 64:	Interação entre θ e $B1$ para ARMA(0,1) para AIC	111
Figura 65:	Interação entre θ e $B0$ para ARMA(0,1) para AIC	111
Figura 66:	Interação entre θ e n para ARMA(0,1) para AIC	111
Figura 67:	Interação entre $B1$ e n e para ARMA(0,1) para AIC	111
Figura 68:	Interação entre $B0$ e n e para ARMA(0,1) para AIC	112
Figura 69:	Gráficos envelope para modelos Beta de AICc de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	114
Figura 70:	Interação entre ϕ e θ para ARMA(1,1) para AICc	115
Figura 71:	Interação entre ϕ e n para ARMA(1,1) para AICc	115
Figura 72:	Interação entre θ e $B1$ para ARMA(1,1) para AICc	115
Figura 73:	Interação entre θ e $B0$ para ARMA(1,1) para AICc	115
Figura 74:	Interação entre θ e n para ARMA(1,1) para AICc	116
Figura 75:	Interação entre ϕ e n para ARMA(1,0) para AICc	117
Figura 76:	Interação entre $B1$ e $B0$ para ARMA(1,0) para AICc	117
Figura 77:	Interação entre θ e $B0$ para ARMA(0,1) para AICc	118
Figura 78:	Interação entre θ e $B1$ para ARMA(0,1) para AICc	118
Figura 79:	Interação entre θ e n para ARMA(0,1) para AICc	118
Figura 80:	Gráficos envelope para modelos Beta de BIC de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	120
Figura 81:	Interação entre ϕ e n para ARMA(1,1) para BIC	121
Figura 82:	Interação entre θ e n para ARMA(1,1) para BIC	121
Figura 83:	Interação entre ϕ e $B0$ para ARMA(1,0) para BIC	122
Figura 84:	Interação entre ϕ e n para ARMA(1,0) para BIC	122
Figura 85:	Interação entre θ e $B1$ para ARMA(0,1) para BIC	123
Figura 86:	Interação entre θ e $B0$ para ARMA(0,1) para BIC	123
Figura 87:	Interação entre θ e n para ARMA(0,1) para BIC	123
Figura 88:	Interação entre $B1$ e $B0$ para ARMA(0,1) para BIC	123
Figura 89:	Gráficos envelope para modelos Beta de T1 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	125
Figura 90:	Interação entre ϕ e θ para ARMA(1,1) para T1	126
Figura 91:	Interação entre ϕ e n para ARMA(1,1) para T1	126
Figura 92:	Interação entre θ e n para ARMA(1,1) para T1	126
Figura 93:	Interação entre ϕ e $B1$ para ARMA(1,0) para T1	127
Figura 94:	Interação entre ϕ e $B0$ para ARMA(1,0) para T1	127
Figura 95:	Interação entre ϕ e n para ARMA(1,0) para T1	127
Figura 96:	Interação entre θ e $B1$ para ARMA(0,1) para T1	129
Figura 97:	Interação entre θ e $B0$ para ARMA(0,1) para T1	129

Figura 98:	Interação entre θ e n para ARMA(0,1) para T1	129
Figura 99:	Interação entre $B1$ e $B0$ para ARMA(0,1) para T1	129
Figura 100:	Gráficos envelope para modelos Beta de T2 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	131
Figura 101:	Interação entre ϕ e θ para ARMA(1,1) para T2	132
Figura 102:	Interação entre ϕ e $B1$ para ARMA(1,1) para T2	132
Figura 103:	Interação entre θ e $B1$ para ARMA(1,1) para T2	132
Figura 104:	Interação entre ϕ e n para ARMA(1,0) para T2	133
Figura 105:	Interação entre θ e $B1$ para ARMA(0,1) para T2	134
Figura 106:	Interação entre θ e $B0$ para ARMA(0,1) para T2	134
Figura 107:	Interação entre θ e n para ARMA(0,1) para T2	134
Figura 108:	Gráficos envelope para modelos Beta de T3 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	136
Figura 109:	Interação entre ϕ e n para ARMA(1,1) para T3	137
Figura 110:	Interação entre θ e n para ARMA(1,1) para T3	137
Figura 111:	Interação entre θ e $B1$ para ARMA(0,1) para T3	138
Figura 112:	Interação entre θ e $B0$ para ARMA(0,1) para T3	138
Figura 113:	Interação entre $B1$ e $B0$ para ARMA(0,1) para T3	138
Figura 114:	Interação entre $B1$ e n para ARMA(0,1) para T3	138
Figura 115:	Interação entre $B0$ e n para ARMA(0,1) para T3	139
Figura 116:	Gráficos envelope para modelos Beta de T4 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	141
Figura 117:	Interação entre ϕ e n para ARMA(1,1) para T4	142
Figura 118:	Interação entre θ e n para ARMA(1,1) para T4	142
Figura 119:	Interação entre ϕ e n para ARMA(1,0) para T4	143
Figura 120:	Interação entre $B1$ e $B0$ para ARMA(1,0) para T4	143
Figura 121:	Interação entre θ e $B1$ para ARMA(0,1) para T4	144
Figura 122:	Interação entre θ e $B0$ para ARMA(0,1) para T4	144
Figura 123:	Interação entre θ e n para ARMA(0,1) para T4	144
Figura 124:	Interação entre $B1$ e $B0$ para ARMA(0,1) para T4	144
Figura 125:	Gráficos envelope para modelos Beta de T5 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	147
Figura 126:	Interação entre ϕ e n para ARMA(1,1) para T5	148
Figura 127:	Interação entre θ e n para ARMA(1,1) para T5	148
Figura 128:	Interação entre ϕ e $B0$ para ARMA(1,0) para T5	149
Figura 129:	Interação entre ϕ e n para ARMA(1,0) para T5	149
Figura 130:	Interação entre θ e $B1$ para ARMA(0,1) para T5	150

Figura 131:	Interação entre θ e $B0$ para ARMA(0,1) para T5	150
Figura 132:	Interação entre θ e n para ARMA(0,1) para T5	150
Figura 133:	Interação entre $B0$ e $B1$ para ARMA(0,1) para T5	150
Figura 134:	Gráficos envelope para modelos Beta de T6 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)	153
Figura 135:	Interação entre ϕ e n para ARMA(1,1) para T6	154
Figura 136:	Interação entre θ e n para ARMA(1,1) para T6	154
Figura 137:	Interação entre ϕ e $B0$ para ARMA(0,1) para T6	155
Figura 138:	Interação entre ϕ e n para ARMA(0,1) para T6	155
Figura 139:	Interação entre $B1$ $B0$ para ARMA(0,1) para T6	155
Figura 140:	Interação entre θ e $B1$ para ARMA(0,1) para T6	156
Figura 141:	Interação entre θ e $B0$ para ARMA(0,1) para T6	156
Figura 142:	Interação entre θ e n para ARMA(0,1) para T6	156
Figura 143:	Interação entre $B1$ e $B0$ para ARMA(0,1) para T6	156
Figura 144:	Boxplot da proporção de acertos para cada Critério de In- formação	171
Figura 145:	Boxplot da proporção de acertos para cada Critério de In- formação	171
Figura 146:	Gráfico Boxplot com os modelos considerados para o critério AIC	173
Figura 147:	Gráfico Boxplot com os modelos considerados para o critério AICc	173
Figura 148:	Gráfico Boxplot com os modelos considerados para o critério BIC	174
Figura 149:	Gráfico Boxplot com os modelos considerados para o critério T1	174
Figura 150:	Gráfico Boxplot com os modelos considerados para o critério T2	174
Figura 151:	Gráfico Boxplot com os modelos considerados para o critério T3	174
Figura 152:	Gráfico Boxplot com os modelos considerados para o critério T4	175
Figura 153:	Gráfico Boxplot com os modelos considerados para o critério T5	175
Figura 154:	Gráfico Boxplot com os modelos considerados para o critério T6	175

Figura 155:	Comportamento do Campo Gradiente associado à função da penalidade de AIC	179
Figura 156:	Comportamento do Campo Gradiente associado à função da penalidade de AICc	180
Figura 157:	Comportamento do $\cos(\psi)$ quando $k = n$ para a penalidade de AICc	182
Figura 158:	Comportamento do Campo Gradiente associado à função da penalidade de BIC	183
Figura 159:	Comportamento do $\cos(\psi)$ para a penalidade de BIC quando $k = n$	184
Figura 160:	Comportamento da curva relativa ao Laplaciano de AICc	193
Figura 161:	Comportamento da curva relativa ao Laplaciano de BIC	194
Figura 162:	Comportamento do campo Gradiente para $c = 0$ e $a = 1.8$	196
Figura 163:	Comportamento do campo Gradiente para $c = 1.1$ e $a = 1.8$	197

Lista de Tabelas

Tabela 1:	Resultados da penalidade de alguns Critérios de Informação	44
Tabela 2:	Valores dos Critérios de Informação AIC, AICc e BIC, respectiva penalidade e LH dos dados analisados.	60
Tabela 3:	Valores do $\ln \left(\left(\frac{M1}{M2} \right)^2 \right)$ para os modelos da demanda de energia com três parâmetros.	62
Tabela 4:	Valores atribuídos aos coeficientes utilizados na geração de modelos $ARMA(p, q)$	93
Tabela 5:	Fórmulas para os critérios para cada E e F (Equação (103)) e para cada n	94
Tabela 6:	Valores de E e F para $C_n^{E;F}$ equivalentes.	95
Tabela 7:	Medidas descritivas da proporção de acertos de cada Critério de Informação para as séries simuladas.	97
Tabela 8:	Proporção média de acertos de cada Critério de Informação por tipo de modelo e ao tamanho da série.	99
Tabela 9:	Resultados do $valor - p$ obtidos pelos testes de Wilcoxon. .	100
Tabela 10:	Resultados do $valor - p$ para os testes de Wilcoxon para todos os tamanhos considerados para o modelo $ARMA(0,1)$.	103
Tabela 11:	Resultados do $valor - p$ para os testes de Wilcoxon para todos os tamanhos considerados para o modelo $ARMA(1,0)$.	103
Tabela 12:	Resultados do $valor - p$ para os testes de Wilcoxon para todos os tamanhos considerados para o modelo $ARMA(1,1)$.	103
Tabela 13:	Resultados da Regressão Beta para o critério AIC.	105
Tabela 14:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação AICc.	113
Tabela 15:	Resultados da Regressão Beta para o critério BIC.	119
Tabela 16:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T1.	124
Tabela 17:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T2.	130
Tabela 18:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T3.	135
Tabela 19:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T4.	140
Tabela 20:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T5.	146

Tabela 21:	Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T6.	152
Tabela 22:	Interações significativas para os critérios BIC, T3, T4 e T6.	158
Tabela 23:	Proporção de acertos em séries simuladas com $n = 50$ para coeficientes ARMA não aleatórios	167
Tabela 24:	Proporção de acertos em séries simuladas com $n = 100$ para coeficientes ARMA não aleatórios	168
Tabela 25:	Proporção de acertos em séries simuladas com $n = 200$ para coeficientes ARMA não aleatórios	169
Tabela 26:	Valores atribuídos aos coeficientes utilizados na geração de modelos $ARMA(p, q)$	170
Tabela 27:	Medidas descritivas da proporção de acertos de cada Critério de Informação para as séries simuladas.	170
Tabela 28:	Proporção média de acertos de cada Critério de Informação por tipo de modelo e ao tamanho da série.	172
Tabela 29:	Resultados do <i>valor - p</i> obtidos pelos testes de Wilcoxon .	173
Tabela 30:	Resultados do <i>valor - p</i> para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(0,1)	176
Tabela 31:	Resultados do <i>valor - p</i> para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,0)	176
Tabela 32:	Resultados do <i>valor - p</i> para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,1)	176
Tabela 33:	Valores das penalidades dos Critérios de Informação w(AIC) e w(AICc) e BIC para $n = 50$, $n = 100$ e $n = 200$	189

Lista de Quadros

Quadro 1: Regras para interpretação dos gráficos da <i>FAC</i> e da <i>FACP</i> ...	31
Quadro 2: Queries utilizadas	38
Quadro 3: Artigos Seleccionados	39

Lista de Símbolos

$2b(g)$ ou w Viés ou Penalidade do Critério de Informação

$[g_{ij}]$ Matriz relativa a métrica

κ_G^{AICc} Curvatura Gaussiana da penalidade de AICc

κ_G^{AIC} Curvatura Gaussiana da penalidade de AIC

$\kappa_G^{BIC(a)}$ Curvatura Gaussiana da penalidade de BIC com coeficiente a

κ_G^{BIC} Curvatura Gaussiana da penalidade de BIC

κ_G Curvatura Gaussiana de uma superfície

∇w Vetor Gradiente relacionado a penalidade w

$\nabla^2 w$ Laplaciano relacionado a penalidade w

ϕ_i Coeficiente da Parcela Autorregressiva na i-ésima Defasagem

ψ Ângulo entre o vetor gradiente e um vetor unitário

θ_j Coeficiente de Médias Móveis na j-ésima Defasagem

ε_t Erro no t-ésimo Tempo

$\vec{\nu}$ Vetor ortogonal a superfície

$\vec{R}$ Forma vetorial de uma superfície

$A, B, C, D, E, F, M, N, P, Q$ Coeficientes da equação quadrática

$C_n^{\gamma, E, F}$ Critério de Informação Generalizado com Coeficientes γ , E e F

$C_n^{E, F}$ Critério de Informação Generalizado com Coeficientes E e F

$Cr(A, B, D, E, F, M, N, P, Q)$ Critério de Informação Generalizado

$D_{\vec{u}} w$ Derivada direcional da penalidade w na direção de um vetor unitário $\vec{u}$

ds Elemento infinitesimal de comprimento

e Base do Logaritmo Neperiano

g Função empírica definida na função de Kullback Leiber

$I(g; f)$	Informação de Kullback Leiber de uma distribuição f
k	Quantidade de Parâmetros utilizados em uma Modelagem
$L(\theta)$	Função de Verossimilhança
n	Tamanho das Amostras
p	Ordem do Processo Autorregressivo
q	Ordem do Processo de Médias Móveis
x_t	Valor Observado no Instante t
β_0	Constante do Modelo <i>ARMA</i>
β_1	Coefficiente de Tendência do Modelo <i>ARMA</i>
μ	Média de um Processo

Lista de Siglas

ADF	Teste de Dickey Fuller Aumentado
AIC	Critério de Informação de Akaike
AICc	Critério de Informação de Akaike Corrigido
AR	Componente Auto-regressiva
ARMA	Modelo Autorregressivo de Médias Móveis
ARIMA	Modelo Integrado Autorregressivo de Médias
BIC	Critério de Informação Bayesiano
CI	Critério de Informação
DIC	Deviance Information Criterion
ESI	Índice de Estabilidade Econômica
FAC	Função de Autocorrelação
FACP	Função de Autocorrelação Parcial
GEV	Distribuição Generalizada de Valor Extremo
HQ	Critério de Hannan-Quinn
IPEADATA	Base de dados gratuita do Instituto de Pesquisa Econômica Aplicada (Ipea)
KIC	Critério de Informação de Kullback-Leibler
KICCVC	Critério de Informação de Kullback-Leibler Corrigido
LH	Logarítmo Neperiano da Máxima Verossimilhança
LRT	Teste de Razão de Verossimilhança
MA	Componente de Média Móvel
MAE	Erro Absoluto Médio
MAI	Índice de Média de Análise Integrada
MAPE	Erro Absoluto Percentual Médio
MML	Comprimento Mínimo de Mensagem
MPE	Erro Percentual Médio
MSE	Erro Quadrático Médio
MUSIC	Método de Classificação de Sinais Múltiplos
NNAR	Redes Neurais Autorregressivas
PP	Philipp and Perron
QSIC	Critérios de Informação Baseado em Superfícies Quadráticas
SARIMA	Modelo Autorregressivo Integrado de Média Móvel Sazonal
WIC	Critério de Informação Média Ponderada
ZA	
PP	Teste de Phillips e Perron
ADF-GLS	Teste de Elliot e Rothenberg

ZA	Teste de Zivot e Andrews
SARIMA	Modelo Sazonal Integrado Autorregressivo de Médias Móveis
FAC	Função de Autocorrelação
FACP	Função de Autocorrelação Parcial

Sumário

1	INTRODUÇÃO	22
1.1	O tema e sua importância	23
1.2	O problema	25
1.3	Hipótese	25
1.4	Justificativa	26
1.5	Objetivos	26
1.6	Estrutura do trabalho	27
2	REFERENCIAL TEÓRICO	28
2.1	Modelagem <i>ARIMA</i>	28
2.2	Critérios de Informação	32
2.2.1	Critério de Informação de Akaike AIC	35
2.2.2	Critério de Informação de Akaike corrigido AICc	35
2.2.3	Critério de Informação Bayesiano de Schwartz BIC	36
2.2.4	$AIC \times AICc \times BIC$	37
2.3	Trabalhos relacionados	37
2.4	Estudo das cônicas e quádricas	45
2.4.1	Cônicas	45
2.4.2	Quádricas	47
3	CONSIDERAÇÕES SOBRE OS CRITÉRIOS DE INFORMAÇÃO	50
3.1	Considerações sobre a penalidade dos critérios de informação AIC, AICc e BIC e as superfícies quádricas	51
3.2	Análise do comportamento da penalidade dos Critérios de In- formação AIC, AICc e BIC em função de k e n	54
3.3	Análise do desempenho dos Critérios de Informação e sua res- pectiva penalidade: um estudo de caso	59
3.4	Equivalência entre as penalidades dos Critérios de Informação AIC, AICc e BIC	67
3.5	Relação entre o critério de informação AICc e as curvas cônicas	70
3.6	Relação entre o Critério de Informação AICc e as superfícies quádricas	75
3.7	Construção de um Critério de Informação adaptável, a partir da equação de uma Superfície Quádrlica.	78
3.8	Análise da construção da penalidade	80

3.9	Formulação empírica	82
3.10	Proposta para o Critério de Informação	86
3.11	Estratégia de análise	91
4	ESTUDO DE SIMULAÇÃO	93
4.1	Metodologia	93
4.2	Resultados	96
4.2.1	Análise descritiva	97
4.2.2	Análise de variância	99
4.2.3	Análise utilizando a regressão Beta	104
5	CONCLUSÃO	159
A	APÊNDICE	166
A.1	Pseudo-código para as simulações	166
A.2	Tabelas de resultados	167
A.3	Resultados da simulação de experimento com coeficientes ARMA não determinísticos	169
A.3.1	Análise de Variância	172
A.4	Resultados encontrados do Gradiente e da Derivada Direcional para os critérios AIC, AICc e BIC	177
A.5	Diferencial e equação diferencial para penalidades	185
A.6	Construção e análise de Penalidades Integrando Hipóteses a Equações Diferenciais	190
A.7	Análise do Laplaciano para a penalidade dos Critérios de In- formação e metodologia complementar para a criação de um Critério de Informação a partir de uma Equação Diferencial Parcial	192
A.8	Análise da curvatura Gaussiana para a penalidade dos critérios AIC AICc e BIC e métrica para BIC(a)	198
A.9	Equações para os Critérios de Informação Baseados em Superfícies Quadráticas (QSIC)	204

1 INTRODUÇÃO

Uma série temporal é um conjunto de observações ordenadas no tempo (Gujarati and Porter, 2011). Conforme apontado por Morettin and Toloi (2018), a modelagem de séries temporais permite entender o processo que gera os dados, identificar padrões e periodicidades, além de prever valores futuros. Entre as etapas cruciais na modelagem está a seleção do modelo mais adequado, o que requer o uso de métricas que ajudem a comparar a qualidade de diferentes modelos.

Dentre os modelos mais utilizados, destacam-se os modelos ARMA (AutoRegressive Moving Average) e ARIMA (AutoRegressive Integrated Moving Average). O ARMA combina dois componentes: a parte autoregressiva (AR), que descreve a dependência linear entre uma observação e seus valores passados, e a parte de média móvel (MA), que modela a relação entre a observação atual e os erros de previsão passados. O modelo ARIMA, por sua vez, é uma extensão do ARMA, incluindo um componente de diferenciação para tornar a série temporal estacionária, ou seja, remover tendências e sazonalidades.

Para selecionar o modelo mais apropriado, é necessário usar métricas que permitam comparar modelos estatísticos, levando em consideração tanto o ajuste aos dados quanto a complexidade do modelo. Os Critérios de Informação (Hurvich and Tsai, 1989; Burnham, 1998) surgem como uma ferramenta útil nesse processo. Esses critérios comparam modelos com base no equilíbrio entre o ajuste e a penalização pela complexidade, favorecendo modelos mais simples, ou seja, aqueles com o menor número de parâmetros, desde que não comprometam a qualidade da previsão (Burnham, 1998). Segundo Vismara (2009), esses critérios não só ajudam a avaliar o desempenho dos modelos, mas também incorporam a incerteza no processo de seleção, penalizando modelos mais complexos, com maior número de parâmetros.

Nesse trabalho buscou-se mostrar que é possível relacionar a estrutura utilizada nas expressões matemáticas que determinam a penalidade de alguns Critérios de Informação já propostos com as Cônicas e Quádricas.

Tal abordagem permitiria a adoção de distintas estratégias para diferentes tipos de modelagens, proporcionando o desenvolvimento de uma estratégia efetiva para identificação dos melhores modelos. Outrossim, ao relacionar o estudo desses critérios com a estrutura de uma superfície Quádrica, a qual é um objeto geométrico, para representar essas penalidades e naturalmente os Critérios de Informação, abre-se a possibilidade de poder agregar outros possíveis conceitos

teóricos relacionados a geometria. Este trabalho teve como foco a análise dessa estrutura em modelos de séries temporais simuladas do tipo Autorregressivas e de Médias Móveis *ARMA*.

1.1 O tema e sua importância

Um modelo é a representação simplificada de um fenômeno que ocorre em um determinado contexto. Na sua geração, não apenas podem ser utilizadas diferentes metodologias, como também podem ser desconsideradas algumas particularidades em função de sua complexidade, razão pela qual é possível encontrar mais de um modelo para representar o comportamento de um mesmo fenômeno em um mesmo contexto. Como não é possível possuir o controle sobre todas as variáveis que podem influenciar o fenômeno em estudo, bem como não é possível mensurá-las com precisão absoluta, são utilizados modelos estocásticos que, por serem probabilísticos, por vezes perderão alguma informação.

Por exemplo, em um banco de dados é comum a presença de dados faltantes ou incertos. Nesse contexto, medidas resumidas, como a média e o desvio padrão, podem ser usadas para entender a tendência central e a dispersão dos dados disponíveis. No entanto, quando aplicadas a conjuntos com dados incompletos, essas medidas podem mascarar variações importantes e levar a uma perda de informação sobre a distribuição real dos dados.

Evidencia-se a necessidade de estabelecer um critério para a escolha do modelo mais adequado, ou seja, aquele que melhor represente o comportamento real do fenômeno estudado. Os autores Anderson et al. (1994) enfatizam a importância de selecionar modelos com base em princípios científicos. Dentre as abordagens propostas pela Estatística para essa seleção, destacam-se as métricas que utilizam a função de verossimilhança.

A função de verossimilhança é uma ferramenta estatística que quantifica a probabilidade de se obter um conjunto de dados a partir de um modelo teórico. Ela representa o valor máximo da probabilidade de observação dos dados dado um conjunto específico de parâmetros do modelo (Sohil et al., 2022). Dentre os critérios que utilizam a função de verossimilhança figuram o Critério de Informação de Akaike (AIC), proposto em Akaike (1974), o Critério de Informação de Akaike corrigido (AICc), proposto em Hurvich and Tsai (1989), e o Critério de Informação Bayesiano (BIC), proposto em Schwarz (1978).

Uma série temporal é considerada como uma entre as diversas trajetórias possíveis de um fenômeno, o que, por analogia, poderia ser comparado a uma

amostra extraída de uma população. Para a sua modelagem, foi proposta uma metodologia, que ficou conhecida como Metodologia de Box-Jenkins (Box, 2013), que faz a modelagem utilizando modelos lineares probabilísticos gerados a partir de valores passados e erros passados. Segundo Morettin and Tolo (2018), essa abordagem tem obtido bons resultados. Entretanto, assim como a maioria dos modelos estatísticos, supõe que as séries temporais são estacionárias e ergódicas Morettin and Tolo (2018).

Um processo estocástico é chamado de fracamente estacionário ou estacionário de segunda ordem se sua média e sua variância não variam ao longo do tempo, e o valor da covariância entre dois períodos depende somente do grau de defasagens entre as observações (Bueno, 2018). Um processo estocástico é dito ergódico quando os momentos amostrais devem convergir em probabilidade para os momentos populacionais (Bueno, 2018).

Vários testes já foram propostos com a finalidade de verificar a hipótese nula de que o processo estocástico tem raiz unitária, sendo não estacionário, cujo desempenho depende de algumas propriedades do processo estocástico analisado, tais como presença de quebra estrutural. Entre eles estão o teste de Dickey Fuller Aumentado (*ADF*), proposto por Dickey and Fuller (1981), um dos mais utilizados na literatura, o teste não paramétrico (*PP*), proposto por Phillips and Perron (1988) e o teste (*ZA*), proposto por Zivot and Andrews (2002), entre outros.

Na metodologia de Box-Jenkins, depois de constatar a presença de estacionariedade, ou aplicar algum artifício matemático para transformá-la de maneira a possuir esta propriedade, é preciso definir a quantidade de valores passados e a quantidade de erros passados a serem considerados no modelo, o que é feito pela análise conjunta dos gráficos das funções de autocorrelação e autocorrelação parcial. Como esta análise gráfica é subjetiva, usualmente determina-se um conjunto de modelos candidatos. Depois de estimar os coeficientes destes modelos e verificar o cumprimento das suposições do método utilizado na estimação dos seus coeficientes, normalmente o Método da Máxima Verossimilhança, são utilizados os Critérios de Informação para estabelecer uma ordem entre os modelos candidatos em função da sua parcimonia, e aqueles que apresentam menores valores para o Critério de Informação, são aqueles que representam melhor o movimento da série temporal modelada. Evidencia-se aqui a importância desta etapa da análise para a escolha do modelo mais adequado.

1.2 O problema

Na literatura, diversos tipos de Critérios de Informação foram propostos, os quais, em tese, permitem quantificar, entre dois modelos, aquele que está mais próximo do modelo teórico em relação ao conjunto de dados a ser modelado. (Hurvich and Tsai, 1989; Emiliano et al., 2010; Dziak et al., 2020). Cada uma dessas métricas possui uma metodologia distinta para tentar penalizar o modelo e encontrar o melhor ajuste (Akaike, 1974; Hurvich and Tsai, 1989; Schwarz, 1978). No entanto, todas essas penalidades são baseadas em generalizações, como sobre a natureza dos dados, sobre a distribuição dos dados ou sobre a forma como o modelo deve se ajustar a eles. No critério de Akaike, por exemplo, é assumido que os erros de modelagem são independentes e identicamente distribuídos (Akaike, 1974), enquanto o critério BIC (Bayesiano) baseia-se na suposição de que o modelo mais adequado é aquele que maximiza a probabilidade a posteriori dos dados (Hurvich and Tsai, 1989; Esteves and Stern, 2017). Além disso, a seleção do modelo mais adequado ainda pode ser afetada por outros fatores, incluindo o tamanho da amostra, a presença ou ausência de outliers e a complexidade do modelo, entre outros.

Acquah (2018) estuda o problema de seleção do modelo de transmissão de preços assimétricos, que envolve o uso de um critério de seleção do modelo para selecionar o modelo mais apropriado entre modelos concorrentes em um estudo de simulação. Nesse estudo, os autores comparam um critério de informação WIC com os critérios AIC e BIC, avaliando os seus desempenhos, em termos de sua capacidade de selecionar o modelo verdadeiro entre três tipos de modelos para transmissão de preços assimétricos, conhecidos como SECM, CECM e HKD para diferentes tamanhos de amostras. Para algumas situações, por exemplo, para o caso HKD o WIC estaria melhor ajustado a esse modelo do que os critérios AIC e BIC.

Nesse sentido, seria interessante explorar a possibilidade de ajuste de um Critério de Informação para um determinado tipo de modelo, criando-se uma estrutura que seja ajustável a este tipo de modelo, o que pode ser permitido ajustando-se determinados coeficientes. Na literatura pesquisada não foram encontrados registros dessa natureza.

1.3 Hipótese

Nesta pesquisa propõe-se a ideia de que através de uma abordagem baseada em uma equação quadrática é possível desenvolver um Critério de In-

formação personalizado para modelos de séries temporais, o que permitiria a sua adaptação a esse tipo de modelo, melhorando sua eficácia por intermédio da otimização de alguns dos coeficientes presentes em sua equação.

1.4 Justificativa

Com base no estudo das métricas utilizadas pelos Critérios de Informação AIC, AICc e BIC, surgiu o interesse em explorar como diferentes metodologias e formulações podem levar a diferentes abordagens e concepções na análise de modelos. Essas métricas foram idealizadas com penalidades distintas para lidar com a complexidade do modelo, e essa diversidade sugere a possibilidade de outras formulações que podem ser exploradas para aprimorar a análise de modelos.

Considerando que esses critérios são formulados a partir de diferentes princípios, como a natureza dos dados ou a sua distribuição, é possível que um determinado modelo, para os dados considerados, fique ligeiramente inflexível a esses Critérios de Informação quando testado em condições para as quais o modelo foi idealizado. Isto sugere que é importante considerar as limitações e pressupostos dos diferentes Critérios de Informação ao se selecionar o modelo mais apropriado para um determinado conjunto de dados.

Nesse sentido, esse estudo permite elaborar e testar a formulação de um critério de informação baseado na equação de uma Quádrica, tornando viável a ideia de ser adaptável a diferentes tipos de modelagens, ou até mesmo a diferentes propriedades dos dados, como tamanho das amostras (Emiliano et al. 2010 ou Acquah 2018) e presença de *Outliers* nos dados.

Portanto, se o mecanismo gerador de dados pode interferir no desempenho dos Critérios de Informação, é plausível adaptar a estrutura de sua penalidade ao tipo de modelagem que está sendo utilizado para identificar o mecanismo gerador dos dados. Assim, esta adaptação serviria como mais uma opção estratégica no que se refere a identificar os melhores ou mais adequados modelos.

1.5 Objetivos

Essa sessão apresenta o objetivo geral e os objetivos específicos desta pesquisa.

Objetivo Geral

Avaliar a possibilidade de, a partir de uma equação Quádrica, formular um Critério de Informação abrangente que seja adaptável ao tipo de modelo escolhido para ser utilizado em modelagens que envolvam séries temporais.

Objetivos Específicos

Para atingir o objetivo geral, os seguintes objetivos específicos deverão ser alcançados:

- a) Demonstrar que o comportamento gráfico da parcela da equação que determina a penalidade de alguns Critérios de Informação têm a forma de uma superfície Quádrica;
- b) Explorar a forma Quádrica da parcela da equação da penalidade do Critério de Informação;
- c) Propor a equação de uma superfície Quádrica para as penalidades dos Critérios de Informação e uma estratégia para ajuste dos coeficientes dessa equação;
- d) Analisar o desempenho da estratégia proposta na identificação dos modelos mais parcimoniosos do tipo *ARMA*.

1.6 Estrutura do trabalho

Para atender os objetivos e discutir as questões que permeiam esta pesquisa, esta tese foi organizada em seis capítulos.

No capítulo 1 são apresentados a contextualização do tema abordado, o problema que motivou o estudo, a justificativa da pesquisa, os objetivos gerais e específicos, além da estrutura adotada para o desenvolvimento deste trabalho.

O capítulo 2 apresenta o referencial teórico necessário para entender os conceitos fundamentais desta pesquisa. São discutidos os modelos ARIMA, os critérios de informação e alguns trabalhos relacionados. Esse capítulo também explora a teoria relacionada às curvas cônicas e superfícies quádricas.

O capítulo 3 foca na análise detalhada dos critérios de informação, especialmente na penalidade dos critérios AIC, AICc e BIC, abordando suas características, suas limitações e fazendo comparações entre esses. Em seguida, é apresentada a análise de desempenho desses critérios no contexto dos modelos ARIMA, as relações entre eles e a relação desses com curvas cônicas e superfícies quádricas. Nesse capítulo também é dada a proposta para a criação de critérios

de informação utilizando superfícies quádricas e uma estratégia para a obtenção dos coeficientes dessa superfície.

No capítulo 4, é apresentado o estudo de simulação realizado para testar a eficácia dos critérios de informação, tanto os tradicionais (AIC, AICc, BIC) quanto os novos critérios propostos. A metodologia utilizada na simulação é explicada detalhadamente, e os resultados obtidos são analisados de forma crítica, comparando o desempenho dos diferentes critérios com base em várias métricas e experimentos com séries temporais geradas a partir de modelos ARMA.

O capítulo 5 trata dos resultados e discussões, apresentando as principais conclusões derivadas das simulações e análises. São discutidas as implicações dos achados, como a adaptação das penalidades dos critérios de informação e a proposta de um novo critério, com base em superfícies quádricas.

O capítulo 6 apresenta as conclusões gerais deste trabalho, resumindo os principais achados e suas implicações práticas. São sugeridas direções para futuras pesquisas, especialmente no que diz respeito à exploração de novos métodos para a seleção de modelos e à adaptação de critérios de informação para contextos mais específicos.

Finalmente, o apêndice apresenta detalhes complementares sobre a metodologia e os resultados da pesquisa. São apresentados o pseudo-código utilizado nas simulações, além de tabelas com os resultados obtidos. Também são discutidos métodos matemáticos, como o uso de equações diferenciais para a formulação de penalidades, assim como são apresentadas propostas para novas abordagens que visam extrair características adicionais relacionadas aos critérios de informação.

2 REFERENCIAL TEÓRICO

O referencial teórico abrange uma série de conceitos essenciais para a compreensão da pesquisa. Inicialmente, são apresentados os princípios fundamentais da modelagem ARIMA (Auto-Regressiva de Média Móvel Integrada), em sequência, são discutidos os Critérios de Informação e em seguida aborda-se a Geometria Analítica. Por fim, estabelece-se uma conexão entre os Critérios de Informação e as Superfícies Quádricas.

2.1 Modelagem *ARIMA*

De acordo com Morettin and Toloi (2018), um processo autorregressivo de médias móveis de ordens p e q , denotado por ARMA (p, q) , é definido por:

$$x_t = \beta_0 + \beta t + \sum_{i=1}^p \phi_i(x_{t-i}) + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t, \quad (1)$$

em que:

β_0 é a constante,

β é o coeficiente da tendência,

x_t é o t -ésimo valor observado, sendo $t = 1, 2, 3, \dots, T$,

x_{t-i} é o t -ésimo valor com defasagem i ,

p é a ordem do processo autorregressivo, sendo $i = 1, 2, 3, \dots, p$,

q é a ordem do processo de médias móveis, sendo $j = 1, 2, 3, \dots, q$,

ϕ_i é o coeficiente da parcela autorregressiva na i -ésima defasagem,

θ_j é o coeficiente de médias móveis na j -ésima defasagem,

$\varepsilon_t \sim RB(0, \sigma^2)$ é o erro no t -ésimo tempo, e

ε_{t-i} é o t -ésimo erro com defasagem i .

Os modelos *ARMA* são construídos a partir da modelagem do comportamento da correlação serial ou autocorrelação em processos estocásticos, estacionários e ergódicos, sendo a determinação de suas ordens p e q feita, usualmente, a partir das funções de autocorrelação (*FAC*) e autocorrelação parcial (*FACP*). Essa etapa é considerada por Morettin and Toloi (2018) a fase mais crítica do processo iterativo entre as etapas de identificação, estimação e diagnóstico.

A função de autocorrelação do processo é definida por estimativas obtidas a partir da autocorrelação amostral (ρ_r), determinada pela Equação (2):

$$\rho_r = \frac{\sum_{t=1}^{t-r} (x_t - \bar{x})(x_{t+r} - \bar{x})}{\sum_{t=1}^t (x_t - \bar{x})^2}, \quad (2)$$

em que $\bar{x}$ é a média aritmética amostral da série e x_t e x_{t+r} são os valores da série no tempo t e no tempo $t+r$.

Além disso, é possível analisar as autocorrelações parciais, que avaliam a relação entre uma observação e uma defasagem específica, eliminando o efeito das defasagens anteriores. Os coeficientes da função de autocorrelação parcial ϕ_{kr} são obtidos a partir das equações de Yule-Walker, que podem ser expressas de forma matricial pela Equação (3):

$$\vec{\rho} = P * \vec{\Phi}, \quad (3)$$

onde:

$$P = \begin{pmatrix} 1 & \rho_1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_2 & \cdots & \rho_{k-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{pmatrix}$$

é a matriz de autocorrelação,

$$\Phi = \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \cdot \\ \cdot \\ \phi_{kr} \end{pmatrix}$$

é o vetor da autocorrelação parcial e

$$\vec{\rho} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \rho_k \end{pmatrix}$$

é o vetor de autocorrelação onde $k = 1, 2, \dots, p$.

Para a resolução dos valores de ϕ_{kr} é considerado o seguinte sistema:

$$\rho_r = \phi_{k1}\rho_{r-1} + \phi_{k2}\rho_{r-2} + \dots + \phi_{kk}\rho_{r-k}, \quad r = 1, 2, 3, \dots, k, \quad (4)$$

sendo ϕ_{kj} o j -ésimo coeficiente de um modelo $AR(k)$ com k parâmetros.

Os resultados obtidos para a autocorrelação e a autocorrelação parcial são disponibilizados em gráficos e interpretados conjuntamente para a determinação das ordens (p, q) , conforme apresentado no Quadro 1.

Como exemplo, tem-se os gráficos da FAC e FACP de um modelo $AR(1)$ apresentados na Figura 1. Esses gráficos decaem lentamente e possuem queda brusca após a primeira defasagem. Já na Figura 2, o gráfico da FAC decai bruscamente após a primeira defasagem, enquanto o da FACP decai lentamente, representando um modelo do tipo $MA(1)$ ou $ARMA(0,1)$. Já a Figura 3 mostra um comportamento senoidal, representando um modelo $ARMA(1,1)$.

Quadro 1: Regras para interpretação dos gráficos da *FAC* e da *FACP* (Adaptado de Morettin, 2011).

	AR(p)	MA(q)	ARMA(p,q)
FAC	Decaimento exponencial ou senoidal, infinita em extensão	Corte brusco após a defasagem q	Decaimento exponencial ou senoidal após a defasagem q-p
FACP	Corte brusco após a defasagem p	Decaimento exponencial ou senoidal, infinita em extensão	Decaimento exponencial ou senoidal, infinita em extensão

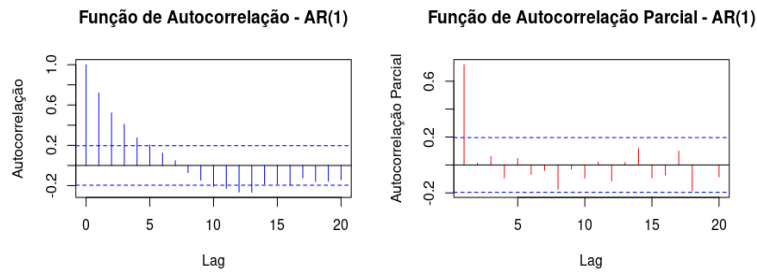


Figura 1: Comportamento FAC (à esquerda) e FACP (à direita) para um modelo AR(1).

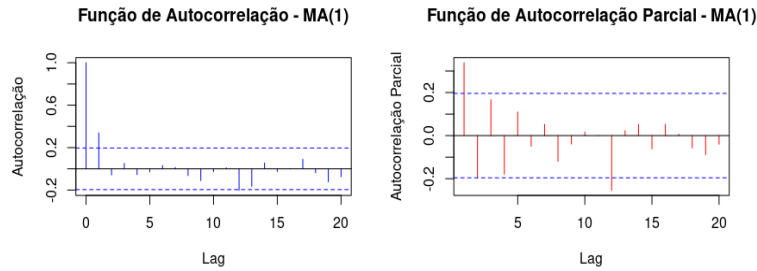


Figura 2: Comportamento FAC (à esquerda) e FACP (à direita) para um modelo MA(1).

No entanto, a modelagem $ARMA(p, q)$ deve ser feita em séries que sejam estacionárias. Caso a série analisada não seja estacionária, há a necessidade de se realizar d diferenças entre observações consecutivas até a série se tornar estacionária. Diante disso, se faz a modelagem $ARIMA(p, d, q)$.

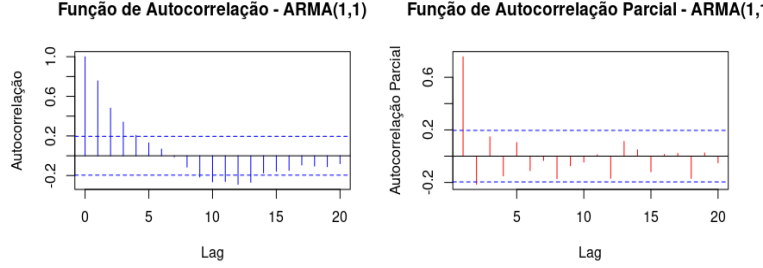


Figura 3: Comportamento FAC (à esquerda) e FACP (à direita) para um modelo ARMA(1,1).

Como a interpretação conjunta dos gráficos das funções *FAC* e *FACP* nem sempre é precisa, uma estratégia que pode ser utilizada consiste em, após a identificação dos possíveis valores atribuídos a p e q , definir um conjunto de modelos candidatos. Nessa situação, considera-se todas as combinações possíveis entre os valores dos inteiros pertencentes ao conjunto definido por $p \pm 1$, $q \pm 1$.

Uma outra estratégia, que também pode ser utilizada para contornar a subjetividade da interpretação gráfica, consiste em definir valores máximos para as ordens p e q e considerar que eles podem assumir os valores inteiros entre o valor mínimo e o valor máximo definido (Tarsitano and Amerise, 2017). Após, utiliza-se a técnica conhecida como *grid search*, onde considera-se todas as combinações possíveis entre os diversos valores possíveis de p e q . Em sequência, são estimados os coeficientes dos modelos e os valores dos Critérios de Informação de cada modelo, o que permite estabelecer uma hierarquia entre os modelos candidatos em função de sua parcimonia (buscar a explicação mais simples e eficaz para um fenômeno, utilizando o menor número possível de variáveis ou parâmetros).

2.2 Critérios de Informação

A informação de Kullback-Leiber, também chamada de função de Kullback-Leiber (Emiliano et al., 2010), é utilizada para comparar modelos definidos a partir de um conjunto de parâmetros $\hat{\theta}$, sendo definida para dados discretos conforme Equação (5):

$$I(g; f) = E_g \left[\ln \left(\frac{g(X)}{f(X|\hat{\theta}_k)} \right) \right], \quad (5)$$

em que: X é a variável analisada, f e g são funções de densidade de probabilidade, $\hat{\theta}_k$ é o vetor de parâmetros do modelo e E_g é a função esperança ou valor esperado.

A Equação (5) sugere que é plausível comparar dois modelos f_1 e f_2 distintos. No entanto, esses modelos irão depender da função densidade de distribuição de probabilidade g dos dados, que geralmente é desconhecida (Emiliano et al., 2010).

Considerando que a modelagem de um conjunto de dados possui um certo grau de incerteza, é necessária a utilização de alguma estratégia que permita identificar alguns modelos, tomados como candidatos, e então selecionar os mais parcimoniosos de acordo com algum tipo de critério.

Segundo Emiliano et al. (2010), uma forma de comparar n modelos f_i definidos em termos de parâmetros $\hat{\theta}_k$: $f_1(X|\hat{\theta}_1)$, $f_2(X|\hat{\theta}_2)$, $\dots$, $f_n(X|\hat{\theta}_n)$ para um determinado conjunto de dados discretos consiste comparar as magnitudes da função suporte maximizada $L(\hat{\theta}_k)$ (função de máxima verossimilhança). Segundo o autor, tal método não fornece uma verdadeira comparação, haja visto que não se conhece o verdadeiro mecanismo gerador de dados, sugerindo a necessidade de consideração de uma penalidade $w = 2b(g)$.

O termo $2b(g)$ surge a partir de uma função empírica g e algumas estimativas e condições pré-estabelecidas, e serve principalmente para corrigir a função suporte da distribuição desconhecida g . Isso sugere que não existe uma forma única para a estrutura da penalidade desses critérios.

Desenvolvendo a Equação (5), obtém-se:

$$I(g; f) = E_g\{-\ln(f(X|\hat{\theta}))\} + E_g\{\ln(g(X))\}. \quad (6)$$

Seja um modelo candidato representado por uma classe paramétrica de funções de densidade dada por $f(X|\hat{\theta})$ e que se deseje identificar o modelo que fornece a melhor aproximação para o desconhecido mecanismo gerador dos dados.

Pode-se definir a função de verossimilhança para os valores de $\hat{\theta}$ conforme Equação (7):

$$L(\hat{\theta}) = f(X|\hat{\theta}). \quad (7)$$

Logo, a Equação (6) pode ser escrita como:

$$I(g; f) = E_g\{-\ln(L(\hat{\theta}))\} + E_g\{\ln(g(X))\}. \quad (8)$$

Multiplicando essa última equação por 2 tem-se uma equação equivalente, ou seja:

$$2I(g; f) = E_g\{-2\ln(L(\hat{\theta}))\} + E_g\{2\ln(g(X))\}. \quad (9)$$

A Equação 9 revela que qualquer classificação de modelos candidatos baseada em $I(g, f)$ é uma classificação que depende unicamente de $E_g\{-2\ln(L(\hat{\theta}))\}$, já que $g(X)$ é desconhecida. Nesse sentido, um modelo candidato com valores estimados de $\hat{\theta}$ pode ser determinado pelo desvio de Kullback-Leiber $d(\hat{\theta})$, ou seja:

$$d(\hat{\theta}) = E_g\{-2\ln(L(\hat{\theta}))\}. \quad (10)$$

Nesse caso, a Equação 9 pode ser reescrita como:

$$I(\hat{\theta}) = E_g\{-2\ln(L(\hat{\theta}))\} + E_g\{2\ln(g(X))\}. \quad (11)$$

Konishi and Kitagawa (2008) propõem que a medida de informação descrita na Equação (11) possa ser generalizada em termos de uma soma ou integral, dependendo da natureza dos dados (discretos ou contínuos). Essa medida é composta por dois valores, sendo elas uma função suporte, que representa a verossimilhança dos dados, e uma penalidade, que leva em consideração algumas características dos modelos e dos dados. Para fundamentar essa abordagem os autores apresentam alguns argumentos e suposições que sustentam sua aplicação em diferentes contextos estatísticos, estabelecendo que a medida de informação $CI(X_n, g)$, ou Critério de Informação, é representada como a diferença entre o valor da função de verossimilhança e uma penalidade $2b(g)$. Sua estrutura é descrita pela Equação (12):

$$CI(X_n, g) = -2\ln(L(\hat{\theta})) + 2b(g), \quad (12)$$

em que: $\ln(L(\hat{\theta}))$ é o logaritmo neperiano da função verossimilhança, X_n é o conjunto de dados e $2b(g)$ é a penalidade dada a partir de uma função g , empírica, baseada nos dados do modelo escolhido. Segundo Konishi and Kitagawa (2008), a penalidade $b(g)$ pode ter várias formas dependendo da relação entre a verdadeira distribuição que gera os dados e o modelo especificado, bem como o método utilizado para construir um modelo estatístico.

2.2.1 Critério de Informação de Akaike AIC

Um dos Critérios de Informação utilizado para avaliar a parcimonia de um modelo é o Critério de Informação de Akaike ou AIC. Ele foi desenvolvido por Hirotugu Akaike em 1974, e tem como objetivo encontrar o modelo que melhor se ajusta aos dados, mas ao mesmo tempo, evite o sobreajuste, que é quando um modelo é muito complexo e se ajusta muito bem aos dados de treinamento, mas não se ajusta tão bem aos dados futuros ou desconhecidos (James et al., 2013).

De acordo com James et al. (2013), um dos principais benefícios de seu uso é que ele permite a comparação entre modelos com diferentes números de parâmetros, apresentando melhor desempenho sobre outras métricas, como R^2 (coeficiente de determinação). Segundo Burnham (1998), o AIC é amplamente utilizado em modelos lineares e não lineares, bem como em modelos de séries temporais. No entanto, para Bozdogan (2000), o AIC não leva em conta a complexidade dos dados e pode ser influenciado pelo tamanho amostral, razão pela qual favorece modelos com maior número de parâmetros quando comparado com os resultados de outros Critérios de Informação.

Segundo Konishi and Kitagawa (2008) o AIC não requer nenhuma derivação analítica da correção do valor da penalidade e não depende da distribuição de probabilidade desconhecida g . Assim sendo, a penalidade associada ao $\ln(L(\hat{\theta}))$ do modelo baseado no método de máxima verossimilhança pode ser aproximado pelo número de parâmetros. Nesse sentido, Akaike (1974) propõe que o Critério de Informação AIC seja dado pela Equação (13):

$$AIC = -2\ln(L(\hat{\theta})) + 2k, \quad (13)$$

impondo para o critério que $2b(g) = 2k$, onde k é o número de parâmetros atribuído ao modelo.

2.2.2 Critério de Informação de Akaike corrigido AICc

O critério AICc é uma variação do AIC que leva em conta o tamanho da amostra ao calcular a seleção de modelos (Burnham, 1998). Ele foi desenvolvido primeiramente por Sugiura (Sugiura, 1978) como uma variação do AIC, e chamado de c-AIC, e futuramente melhorado por Hurvich and Tsai (1989), sendo chamado de AICc.

O critério AICc é calculado a partir da Equação (14):

$$AICc = -2\ln(L(\hat{\theta})) + 2k + \frac{2k(k+1)}{n-k-1}, \quad (14)$$

onde k é o número de parâmetros do modelo e n é o tamanho da amostra.

Assim como o AIC, o AICc é utilizado para selecionar entre diferentes modelos. Ele tem como objetivo encontrar o modelo que melhor se ajusta aos dados, mas ao mesmo tempo, evita o sobreajuste. Entretanto, ele leva em conta o tamanho da amostra, o que não ocorre com o AIC. Nesse sentido, comparando o critério AIC com o AICc, o critério AICc é muito mais conservador do que que AIC, pois o mesmo penaliza modelos com mais parâmetros de forma mais intensa que do que AIC (Burnham, 1998).

Para utilizar o AICc é preciso calcular o valor para cada modelo e compará-los, assim como ocorre com AIC, escolhe-se o modelo com o menor valor de AICc.

2.2.3 Critério de Informação Bayesiano de Schwartz BIC

Um outro Critério de Informação bastante importante na literatura é o chamado critério BIC. Esse critério é baseado na teoria Bayesiana (Esteves and Stern, 2017) e leva em conta a qualidade do ajuste do modelo aos dados, o número de parâmetros do modelo e o tamanho da amostra.

A Equação (15) define a forma como esse critério penaliza um determinado modelo, assim:

$$BIC = -2\ln(L(\hat{\theta})) + k\ln(n), \quad (15)$$

em que: $\ln(L(\hat{\theta}))$ é o logaritmo neperiano da função de verossimilhança $L(\hat{\theta})$, k é o número de parâmetros do modelo e n o tamanho da série a ser modelada.

A ideia por trás do BIC é que um modelo com muitos parâmetros e poucos dados tem uma verossimilhança menor do que um modelo com menos parâmetros e mais dados. Tal fato instiga a ideia de que ele penaliza modelos com muitos parâmetros de forma mais intensa do que os critérios anteriores (Burnham, 1998). Essa característica do BIC é simples de ser visualizada quando compara-se a sua equação com a equação do critério AIC, por exemplo, dado que o BIC incorpora um termo de penalidade maior do que o AIC que é proporcional ao número de parâmetros do modelo e à quantidade de dados. Essa característica acaba refletindo uma preferência mais forte por modelos mais simples, devido ao princípio da parcimônia, que sugere que modelos mais

simples são preferíveis.

2.2.4 AIC $\times$ AICc $\times$ BIC

É interessante observar que os três critérios citados anteriormente possuem dependência da função de máxima verossimilhança, e que os critérios AICc e BIC podem ser escritos em termos do critério AIC, conforme as Equações (16) e (17), respectivamente.

$$AICc = AIC + 2\frac{k(k+1)}{n-k-1}, \quad (16)$$

$$BIC = AIC + k(\ln(n) - 2k). \quad (17)$$

Essas expressões são interessantes pois, se houver necessidade de calcular os três critérios, se determina somente uma vez o valor de $\ln(L(\theta))$ a partir de AIC.

As penalidades $w = 2b(g)$ atribuídas aos critérios AIC, AICc e BIC são representadas pelas Equações (18), (19) e (20), respectivamente.

$$w = 2k, \quad (18)$$

$$w = 2k + \frac{2k(k+1)}{n-k-1}, \quad (19)$$

$$w = k\ln(n), \quad (20)$$

onde a penalidade w é uma função apenas de k ou de k e n . Alguns autores como James et al. (2013), também colocam o número de parâmetros como um indicativo da complexidade de um modelo, já que aparentemente quanto mais parâmetros um modelo possui, mais ajustado aos dados ele se torna. Entretanto, fica mais sujeito ao chamado *overfitting* (sobreajuste), não se adaptando, por exemplo, a novos dados.

2.3 Trabalhos relacionados

Para completar o embasamento teórico fez-se uma busca a partir da base de dados dos periódicos *scopus* e *web of Science*. Nessas bases de dados foram utilizados algumas Queries, que trouxeram um número de documentos para cada combinação, conforme explicitado no Quadro 2.

Quadro 2: Queries utilizadas

Queries	Número de documentos		
	Scopus	Web of Science	Total
(information criteria) AND (simulation)	19481	15169	34650
(information criteria) AND (simulation) AND (performance)	7217	5836	13053
(information criteria) AND (simulation) AND (performance) AND (time series)	243	273	516
((information AND criteria) AND (quadric AND surfaces))	0	0	0
Total	26941	21278	50219

As Queries (*information criteria*) and (*simulation*) and (*performance*) and (*time series*) geraram um total de 516 documentos, sendo 243 para *Scopus* e 273 para *Web of Science*. A fim de filtrar os documentos que necessariamente abordariam o assunto sobre os Critérios de Informação em relação a sua acurácia, fez-se um filtro sobre esses documentos a partir da combinação *performance of information aic, aicc, bic in time series*, gerando então um total de 40 documentos. Desses 40 documentos, sendo alguns deles repetidos em ambas plataformas, fez-se a leitura dos resumos e de trabalhos futuros. Os artigos referência utilizados, retirados a partir desses filtros estão listados por autor, título e ano conforme Quadro 3, totalizando 15 documentos.

Quadro 3: Artigos selecionados

Autor	Título	Ano
Harbecke, Jens, Jonas Grunau, and Philip Samanek	Are the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) Applicable in Determining the Optimal Fit and Simplicity of Mechanistic Models?	2024
Silveira, Anderson Garcia	Estratégia para definição do número de defasagens na aplicação do teste ADF	2023
Xavier, ACF; Blain, GC; de Moraes, MVB; Sobierajski, GD	Selecting the best nonstationary Generalized Extreme Value (GEV) distribution: on the influence of different numbers of GEV-models	2019
Wu, Shaomin	A failure process model with the exponential smoothing of intensity functions	2019
Acquah, H.D.G.	Weighted average information criterion for selection of an asymmetric price relationship	2018
Kim, H; Kim, S; Shin, H; Heo, JH	Appropriate model selection methods for nonstationary generalized extreme value models	2017
Azeez, A., Obaromi, D., Odeyemi, A., Ndege, J., Muntabayi, R.	Seasonality and trend forecasting of tuberculosis prevalence data in Eastern Cape, South Africa, using a hybrid model	2016
Brewer, Mark J and Butler, Adam and Cooksley, Susan L	The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity	2016
Lu, D; Ye, M; Meyer, PD; Curtis, GP; Shi, XQ; Niu, XF; Yabusaki, SB	Effects of error covariance structure on estimation of model averaging weights and predictive performance	2013
Pappas, SS; Ekonomou, L; Karampelas, P; Karamousantas, DC; Katsikas, SK; Chatzarakis, GE; Skafidas, PD	Electricity demand load forecasting of the Hellenic power system using an ARMA model	2010
Whittaker, Tiffany A and Furlow, Carolyn F	The comparison of model selection criteria when selecting among competing hierarchical linear models	2009
Chen, P; Wu, TJ; Yang, J	A comparative study of model selection criteria for the number of signals	2008
De Gooijer, J.G.	Cross-validation criteria for Setar model selection	2001
Emiliano, PC; Vivanco, MJF; de Menezes, FS	Information criteria: How do they behave in different models?	2014
Fitzgibbon, L.J., Dowe, D.L., Vahid, F.	Minimum Message Length Autoregressive Model Order Selection	2004
Kundu, D.;Mitra,A.	Estimating the number of signals of the damped exponential models. Computational statistics & data analysis	2001

Harbecke et al. (2024) desenvolveram um estudo para verificar se as métricas utilizadas na avaliação da parcimonia de um modelo têm alguma significância dentro das explicações mecanicistas, investigando o desempenho de alguns Critérios de Informação, entre os quais AIC e BIC, que buscam um ponto de equilíbrio entre ajuste e simplicidade. Os autores concluem que esses critérios não ajudam muito, sendo necessário um tipo de abordagem diferente.

O estudo realizado por Silveira (2023) teve como objetivo identificar a melhor maneira de selecionar a quantidade ideal de defasagens a ser considerado no teste ADF, bem como determinar ideal de defasagens, considerando quatro fatores: influência da estrutura do modelo da série temporal, critério de informação utilizado para identificar o modelo mais parcimonioso, critério de determinação da defasagem máxima e nível de significância. De acordo com o autor, os melhores resultados foram encontrados quando o critério de seleção de defasagens máxima foi determinado pelo critério de defasagem l_4 e o número de defasagens foi selecionado pelo critério de informação BIC.

Xavier et al. (2019) utilizaram uma distribuição de probabilidade chamada Generalized Extreme Value (GEV), que é uma distribuição usada para modelar dados que podem apresentar valores extremos como enchentes, desastres naturais ou mercados financeiros. Nesse trabalho, os autores compararam os critérios AIC, AICc e BIC e, concluíram que AIC e AICc devem ser preferidos em relação ao BIC nesses casos.

Wu (2019) testou um método de estimação de parâmetros chamados ESI e MAI e comparou os seus desempenhos a partir de modelos conhecidos utilizando dados gerados de forma artificial e conjuntos de dados reais. Nesse artigo, utilizando os Critérios de Informação AIC, AICc e BIC, o autor mostrou que os novos modelos superam alguns dos modelos conhecidos em termos desses critérios.

Acquah (2018) avaliou o desempenho do critério de seleção de modelo (WIC) em comparação com AIC e BIC em termos de sua capacidade de recuperar o verdadeiro processo assimétrico de geração de dados. Utilizando de simulação de Monte Carlo, o autor indica que o desempenho dos métodos de seleção de modelo depende do tamanho da amostra, da diferença nos parâmetros de ajuste assimétricos e da quantidade de ruído no modelo usado na aplicação. Ele conclui que WIC supera o AIC e o BIC em condições estáveis, sugerindo que o WIC é um critério muito confiável e útil na seleção de modelo no estudo realizado.

Kim et al. (2017), a partir do estudo de dados não-estacionários utilizaram para a seleção desses modelos, em modelos simulados, o critério de informação

de Akaike (AIC), o critério de informação corrigido de Akaike (AICc), o critério de informação bayesiano (BIC) e o teste de razão de verossimilhança (LRT). Nesse estudo eles utilizaram simulações de Monte Carlo para comparar o desempenho desses critérios em relação as distribuições generalizadas extremas (GEV) estacionárias e não-estacionárias. Eles demonstraram que o BIC resultou no melhor desempenho em relação aos modelos GEV estacionários. No caso de modelos GEV não-estacionários, o AIC provou ser melhor do que os outros três métodos, quando tamanhos de amostra relativamente pequenos foram considerados. Com tamanhos de amostra maiores, o AIC, BIC e LRT apresentaram os melhores desempenhos para modelos GEV que possuem parâmetros de localização e/ou escala não-estacionários, respectivamente.

Azeez et al. (2016) utilizaram a combinação de dois modelos de séries temporais, sendo elas o SARIMA e o SARIMA-NNAR para analisar e prever os dados relacionados a tuberculose na África do Sul. Para as comparações utilizaram os Critérios de Informação AIC, AICc e BIC, além das métricas de erros MAE, MAPE, MSE, MPE e RMSE. Concluíram que a aplicação de modelos híbridos para prever dados de séries temporais é mais eficaz e eficiente do que um único modelo na geração de previsões para os casos de incidência de tuberculose.

Lu et al. (2013) conduziram um estudo relacionado a média de modelos para avaliar a incerteza do modelo conceitual de águas subterrâneas. Utilizaram os critérios de seleção do modelo, como AIC, AICc, BIC e KIC (Critério de Informação de Akaike, Critério de Informação de Akaike Corrigido, Critério de Informação Bayesiano e Critério de Informação de Kashyap, respectivamente). Os autores supuseram que esse método muitas vezes leva a uma situação irreal em que o melhor modelo recebe um peso de média esmagadoramente grande. Nesse estudo descobriram que esse problema foi causado pelo uso da matriz de covariância CE, de erros de medição para estimar a função de log-verossimilhança negativa comum a todos os critérios de seleção de modelo, e resolveram esse problema usando de uma matriz de covariância. Eles concluíram em primeira instância que essas correções melhoram o desempenho preditivo dos modelos individuais ao fornecer intervalos de confiança mais amplos.

Pappas et al. (2010) apresentaram um novo método para a previsão de carga de demanda de eletricidade usando a teoria de particionamento de vários modelos e compararam seu desempenho com três outras técnicas de análise de séries temporais bem estabelecidas, a saber, Critério de Informação de Akaike Corrigido (AICc), Critério de Informação de Akaike (AIC) e Critério de Informação Bayesiano de Schwarz (BIC), mostrando que os métodos apresentaram um de-

sempenho satisfatório no que se refere à previsão de carga. No entanto, MMPF foi capaz de obter um desempenho ligeiramente melhor em termos do erro percentual médio absoluto.

Whittaker and Furlow (2009) por meio de simulações compararam alguns Critérios de Informação como AIC, BIC e DIC (Deviance Information Criterion). Os autores concluíram que DIC é geralmente melhor do que AIC e BIC em selecionar modelos hierárquicos lineares, especialmente quando há muitos parâmetros e alta incerteza nos dados. Descobriram ainda que DIC é menos suscetível a selecionar modelos com parâmetros exagerados do que BIC. Eles sugerem que DIC é uma escolha preferível para seleção de modelos em situações hierárquicas lineares.

Chen et al. (2008) compararam em seu estudo o desempenho de seis Critérios de Informação quando aplicados ao problema do número de sinais no método de classificação de sinais múltiplos (MUSIC), sendo esses o Critério AIC, o Critério de Hannan e Quinn, o BIC, o AICc, de Kullback Vetorial Corrigido (KICvc) e o Critério de Informação Média Ponderada (WIC) utilizando vários conjuntos de dados simulados. Os autores mostraram que no modelo de ruído branco aditivo do MUSIC, para um tamanho de amostra pequeno n , o WIC apresenta desempenho quase tão bom quanto o AICc e supera outros critérios, e para um n moderadamente grande a grande, o WIC apresenta desempenho quase tão bom quanto o BIC e supera outros critérios.

Fitzgibbon et al. (2004) derivaram um estimador de comprimento mínimo de mensagem (MML) para modelos autorregressivos estacionários e não estacionários. O desempenho de seleção do modelo do estimador MML é comparado empiricamente com os métodos AIC, AICc, BIC e HQ em um experimento de Monte Carlo, com amostragem uniforme da região de estacionariedade autorregressiva. Os autores mostraram a partir de resultados experimentais que o estimador MML apresenta a melhor média geral de erro quadrático de predição e a melhor capacidade de escolher a verdadeira ordem do modelo quando comparados com os métodos escolhidos.

Kundu and Mitra (2001) consideraram o problema de estimar o número de sinais dos modelos exponenciais amortecidos usando diferentes critérios de teoria da informação para detectar o número de sinais a fim de comparar o seu desempenho em amostras pequenas por meio de simulações de Monte Carlo. Apesar de nesse artigo eles não terem considerado os três critérios AIC, AICc e BIC em suas simulações, o critério AIC foi utilizado para a detecção. Os autores concluíram que de forma geral o critério AIC não forneceu estimativas

consistentes em suas simulações.

Nessas comparações uma das formas pelas quais se pode comparar esses modelos seria a partir de modelos pré-selecionados em um estudo de simulação, como modelos do tipo *ARIMA* como realizado pelos autores Emiliano et al. (2010), ou de modelos hierárquicos lineares como realizado pelos autores Whittaker and Furlow (2009).

No trabalho dos autores Emiliano et al. (2010), em uma de suas vertentes, realizaram um estudo detalhado sobre o comportamento desses critérios com relação aos diferentes modelos *ARIMA* a partir de simulações. Segundo esses mesmos autores, o critério BIC foi superior para modelos com n grande e similar aos critérios AIC e AICc no caso de amostras com n pequeno. Esses autores sugerem o uso do critério BIC em modelagem de séries temporais.

Ludden et al. (1994) compararam os critérios AIC, AICc e BIC em um estudo de caso, concluindo que o critério BIC é aparentemente mais consistente no que diz respeito a seleção de modelos.

Em contrapartida, Brewer et al. (2016) sugerem que a melhor forma de realizar a seleção de modelos baseados em AIC e BIC é em função de seu desempenho alcançado, especialmente quando o objetivo é realizar previsão. Em contraste, Murtaugh (2009) conclui, após estudos de simulação, que a escolha da abordagem de seleção de modelo parece ter pouca influência nos resultados, seja para identificação de variáveis importantes, seja para fazer previsão.

Dziak et al. (2020) supuseram que, nesses Critérios de Informação (AIC, AICc e BIC), o peso da penalidade interfere na parcimônia do modelo a ser utilizado para previsão. Os autores determinaram valores para essa penalidade para alguns valores de n , conforme observado na Tabela 1. Eles afirmam que em amostras com n pequeno, o desempenho dos Critérios de Informação AIC ou AICc é melhor, mas para n muito grande, consideram o Critério de Informação BIC como o melhor para identificar o modelo mais parcimonioso.

Tabela 1: Resultados da penalidade de alguns Critérios de Informação

n	AIC	ABIC	BIC	CAIC
10	2	-0,6931	2,3026	3,3026
50	2	0,7732	3,912	4,912
100	2	1,4469	4,6052	506052
500	2	3,0405	6,2146	7,2146
1000	2	3,7317	6,9078	7,9078
5000	2	5,3395	8,5172	9,5172
10000	2	6,0325	9,2103	10,2103
100000	2	8,3349	11,5129	12,5129

Fonte:Dziak et al. (2020)

Portanto, de acordo com esses estudos, a acurácia do Critério de Informação BIC foi ligeiramente maior, em geral, que a dos Critérios de Informação AICc e AIC, o que pode estar relacionado ao tamanho da série utilizada na modelagem.

Em muitos desses trabalhos foram feitas comparações entre diferentes critérios de seleção de modelos para determinadas situações. Nesse processo, os autores analisaram em quais cenários cada critério apresentou melhor resultado. No entanto, uma possível interpretação implícita poderia levar a uma segunda leitura, ou seja, sugerindo que os critérios podem ter um desempenho diferente dependendo do tipo de modelo utilizado.

Apesar da interpretação de que os critérios se saem melhor em determinadas situações não estar incorreta, pode-se pensar que o que se está a tentar realizar é a adaptação de um determinado critério de informação para um determinado tipo de modelo. Nesse sentido, se um determinado critério se sai melhor para uma determinada situação em relação a outros que foram utilizados para comparação, pode-se dizer que este critério estaria mais adaptado àquela situação do que os outros que não se saíram melhor. Os autores Emiliano et al. (2010) ou os autores Whittaker and Furlow (2009), apesar de não deixarem explícitos essa interpretação, em seus trabalhos fazem exatamente isso, ou seja, adaptam um determinado critério a um ou mais modelos distintos, já que estão a determinar qual dos critérios obtém o maior número de acertos para esses modelos em situações simuladas. Para o caso do trabalho dos autores Emiliano et al. (2010), poder-se-ia concluir que, por exemplo, BIC é melhor adaptado a modelos que envolvem simulações do tipo ARIMA, ou situações que envolvam séries temporais. No entanto, é plausível tentar encontrar um determinado tipo de critério que permita sua adaptação para modelos de modo que a estrutura desse critério pudesse ser escrita em uma forma generalizada.

Em relação aos critérios de informação, ainda não há um consenso sobre um critério universal que possa ser utilizado em todas as situações possíveis. Além disso, não há uma formulação única para sua construção, o que torna sua escolha ainda mais desafiadora. No entanto, existem diversos critérios de informação disponíveis, como o AIC (Akaike Information Criterion), o AICc (Akaike corrigido) e o BIC (Bayesian Information Criterion), que são amplamente utilizados na modelagem estatística para selecionar o melhor modelo dentre várias opções, assim como para tentar determinar a eficácia de novos critérios (Acquah, 2018).

Embora a criação de critérios para avaliar a qualidade dos modelos seja uma tarefa complexa, eles são fundamentais para a seleção dos melhores modelos em diferentes áreas.

Entretanto, esses critérios podem ser limitados em alguns casos, como quando é necessário adaptar o modelo a diferentes conjuntos de dados ou quando a complexidade do modelo é um fator crítico a ser considerado. Nesses casos, a construção de um critério mais flexível pode se tornar uma solução mais viável.

2.4 Estudo das cônicas e quádricas

A seguir, será abordado o estudo das cônicas e quádricas, que são duas classes de figuras geométricas descritas por equações algébricas de segundo grau. As cônicas são curvas planas que podem ser representadas por uma equação quadrática no plano cartesiano, enquanto as quádricas correspondem a superfícies no espaço tridimensional com equações similares.

2.4.1 Cônicas

Segundo Steinbruch and Winterle (1987), chama-se Curva Cônica a todo conjunto de pontos do plano cujas coordenadas k_1 e n_1 , em relação à base canônica, satisfazem a equação do segundo grau representada aqui pela Equação (21).

$$Ak_1^2 + Bn_1^2 + 2Ck_1n_1 + Dk_1 + En_1 + F = 0, \quad (21)$$

em que A, B e C não são todos nulos.

Um dos propósitos dessa generalização no estudo das cônicas está em fazer o reconhecimento da mesma. Segundo Steinbruch and Winterle (1987), isto se dá em basicamente duas etapas, onde a primeira se refere a uma rotação dos eixos coordenados, sendo essa dividida em três passos que são:

1º passo: Eliminação do termo em kn ;

2º passo: Cálculo dos autovalores λ_1, λ_2 e respectivos autovetores $\vec{v}_1, \vec{v}_2$ da matriz M_C dos coeficientes;

3º passo: Substituição da nova equação na forma quadrática a partir desses autovalores e autovetores e reconhecimento da cônica.

A segunda etapa se refere a translação dos eixos, que nada mais é do que simplificar a expressão a fim de que se recaia em sua forma reduzida.

No 1º passo da etapa 1, escreve-se a Equação (21) em forma matricial, assim:

$$\begin{pmatrix} k_1 & n_1 \end{pmatrix} \begin{pmatrix} A & C \\ C & B \end{pmatrix} \begin{pmatrix} k_1 \\ n_1 \end{pmatrix} + \begin{pmatrix} D & E \end{pmatrix} \begin{pmatrix} k_1 \\ n_1 \end{pmatrix} + F = 0. \quad (22)$$

Sendo a matriz M_C dos coeficientes dada por:

$$M_C = \begin{pmatrix} A & C \\ C & B \end{pmatrix}.$$

No 2º passo da etapa 1, determinam os autovalores da matriz M_C (λ_1, λ_2) e seus respectivos autovetores unitários, sendo eles:

$$\vec{u}_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}. \quad (23)$$

No 3º passo da etapa 1, substitui-se a Equação (22) pela Equação (24):

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + F = 0. \quad (24)$$

A Equação (24) se refere à cônica escrita na base dos autovetores enquanto a Equação (22) está escrita na forma canônica.

Na segunda etapa pode ser realizado o Completamento de Quadrados afim de obter a Curva cônica no sistema xOy , podendo se chegar aos seguintes casos gerais:

- Parábola, cujas equações reduzidas apresentam as formas:

$$(x - h)^2 = 2p_v(y - m) \quad (25)$$

ou

$$(y - m)^2 = 2p_v(x - h); \quad (26)$$

- Hipérbole, cujas equações reduzidas têm as formas:

$$\frac{(x - h)^2}{a^2} - \frac{(y - m)^2}{b^2} = 1 \quad (27)$$

ou

$$\frac{(y - m)^2}{b^2} - \frac{(x - h)^2}{a^2} = 1; \quad (28)$$

- Elipse, cuja equação reduzida tem a forma:

$$\frac{(x - h)^2}{a^2} + \frac{(y - m)^2}{b^2} = 1, \quad (29)$$

onde a , b , h , m e p_v (distância entre o vértice e a reta diretriz da parábola) são constantes.

Em termos dos autovalores, se ambos tiverem o mesmo sinal, caracterizaria uma elipse, sinais diferentes uma hipérbole e quando um deles for nulo caracteriza uma parábola (Steinbruch and Winterle, 1987).

2.4.2 Quádricas

Segundo Steinbruch and Winterle (1987), chama-se Superfície Quádrica a todo conjunto de pontos do espaço tridimensional de coordenadas k , n e w em relação à base canônica, satisfazendo a equação do segundo grau representada aqui pela Equação (30).

$$Ak_1^2 + Bn_1^2 + Cw^2 + 2Dk_1n_1 + 2Ek_1w + 2Fn_1w + Mk_1 + Nn_1 + Pw + Q = 0, \quad (30)$$

em que A, B, C, D, E, F, M, N e P não são todos nulos.

Em forma matricial, a Equação (30) pode ser escrita como segue:

$$\begin{pmatrix} k_1 & n_1 & w \end{pmatrix} \begin{pmatrix} A & D & E \\ D & B & F \\ E & F & C \end{pmatrix} \begin{pmatrix} k_1 \\ n_1 \\ w \end{pmatrix} + \begin{pmatrix} M & N & P \end{pmatrix} \begin{pmatrix} k_1 \\ n_1 \\ w \end{pmatrix} + Q = 0. \quad (31)$$

Sendo a matriz M_A dos coeficientes dadas por:

$$M_A = \begin{pmatrix} A & D & E \\ D & B & F \\ E & F & C \end{pmatrix}.$$

Assim, como no estudo das Curvas Cônicas, é interessante identificar o tipo de Superfície Quádrica a partir da transformação dos eixos. Nesse caso, a identificação é feita da mesma forma que nas Cônicas, com a diferença que agora a matriz dos coeficientes M_c é do tipo 3×3 ao invés de 2×2 . Os tipos de quádricas que podem ser obtidas a partir do processo de Completamento de Quadrados são:

- Elipsoide, cuja equação reduzida tem a forma:

$$\frac{(x-h)^2}{a^2} + \frac{(y-m)^2}{b^2} + \frac{(z-l)^2}{c^2} = 1; \quad (32)$$

- Cone, cujas equações reduzidas têm as formas:

$$\frac{(x-h)^2}{a^2} = \frac{(y-m)^2}{b^2} + \frac{(z-l)^2}{c^2} \quad (33)$$

ou

$$\frac{(y-m)^2}{b^2} = \frac{(x-h)^2}{a^2} + \frac{(z-l)^2}{c^2} \quad (34)$$

ou

$$\frac{(z-l)^2}{c^2} = \frac{(y-m)^2}{b^2} + \frac{(x-h)^2}{a^2}; \quad (35)$$

- Parabolóide Elíptico, cujas equações reduzidas têm as formas:

$$\frac{(x-h)}{a} = \frac{(y-m)^2}{b^2} + \frac{(z-l)^2}{c^2} \quad (36)$$

ou

$$\frac{(y-m)}{a} = \frac{(x-h)^2}{b^2} + \frac{(z-l)^2}{c^2} \quad (37)$$

ou

$$\frac{(z-l)}{c} = \frac{(y-m)^2}{b^2} + \frac{(x-h)^2}{a^2}; \quad (38)$$

- Hiperboloide Elíptico (ou Hiperboloide de uma folha), cujas equações reduzidas têm as formas:

$$\frac{(x-h)^2}{a^2} + \frac{(y-m)^2}{b^2} - \frac{(z-l)^2}{c^2} = 1 \quad (39)$$

ou

$$\frac{(y-m)^2}{b^2} + \frac{(z-l)^2}{c^2} - \frac{(x-h)^2}{a^2} = 1 \quad (40)$$

ou

$$\frac{(x-h)^2}{a^2} + \frac{(z-l)^2}{c^2} - \frac{(y-m)^2}{b^2} = 1 \quad (41)$$

- Paraboloide Hiperbólico, cujas equações reduzidas têm as formas:

$$\frac{(x-h)^2}{a^2} = \frac{(y-m)^2}{b^2} - \frac{(z-l)^2}{c^2} \quad (42)$$

ou

$$\frac{(y-m)^2}{b^2} = \frac{(x-h)^2}{a^2} - \frac{(z-l)^2}{c^2} \quad (43)$$

ou

$$\frac{(z-l)^2}{c^2} = \frac{(y-m)^2}{b^2} - \frac{(x-h)^2}{a^2}; \quad (44)$$

- Hiperboloide de duas folhas, cujas equações reduzidas têm as formas:

$$\frac{(x-h)^2}{a^2} - \frac{(y-m)^2}{b^2} - \frac{(z-l)^2}{c^2} = 1 \quad (45)$$

ou

$$-\frac{(x-h)^2}{a^2} - \frac{(y-m)^2}{b^2} + \frac{(z-l)^2}{c^2} = 1 \quad (46)$$

ou

$$-\frac{(x-h)^2}{a^2} + \frac{(y-m)^2}{b^2} - \frac{(z-l)^2}{c^2} = 1. \quad (47)$$

A Figura 4 mostra algumas formas gráficas dessas superfícies tomando h, m e l como tendo valor zero.

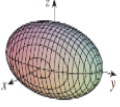
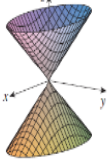
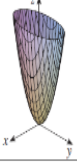
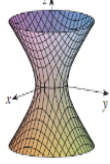
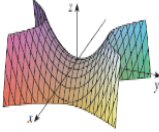
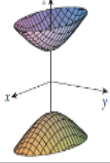
Superfície	Equação	Superfície	Equação
Elipsoide 	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ Todos os cortes são elipses. Se $a = b = c$, o elipsoide é uma esfera.	Cone 	$\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Cortes horizontais são elipses. Cortes verticais nos planos $x = k$ e $y = k$ são hipérboles se $k \neq 0$, mas são um par de retas quando $k = 0$.
Parabolóide Elíptico 	$\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Cortes horizontais são elipses. Cortes verticais são parábolas. A variável elevada à primeira potência indica o eixo do parabolóide.	Hiperbolóide de Uma Folha 	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ Cortes horizontais são elipses. Cortes verticais são hipérboles. O eixo de simetria corresponde à variável cujo coeficiente é negativo.
Parabolóide Hiperbólico 	$\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ Cortes horizontais são hipérboles. Cortes verticais são parábolas. O caso aqui ilustrado corresponde a $c < 0$.	Hiperbolóide de Duas Folhas 	$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ Cortes horizontais em $z = k$ são elipses se $k > c$ ou se $k < -c$. Cortes verticais são hipérboles. Os dois sinais de menos indicam duas folhas.

Figura 4: Comportamento Gráfico das Superfícies Quádricas em termos do eixo z (Fonte: Stewart (2013)).

Em termos dos autovalores, tem-se um elipsoide quando todos os autovalores são positivos, um hiperbolóide de uma folha quando dois forem positivos e um negativo, hiperbolóide de duas folhas quando houver um positivo e 2 negativos, parabolóide elíptico quando dois são positivos e um é nulo e um parabolóide hiperbólico quando dois autovalores tem sinais contrários somente e o outro é nulo.

3 CONSIDERAÇÕES SOBRE OS CRITÉRIOS DE INFORMAÇÃO

A seguir, serão discutidas as considerações sobre os Critérios de Informação AIC, AICc e BIC, onde será explorado a relação desses critérios com as Superfícies Quádricas.

3.1 Considerações sobre a penalidade dos critérios de informação AIC, AICc e BIC e as superfícies quádricas

Os critérios de Akaike, Akaike corrigido e Bayesiano definem a penalidade em função do número de parâmetros do modelo e/ou do tamanho da série, conforme foi observado nas Equações (13), (14) e (15).

Essas penalidades, representadas em um espaço tridimensional como função de duas variáveis $w = w(k, n)$, assumem a forma das superfícies apresentadas nas figuras 5, 6 e 7.

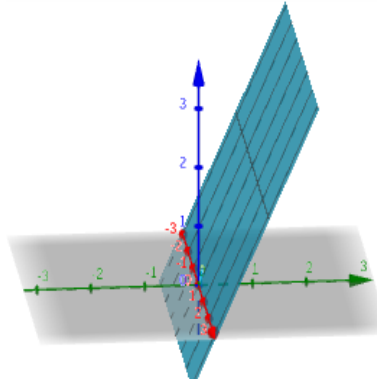


Figura 5: Representação gráfica da função $w_{AIC} = w_{AIC}(n, k) = 2k$ utilizada na determinação do Critério de Informação AIC

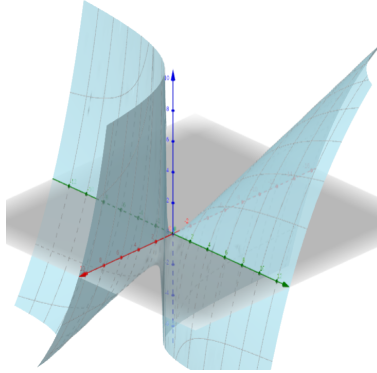


Figura 6: Representação gráfica da função $w_{AICc} = w_{AICc}(n, k) = 2k + \frac{2k(k+1)}{n-k-1}$ utilizada na determinação do Critério de Informação AICc

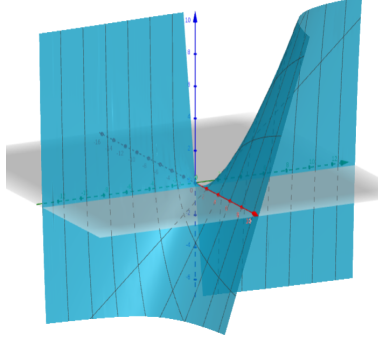


Figura 7: Representação gráfica da função $w = w_{BIC}(n, k) = k \ln(n)$ utilizada na determinação do Critério de Informação BIC

Observando-se as figuras 5, 6 e 7, constata-se que a superfície gerada pela penalidade utilizada na determinação do AIC é um plano, enquanto as superfícies geradas pelas penalidades utilizadas na determinação dos critérios AICc e BIC assemelham-se a uma superfície Quádrica, descritas pela Equação (30).

Tais superfícies podem assumir formas distintas, conforme os valores dos coeficientes A, B, C, D, E, F, M, N, P e Q. Dentre as superfícies geradas, observa-se que aquela relacionada a penalidade do Critério de Informação AICc possui coeficientes que possibilitariam, utilizando os recursos da Geometria Analítica, identificar mais precisamente o tipo de Superfície Quádrica dada pela Equação (30). O que possibilitaria motivar uma possível abordagem que permitiria efetuar uma correspondência entre os estudos relacionados da Geometria Analítica e o estudo da penalidade dos Critérios de Informação.

É importante destacar que, embora neste estudo tenha sido abordada apenas a penalidade em uma análise de superfícies, é possível considerar os próprios critérios como superfícies em termos de uma função de duas variáveis. A motivação para isso pode derivar das seguintes hipóteses. Primeiramente, critérios como AIC, AICc e BIC incluem em suas equações uma parte correspondente à função de verossimilhança e outra que depende das variáveis k e n (generalizados pela Equação (12)). Nesse contexto, podemos visualizar as superfícies da penalidade dos critérios como uma estrutura, na qual, o valor da verossimilhança é adicionado em cada ponto (k, n) dessa superfície (tomando k e n como um sistema ortogonal de eixos). Nos contextos teóricos convencionais não existem registros de estudos que analisem os critérios de informação sob uma perspectiva

geométrica relacionando-os a superfícies. Neste estudo, adotou-se uma abordagem dando destaque a penalidade (mais precisamente a sua representação geométrica) e tomando a verossimilhança como o valor a ser atribuído em cada ponto dessa superfície. Esse conceito pode ser idealizado, especialmente no que se refere à eficiência de critérios desse tipo. Por exemplo, conforme abordado em Emiliano et al. (2010), foi realizada uma análise objetiva do comportamento dos critérios AIC, AICc e BIC em relação a sua eficiência. Dado que esses critérios diferem apenas no valor da penalidade, é razoável concluir que, nesse tipo de teste, se está avaliando a estrutura da penalidade com relação a sua eficiência, uma vez que a distinção entre eles se dá exclusivamente nessa estrutura.

Para simplificar, consideremos a ideia de visualizar a superfície de uma penalidade em termos das variáveis k e n , de modo que essa penalidade seja sempre zero, independentemente dos valores de k e n . Em um espaço tridimensional, o eixo z representaria o valor do critério de informação (onde o critério de informação incluiria apenas a parte relacionada à função de verossimilhança) para cada par de pontos (k,n) , conforme ilustrado na Figura 8.

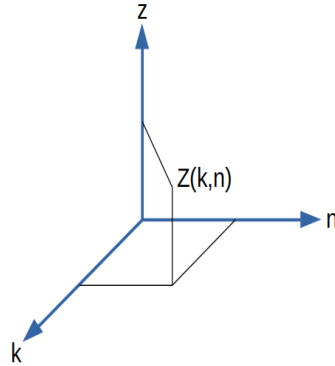


Figura 8: Representação de um ponto particular de um critério $Z=Z(k,n)$

Como os valores da penalidade são sempre nulos, os valores de $Z(k,n)$ ficam restritos ao tipo de dados disponíveis, uma vez que, nesse caso, o valor da verossimilhança está embutido no Critério de Informação Z . Quando a penalidade é $w(k,n) = 0$, é como se estivéssemos "puxando" valores a partir do plano $z = 0$ para cada ponto (k,n) considerado. No entanto, para qualquer superfície de penalidade $w = w(k,n)$, a ideia é semelhante, ou seja, temos o valor da penalidade em um dado ponto (k,n) e, a partir desse ponto, é adicionado o valor atribuído da verossimilhança, modificando assim a geometria original de sua superfície.

O que se perceberia nesses casos é que a superfície considerada da penalidade é imutável, o que possibilita seu estudo de forma isolada, como feito nesta tese.

3.2 Análise do comportamento da penalidade dos Critérios de Informação AIC, AICc e BIC em função de k e n

Nesta seção foi analisada a relação entre a penalidade e a quantidade de parâmetros k , fazendo k variar entre 0 e 50, com $n = 10$. Graficamente se percebe um comportamento linear para a penalidade dos Critérios de Informação AIC e BIC, enquanto se observa um comportamento hiperbólico para a penalidade do AICc, conforme visto na Figura 9.

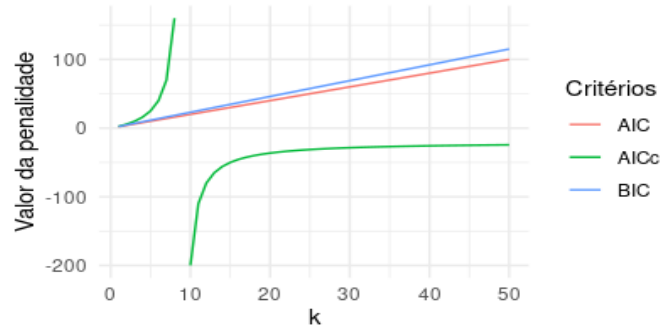


Figura 9: Comportamento gráfico da penalidade quando $n = 10$

Esses comportamentos eram esperados já que, para $n = 10$ (constante), teríamos a interseção de um plano de equação $n = 10$ interceptando as superfícies dadas pelas figuras 5, 6 e 7 da penalidade de AIC, AICc e BIC, respectivamente. No caso da penalidade de AIC, temos como interseção uma única reta, enquanto para a penalidade de AICc tem-se diferentes Hipérboles, e para a penalidade do BIC, tem-se uma coleção de retas, uma para cada valor de $\ln(n)$. Tais curvas da interseção podem ser vistas nas figuras 10, 11 e 12 para a penalidade de AIC, AICc e BIC, respectivamente, sendo $n = 2$.

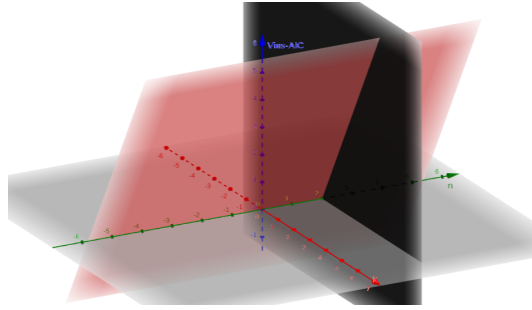


Figura 10: Reta formada pela interseção do plano $n = 2$ (destacado em preto) com o gráfico da penalidade de AIC (em vermelho)

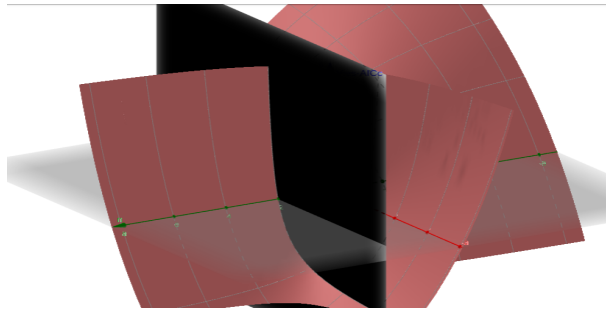


Figura 11: A Hipérbole destacada na interseção do plano $n = 2$ (em preto) com o gráfico da penalidade de AICc (em vermelho)

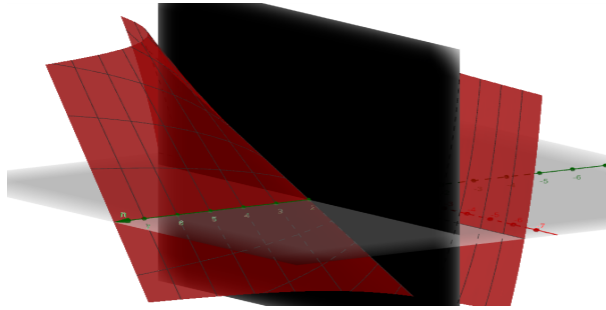


Figura 12: Reta gerada pela interseção do plano $n = 2$ (em preto) com o gráfico da penalidade de BIC (em vermelho)

Cabe ressaltar que na penalidade de BIC esta intersecção é representada por uma reta que aumenta sua declividade conforme n aumenta conforme visto na Figura 13, para $n = 10$, $n = 20$ e $n = 30$.

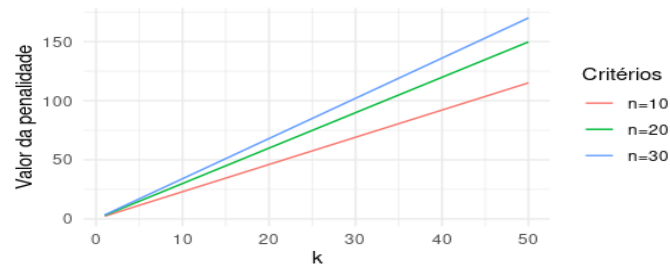


Figura 13: Comportamento da penalidade de BIC para $n = 10$, $n = 20$ e $n = 30$

Para o Critério de Informação AICc, o comportamento das Hipérboles para $n = 10$, $n = 20$ e $n = 30$ podem ser vistas na Figura 14.

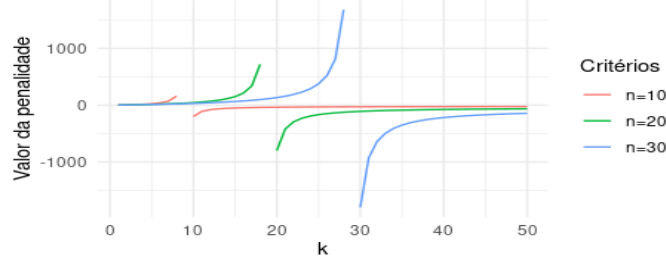


Figura 14: Comportamento da penalidade de AICc para $n = 10$, $n = 20$ e $n = 30$

É possível observar que, a partir de um determinado valor de n , a penalidade de AICc tende a ficar mais próxima a de uma linha reta, ou mais precisamente, tende a ter o comportamento da penalidade de AIC, perdendo sua característica Hiperbólica a partir de um determinado n , o que pode ser visto na Figura 15.

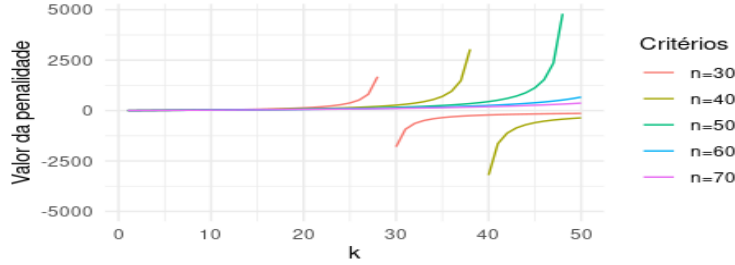


Figura 15: Comportamento da penalidade de AICc para $n = 30$, $n = 40$, $n = 50$, $n = 60$ e $n = 70$

Analisando por outro lado, se for determinada a taxa de variação de w em termos de k para a penalidade de AIC, dado pela Equação 18, verifica-se que esta taxa de variação é sempre 2, independente de n , pois:

$$\frac{\partial w_{AIC}}{\partial k} = 2. \quad (48)$$

Esse valor explica o padrão encontrado pelo autor Dziak et al. (2020) (relacionada ao critério AIC (Tabela 1)) e o padrão apresentado na coluna que mostra a penalidade de AIC na Tabela 2. No primeiro padrão, k é considerado como sendo 1, mostrando o resultado 2 se repetindo, enquanto no segundo padrão k

se movimentando dobrando o seu valor para cada situação.

Determinando a taxa de variação em termos de k para a penalidade de AICc, dado pela Equação (19), essa taxa resulta em:

$$\frac{\partial w_{AICc}}{\partial k} = \frac{2n(n-1)}{(n-k-1)^2}. \quad (49)$$

Essa equação revela como se dá o padrão encontrado na Tabela 2 (logo a seguir), relacionada a penalidade de AICc quando k se movimenta. Essa mesma equação também revela que a medida que n cresce, esta taxa tende a se estabilizar em torno de 2, igualando-se a penalidade de AIC pois:

$$\lim_{n \rightarrow \infty} \frac{\partial w_{AICc}}{\partial k} = \lim_{n \rightarrow \infty} \frac{2n^2}{n^2} = 2. \quad (50)$$

Para a penalidade do Critério de Informação BIC, dado pela Equação (20), a taxa de variação de w em termos de k fornece:

$$\frac{\partial w_{BIC}}{\partial k} = \ln(n), \quad (51)$$

que revela que a medida que n cresce, a taxa de variação do critério BIC também cresce, aumentando o valor da penalidade a medida que o valor de n também aumenta, sendo essas punições múltiplos de $\ln(n)$. Essa propriedade de BIC revela que, diferentemente dos outros critérios, a medida que n aumenta também aumenta o valor da penalidade de BIC, explicando assim o padrão encontrado na Tabela 2, relacionada a penalidade desse mesmo critério. Uma outra interpretação é que em termos da variável k , $\ln(n)$ se comporta como a variação do coeficiente angular de uma reta sem coeficiente linear.

Também é possível fazer uma análise sobre o que ocorre com as penalidades ao se variar n e fixar o valor de k . O que se verificaria é que para o Critério de Informação AIC, se veria um comportamento linear, mais especificamente, retas de equação $2k$ paralelas ao eixo n uma para cada valor de k . Para o critério BIC, verifica-se um comportamento representado por uma coleção de Hipérboles de equação $w = k \ln(n)$, uma para cada valor de k . Já para o Critério de Informação AICc, este comportamento é representado por uma família de Hipérboles de Equação $w = \frac{2kn}{n-k-1}$ uma para cada valor de k . O gráficos da Figura 16 mostram esses comportamentos para a penalidade dos critérios AIC, AICc e BIC considerando $k = 1$ e variando n . Evidenciando mais uma vez que na medida que n aumenta o critério AICc tende a ficar mais próximo ao critério AIC.

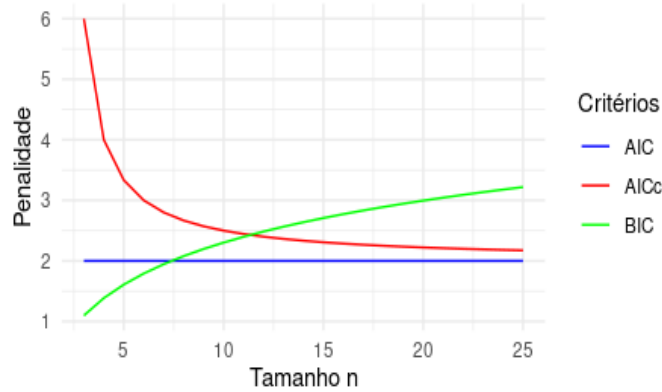


Figura 16: Comportamento da penalidade de AIC, AICc e BIC com $k = 1$

3.3 Análise do desempenho dos Critérios de Informação e sua respectiva penalidade: um estudo de caso

Em uma análise inicial avaliou-se alguns comportamentos numéricos dos Critérios de Informação AIC, AICc e BIC, e a respectiva penalidade, além de (Logaritmo Neperiano da Verossimilhança, denotado por LH).

Realizou-se a modelagem de uma série temporal com dados anuais da demanda de energia elétrica do Brasil (IPEADATA, 2021), no período de 1961 até 2021. Utilizou-se a modelagem $ARIMA(p, d, q)$ no período 1961-2014 para o ajuste, ficando o restante da série para validação. Para o período de ajuste, depois de identificar a ordem de integração necessária para se tornar estacionária, foram estimados os coeficientes de todos os modelos possíveis considerando p e q variando entre 0 e 4, o que levou a 25 modelos candidatos (Tarsitano and Amerise, 2017). Para cada um encontrou-se os valores de AIC, AICc, BIC, LH, as penalidades de AIC, AICc e de BIC, que podem ser visualizados na Tabela 2.

Tabela 2: Valores dos Critérios de Informação AIC, AICc e BIC, respectiva penalidade e LH dos dados analisados.

Modelos	Critérios e LH				Penalidade		
	AIC	AICc	BIC	LH	AIC	AICc	BIC
ARIMA(0,1,0)	-121,43	-121,35	-119,44	61,71	2	2,08	3,97
ARIMA(0,1,1)	-159,75	-159,51	-155,77	81,87	4	4,24	7,94
ARIMA(0,1,2)	-171,61	-171,13	-165,64	88,8	6	6,49	11,91
ARIMA(0,1,3)	-179,65	-178,83	-171,69	93,82	8	8,83	15,88
ARIMA(0,1,4)	-178,56	-177,31	-168,61	94,28	10	11,28	19,85
ARIMA(1,1,0)	-197,94	-197,7	-193,96	100,97	4	4,24	7,94
ARIMA(1,1,1)	-200,15	-199,67	-194,19	103,08	6	6,49	11,91
ARIMA(1,1,2)	-201,62	-200,81	-193,67	104,81	8	8,83	15,88
ARIMA(1,1,3)	-200,16	-198,91	-190,21	105,08	10	11,28	19,85
ARIMA(1,1,4)	-198,03	-196,25	-186,1	105,01	12	13,83	23,82
ARIMA(2,1,0)	-198,44	-197,96	-192,47	102,22	6	6,49	11,91
ARIMA(2,1,1)	-202,26	-201,45	-194,31	105,13	8	8,83	15,88
ARIMA(2,1,2)	-200,04	-198,79	-190,1	105,02	10	11,28	19,85
ARIMA(2,1,3)	-197,98	-196,2	-186,05	104,99	12	13,83	23,82
ARIMA(2,1,4)	-197,47	-195,03	-183,55	105,73	14	16,49	27,79
ARIMA(3,1,0)	-197,95	-197,13	-189,99	102,97	8	8,83	15,88
ARIMA(3,1,1)	-196,22	-194,97	-186,27	103,11	10	11,28	19,85
ARIMA(3,1,2)	-199,2	-197,42	-187,27	105,6	12	13,83	23,82
ARIMA(3,1,3)	-198,79	-196,35	-184,86	106,39	14	16,49	27,79
ARIMA(3,1,4)	-195,02	-191,82	-179,11	105,51	16	19,27	31,76
ARIMA(4,1,0)	-196,5	-195,25	-186,55	103,25	10	11,28	19,85
ARIMA(4,1,1)	-194,5	-192,72	-182,57	103,25	12	13,83	23,82
ARIMA(4,1,2)	-197	-194,56	-183,08	105,5	14	16,49	27,79
ARIMA(4,1,3)	-195,26	-192,06	-179,35	105,63	16	19,27	31,76
ARIMA(4,1,4)	-195,46	-191,37	-177,56	106,73	18	22,19	35,73

A Figura 17 apresenta a penalidade dos Critérios de Informação AIC, AICc e BIC, evidenciando que a penalidade do BIC é sempre maior do que as demais. Além disso, observa-se um comportamento periódico na penalidade desses três critérios, o que é esperado devido ao tamanho n da série ser fixo, variando somente a quantidade de parâmetros $p + q + 1$ do modelo ARIMA, pois, por exemplo, os modelos $ARMA(1, 2)$, $ARMA(2, 1)$, $ARMA(0, 3)$ e $ARMA(3, 0)$ terão o mesmo valor de penalidade para um dado n .

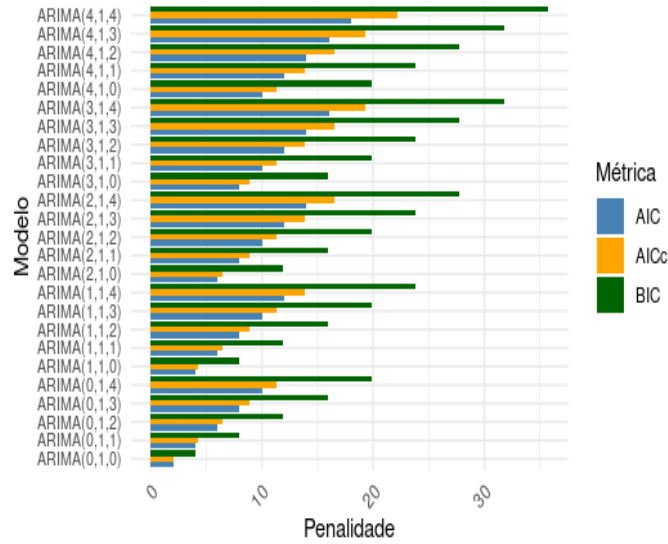


Figura 17: Resultados da penalidade encontrada nos Critérios de Informação AIC, AICc e BIC para os dados analisados

Essa característica também mostra que para modelos candidatos (utilizando um mesmo critério de informação) que possuem o mesmo número de parâmetros, o que os diferencia é somente o $\ln$ da razão ao quadrado da verossimilhança de um modelo em relação ao outro, independente do critério escolhido. Para visualizar essa característica tomou-se o critério BIC para dois modelos ARIMA distintos, mas com o mesmo número de parâmetros e mesmo n , assim:

$$BIC_1 = -2\ln(L(\hat{\theta}_1)) + k\ln(n),$$

$$BIC_2 = -2\ln(L(\hat{\theta}_2)) + k\ln(n),$$

onde: BIC_1 é o critério de um modelo 1 e BIC_2 é o critério de um modelo 2. Fazendo a diferença entre esses modelos, tem-se:

$$BIC_1 - BIC_2 = -2\ln(L(\hat{\theta}_1)) + 2\ln(L(\hat{\theta}_2)),$$

$$BIC_1 - BIC_2 = \ln \left(\left(\frac{L(\hat{\theta}_2)}{L(\hat{\theta}_1)} \right)^2 \right). \quad (52)$$

A Equação (52) revela que os valores atribuídos aos coeficientes $\hat{\theta}$ de um modelo são importantes para decidir qual o melhor modelo entre eles. Para modelos ARIMA, que possuem coeficientes Autorregressivos e de Médias Móveis, a Equação (52) revela que os valores atribuídos a esses coeficientes serão importantes para classificar entre dois modelos que possuem o mesmo número de parâmetros, já que $L(\hat{\theta})$ (verossimilhança) é determinada da mesma forma tanto para um modelo quanto para o outro.

Para os dados da demanda de energia, a Tabela 3 apresenta os valores de $\ln \left(\left(\frac{M1}{M2} \right)^2 \right)$ ($M1 = L(\hat{\theta}_1)$ e $M2 = L(\hat{\theta}_2)$) para os modelos que possuem três parâmetros. Naturalmente, observa-se que valores negativos mostram que M2 possui valor maior do que M1 para o valor da verossimilhança (validando naturalmente a sua recíproca).

Tabela 3: Valores do $\ln \left(\left(\frac{M1}{M2} \right)^2 \right)$ para os modelos da demanda de energia com três parâmetros.

M1	M2	$\ln \left(\left(\frac{M1}{M2} \right)^2 \right)$
ARIMA(0,1,3)	ARIMA(2,1,1)	-22,61
ARIMA(0,1,3)	ARIMA(3,1,0)	-18,29
ARIMA(2,1,1)	ARIMA(3,1,0)	4,32

Para avaliar a relação entre os modelos considerados os mais parcimoniosos e os demais, foi determinado um quociente entre o valor do respectivo critério de cada um, e aquele considerado como mais parcimonioso pelo critério. Os resultados obtidos estão apresentados nas figuras 18, 19 e 20, para os Critérios de Informação AIC, AICc e BIC, respectivamente.

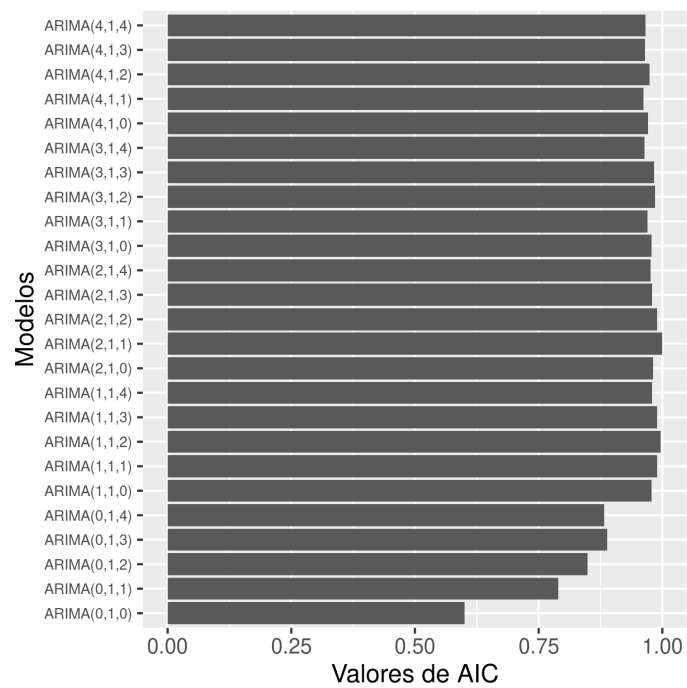


Figura 18: Comportamento da razão dos valores de AIC entre os modelos e o mais parcimonioso.

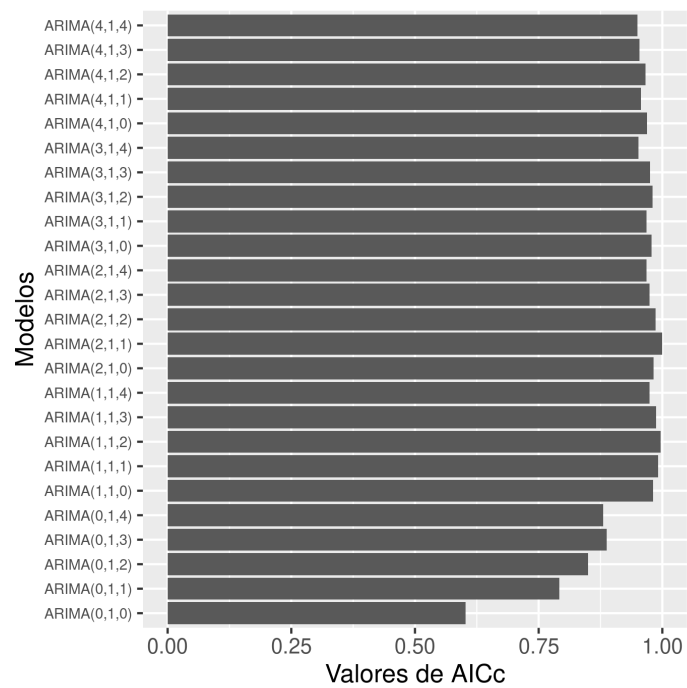


Figura 19: Comportamento da razão entre o modelo mais parcimonioso e os demais, ordenado pelo AICc para os dados analisados.

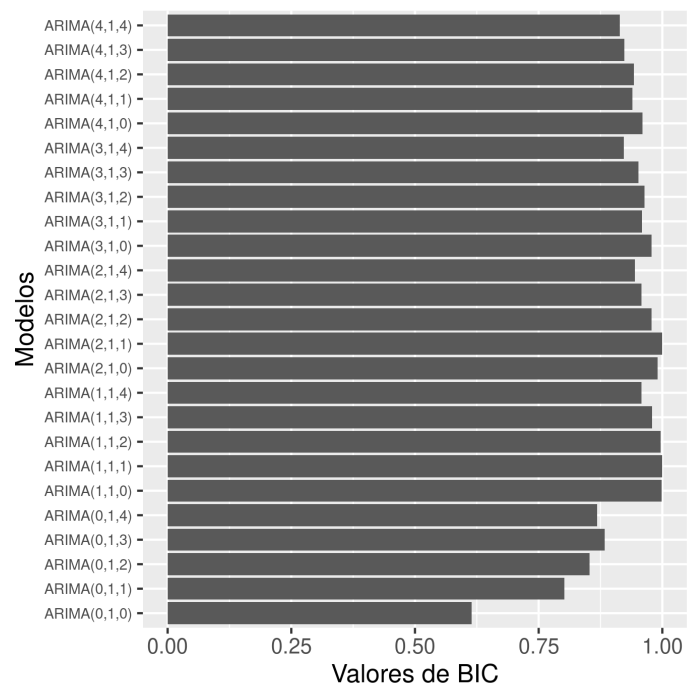


Figura 20: Comportamento da razão entre o modelo mais parcimonioso e os demais, ordenado pelo BIC para os dados analisados.

Vale ressaltar que ambos Critérios de Informação classificam como melhor modelo o modelo ARIMA(2,1,1).

Se fizermos os gráficos dos valores de $-2 * LH$ e dos critérios de informação AIC, AICc e BIC para os 25 modelos, conforme os dados encontrados na Tabela 3 tem-se a Figura 21, onde pode-se perceber a diferença entre os comportamentos desses critérios, assim como observar a proximidade entre AIC e AICc.

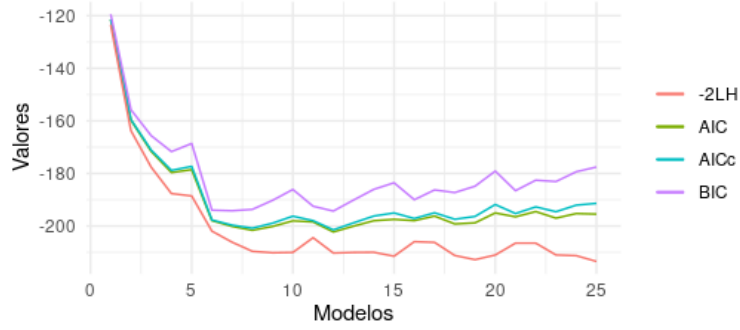


Figura 21: Comportamento de $-2*LH$, AIC, AICc e BIC para os 25 modelos candidatos

A distância entre os pares desses critérios para cada modelo candidato é dado somente entre os valores que calculam as penalidades desses critérios, sendo elas dadas em módulo, segundo as Equações (53), (54) e (55).

$$|AIC - AICc| = \left| 2k - \frac{2kn}{n - k - 1} \right|, \quad (53)$$

$$|AIC - BIC| = |2k - k \ln(n)|, \quad (54)$$

$$|AICc - BIC| = \left| \frac{2kn}{n - k - 1} - k \ln(n) \right|. \quad (55)$$

Para os dados da demanda de energia, as distâncias para o modelo mais parcimonioso, dado que o menos parcimonioso é zero, são dadas a partir das Equações (53), (54) e (55) são:

$$|AIC - AICc| = 0,48, \quad (56)$$

$$|AIC - BIC| = 5,96, \quad (57)$$

$$|AICc - BIC| = 5,98. \quad (58)$$

É interessante observar, combinando o exposto nessa seção e na anterior, que esses critérios de informação são uma adição de uma função logarítmica dada pela expressão que envolve a função de verossimilhança e uma curva cônica hiperbólica, que será mostrada em seções posteriores, sendo exceção esta ocorrendo somente com o Critério de Informação AIC.

3.4 Equivalência entre as penalidades dos Critérios de Informação AIC, AICc e BIC

Nesta seção avalia-se a possibilidade de encontrar a intersecção entre as curvas que representam a penalidade dos Critérios de Informação analisados.

Neste ponto, é interessante avaliar a curva formada pela intersecção destas penalidades, já que o par de pontos n e k não serão únicos para alguns casos. Inicialmente analisa-se a intersecção entre as curvas da penalidade do Critério de Informação AIC com a penalidade do Critério de Informação AICc, representada na Figura 22.

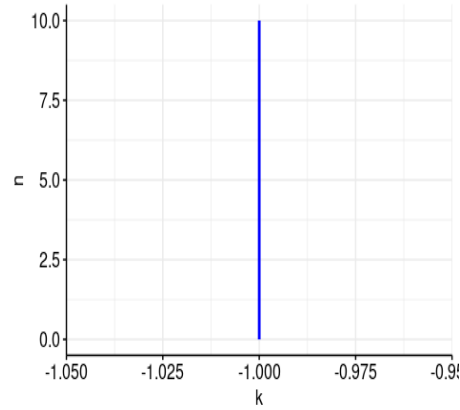


Figura 22: Comportamento da intersecção da penalidade de AIC com a penalidade de AICc

Como esperado, essa intersecção resulta em uma reta vertical de equação $k = -1$, pois:

$$2k = \frac{2kn}{n - k - 1}$$

levando a:

$$k = -1$$

Essa reta não faz sentido para o caso, pois está se supondo que k tenha que ser negativo, o que é uma impossibilidade. O outro ponto no qual as duas equações se equivalem é quando $k = 0$, o que é natural, pois é onde todas as penalidades se anulam.

O segundo caso ocorre quando se compara a penalidade de AIC com o do BIC, resultando o seguinte comportamento gráfico, conforme observado na Figura 23.

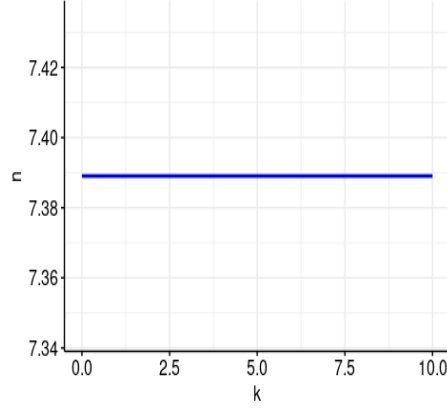


Figura 23: Comportamento da interseção da penalidade de AIC com a penalidade de BIC

Este comportamento nos diz que a penalidade do AIC e a penalidade do BIC terão valores idênticos, para qualquer valor de k , desde que n seja o valor e^2 (o que daria uma amostra em torno de $n = 7$), pois:

$$\begin{aligned} 2k &= k \ln(n), \\ \ln(n) &= 2, \\ n &= e^2 \end{aligned} \tag{59}$$

Para a penalidade de AICc e BIC, o comportamento é dado segundo o gráfico da Figura 24, que representa uma curva escrita segundo a Equação (60), ou seja:

$$(2 - \ln(n))n + \ln(n)k + \ln(n) = 0. \tag{60}$$

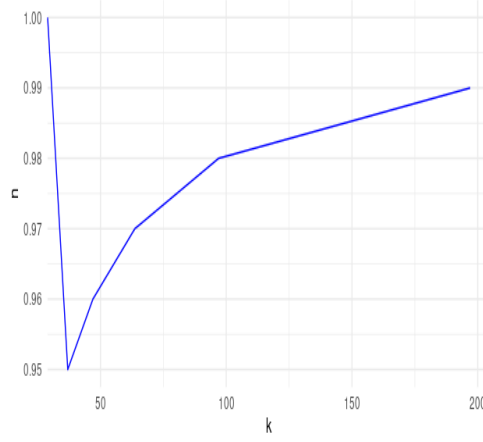


Figura 24: Comportamento da interseção da penalidade de AICc com a penalidade de BIC

O comportamento do gráfico é de uma Hipérbole, onde a parte da curva à direita não pode assumir valores negativos devido à restrição do $\ln(n)$. É interessante observar que estes dois critérios somente teriam valores equivalentes para n próximo ao valor 1 (o inteiro mais próximo), o que ocorreria para quando k fosse grande, o que não faria sentido. Quando $k = 0$ a solução da Equação (60) é dada por:

$$n^{\frac{n-1}{n}} = e^2, \quad (61)$$

que sugere novamente uma relação interessante entre o valor de (n) e a constante e^2 (que parece surgir naturalmente a partir da comparação entre esses critérios). Os valores de n , no entanto, que satisfazem essa constante teriam de ser $n \approx 9,38$ e $n \approx 0,34$, que são os valores vistos no gráfico, apesar do valor de ambas as penalidades, para $k = 0$, serem independentes de n .

3.5 Relação entre o critério de informação AICc e as curvas cônicas

Para evidenciarmos que a parcela relacionada à penalidade no AICc corresponde, na realidade, a uma superfície construída a partir de hipérboles, podemos simplificar, por conveniência, assumindo que essa penalidade, expressa pela

Equação (19), seja igual a unidade. Assim para $w_{AIC_c} = 1$ tem-se:

$$2k + \frac{2k(k+1)}{n-k-1} = 1. \quad (62)$$

Manipulando algebricamente essa equação, tem-se:

$$n - 2kn - k - 1 = 0. \quad (63)$$

A Equação (63) gera uma curva no plano que se assemelha a uma hipérbole rotacionada e transladada, conforme observado na Figura 25.

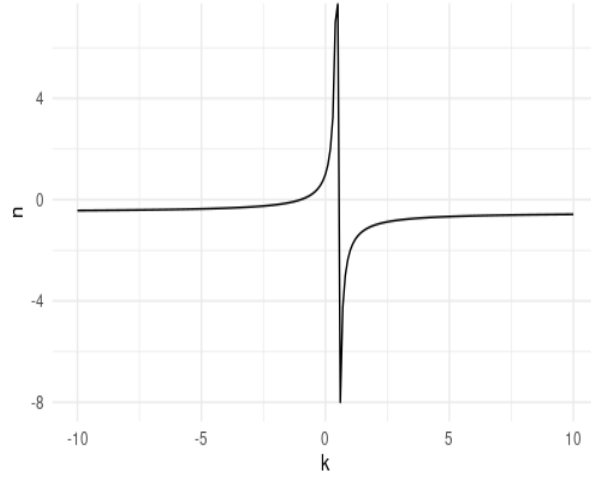


Figura 25: Gráfico da função descrita pela Equação (63)

Comparando a Equação (63) com a Equação (21), chega-se aos seguintes valores para os coeficientes da forma geral:

$$A = 0, B = 0, C = -1, D = -1, E = 1, F = -1.$$

Se for utilizada a forma matricial, dada pela Equação (22), tem-se a Equação (64):

$$\begin{pmatrix} k & n \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} k \\ n \end{pmatrix} + \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} k \\ n \end{pmatrix} - 1 = 0. \quad (64)$$

Para identificar o tipo de Cônica que descreve a penalidade desse critério de

informação, deve-se primeiro determinar os autovalores e autovetores associados da matriz M_C (Steinbruch and Winterle, 1987) e colocar a Equação (64) em sua forma canônica. Assim, tem-se:

$$M_C = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \quad (65)$$

cujos autovalores são:

$$\lambda_1 = 1, \lambda_2 = -1.$$

Os autovetores associados são:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Em sequência, são determinados os vetores unitários nas direções dos autovetores encontrados. Tem-se para tanto:

$$\vec{u}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}.$$

A matriz P_C formada pelos autovetores da matriz M_C , que deve apresentar determinante igual a 1 (Steinbruch and Winterle (1987)), é:

$$P_C = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}. \quad (66)$$

Sua transformação por rotação da Cônica é dada por:

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -1 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - 1 = 0, \quad (67)$$

que assume a forma reduzida apresentada na Equação (68):

$$\frac{x^2}{\frac{2}{4}} - \frac{(y - \frac{\sqrt{2}}{2})^2}{\frac{2}{4}} = 1. \quad (68)$$

Essa equação mostra que a curva descrita pela Equação (63) é de fato uma Hipérbole gerada a partir da Equação (68) por uma rotação e translação de eixos, conforme evidenciado também na Figura 26.

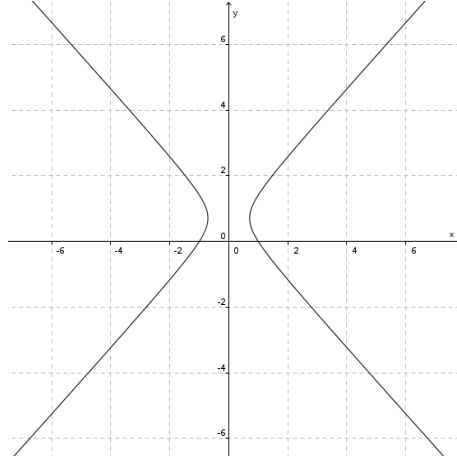


Figura 26: Gráfico da função descrita pela Equação (68)

Considerando que a Hipérbole possui a equação reduzida descrita pela Equação (27), seus coeficientes assumem os valores $a = \frac{\sqrt{2}}{2}$, $b = \frac{\sqrt{2}}{2}$, $h = 0$ e $m = \frac{\sqrt{2}}{2}$.

Uma propriedade importante na identificação de uma Hipérbole é a sua excentricidade (LEHMAN, 1980), dada pela razão $e_x = c/a$, sendo $c = \sqrt{a^2 + b^2}$. Nesse caso analisado, encontra-se o valor $e_x = \sqrt{2}$, um número positivo maior que a unidade, confirmando assim a característica Hiperbólica da penalidade do critério de informação AICc.

Na sequência, a Equação (68) pode ser novamente transformada na Equação (63) por meio da Equação (69) (Steinbruch and Winterle, 1987), ou seja:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\hat{\varphi}) & -\sin(\hat{\varphi}) \\ \sin(\hat{\varphi}) & \cos(\hat{\varphi}) \end{pmatrix} \begin{pmatrix} k \\ n \end{pmatrix}. \quad (69)$$

Para encontrar o ângulo de rotação $\hat{\varphi}$ podemos utilizar a Equação (70) (LEHMAN, 1980), assim:

$$\tan(2\hat{\varphi}) = \frac{C}{A - B}, \quad (70)$$

sendo $\hat{\varphi} = \pi/4$ quando $A - B = 0$ e C diferente de zero.

Para o caso em que $A = 0$, $B = 0$ e $C = -1$, a rotação $\hat{\varphi}$ é de 45° , sendo que a Equação (69) se transforma na Equação (71):

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} k \\ n \end{pmatrix}. \quad (71)$$

Assim, é possível expressar x e y em termos de k e n , respectivamente pelas Equações (72) e (73), ou seja:

$$x = \frac{\sqrt{2}}{2}k - \frac{\sqrt{2}}{2}n \quad (72)$$

e

$$y = \frac{\sqrt{2}}{2}k + \frac{\sqrt{2}}{2}n. \quad (73)$$

Esses resultados satisfazem a Equação (63), mostrando mais uma vez a forma Cônica do tipo Hipérbole da penalidade do Critério de Informação AICc.

Se tomarmos o maior e o menor valor para o critério de informação AICc (melhor e pior modelo) para os dados da demanda de Energia e substituirmos pelo valor 1 da Equação (68), o comportamento das Hipérboles (para qualquer par (x, y) de valores terão a forma dada nas figuras 27 e 28.

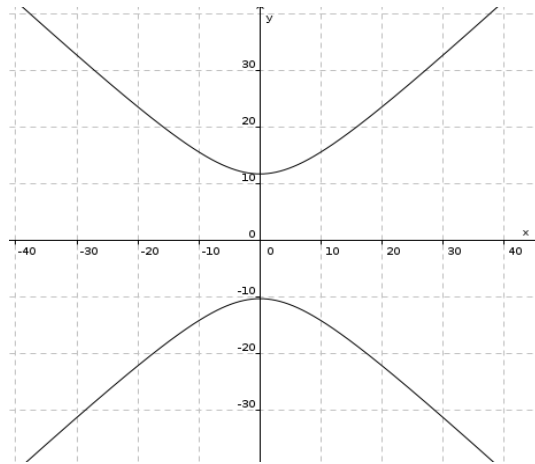


Figura 27: Comportamento da Hipérbole sobre o menor valor de AICc (melhor modelo)

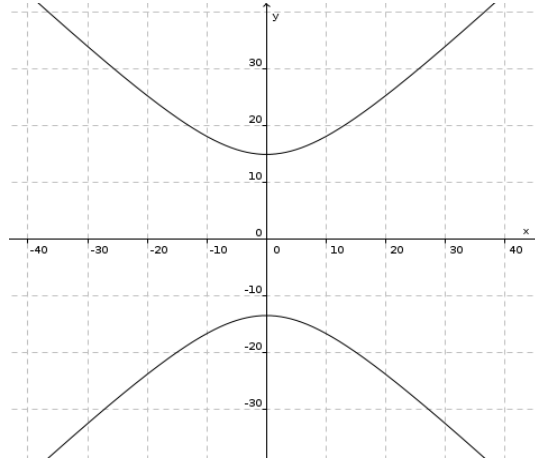


Figura 28: Comportamento da Hipérbole sobre o maior valor de AICc (pior modelo)

O que se observou nesses dois gráficos é que quanto melhor o modelo, menor será o valor do eixo real a para a Hipérbole, e quanto pior o modelo, maior será o valor do eixo real a da Hipérbole. Poderia se supor que a partir da Equação (27) seja o eixo b , mas como essas Hipérboles possuem o mesmo valor para ambos os eixos se torna indiferente a análise nesse caso.

3.6 Relação entre o Critério de Informação AICc e as superfícies quádricas

Para demonstrar-se que a parcela referente à penalidade no AICc representa, de fato, uma superfície quádrica, basta considerar que a manipulação algébrica da Equação (19) leva à Equação (74).

$$wk - wn + 2kn - w = 0. \quad (74)$$

Comparando a Equação (74) com a Equação da Quádrica, dada neste texto pela Equação (30), obtém-se os seguintes valores para os seus coeficientes:

$$A = 0, B = 0, C = 0, D = 1, E = 1/2, F = -1/2, M = 0, N = 0, P = 1, Q = 0.$$

Se for utilizada a forma matricial, conforme o apresentado na Equação (31), tem-se a Equação (75):

$$\begin{pmatrix} k & n & w \end{pmatrix} \begin{pmatrix} 0 & 1 & 1/2 \\ 1 & 0 & -1/2 \\ 1/2 & -1/2 & 0 \end{pmatrix} \begin{pmatrix} k \\ n \\ w \end{pmatrix} + \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} k \\ n \\ w \end{pmatrix} + 0 = 0. \quad (75)$$

Para identificar o tipo de superfície Quádrica que descreve a penalidade desse Critério de Informação, deve-se encontrar os autovalores e autovetores associados da matriz M_A (Steinbruch and Winterle, 1987), dada na Equação (75), assim tem-se:

$$M_A = \begin{pmatrix} 0 & 1 & 1/2 \\ 1 & 0 & -1/2 \\ 1/2 & -1/2 & 0 \end{pmatrix}, \quad (76)$$

cujos autovalores são:

$$\lambda_1 = 1, \lambda_2 = -\frac{\sqrt{3}-1}{2}, \lambda_3 = \frac{\sqrt{3}-1}{2}.$$

Os autovetores associados são:

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \vec{v}_2 = \begin{pmatrix} \frac{-\sqrt{3}-1}{2} \\ \frac{\sqrt{3}+1}{2} \\ 1 \end{pmatrix}, \vec{v}_3 = \begin{pmatrix} \frac{\sqrt{3}-1}{2} \\ \frac{-\sqrt{3}+1}{2} \\ 1 \end{pmatrix}.$$

Em sequência, são determinados vetores unitários desses autovetores, assim tem-se:

$$\vec{u}_1 = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} \frac{-\sqrt{3}-1}{2\sqrt{3+\sqrt{3}}} \\ \frac{\sqrt{3}+1}{2\sqrt{3+\sqrt{3}}} \\ \frac{1}{\sqrt{3+\sqrt{3}}} \end{pmatrix}, \vec{u}_3 = \begin{pmatrix} \frac{\sqrt{3}-1}{2\sqrt{3-\sqrt{3}}} \\ \frac{-\sqrt{3}+1}{2\sqrt{3-\sqrt{3}}} \\ \frac{1}{\sqrt{3-\sqrt{3}}} \end{pmatrix}.$$

A matriz P_A com os autovetores unitários que diagonalizam a matriz M_A devem apresentar determinante igual a 1 (Steinbruch and Winterle, 1987). Logo, tem-se:

$$P_A = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}-1}{2\sqrt{3+\sqrt{3}}} & \frac{\sqrt{3}-1}{2\sqrt{3-\sqrt{3}}} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}+1}{2\sqrt{3+\sqrt{3}}} & \frac{-\sqrt{3}+1}{2\sqrt{3-\sqrt{3}}} \\ 0 & \frac{1}{\sqrt{3+\sqrt{3}}} & \frac{1}{\sqrt{3-\sqrt{3}}} \end{pmatrix}. \quad (77)$$

A transformação por rotação da Quádrica fornece a seguinte expressão:

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{-\sqrt{3}-1}{2} & 0 \\ 0 & 0 & \frac{\sqrt{3}-1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \\ \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{-\sqrt{3}-1}{2\sqrt{3+\sqrt{3}}} & \frac{\sqrt{3}-1}{2\sqrt{3-\sqrt{3}}} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{3}+1}{2\sqrt{3+\sqrt{3}}} & \frac{-\sqrt{3}+1}{2\sqrt{3-\sqrt{3}}} \\ 0 & \frac{1}{\sqrt{3+\sqrt{3}}} & \frac{1}{\sqrt{3-\sqrt{3}}} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + 0 = 0,$$

que em sua forma geral pode ser reescrita como:

$$x^2 - \frac{\left(y - \frac{1}{(\sqrt{3}+1)(\sqrt{3+\sqrt{3}})}\right)^2}{\frac{1}{\frac{\sqrt{3}+1}{2}}} + \frac{\left(z + \frac{1}{(\sqrt{3}-1)(\sqrt{3-\sqrt{3}})}\right)^2}{\frac{1}{\frac{\sqrt{3}-1}{2}}} = 1. \quad (78)$$

A superfície dada pela Equação (78), representada graficamente na Figura 29, mostra que a penalidade de AICc pode ser representada por uma Superfície Quádrica do tipo Hiperboloide.

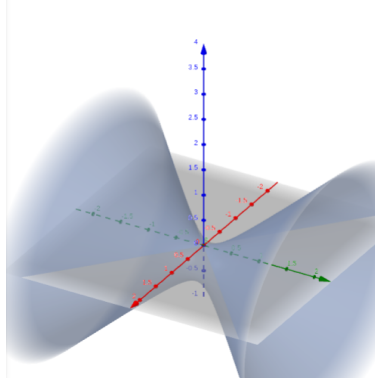


Figura 29: Gráfico da Hiperboloide obtida a partir da penalidade do Critério de Informação AICc

3.7 Construção de um Critério de Informação adaptável, a partir da equação de uma Superfície Quádrica.

A expressão da penalidade (w) de AICc, dado pela Equação (19), pode ser manipulada algebricamente transformando-se na Equação (79), apresentada a seguir:

$$w_{AICc} = \frac{2kn}{n - k - 1}. \quad (79)$$

O termo n é sempre maior do que k , já que n é o tamanho da amostra e k o número de parâmetros do modelo. Nesse sentido, o aumento do tamanho da amostra tende a diminuir o valor de w , fazendo com que, para valores de n muito altos, a penalidade de AICc se torne igual a penalidade AIC, pois:

$$\lim_{n \rightarrow \infty} w = \lim_{n \rightarrow \infty} \frac{2kn}{n - k - 1}, \quad (80)$$

$$\lim_{n \rightarrow \infty} w = \lim_{n \rightarrow \infty} \frac{n2k}{n(1 - \frac{k}{n} - \frac{1}{n})} = 2k. \quad (81)$$

Considerando que a penalidade de AICc possui características de uma Superfície Quádrica, descrita conforme Equação (30), onde o termo C é nulo, pode-se pensar em estruturar uma penalidade como sendo escrita a partir da Equação (82), ou seja:

$$w = \frac{-Ak^2 - Bn^2 - 2Dkn - Mk - Nn - Q}{2Ek + 2Fn + P}. \quad (82)$$

Nesse caso, se teria a equação de uma Superfície Quádrica, a qual é uma equação do segundo grau, expressa em termos de n e k , para representar a penalidade de um Critério de Informação. É possível que se possa estruturar uma penalidade a partir de todos os coeficientes da Equação (30). Chama-se a atenção para o fato do coeficiente C ter sido considerado nulo. Esse fato é devido a se estar estudando a penalidade w , portanto não faria sentido, inicialmente, esse coeficiente estar presente na Equação (30), visto que figuraria um termo do tipo w^2 .

Nesse estudo, a parcela da penalidade do Critério de Informação será generalizada a partir da Equação (82), sendo a penalidade de AICc um caso particular, em que: $A=0$, $B=0$, $D=-1$, $E=-1/2$, $F=1/2$, $M=0$, $N=0$, $Q=0$ e $P=-1$.

Essa generalização não prejudica a propriedade que a penalidade de AICc possui de se tornar igual a penalidade do AIC (desde que $A = 0$, $B = 0$, $D = -1$, $E = -1/2$ e $F = 1/2$, $P = -1$ e o restante nulo) quando $n \rightarrow \infty$. Nesse caso se tem uma superfície generalizada, possibilitando para cada combinação de valores de A , B , D , E , F , M , N , P e Q construir um Critério de Informação na forma:

$$Cr(A, B, D, E, F, M, N, P, Q) = -2\ln(L(\hat{\theta})) + w, \quad (83)$$

onde w é dado pela Equação (82).

Nesse sentido, o que se quer determinar é a Superfície Quádrica que apresenta o melhor desempenho na identificação do modelo mais parcimonioso, ou seja, encontrar a combinação ótima entre os coeficientes A , B , D , E , F , M , N , P e Q , adaptado ao tipo de modelagem.

A princípio, nada impede que esses coeficientes sejam funções do tipo $g_i(n, k)$, o que transformaria, ou moldaria, a Superfície Quádrica dada na Equação (82).

Logo, se estaria buscando não somente coeficientes reais, mas um conjunto de funções para a penalidade de um Critério de Informação, o que resultaria em um conjunto de situações distintas e apreciadas de acordo com cada tipo de função (exponenciais, lineares, logarítmicas, trigonométricas, etc) ou combinação entre elas.

3.8 Análise da construção da penalidade

Ao se propor a determinação da penalidade de um Critério de Informação na forma da Equação (82), é interessante que a mesma continue com as propriedades observadas nos Critérios de Informação já conhecidos.

Tomamos primeiramente a derivada da Equação (82) em relação a k , da mesma forma que foi feito com a penalidade dos Critérios AIC, AICc e BIC anteriormente. Assim, temos:

$$\begin{aligned} \frac{\partial w}{\partial k} &= \frac{(-2Ak - 2Dn - M)(2Ek + 2Fn + P)}{(2Ek + 2Fn + P)^2} \\ &\quad - \frac{2E(-Ak^2 - Bn^2 - 2Dkn - Mk - Nn - Q)}{(2Ek + 2Fn + P)^2}. \end{aligned} \quad (84)$$

Se tomarmos o limite da Equação (84) quando $n \rightarrow \infty$, teremos:

$$\lim_{n \rightarrow \infty} \frac{\partial w}{\partial k} = \frac{-2DF + EB}{2F^2}. \quad (85)$$

Para a penalidade de AICc, os coeficientes da superfície Quádrica assumem os valores $B=0$, $D=-1$, $E=-1/2$, $F=1/2$, conforme visto na seção 3.7 e o limite calculado na Equação (85) tem como resultado o valor 2, concordando com o resultado encontrado anteriormente.

Para o critério de informação AIC existem várias combinações que levam a estrutura da Equação (85) a se tornar sua penalidade. Entretanto, algumas ficam inviabilizadas por resultar em uma indeterminação matemática ou em um limite nulo. A combinação entre os valores $D=-1$, $F=1/2$, considerando os restantes nulos, levam ao valor 2 como limite. É interessante observar que a penalidade de AIC nesse caso seria escrito segundo a Equação (82) como sendo:

$$w = \frac{-2(-1)kn}{2 * \frac{1}{2}n}. \quad (86)$$

Nesse caso, a Equação (86) retorna a penalidade do Critério de Informação AIC, já que sempre haverá a simplificação dos valores de n .

Argumento semelhante feito para AIC pode ser estruturado para o BIC. Para o caso, a Equação (87) retorna a penalidade do Critério de Informação BIC.

$$w = \frac{-2(1/2 \ln(n)kn)}{2 \left(\frac{-1}{2}n\right)}. \quad (87)$$

Isto sugere, para não haver conflito com o limite da Equação (85), que $D =$

$\frac{1}{2}\ln(n)$, $F = -1/2$ e o restante dos coeficientes nulos para que o limite retorne $\ln(n)$ como esperado.

A partir do exposto, mostrou-se que é possível considerar a penalidade dos Critérios de Informação como um caso particular de uma Superfície Quádrica descrita pela Equação (82). A relação da Equação (82) para o $\cos(\psi)$ (vide Apêndice A.4), em relação ao gradiente, fica dado pela Equação (88), ou seja:

$$\cos(\psi) = \frac{\sqrt{2}}{2} \frac{((A_1 + A_3)A_2 + 2B_1(E + F))}{\sqrt{(A_1A_2 + 2EB_1)^2 + (A_3A_2 + 2FB_1)^2}}. \quad (88)$$

sendo:

$$A_1 = -2Ak - 2Dn - M,$$

$$A_2 = 2Ek + 2Fn + P,$$

$$A_3 = -2Bn - 2Dk - N,$$

$$B_1 = Ak^2 + Bn^2 + 2Dkn + Mk + Nn + Q.$$

A Equação (88) também sugere uma estrutura geométrica onde se poderia inferir o seguinte triângulo retângulo, conforme Figura 30, que possibilitaria encontrar as outras relações trigonométricas a partir desse triângulo, como por exemplo, encontrar valores que ficam entre $\pm\infty$ no caso do uso da função tangente. Nesse trabalho restringiremos ao cosseno somente.

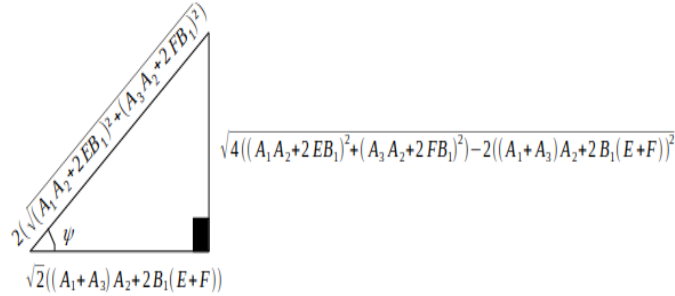


Figura 30: Triângulo Retângulo para ψ

O limite do $\cos(\psi)$ quando n se torna muito grande é dependente somente dos coeficientes da Equação (82), conforme pode ser visto na Equação (89), ou seja:

$$\cos(\psi) = \frac{\sqrt{2}}{2} \frac{2B(E - F) - 4DF}{\sqrt{(4B^2 + 16D^2)F^2 + (4B^2E - 16DFB)E}}. \quad (89)$$

Esse resultado retorna aos casos particulares vistos no Apêndice A.4 para os

critérios AIC e AICc. Por exemplo, para AIC, onde para $D = -1$ e $F = 1/2$ e o restante dos coeficientes nulos, a Equação (89) retorna para o valor $\sqrt{2}/2$. Nesse sentido, independente de qual razão trigonométrica seja utilizada, seno, cosseno, tangente, etc, a grandeza relacionada aos coeficientes dessa equação como os da Equação (88) é adimensional.

3.9 Formulação empírica

Como os Critérios de Informação AIC e AICc possuem características similares, é interessante fazer algumas análises e construções para que se possa obter subsídios para estimar os coeficientes da Superfície Quádrica que definem as estruturas da penalidade.

Essas construções foram realizadas também com o intuito de se poder resolver o problema de forma que as simulações não tivessem um peso computacional muito elevado, pois a Equação (82) possui muitos coeficientes e estes podem ser números reais. Nesse sentido, pode ser interessante já se ter alguma constante presente nas equações, como a constante 2 presente na penalidade do Critério de Informação AIC ou AICc.

Podemos supor que haja, empiricamente, uma penalidade equivalente ao do Critério AIC. Essa equação, apesar de construída, deve retornar ao critério AIC como um caso particular. Assim sendo, pode se tentar impor esse critério equivalente a partir da Equação (90), o que torna a equação da penalidade do Critério de Informação AIC como um caso particular dessa equação quando $\alpha = 1$, assim:

$$w = 2^\alpha k, \quad (90)$$

onde α define o grau de penalização dada pela penalidade $w = 2^\alpha k$

Anteriormente foi visto que os Critérios de Informação AIC e BIC se tornam equivalentes quando $n = e^2$. Podemos impor agora, de forma também empírica, baseada na ideia de $n = e^2$, conforme observado na Seção 3.4 (Equação: (59)), uma equação que encapsule duas constantes relevantes, sendo elas o 2 e o e , ao mesmo tempo em que conecta o critério equivalente $w = 2^\alpha k$ a essas constantes. Dessa maneira, pode-se construir essa relação conforme a Equação (91), logo:

$$n = \frac{2^{\alpha-1} e^2}{\gamma}. \quad (91)$$

Para o caso especial de $\alpha = 1$ e $\gamma = 1$, a Equação (91) recai na situação

de comparação da penalidade do Critério BIC com do AIC, vista anteriormente (naturalmente fazendo com que a penalidade do Critério de Informação, equivalente $w = 2^\alpha k$, recaia para a penalidade do Critério de Informação AIC). Os valores α e γ são úteis para conectar as penalidades dos Critérios de Informação sobre esses coeficientes em termos de comparação. Sendo o coeficiente γ um coeficiente real, manipula-se α de tal maneira que se mantenha n inteiro, da mesma maneira se α for real pode se manipular γ mantendo-se n inteiro (caso se queira manter esta característica).

A partir dessas análises pode-se impor que a Equação (85) tenha como referência o valor limite dado pela Equação (92), ou seja:

$$\lim_{n \rightarrow \infty} \frac{\partial w}{\partial k} = \frac{-2DF + EB}{2F^2} = 2^\alpha = \frac{2\gamma n}{e^2}, \quad (92)$$

onde α seria dado pela Equação (93), que impõe que γ deve ser um número positivo, já que n , nesse contexto, também o é.

$$\alpha = \log_2 \left(\frac{2\gamma n}{e^2} \right). \quad (93)$$

Da mesma forma que foi realizada com a penalidade do critério AIC, pode-se encontrar a equivalência do critério auxiliar $w = 2^\alpha k$ com o critério BIC em termos de α , o que ocorre a partir da relação dada por:

$$\alpha = \log_2(\ln(n)). \quad (94)$$

Nesse caso, a sua taxa de variação em relação a n é dada por:

$$\frac{d\alpha}{dn} = \frac{\log_2 e}{n \ln(n)}, \quad (95)$$

que nos revela que a medida que n aumenta, essa taxa diminui, conforme evidenciado na Figura 31.

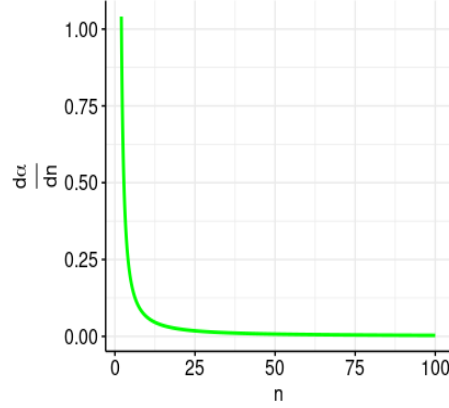


Figura 31: Comportamento gráfico para $\frac{d\alpha}{dn}$ em termos de n

Se a Equação (94) for comparada com a Equação (93), tem-se a relação de comparação com a penalidade de BIC, em termos de γ , dada pela Equação (96).

$$\gamma = \frac{e^2 \ln(n)}{2n}. \quad (96)$$

Essa equação nos retorna os valores de γ a partir de n quando se quer equiparar o critério $w = 2^\alpha$ com o critério BIC diretamente a partir da Equação (92).

A taxa de variação da Equação (96) com n é dada por:

$$\frac{d\gamma}{dn} = \frac{e^2(1 - \ln(n))}{2n^2}, \quad (97)$$

que nos revela que essa taxa tende a se anular a medida que n aumenta, conforme visualizado na Figura 32.

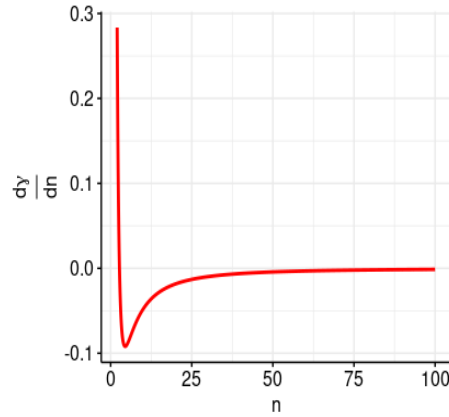


Figura 32: Comportamento gráfico para $\frac{d\gamma}{dn}$ em termos de n

Os comportamentos gráficos de α e γ em termos de n podem ser visualizados na Figura 33. O valor de α aumenta lentamente enquanto o de γ diminui lentamente conforme n aumenta, sugerindo que γ deve ser menor que 1 para n acima de 7, quando comparado a penalidade $w = 2^\alpha$ com a penalidade do critério BIC.

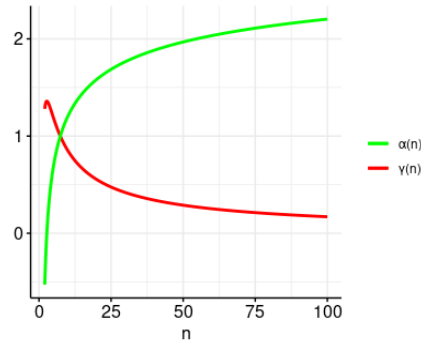


Figura 33: Comportamento gráfico para α (em verde) e γ (em vermelho) em termos de n

Essas relações serão importantes para a construção de uma estratégia para determinar os coeficientes da Superfície Quádrica a serem levados em consideração na determinação da penalidade. Além disso, algumas dessas construções empíricas servirão como base para estruturar uma metodologia que visa

a redução do número de coeficientes, assim como incorporar a presença das constantes 2 e e .

3.10 Proposta para o Critério de Informação

A Equação (82), apesar de proporcionar uma possível generalização para a penalidade de um Critério de Informação na forma de uma Superfície Quádrica, também traz consigo o problema de possuir muitos coeficientes, o que poderia acarretar um grande trabalho computacional. Para tanto, se faz necessário buscar algum tipo de estratégia que permita reduzir esse peso computacional, como a redução do número de coeficientes, tomando alguns deles nulos.

Existem possibilidades estratégicas que podem ser adotadas sobre os valores dos coeficientes e algumas restrições que devem ser obedecidas (como não se ter divisão por zero em uma penalidade). Embora cada uma dessas restrições possa ser estudada isoladamente, neste trabalho será explorado dois tipos de estruturas para a penalidade, sendo uma delas dada pela Equação (98) e a outra pela Equação (99):

$$w = \frac{-Dk}{F}, \quad (98)$$

$$w = \frac{-Dkn}{Ek + Fn}. \quad (99)$$

A estrutura dada na Equação (98) envolve somente dois coeficientes, sendo sua superfície um plano. Essa equação permite estruturar Critérios mais simples, como os Critérios de Informação AIC e o BIC, por exemplo.

A estrutura da segunda equação envolve três coeficientes, sendo eles D, E e F. Essas escolhas se justificam pelo fato desses coeficientes aparecerem no limite dado pela Equação (85). O coeficiente B , que também aparece na Equação (82), envolve o termo n^2 , razão pela qual foi anulado. É possível acrescentar nas Equações (98) e (99) uma constante de integração adicional (Apêndice A.5), como explicitado por exemplo nas Equações (133), (134) e (135), no entanto, optou-se por considerá-la nula.

Em sequência, foram determinadas as relações dessas estruturas propostas a partir das relações empíricas dadas pelas Equações (90) e (91). Assim, tem-se para a estrutura proposta pelas Equações (98) e (99) a seguinte relação:

$$\frac{-D}{F} = 2^\alpha = \frac{2\gamma n}{e^2}. \quad (100)$$

Dada essa condição, pode-se construir um Critério de informação conforme a Equação (101):

$$C_n^{\gamma, E, F} = -2\ln(L) + \frac{2F\gamma n^2 k}{e^2(Ek + Fn)}. \quad (101)$$

É interessante observar que a Equação (101) tende a ficar igual ao critério equivalente dado por $w = 2^\alpha$ quando n é muito grande (Equação (90)). Nesse caso, mantém-se o mesmo padrão observado nos Critérios de Informação AIC e AICc, preservando assim a mesma característica.

Para analisar o possível comportamento de γ , E e F , a penalidade da Equação (101) foi comparada com a penalidade do critério BIC. Assim, tem-se:

$$\frac{2F\gamma n^2 k}{e^2(Ek + Fn)} = k\ln(n),$$

que pode ser modificada algebricamente, definindo a Equação (102) para γ :

$$\gamma = \frac{e^2(Ek + Fn)\ln(n)}{2Fn^2}. \quad (102)$$

A Equação (102) depende dos valores desconhecidos E e F (com excessão da razão $\frac{2}{e^2} \approx 0,27$). A Figura 34 mostra o comportamento de γ para um valor fixo de k ($k = 1$) e diferentes valores de n . As superfícies foram geradas para $n = 25$, $n = 50$, $n = 100$ e $n = 200$ e exibem um padrão semelhante de comportamento.

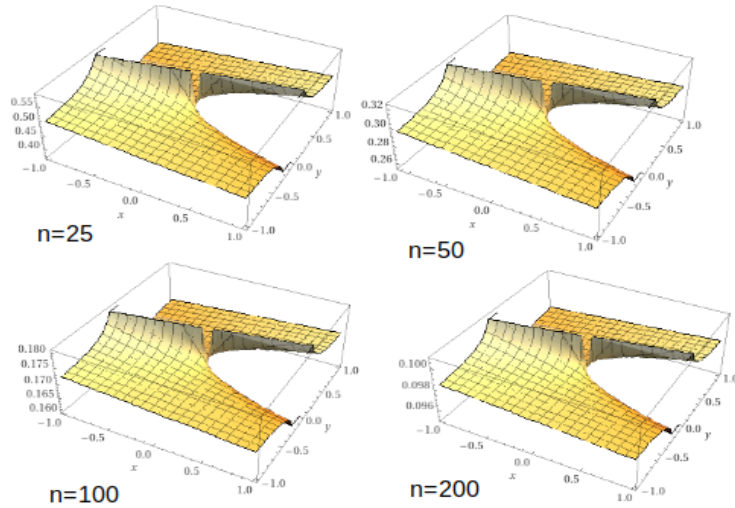


Figura 34: Comportamento gráfico para γ , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma, E, F}$ com a penalidade do BIC para diferentes valores de n

Entretanto, os gráficos da Figura 34 mostram que γ assume valores entre 0.40 e 0.55 para $n = 25$, assume valores que ficam entre 0.26 e 0.32 para $n = 50$, assume valores entre 0.16 e 0.18 para $n = 100$ e assume valores entre 0.09 e 0.1 para para $n = 200$. Já os valores de E e F estão entre -1 e 1. Se fizermos a mesma comparação com os critérios AIC ou AICc, considerando os gráficos para $n = 25$, $n = 50$, $n = 100$ e $n = 200$, observa-se o mesmo padrão semelhante de comportamento quando comparados ao Critério de Informação BIC, conforme visto nas figuras 35 e 36.

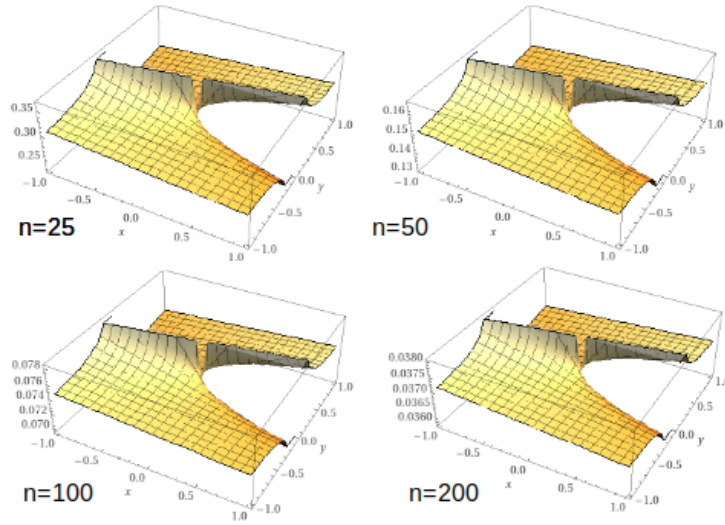


Figura 35: Comportamento gráfico para γ_1 , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma, E, F}$ com a penalidade do AIC para diferentes valores de n

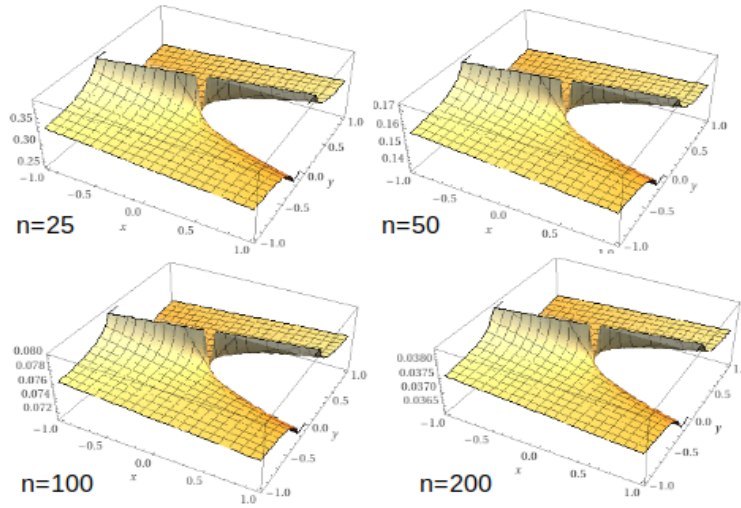


Figura 36: Comportamento gráfico para γ_1 , $E = x$ e $F = y$ quando comparados a penalidade de $C_n^{\gamma, E, F}$ com a penalidade do AICc para diferentes valores de n

Diante do exposto, conclui-se que AIC, AICc e BIC podem ser considerados também como um caso particular do Critério de Informação dado pela

Equação (101), que possui uma Superfície Quádrica que define a penalidade, diferenciando-se pelos valores adotados pelos coeficientes γ , E e F. Por exemplo, para $E = 0.5$, $F = 0.5$, $k = 1$ e $\gamma = 0.4947$, teríamos o valor 3.218, que é o valor aproximado da penalidade de BIC para $n = 25$ e $k = 1$.

Como na Equação (101), se torna necessário otimizar γ , E e F, e entendendo que muitos textos sugerem que o BIC é melhor do que os Critérios de AIC e AICc em muitos casos (Emiliano et al., 2010), se torna interessante tomar γ conforme Equação (96), fazendo com que a Equação (101) tome a forma mais simples, ou seja:

$$C_n^{E;F} = -2\ln(L) + \frac{Fkn.\ln(n)}{Ek + Fn}. \quad (103)$$

A Equação (103) permite a redução de mais um coeficiente, tendo agora somente os coeficientes E e F para serem otimizados, além de permitir uma certa fusão entre os 3 critérios conhecidos. Nesse sentido, pode-se pensar em duas afirmações acerca dessa equação. A primeira é que se está tentando otimizar o BIC, pois o mesmo se modifica conforme os valores de E e F dados na Equação (103), enquanto a segunda é que se está tomando como ponto de partida o valor do critério BIC para dado n e k . Nesse trabalho será considerado a segunda afirmação, ou seja, estabelecer uma estratégia que permita nos dar um ponto de partida, evitando assim um possível desvio dos objetivos pré-estabelecidos. Contudo, pode-se visualizar a primeira afirmação escrevendo a equação anterior em uma forma alternativa, introduzindo o valor do critério BIC na Equação (103), ou seja:

$$C_n^{E;F} = BIC + k\ln(n) \left(\frac{Fn}{Ek + Fn} - 1 \right). \quad (104)$$

A Equação (104) revela o fator de correção ν , que se adiciona ao valor de BIC, dado nesse caso pela Equação (105).

$$\nu = k.\ln(n) \left(\frac{Fn}{Ek + Fn} - 1 \right). \quad (105)$$

Para n grande, ν tende a ficar mais próximo de zero, fazendo com que o Critério de Informação tenda a ficar mais próximo do Critério de Informação BIC, como esperado.

Nesse trabalho será definida uma estratégia para otimizar o Critério de Informação dado pela Equação (103). Todavia, é possível otimizar diretamente a Equação (101), ou a Equação (82), ou seguir outros tipos de estruturas. Optou-

se pela configuração dada pela estrutura proposta pela Equação (103), deixando outras condições e interpretações para possíveis trabalhos futuros.

3.11 Estratégia de análise

Nesse trabalho os dados simulados apresentarão as propriedades de um modelo $ARMA(p, q)$, possuindo autocorrelação com valores passados e/ou erros passados.

A estratégia de ajuste dos coeficientes que compõe a estrutura da penalidade da Equação $C_n^{E,F}$ (Equação (103)) é realizada em duas fases, descritas a seguir:

FASE 1:

i) Determinar um número m de séries temporais $Y = \{\vec{y}_1, \vec{y}_2, \dots, \vec{y}_m\}$ a serem geradas segundo um modelo do tipo $ARMA(p, q)$, sendo cada série representada por um vetor de tamanho n , que contém os dados $\vec{y}_m$;

ii) Escolher um conjunto de valores para o coeficiente E e um outro conjunto para o coeficiente F , que deverão estar contidos no conjunto de valores $[-1, 1]$, excluindo-se o zero.

FASE 2:

i) Para cada uma das séries geradas, fazer a modelagem dos dados de $\vec{y}_m$ para o tamanho de amostra n , definindo um conjunto de modelos candidatos $ARMA(p, q)$ a partir de todas as combinações possíveis entre os valores de p e q escolhidos.

ii) Para cada um dos modelos gerados, avaliar a estacionariedade, estimar seus coeficientes e determinar o valor do Critério de Informação $C_n^{E,F}$ para cada combinação dos valores de E e F escolhidos.

iii) Encontrar o modelo candidato com o menor valor de $C_n^{E,F}$ e comparar com o modelo $ARMA(p, q)$ simulado de $\vec{y}_m$ gerado. Se o modelo $ARMA(p, q)$ que gerou os dados for igual ao modelo com o menor valor de $C_n^{E,F}$, então o critério obteve êxito.

iv) Contabilizar a quantidade de êxitos totais entre os m modelos para cada $ARMA(p, q)$.

A quantidade de acertos A_c , ou êxitos, vai dizer o quanto o critério $C_n^{E,F}$ foi satisfatório com relação ao total de modelos simulados para todas combinações p e q , identificando o melhor valor para o par E e F .

As fases descritas anteriormente devem ser realizadas para cada série gerada em condição experimental analisada, contabilizando assim o total de acertos A_c para cada caso. Como a contabilização de acertos está relacionada ao critério

$C_n^{E,F}$ dado pela Equação (103), é interessante comparar, nas mesmas condições, com os acertos dados por BIC, AIC e AICc. Essas fases podem ser visualizadas conforme fluxograma apresentado na Figura 37.

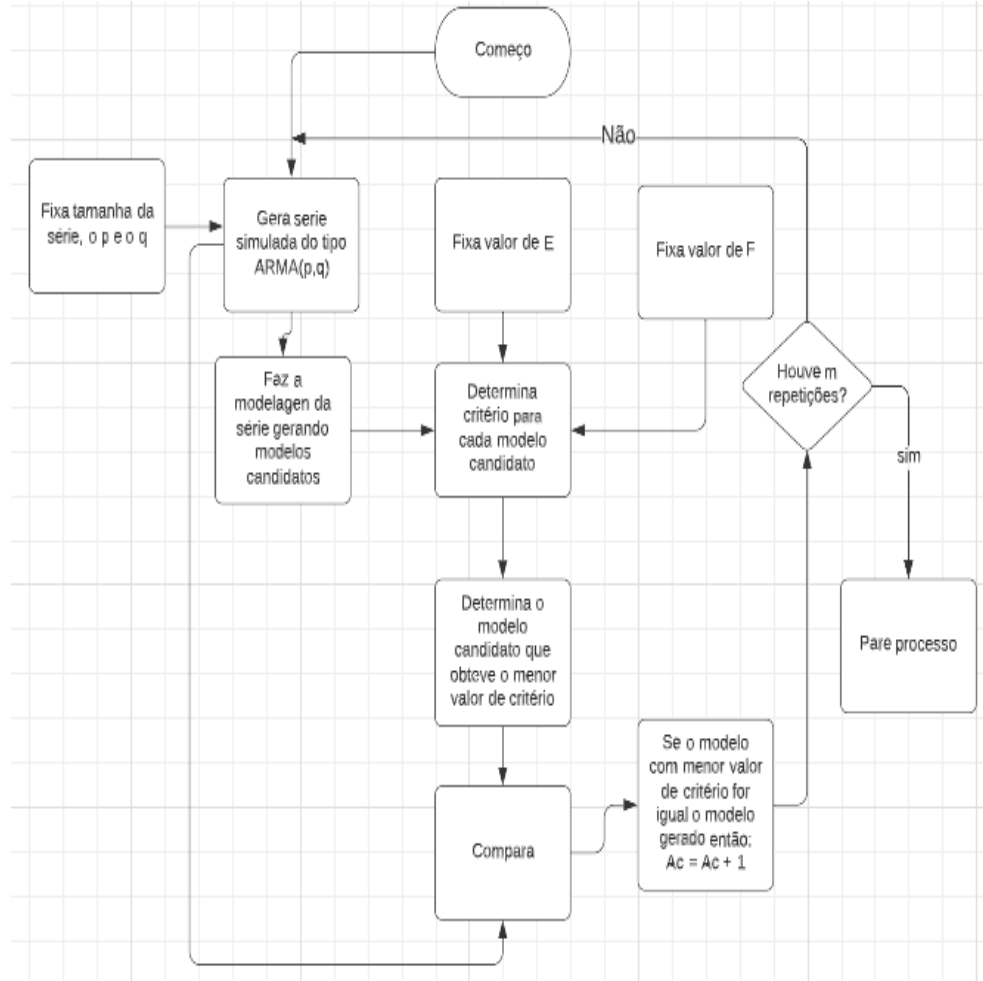


Figura 37: Fluxograma para a estratégia proposta

Apesar do fluxograma mostrar a estratégia para o critério proposto, esse serve também para os critérios AIC, AICc e BIC.

4 ESTUDO DE SIMULAÇÃO

Neste capítulo são detalhados os procedimentos para a realização das simulações que avaliam os Critérios de Informação tradicionais além dos propostos neste estudo, com base na acurácia dos resultados obtidos ao utilizar modelos do tipo $ARMA(p, q)$.

4.1 Metodologia

Os dados simulados apresentarão as propriedades de um modelo $ARMA(p, q)$, possuindo autocorrelação com valores passados e/ou erros passados, gerados conforme Equação (1).

Seus coeficientes são: β_0 representando o coeficiente escalar, β_1 representando coeficiente angular, ϕ_i ($i = 1, 2 \dots p$), representando os coeficientes dos termos autorregressivos, e θ_j ($j = 1, 2 \dots q$) associado às parcelas de médias móveis, além do erro ϵ_t representando o efeito aleatório. Além disso, também serão geradas séries de diferentes tamanhos n , sendo esses 50, 100 e 200.

Os coeficientes β_0 e β_1 podem estar presentes ou não e podem assumir valores iguais a 10 e 1, respectivamente, quando presentes. Os coeficientes ϕ_i e θ_j serão considerados fixos. O erro ϵ_t se comportará como ruído branco gaussiano.

Uma melhor descrição dos fatores e seus níveis está apresentada na Tabela 4, sendo que para cada condição experimental foram geradas $m = 3000$ séries temporais.

Tabela 4: Valores atribuídos aos coeficientes utilizados na geração de modelos $ARMA(p, q)$.

Fatores	Níveis
ϕ_1	0,3 e 0,7
θ_1	0,3 e 0,7
β_0	0 e 10
β_1	0 e 1

Os modelos $ARMA(p, q)$ utilizados na geração das séries são $ARMA(0, 1)$, $ARMA(1, 0)$ e $ARMA(1, 1)$, enquanto que, para compor o conjunto de modelos candidatos na modelagem de cada série gerada, são considerados os modelos $ARMA(0, 1)$, $ARMA(0, 2)$, $ARMA(1, 0)$, $ARMA(1, 1)$, $ARMA(1, 2)$, $ARMA(2, 0)$, $ARMA(2, 1)$, $ARMA(2, 2)$, conforme ilustrado na Figura 38 para $n = 50$.

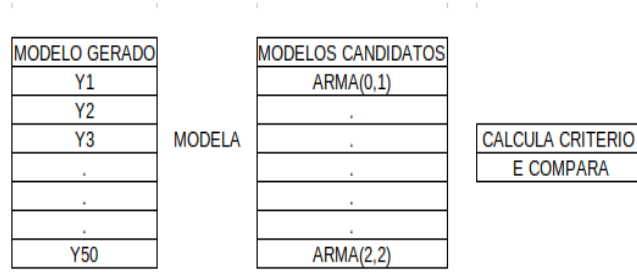


Figura 38: Metodologia usada modelar uma série temporal com $n = 50$

Na identificação do modelo mais parcimonioso são utilizados os critérios tradicionais AIC, AICc e BIC, além dos propostos nesta tese.

Conforme descrito na Seção 3.10, estes critérios propostos, determinados pela Equação (103), estão apresentados na Tabela 5 para cada valor de n considerado. Cada uma destas equações apresenta duas variáveis (E e F), que, nessa tese, irão assumir valores iguais a -1.0, -0.5, 0.5 e 1.0, totalizando 16 diferentes combinações, algumas das quais levam a critérios equivalentes. Para cada par de valores E e F , os critérios equivalentes estão listados na Tabela 6, onde são identificados 6 diferentes critérios de informação, denominados T1, T2, T3, T4, T5 e T6 (Apêndice A.9).

Tabela 5: Fórmulas para os critérios para cada E e F (Equação (103)) e para cada n .

n	CRITÉRIOS
50	$C_{50}^{E,F} = -2\ln(L) + \frac{195,60kF}{Ek+50F}$
100	$C_{100}^{E,F} = -2\ln(L) + \frac{460,51kF}{Ek+100F}$
200	$C_{200}^{E,F} = -2\ln(L) + \frac{1059,66kF}{Ek+200F}$

Tabela 6: Valores de E e F para $C_n^{E;F}$ equivalentes.

Notação	Critério $C_n^{E;F}$	Pares (E;F) Equivalentes
T1	$C_n^{-1;-1}$	$(-1; -1), (-0.5; -0.5), (0, 5; 0.5), (1, 1)$
T2	$C_n^{-1;-0.5}$	$(-1; -0.5), (1; 0.5)$
T3	$C_n^{-1;0.5}$	$(-1; 0.5), (1; -0.5)$
T4	$C_n^{-1;1}$	$(-1, 1), (-0.5; 0.5), (0.5; -0.5), (1, -1)$
T5	$C_n^{-0.5;-1}$	$(-0.5; -1), (0.5; 1)$
T6	$C_n^{-0.5;1}$	$(-0.5; 1), (0.5; -1)$

Por fim, de acordo com cada Critério de Informação, identifica-se o menor valor obtido pelo critério para cada um dos modelos candidatos, que será considerado o modelo mais parcimonioso e, conseqüentemente, o que apresenta melhor ajuste. Se este coincidir com o modelo usado na geração da série temporal, considera-se que o critério o identificou corretamente. Espera-se que o Critério de Informação avalie como mais parcimonioso o mesmo modelo que foi usado na geração da série simulada. Nesse trabalho irá se determinar a proporção de acertos para cada tipo de critério considerado, para 3000 modelos simulados da forma $ARMA(p, q)$.

Todas as análises serão realizadas com o auxílio do software *R* (de Micheaux et al., 2013), utilizando alguns de seus pacotes, como o *urca*, (Pfaff et al., 2016), *betareg*, (Grün et al., 2011) e *forecast* (Hyndman and Khandakar, 2008). A saída dos resultados em termos de acertos e erros ficará armazenada conforme ilustra a Figura 39 para um modelo simulado de tamanho 50, o que permite ser utilizado por qualquer tipo de *software* (como planilhas eletrônicas e o próprio R).

modelo	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC	phi	theta	B1	B0	n= 50
arma01										0	0.3	0	0	50
arma01										0	0.3	0	10	50
arma01										0	0.3	1	0	50
arma01										0	0.3	1	10	50
arma01										0	0.7	0	0	50
arma01										0	0.7	0	10	50
arma01										0	0.7	1	0	50
arma01										0	0.7	1	10	50
arma10										0.3	0	0	0	50
arma10										0.3	0	0	10	50
arma10										0.3	0	1	0	50
arma10										0.3	0	1	10	50
arma11										0.3	0.3	0	0	50
arma11										0.3	0.3	0	10	50
arma11										0.3	0.3	1	0	50
arma11										0.3	0.3	1	10	50
arma11										0.3	0.7	0	0	50
arma11										0.3	0.7	0	10	50
arma11										0.3	0.7	1	0	50
arma11										0.3	0.7	1	10	50
arma10										0.7	0	0	0	50
arma10										0.7	0	0	10	50
arma10										0.7	0	1	0	50
arma10										0.7	0	1	10	50
arma11										0.7	0.3	0	0	50
arma11										0.7	0.3	0	10	50
arma11										0.7	0.3	1	0	50
arma11										0.7	0.3	1	10	50
arma11										0.7	0.7	0	0	50
arma11										0.7	0.7	0	10	50
arma11										0.7	0.7	1	0	50
arma11										0.7	0.7	1	10	50

Figura 39: Saída dos resultados armazenados em uma planilha eletrônica para modelos de tamanho 50

A análise dos resultados obtidos nas simulações será realizada utilizando a metodologia de análise de variância e a regressão Beta. A análise de variância permitirá identificar se há diferenças significativas entre os grupos analisados, enquanto a regressão Beta fornecerá informações sobre a relação entre as variáveis, ajudando a entender como as mudanças em uma variável podem influenciar as demais.

4.2 Resultados

Inicialmente foi realizada uma análise descritiva das proporções de acertos obtidas por cada Critério de Informação, utilizando medidas descritivas e gráficos. Em seguida, realizou-se uma Análise de Variância para comparar as proporções de acertos obtidas pelos diferentes critérios. Após, fez-se a modelagem dos resultados pela regressão Beta e discussão das conclusões e implicações

derivadas dessa análise. Este último procedimento foi realizado apenas nos dados do primeiro experimento.

4.2.1 Análise descritiva

Inicialmente foram determinadas as proporções de acertos de cada um dos critérios AIC, AICc e BIC, além dos seis propostos neste estudo. Considera-se acerto identificar corretamente o tipo de modelo ARMA(p,q) gerado na simulação. Isso foi feito a partir das 3000 séries simuladas para cada uma das 32 condições experimentais.

As medidas descritivas dos resultados encontrados que incluem Mínimo, Primeiro Quartil, Mediana, Média, Terceiro Quartil, Máximo, Desvio Padrão, Coeficiente de Curtose e Coeficiente de Assimetria estão detalhadas na Tabela 7.

Tabela 7: Medidas descritivas da proporção de acertos de cada Critério de Informação para as séries simuladas.

Medidas	Critérios								
	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC
Mínimo	0,1003	0,1537	0,0000	0,0060	0,0727	0,0213	0,2523	0,2440	0,0440
1º Quartil	0,576	0,5505	0,6096	0,6042	0,5903	0,6016	0,4486	0,4958	0,5999
Mediana	0,6724	0,6674	0,7153	0,698	0,6786	0,6917	0,5315	0,5440	0,6865
Média	0,6828	0,6718	0,7076	0,7010	0,6880	0,6969	0,5368	0,5676	0,6923
3º Quartil	0,8980	0,8723	0,9614	0,9412	0,9078	0,9291	0,6618	0,7024	0,9172
Máximo	0,9687	0,9663	0,9773	0,9737	0,9697	0,9723	0,7620	0,8120	0,9713
Desvio Padrão	0,2306	0,2155	0,2712	0,2604	0,2384	0,2539	0,1295	0,1390	0,2466
Curtose	3,2724	2,7841	3,8854	3,8808	3,4913	3,8101	2,2191	2,3594	3,6831
Assimetria	-0,8399	-0,5911	-1,1515	-1,1257	-0,9437	-1,0875	-0,1621	-0,2224	-1,0288

Obs: Maiores resultados = negrito itálico; Menores resultados = negrito.

Na Tabela 7 observa-se que, em geral, os maiores valores para todas as medidas estão associados ao critério Q3. Entretanto, em relação ao valor mínimo, acontece exatamente o contrário: este critério apresenta o menor resultado. Observa-se também que os menores valores são encontrados no critério AIC, com exceção do valor mínimo, que é o maior.

Quanto à curtose, todas as distribuições foram classificadas como leptocúrticas, o que significa que elas têm mais caudas pesadas, com um pico mais acentuado em torno da média do que uma distribuição Normal.

No que diz respeito ao coeficiente de assimetria, todas as distribuições foram identificadas como assimétricas negativas. Isso implica na existência de uma maior concentração dos valores maiores, com uma cauda mais longa à esquerda.

Na Figura 40 é possível ver os gráficos Boxplots para as proporções de acertos

de cada critério estudado, enquanto que na Figura 41 os gráficos Boxplots desses mesmos critérios por modelo ARMA. Na Figura 40 nota-se que os critérios BIC, T1, T2, T3, T4, T5 e T6 apresentam comportamentos semelhantes, e além disso, alguns deles (BIC, T3, T4, T5 e T6) exibem valores que se situam fora do padrão. Observa-se também que os critérios AIC e AICc apresentam comportamentos semelhantes. Já na Figura 41 nota-se que a proporção de acertos para o modelo ARMA(1,1) é bem menor do que para os modelos ARMA(1,0) e ARMA(0,1).

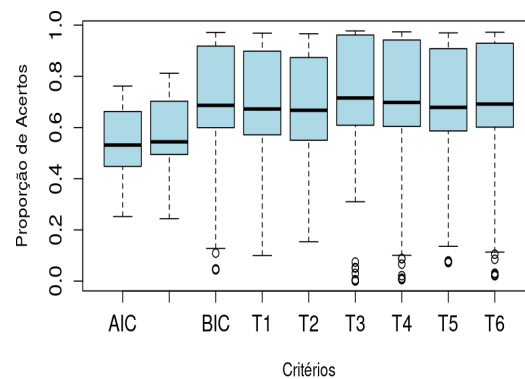


Figura 40: Boxplot das proporções de acertos para cada critério de informação

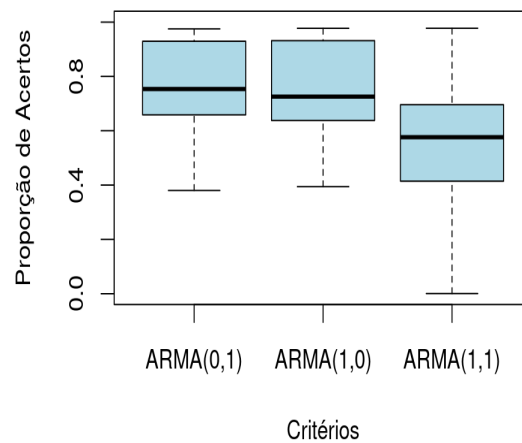


Figura 41: Boxplot das proporções de acertos por modelo ARMA

A Tabela 8 mostra os valores de proporção média para os critérios utilizados, subdividido-os em relação ao tipo de modelo e ao tamanho da série. Nesta

tabela observa-se que a proporção média de acertos dos Critérios de Informação varia conforme o tipo de modelo ARMA e o tamanho da série. Em geral, os modelos ARMA(1,1) apresentaram uma proporção média de acertos inferior em comparação com os modelos ARMA(0,1) e ARMA(1,0). Entre os Critérios de Informação, o AIC mostrou a menor proporção média de acertos, enquanto o critério T3 apresentou a maior.

Tabela 8: Proporção média de acertos de cada Critério de Informação por tipo de modelo e ao tamanho da série.

n	ARMA(p,q)	AIC	AICc	BIC	Critério de Informação					
					T1	T2	T3	T4	T5	T6
50	ARMA(0,1)	0,5329	0,6004	0,7598	0,7163	0,6691	0,8260	0,7972	0,7392	0,7794
	ARMA(1,0)	0,5504	0,6165	0,7535	0,7193	0,6796	0,8048	0,7828	0,7379	0,7694
	ARMA(1,1)	0,4666	0,5024	0,5032	0,4985	0,4932	0,5173	0,5110	0,5017	0,5067
100	ARMA(0,1)	0,5610	0,5953	0,8172	0,8029	0,7869	0,8415	0,8303	0,8104	0,8240
	ARMA(1,0)	0,5732	0,6055	0,7998	0,7879	0,7752	0,8203	0,8107	0,7940	0,8050
	ARMA(1,1)	0,5010	0,5230	0,5647	0,5680	0,5709	0,5578	0,5612	0,5668	0,5633
200	ARMA(0,1)	0,5943	0,6111	0,8590	0,8539	0,8480	0,8684	0,8640	0,8560	0,8615
	ARMA(1,0)	0,6043	0,6198	0,8468	0,8433	0,8388	0,8555	0,8514	0,8449	0,8488
	ARMA(1,1)	0,5454	0,5560	0,6678	0,6684	0,6682	0,6623	0,6659	0,6681	0,6671
Todos	ARMA(0,1)	0,5628	0,6023	0,8120	0,7910	0,7680	0,8453	0,8305	0,8018	0,8216
	ARMA(1,0)	0,5760	0,6139	0,8000	0,7835	0,7645	0,8269	0,8149	0,7923	0,8077
	ARMA(1,1)	0,5043	0,5271	0,5786	0,5783	0,5774	0,5791	0,5794	0,5788	0,5790

Obs: Maiores resultados = negrito itálico; Menores resultados = negrito.

4.2.2 Análise de variância

Com o intuito de comparar estatisticamente as proporções de acertos obtidas pelos Critérios de Informação analisados, foi aplicada uma Análise da Variância com blocos. Inicialmente foi utilizada uma análise paramétrica com o teste F, mas as suposições de independência, normalidade e homocedasticidade não foram verificadas. Diante disso, foi utilizada uma Análise de Variância não paramétrica, reconhecendo a necessidade de abordagens que sejam mais flexíveis em relação às distribuições dos dados. Para isso utilizou-se o teste de Friedman, obtendo-se $\chi^2 = 459,69$, com 8 graus de liberdade e $\text{valor} - p = 2,10^{-16}$. Esse resultado indica que há evidência estatística para concluir que há pelo menos uma diferença significativa entre as proporções medias de acertos obtidas pelos critérios estudados.

Para analisar mais detalhadamente essas diferenças detectadas, foram realizadas comparações múltiplas utilizando o teste de Wilcoxon. Os resultados podem ser observados na Tabela 9.

Tabela 9: Resultados do *valor - p* obtidos pelos testes de Wilcoxon.

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	1,8E-15	-	-	-	-	-	-	-
BIC	8,4E-11	2,5E-09	-	-	-	-	-	-
T1	6,9E-12	2,4E-10	4,7E-06	-	-	-	-	-
T2	4,4E-13	2,1E-11	3,5E-06	2,1E-06	-	-	-	-
T3	9,3E-10	2,0E-08	4,7E-06	4,7E-06	4,7E-06	-	-	-
T4	6,3E-10	1,3E-08	4,7E-06	4,7E-06	4,7E-06	4,7E-06	-	-
T5	1,8E-11	6,4E-10	4,7E-06	4,7E-06	2,6E-06	4,7E-06	4,7E-06	-
T6	2,1E-10	6,7E-09	4,7E-06	4,7E-06	3,7E-06	4,7E-06	4,7E-06	4,7E-06

A Tabela 9 indica que há evidência de diferença entre todas as proporções de acertos obtidas por todos os Critérios de Informação em função dos baixos valores de probabilidade (*valor - p* < 0.05). De acordo com a análise realizada, o melhor desempenho foi obtido pelo T3(0.7076). Em sequência vieram T4(0.7010), T6(0.6968), BIC (0.6922), T5(0.6879), T1(0.6827), AICc (0.5676) e AIC(0.5368).

No entanto, parece crucial considerar que essas proporções podem estar sendo influenciadas por alguns dos fatores considerados no experimento, especialmente o tipo de modelo ARMA, objeto de estudo desta análise. Os Boxplots apresentados nas figuras 42, 43, 45, 46, 47, 48, 49 e 50 sugerem que se confirma a suposição inicial de que a estrutura do modelo ARMA, em termos de apresentar ou não termos autorregressivos e erros passados, interfere no desempenho dos critérios, inclusive nos propostos. Além disso, essa interferência parece ser menor em AIC e AICc.

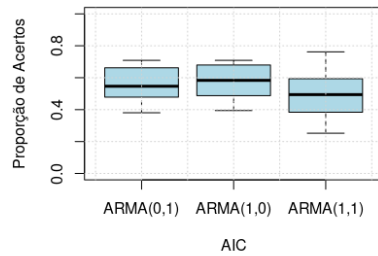


Figura 42: Gráfico Boxplot com os modelos considerados para o critério AIC

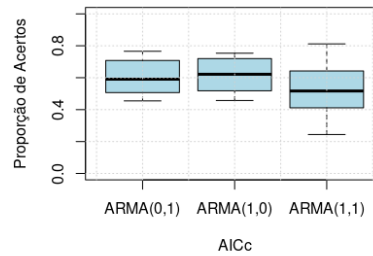


Figura 43: Gráfico Boxplot com os modelos considerados para o critério AICc

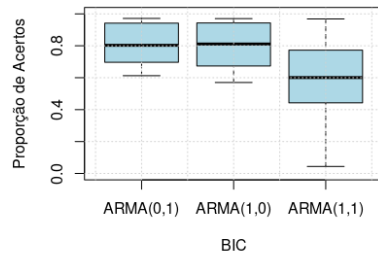


Figura 44: Gráfico Boxplot com os modelos considerados para o critério BIC

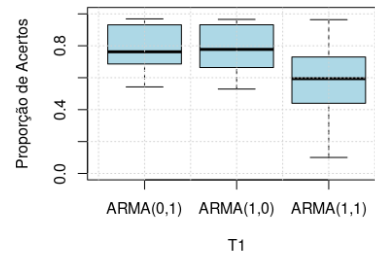


Figura 45: Gráfico Boxplot com os modelos considerados para o critério T1

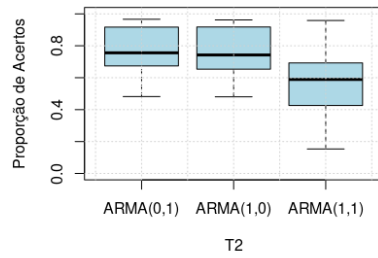


Figura 46: Gráfico Boxplot com os modelos considerados para o critério T2

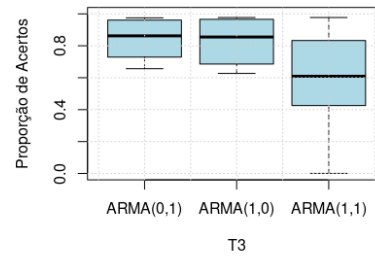


Figura 47: Gráfico Boxplot com os modelos considerados para o critério T3

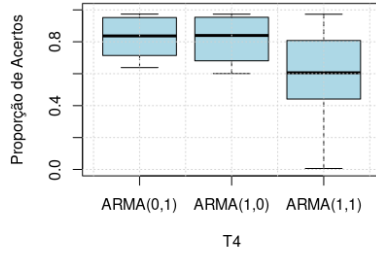


Figura 48: Gráfico Boxplot com os modelos considerados para o critério T4

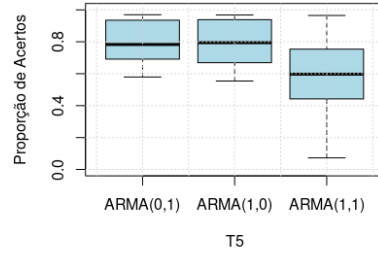


Figura 49: Gráfico Boxplot com os modelos considerados para o critério T5

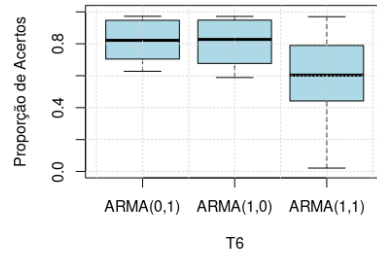


Figura 50: Gráfico Boxplot com os modelos considerados para o critério T6

Nos gráficos de boxplot, observa-se uma maior variabilidade nos modelos ARMA(1,1) para todos os critérios, enquanto o modelo ARMA(0,1) e ARMA(1,0) são os modelos que apresentam a menor variabilidade.

Novos testes de Friedman foram realizados para cada um dos tipos de modelo (ARMA(0,1), ARMA(1,0), ARMA(1,1)), encontrando-se $\text{valor} - p = 2,10^{-16}$ com 8 graus de liberdade nas três aplicações. Os resultados das comparações múltiplas, utilizando o teste de Wilcoxon, podem ser observados nas Tabelas 10, 11 e 12.

Tabela 10: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(0,1).

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	3,2E-04	-	-	-	-	-	-	-
BIC	4,2E-06	4,2E-06	-	-	-	-	-	-
T1	4,2E-06	4,2E-06	3,2E-04	-	-	-	-	-
T2	4,2E-06	4,2E-06	4,2E-06	3,2E-04	-	-	-	-
T3	4,2E-06	4,2E-06	3,2E-04	4,2E-06	4,2E-06	-	-	-
T4	4,2E-06	3,2E-04	3,2E-04	4,2E-06	4,2E-06	3,2E-04	-	-
T5	4,2E-06	4,2E-06	3,2E-04	3,2E-04	4,2E-06	3,2E-04	3,2E-04	-
T6	4,2E-06	4,2E-06	3,2E-04	3,2E-04	3,2E-04	3,2E-04	3,2E-04	3,2E-04

Tabela 11: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,0).

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	4,2E-04	-	-	-	-	-	-	-
BIC	4,2E-06	4,2E-06	-	-	-	-	-	-
T1	4,2E-04	4,2E-04	4,2E-04	-	-	-	-	-
T2	4,2E-06	4,2E-06	4,2E-04	4,2E-04	-	-	-	-
T3	4,2E-04	4,2E-06	4,2E-04	4,2E-06	4,2E-06	-	-	-
T4	4,2E-06	4,2E-06	4,2E-04	4,2E-04	4,2E-04	4,2E-04	-	-
T5	4,2E-06	4,2E-06	4,2E-04	4,2E-04	4,2E-06	4,2E-04	4,2E-04	-
T6	4,2E-04	4,2E-04	4,2E-04	4,2E-06	4,2E-06	4,2E-04	4,2E-04	4,2E-04

Tabela 12: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,1).

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	3,4E-07	-	-	-	-	-	-	-
BIC	7,6E-02	2,0E-01	-	-	-	-	-	-
T1	2,7E-02	1,0E-01	>0,99	-	-	-	-	-
T2	2,1E-03	2,0E-02	>0,99	>0,99	-	-	-	-
T3	2,0E-01	3,8E-01	>0,99	>0,99	>0,99	-	-	-
T4	1,9E-01	3,7E-01	>0,99	>0,99	>0,99	>0,99	-	-
T5	5,1E-02	1,4E-01	>0,99	>0,99	>0,99	>0,99	>0,99	-
T6	1,0E-01	3,0E-01	>0,99	>0,99	>0,99	>0,99	>0,99	>0,99

No modelo ARMA(0,1) foram obtidas as seguintes proporções de acertos: T3(0.8453), T4(0.8305), T6(0.8216), BIC(0.8120), T5(0.8018), T1(0.7910), T2 (0.7680), AICc(0.6023) e AIC(0.5628). Já no modelo ARMA(1,0) foram obtidas as seguintes proporções: T3(0.8269), T4(0.8149), T6(0.8077), BIC(0.8000), T5(0.7923), T1(0.7835), T2(0.7646), AICc(0.6139) e AIC(0.5760). No modelo ARMA(1,1), existem diferenças significativas entre AIC e AICc em relação aos demais, considerando o nível de significância de 0.05. Já entre BIC, T1, T2, T3, T4, T5, e T6 não foram encontradas evidências de proporções de acertos diferentes. Além disso, para a estrutura do modelo, a hierarquia entre os

resultados de desempenho, embora assumindo proporções bem mais baixas, segue a mesma: T3(0.5791), T4(0.5794), T6(0.5790), BIC(0.5786), T5(0.5788), T1(0.5783), T2(0.5774), AICc(0.5271) e AIC(0.5043).

Assim sendo, o que se conclui é que esses Critérios de Informação apresentam desempenhos diferentes em função do tipo de modelo, mas todos os critérios analisados foram afetados da mesma maneira por essa estrutura.

4.2.3 Análise utilizando a regressão Beta

Para se obter um entendimento mais profundo sobre a relação as variáveis independentes consideradas no experimento com a proporção de acertos, utilizou-se a Regressão Beta. Nessa análise foram consideradas, além do efeito principal, as interações de até segunda ordem dos fatores ϕ (coeficiente AR), θ (coeficiente MA), $B1$ (coeficiente da tendência), $B0$ (constante) e n (tamanho da amostra), o que foi feito em cada um dos tipos de modelos para cada um dos critérios.

Para a modelagem Beta, adotou-se uma abordagem iterativa para otimizar o modelo, iniciada com um modelo com todas as variáveis independentes e todas as interações até segunda ordem. Em cada etapa, avaliou-se a porcentagem de pontos localizados dentro do gráfico envelope para confirmar a adequação do método utilizado. Mesmo em casos onde essa porcentagem fosse inferior a 5%, analisou-se o valor obtido pelo critério BIC (buscando-se sempre o menor valor). Considerou-se também os valores-p que informam sobre a significância de um coeficiente. O modelo selecionado foi aquele no qual se obteve o menor valor BIC e gráfico envelope com no máximo 5% dos pontos fora dos limites de controle.

A seguir apresentam-se os resultados obtidos para cada um dos Critérios de Informação analisados.

a) Critério AIC: A Tabela 13 mostra os resultados obtidos para o critério AIC, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 51 mostra os respectivos gráficos envelope.

Tabela 13: Resultados da Regressão Beta para o critério AIC.

Componente	Coefficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-1.0024	0.051	-19.6619	<0,01	ARMA(0,1)
θ	2.6451	0.0897	29.4911	<0,01	
$B1$	-0.0457	0.0475	-0.9623	0.3359	
$B0$	2e-04	0.0048	0.0361	0.9712	
n	0.0033	0.0003	9.9952	<0,01	
$\theta:B1$	-0.3154	0.0701	-4.5021	<0,01	
$\theta:B0$	-0.044	0.007	-6.2745	<0,01	
$\theta: n$	-0.0045	0.0006	-7.9476	<0,01	
$B1:B0$	-0.0147	0.0028	-5.2584	<0,01	
$B1: n$	5e-04	0.0002	2.1581	0.0309	
$B0: n$	1e-04	0.0020	2.4703	<0,01	
(Intercept)	-0.9027	0.0320	-28.1756	<0,01	ARMA(1,0)
ϕ	2.426	0.0554	43.7637	<0,01	
$B1$	-0.1513	0.0245	-6.1749	<0,01	
$B0$	-0.0108	0.0015	-7.2457	<0,01	
n	0.0034	0.0004	14.3547	<0,01	
$\phi: n$	-0.0043	0.0004	-10.1576	<0,01	
$B1:B0$	-0.0127	0.0021	-6.0617	<0,01	
$B1: n$	4e-04	0.0002	2.3405	<0,01	
(Intercept)	-2.4078	0.0759	-31.7113	<0,01	ARMA(1,1)
ϕ	2.0832	0.1148	18.1406	<0,01	
θ	2.1649	0.1149	18.8452	<0,01	
$B1$	0.0674	0.0563	1.1973	0.2312	
$B0$	0.0159	0.005	3.1792	<0,01	
n	0.0041	0.0004	10.2212	<0,01	
$\phi:\theta$	0.8943	0.1607	5.5648	<0,01	
$\phi:B1$	-0.1367	0.0642	-2.1279	0.0333	
$\phi:B0$	-0.0166	0.0064	-2.581	<0,01	
$\phi: n$	-0.0023	0.0005	-4.496	<0,01	
$\theta:B1$	-0.3538	0.0642	-5.508	<0,01	
$\theta:B0$	-0.0373	0.0064	-5.8021	<0,01	
$\theta: n$	-0.0019	0.0005	-3.6613	<0,01	
$B1:B0$	-0.0132	0.0026	-5.1406	<0,01	
$B1: n$	5e-04	0.0002	2.2199	0.0264	

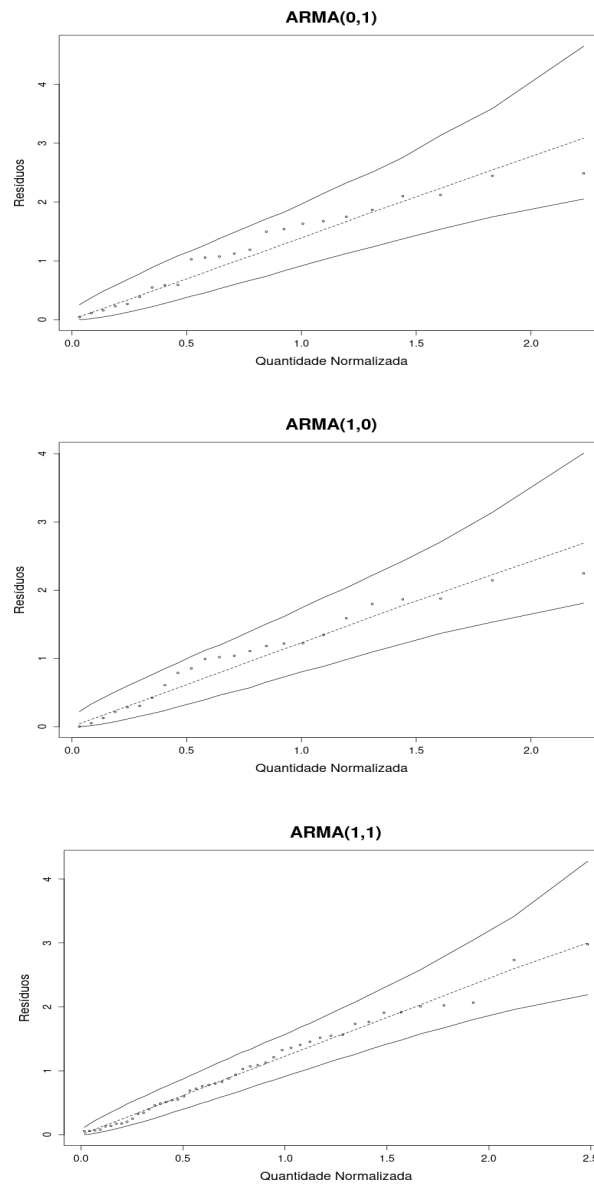


Figura 51: Gráficos envelope para modelos Beta de AIC de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

a.1) ARMA(1,1): De acordo com a Tabela 13, para o modelo ARMA(1,1), apenas o efeito principal de $B1$ não foi significativo. No entanto, as interações de $B1$ com as outras componentes do modelo Beta se mostraram significativas.

Essa tabela também mostra que existem evidências de que as proporções de acertos do critério AIC são maiores para $\phi = 0.7$ (0,6033), $\theta = 0.7$ (0,5993), $B0 = 10$ (0,4843) e $n = 200$ (0,6032).

Ainda de acordo com as informações desta mesma tabela, com relação às interações, existem evidências de que os resultados das proporções de acertos do AIC são afetados pela interação de ϕ com θ , de ϕ com $B1$, de ϕ com $B0$, de ϕ com n , de θ com $B1$, de θ com $B0$, de θ com n , de $B1$ com $B0$ e de $B1$ com n conforme observado nas figuras 52, 53, 54, 55, 57, 58, 59 e 60.

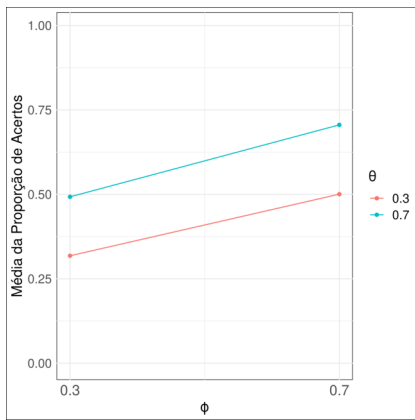


Figura 52: Interação entre ϕ e θ para ARMA(1,1) para AIC

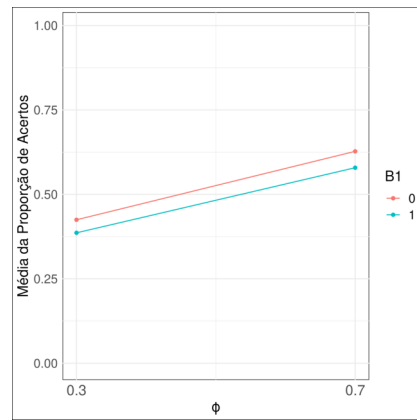


Figura 53: Interação entre ϕ e $B1$ para ARMA(1,1) para AIC

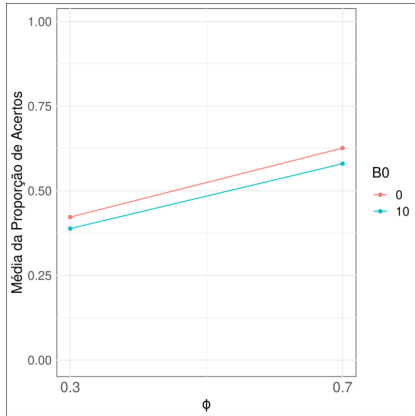


Figura 54: Interação entre ϕ e $B0$ para ARMA(1,1) para AIC

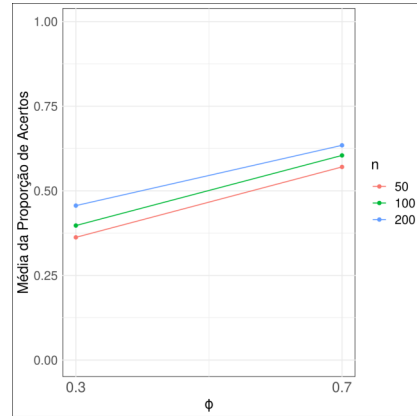


Figura 55: Interação entre ϕ e n para ARMA(1,1) para AIC

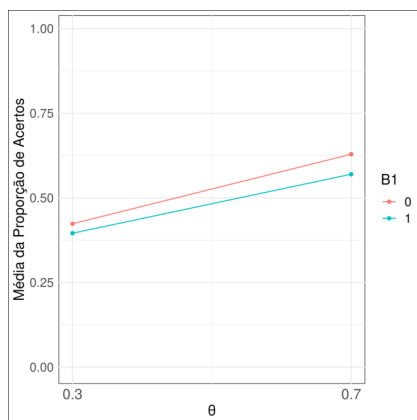


Figura 56: Interação entre θ e $B1$ para ARMA(1,1) para AIC

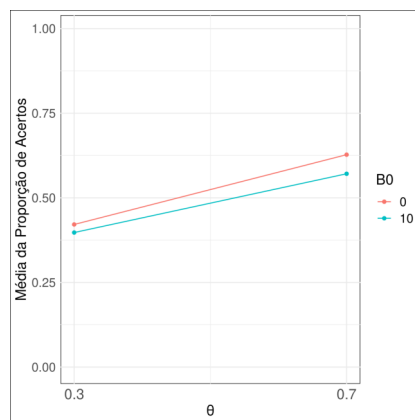


Figura 57: Interação entre θ e $B0$ para ARMA(1,1) para AIC

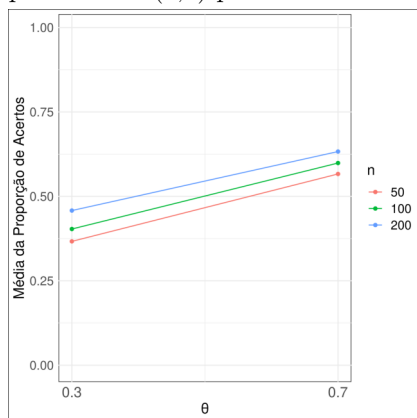


Figura 58: Interação entre θ e n para ARMA(1,1) para AIC

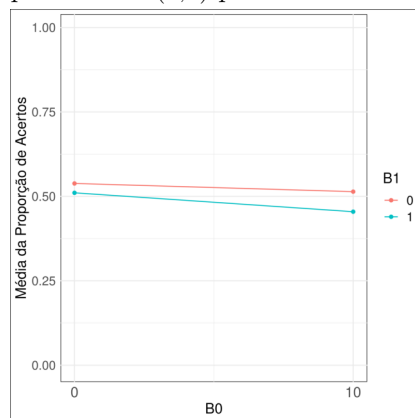


Figura 59: Interação entre $B1$ e $B0$ para ARMA(1,1) para AIC

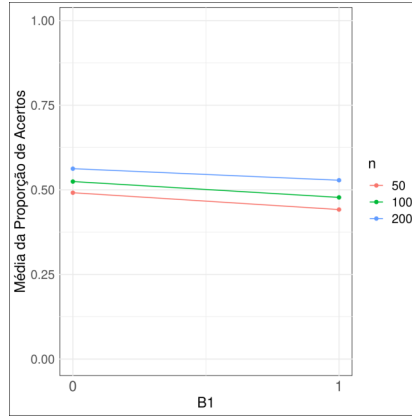


Figura 60: Interação entre $B1$ e n para ARMA(1,1) para AIC

No gráfico da Figura 52, a diferença entre as proporções de acertos para $\theta = 0.3$ e $\theta = 0.7$ é menor quando $\phi = 0.3$. No gráfico da Figura 53, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $\phi = 0.3$. No gráfico da Figura 54, a diferença entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ é menor quando $\phi = 0.3$. No gráfico da Figura 55, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi = 0.7$. No gráfico da Figura 56, as diferenças entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ são menores quando $\theta = 0.3$. No gráfico da Figura 57, as diferenças entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ são menores quando $\theta = 0.3$. No gráfico da Figura 58, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\theta = 0.7$. No gráfico da Figura 59, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$. No gráfico da Figura 60, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $B1 = 0$.

a.2) ARMA(1,0): De acordo com a Tabela 13, se for retirado o termo de Médias Móveis, ficando um modelo ARMA(1,0), os efeitos principais de ϕ , $B0$, $B1$ e n são todos significativos. Existem evidências de que as proporções de acertos do critério AIC são maiores para $\phi = 0.7$ (0,6685), $B0 = 0$ (0,5961), $B1 = 0$ (0,5959) e $n = 200$ (0,6043). Com relação às interações, existem evidências de que os resultados das proporções de acertos do AIC são afetados pela interação de ϕ com n , de $B1$ com $B0$ e de $B1$ com n , conforme observado nas figuras 61,

62 e 63.

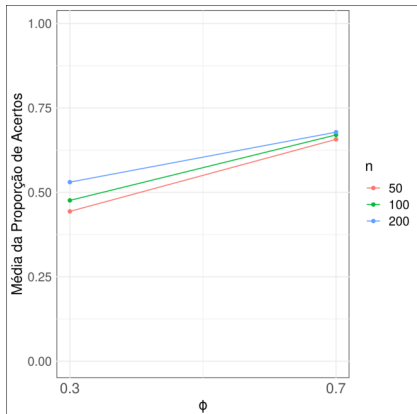


Figura 61: Interação entre ϕ e n para ARMA(1,0) para AIC

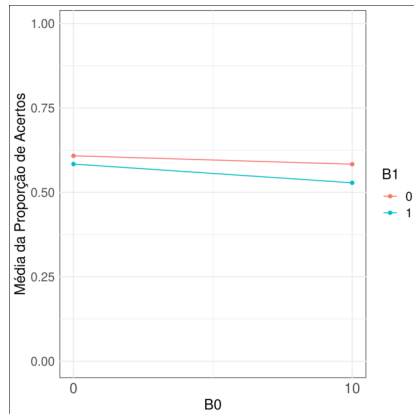


Figura 62: Interação entre $B1$ e $B0$ para ARMA(1,0) para AIC

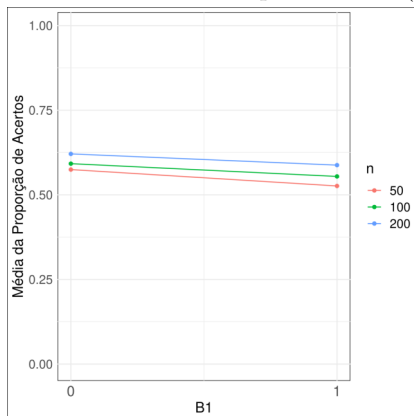


Figura 63: Interação entre $B1$ e n para ARMA(1,0) para AIC

O gráfico da Figura 61 mostra que neste tipo de modelo o efeito da interação entre ϕ e n é mais acentuado, sendo as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ quase iguais quando $\phi = 0.7$.

O gráfico da Figura 62 é muito semelhante ao gráfico da Figura 59 para o modelo ARMA(1,1), o mesmo acontecendo para o gráfico da Figura 63 em relação ao gráfico da Figura 59. Isso indica que esses fatores interagem de maneira semelhante em modelos ARMA(1,0) e ARMA(1,1).

a.3) ARMA(0,1) Ainda de acordo com a Tabela 13, se for retirado apenas o termo Autoregressivo do modelo ARMA(1,1), ficando o modelo do tipo ARMA(0,1), os efeitos principais de $B0$ e $B1$ deixam de ser significativos, embora eles possam interagir com outros fatores significativos. Existem evidências de que as proporções de acertos do critério AIC são maiores para $\theta = 0.7$ (0,6469) e $n = 200$ (0,5943). Além disso, θ interagiu com $B1$, com $B0$ e com n , $B1$ interagiu com $B0$ e n e $B0$ interagiu com n , conforme o evidenciado nas Figuras 64, 65, 66, 67 e 68. Neste caso, também observou-se que a interação entre $B0$ e $B1$ também foi significativa, mas como os efeitos principais desses fatores não o foram, a análise dessa interação não foi considerada.

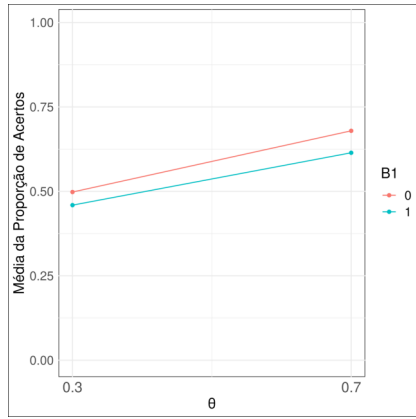


Figura 64: Interação entre θ e $B1$ para ARMA(0,1) para AIC

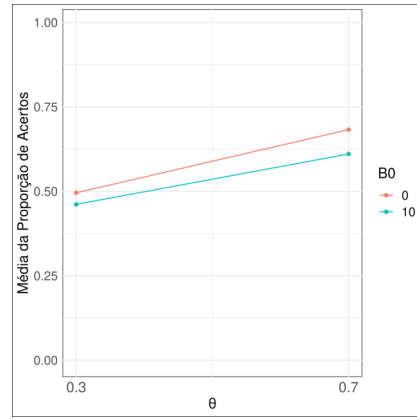


Figura 65: Interação entre θ e $B0$ para ARMA(0,1) para AIC

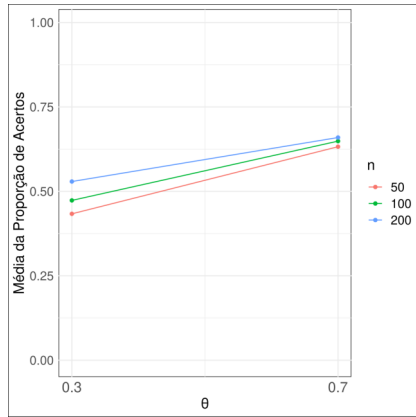


Figura 66: Interação entre θ e n para ARMA(0,1) para AIC

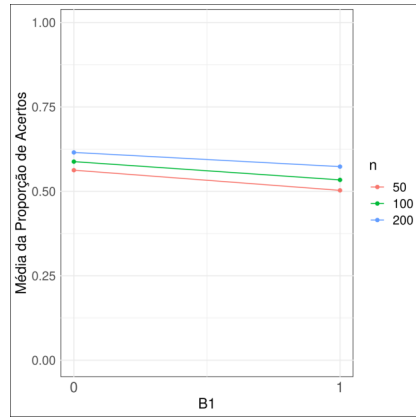


Figura 67: Interação entre $B1$ e n para ARMA(0,1) para AIC

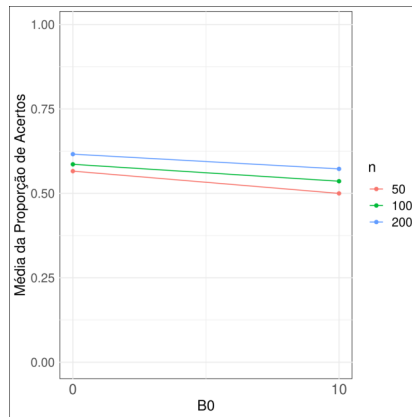


Figura 68: Interação entre $B0$ e n e para ARMA(0,1) para AIC

Neste modelo, os efeitos das interações de θ com $B1$, $B0$ e n , representadas nas figuras 64, 65 e 66, respectivamente, são semelhantes aos ocorridos no modelo ARMA(1,1), conforme figuras 56, 57 e 58, respectivamente. Já o efeito de $B1$ com n , representado no gráfico da Figura 67, é similar ao ocorrido nos modelos ARMA(1,1) e ARMA(1,0) conforme figuras 60 e 63, respectivamente. Diferentemente, o efeito da interação $B0$ com n ocorreu apenas neste tipo de estrutura, sendo a diferença entre as proporções acertos para $n = 50$, $n = 100$ e $n = 200$ menores quando $B0 = 0$, conforme observado na Figura 68.

a.4) Considerações: O critério AIC encontrado pelo modelo ARMA(1,0) apresenta comportamento mais robusto. O efeito principal de $B1$ não foi significativo em dois modelos (ARMA(0,1) e ARMA(1,1)), mas em ARMA(0,1) ele interagiu com θ e com n e em ARMA(1,1) interagiu com ϕ com θ e com n . O coeficiente $B0$, por sua vez, não foi significativo em ARMA(0,1), mas interagiu com θ e com n .

b) Critério AICc: A Tabela 14 mostra os resultados finais obtidos para o critério AICc, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 69 mostra os respectivos gráficos envelope.

Tabela 14: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação AICc.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-0.9318	0.0558	-16.7042	<0.01	ARMA(0,1)
θ	2.976	0.103	28.8957	<0.01	
$B1$	0.0515	0.0445	1.1564	0.2475	
$B0$	0.0081	0.0053	1.5139	0.1301	
n	0.0031	4e-04	8.7653	<0.01	
$\theta:B1$	-0.3143	0.08	-3.9292	<0.01	
$\theta:B0$	-0.0414	0.008	-5.1776	<0.01	
$\theta:n$	-0.0059	6e-04	-9.1845	<0.01	
$B1:B0$	-0.015	0.0032	-4.7107	<0.01	
$B0:n$	0,0003	0,0003	0.5763	0.5644	
(Intercept)	-0.8369	0.0413	-20.2481	<0.01	ARMA(1,0)
ϕ	2.8509	0.0762	37.4046	<0.01	
$B1$	-0.0721	0.0203	-3.5556	<0.01	
$B0$	-0.0081	0.002	-3.9825	<0.01	
n	0.0031	3e-04	10.3412	<0.01	
$\phi:n$	-0.0061	6e-04	-10.5767	<0.01	
$B1:B0$	-0.0128	0.0028	-4.4917	<0.01	
(Intercept)	-2.6256	0.084	-31.2576	<0.01	ARMA(1,1)
ϕ	2.2462	0.1251	17.9505	<0.01	
θ	2.4909	0.1354	18.4018	<0.01	
$B1$	0.0474	0.043	1.1041	0.2695	
$B0$	0.0056	0.0043	1.3094	0.1904	
n	0.0064	5e-04	13.8527	<0.01	
$\phi:\theta$	1.3452	0.1910	7.0428	<0.01	
$\phi:n$	-0.0049	6e-04	-8.0607	<0.01	
$\theta:B1$	-0.3073	0.0759	-4.050	<0.01	
$\theta:B0$	-0.0277	0.0076	-3.6477	<0.01	
$\theta:n$	-0.0049	6e-04	-8.0176	<0.01	
$B1:B0$	-0.0137	0.003	-4.5211	<0.01	

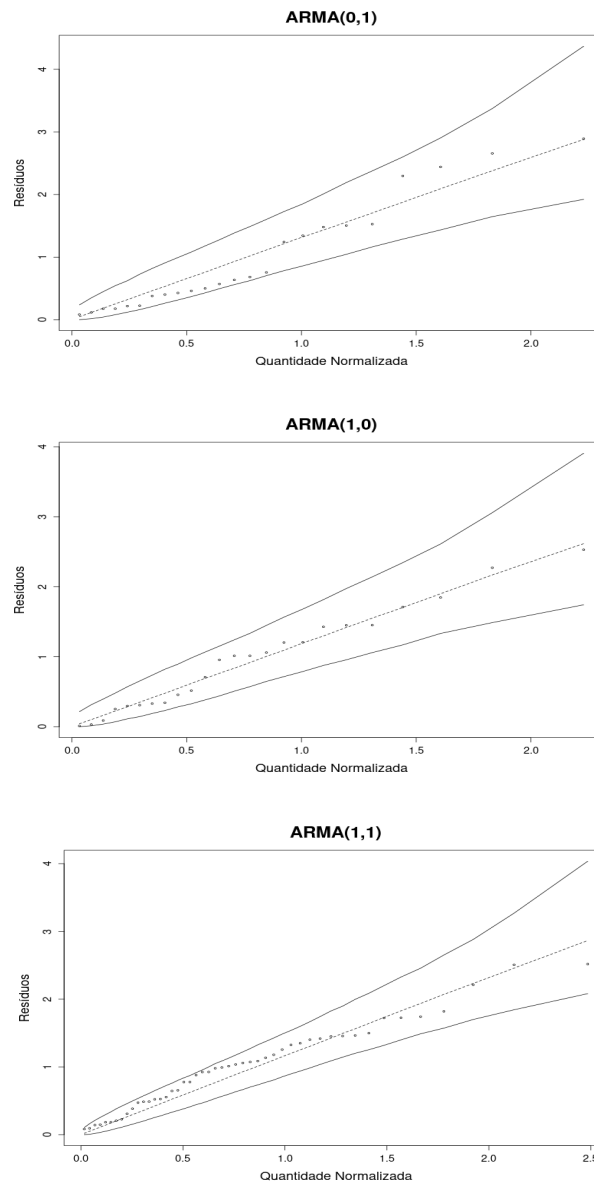


Figura 69: Gráficos envelope para modelos Beta de AICc de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

b.1) ARMA(1,1): De acordo com a Tabela 14, para o critério AICc para o modelo ARMA(1,1), apenas os efeitos principais de B_0 e B_1 não foram significativos. Existem evidências de que as proporções de acertos do critério AICc são

maiores para $\theta = 0.7$ (0,6329), $\phi = 0.7$ (0,6352) e $n = 200$ (0,556). Com relação às interações, existem evidências de que os resultados das proporções de acertos do AICc são afetados pela interação de ϕ com θ e n , de θ com $B1$, e de $B0$ com n conforme figuras 70, 71, 72, 73 e 74.

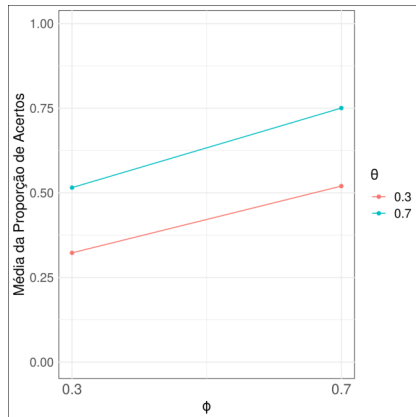


Figura 70: Interação entre ϕ e θ para ARMA(1,1) para AICc

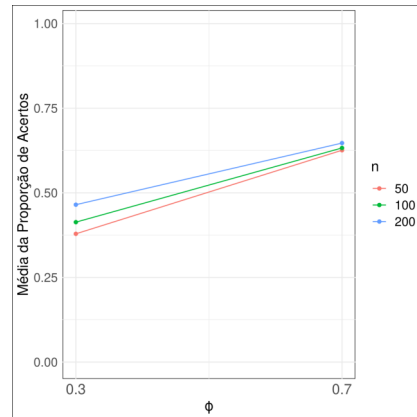


Figura 71: Interação entre ϕ e n para ARMA(1,1) para AICc

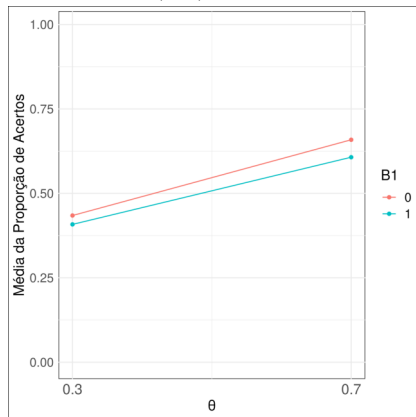


Figura 72: Interação entre θ e $B1$ para ARMA(1,1) para AICc

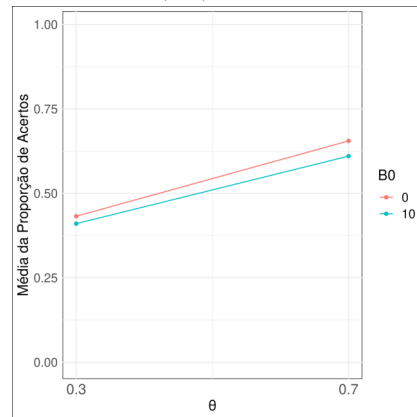


Figura 73: Interação entre θ e $B0$ para ARMA(1,1) para AICc

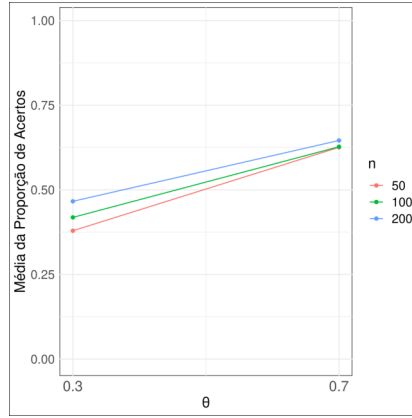


Figura 74: Interação entre θ e n para ARMA(1,1) para AICc

No gráfico da Figura 70, a diferença entre as proporções de acertos para $\theta = 0.3$ e $\theta = 0.7$ é menor quando $\phi = 0.3$. Da mesma maneira, no gráfico da Figura 71, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi = 0.7$. Já no gráfico da Figura 72, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\theta = 0.3$. Enquanto que no gráfico da Figura 73, as diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta = 0.3$. Por fim, no gráfico da Figura 74, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\theta = 0.7$.

b.2) ARMA(1,0): De acordo com a Tabela 14, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), os efeitos principais são todos significativos. Existem evidências que a média das proporções é maior para $\phi = 0.7$ (0.7134), para $B0=10$ (0.5974) e para $B1=0$ (0.5984). Com relação às interações, existem evidências de que os resultados das proporções de acertos do AICc são afetados pela interação de ϕ com n e de $B1$ com $B0$, conforme observado nas figuras 75 e 76.

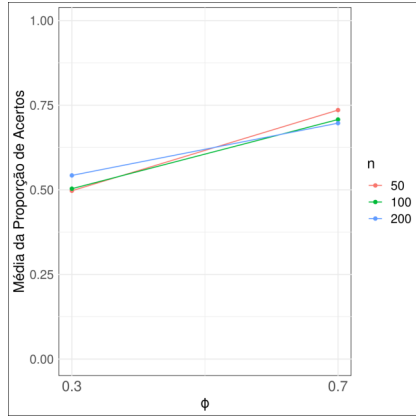


Figura 75: Interação entre ϕ e n para ARMA(1,0) para AICc

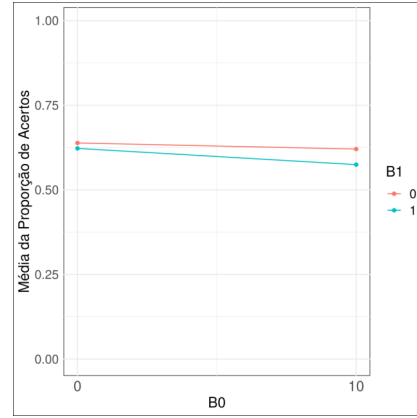


Figura 76: Interação entre $B1$ e $B0$ para ARMA(1,0) para AICc

No gráfico da Figura 75, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\phi=0.7$. Além disso, em $\phi=0.3$, a maior proporção ocorreu para $n=200$, enquanto que para $\phi=0.7$, foi para $n=50$. Já no gráfico da Figura 76, a diferença entre as proporções de acertos para $B1=0$ e $B1=1$ é menor quando $B0=0$.

b.3) ARMA(0,1): De acordo com a Tabela 14, para o critério AICc se for retirado apenas o termo Autorregressivo ficando o modelo do tipo ARMA(0,1) os efeitos principais de $B0$ e $B1$ deixam de ser significativos. Existem evidências de que as proporções de acertos do critério AICc são maiores para $\theta=0.7$ (0,6925) e $n=200$ (0,6111). Com relação às interações, existem evidências de que os resultados das proporções de acertos do AICc são afetados pela interação de θ com $B0$, de $B1$ com n , e de $B0$ com $B1$, conforme observado nas figuras 77, 78 e 79. Ainda de acordo com a Tabela 14, a interação entre $B0$ e $B1$ também foi significativa, mas como os efeitos principais desses fatores não o foram, a análise dessa interação não foi considerada.

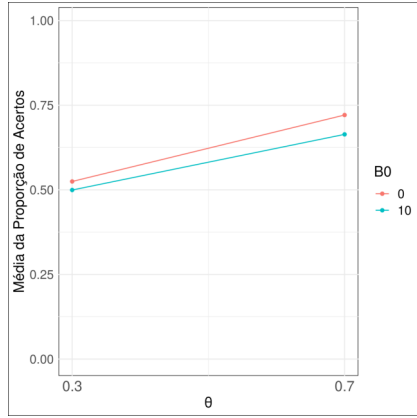


Figura 77: Interação entre θ e $B0$ para ARMA(0,1) para AICc

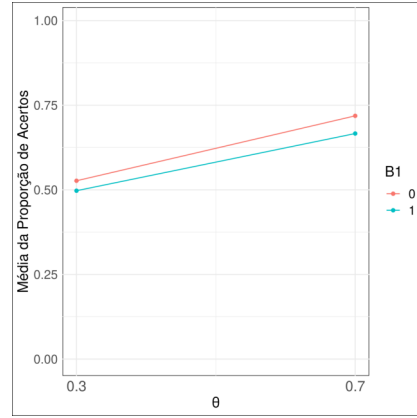


Figura 78: Interação entre θ e $B1$ para ARMA(0,1) para AICc

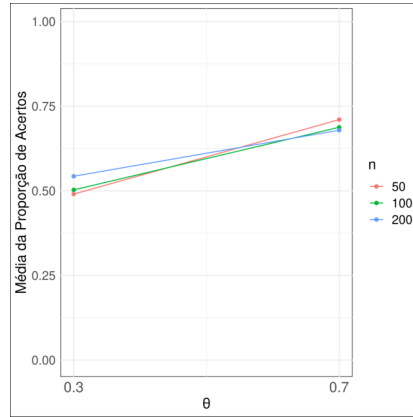


Figura 79: Interação entre θ e n para ARMA(0,1) para AICc

Neste modelo, os efeitos das interações de θ com $B1$, $B0$ e n , representadas nas figuras 77, 78 e 79, respectivamente, são semelhantes aos ocorridos no modelo ARMA(1,1), conforme figuras 72, 73 e 74. Na análise da interação entre θ e n , observou-se um cruzamento dessas variáveis. Para $n = 50$, isso sugere que, ao aumentar o valor de θ de 0.3 para 0.7, a proporção média de acertos tende a se tornar maior quando comparada a proporção de $n = 100$ e $n = 200$.

b.4) Considerações: O critério AICc encontrado pelo modelo ARMA(1,0) apresenta comportamento mais robusto. Os efeitos principais de $B1$ e $B0$ não foram significativos nos modelos ARMA(0,1) e ARMA(1,1). No entanto, no modelo ARMA(0,1), $B1$ e $B0$ interagiram com θ em ambos os modelos.

Tanto o critério AIC quanto o AICc apresentaram comportamentos semelhantes. Em ambos os casos, a maioria dos casos analisados mostrou que, apesar das interações serem significativas, os gráficos indicam que sua magnitude pode ser relativamente baixa. Além disso, observa-se que a diferença nas proporções entre os níveis de um fator é ligeiramente reduzida em um dos níveis do outro fator considerado na interação.

c) Critério BIC: A Tabela 15 mostra os resultados finais obtidos para o critério BIC, para cada um dos modelos Arma(0,1), Arma(1,0) e Arma(1,1), enquanto a Figura 80 mostra os respectivos gráficos envelope.

Tabela 15: Resultados da Regressão Beta para o critério BIC.

Componente	Coefficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-1.0291	0.1052	-9.7862	<0.01	ARMA(0,1)
θ	4.5166	0.2437	18.5357	<0.01	
$B1$	0.4121	0.0882	4.673	<0.01	
$B0$	0.033	0.0088	3.7545	<0.01	
n	0.0018	7e-04	2.5223	0.0117	
$\theta:B1$	-0.9557	0.193	-4.9521	<0.01	
$\theta:B0$	-0.0847	0.0193	-4.3874	<0.01	
$\theta:n$	0.0075	0.0017	4.4126	<0.01	
$B1:B0$	-0.0197	0.0066	-2.9842	0.0028	
(Intercept)	-0.9201	0.0738	-12.4635	<0.01	ARMA(1,0)
ϕ	4.2957	0.1694	25.3578	<0.01	
$B1$	-0.0661	0.0353	-1.8742	0.0609	
$B0$	0.0032	0.0067	0.4736	0.6358	
n	0.0014	5e-04	2.6154	0.0089	
$\phi:B0$	-0.0333	0.0148	-2.2502	0.0244	
$\phi:n$	0.0079	0.0013	6.0088	<0.01	
$B1:B0$	-0.0139	0.0049	-2.8474	0.0044	
(Intercept)	-7.3531	0.3687	-19.9431	<0.01	ARMA(1,1)
ϕ	7.2072	0.4224	17.0608	<0.01	
θ	7.2972	0.4226	17.2671	<0.01	
$B1$	-0.1403	0.0675	-2.0779	0.0377	
$B0$	-0.013	0.0068	-1.9266	0.0540	
n	0.022	0.0024	9.1025	<0.01	
$\phi:n$	-0.0164	0.003	-5.5045	<0.01	
$\theta:n$	-0.0161	0.003	-5.416	<0.01	

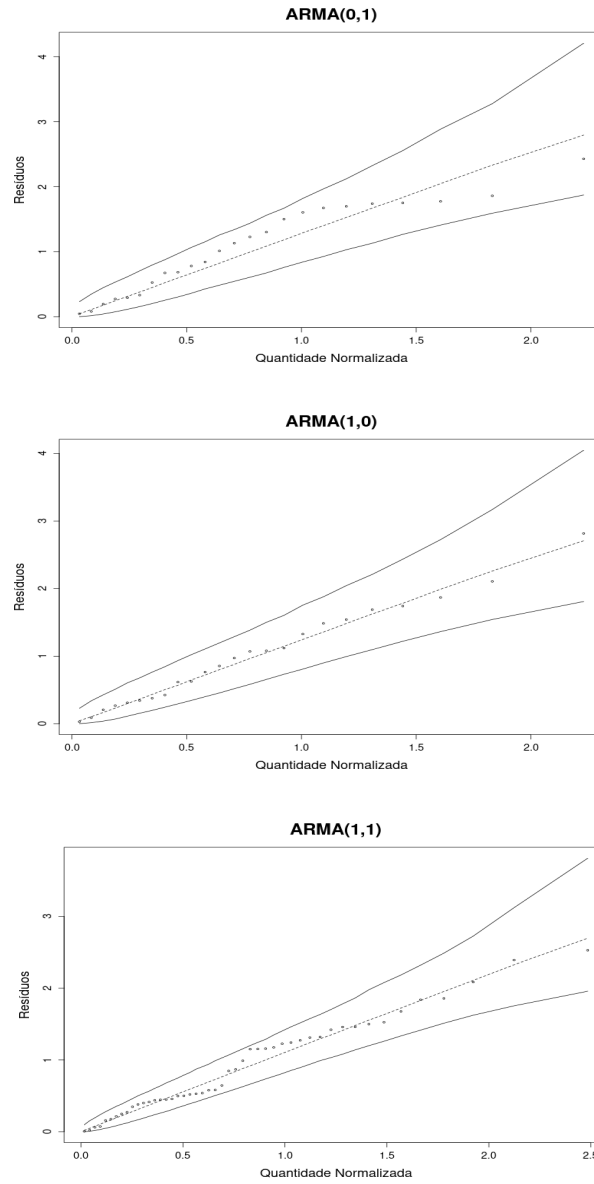


Figura 80: Gráficos envelope para modelos Beta de BIC de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

c.1) ARMA(1,1): De acordo com a Tabela 15, para o modelo ARMA(1,1), apenas o efeito principal de B_0 não foi significativo. Existem evidências de que as proporções de acertos do critério BIC são maiores para $\theta=0.7$ (0,7677), para

$\phi=0,7$ (0,7622) e para $n=200$ (0,6678). Com relação às interações, existem evidências de que os resultados das proporções de acertos do BIC são afetados pela interação de ϕ com n e de θ com n , conforme observado nas figuras 81 e 82.

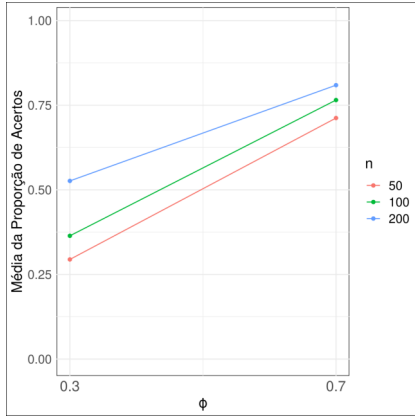


Figura 81: Interação entre ϕ e n para ARMA(1,1) para BIC

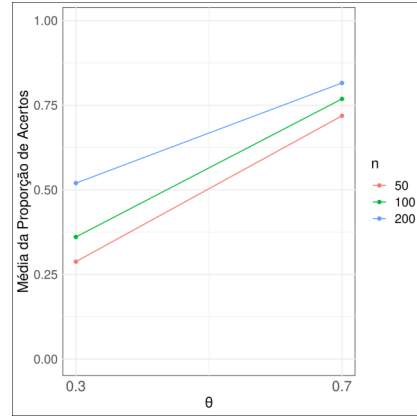


Figura 82: Interação entre θ e n para ARMA(1,1) para BIC

No gráfico da Figura 81, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\phi=0.7$.

No gráfico da Figura 82, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\theta=0.7$.

c.2) ARMA(1,0): De acordo com a Tabela 15, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), os efeitos principais dos fatores $B1$ e $B0$ não são significativos. Existem evidências de que as proporções de acertos do critério BIC são maiores para $\phi=0.7$ (0,9332) e para $n=200$ (0,8468). Além disso, ϕ interagiu com $B0$, ϕ interagiu com n , e $B1$ interagiu com $B0$ conforme evidenciado nas figuras 83 e 84. Como os efeitos principais de $B1$ e de $B0$ não foram significativos, foram retirados da análise.

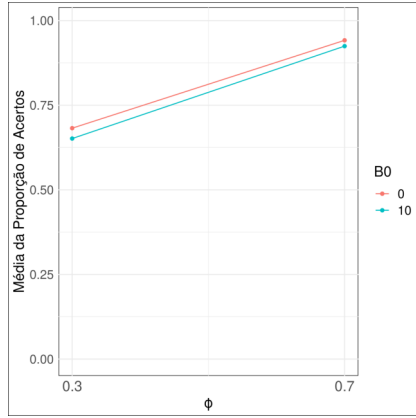


Figura 83: Interação entre ϕ e B_0 para ARMA(1,0) para BIC

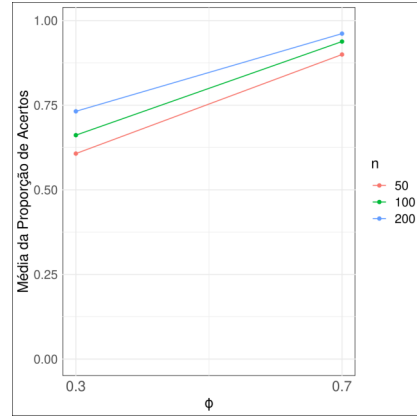


Figura 84: Interação entre ϕ e n para ARMA(1,0) para BIC

No gráfico da Figura 83, a diferença entre as proporções de acertos para $B_0=0$ e $B_0=10$ é menor quando $\phi=0.7$. Para o gráfico da Figura 84, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi=0.7$, semelhante ao que ocorre com o gráfico da Figura 81 para ARMA(1,1).

c.3) Critério ARMA(0,1): De acordo com a Tabela 15, se for retirado apenas o termo Autorregressivo, ficando o modelo do tipo ARMA(0,1), todos os efeitos principais se tornam significativos. Existem evidências de que as proporções de acertos do critério BIC são maiores para $\theta=0.7$ (0,9266), $B_0=0$ (0,8194), $B_1=0$ (0,8169) e para $n = 200$ (0,8590). Com relação às interações, existem evidências de que os resultados das proporções de acertos do BIC são afetados pela interação de θ com B_1 , de θ com B_0 , de θ com n e de B_1 com B_0 , conforme observado nas figuras 85, 86, 87 e 88.

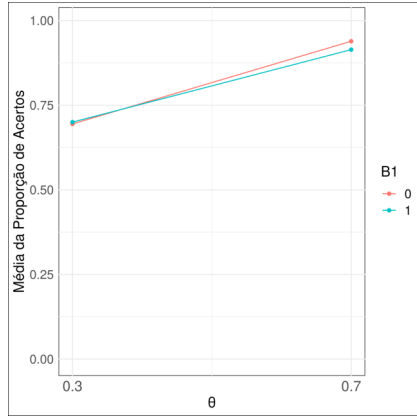


Figura 85: Interação entre θ e $B1$ para ARMA(0,1) para BIC

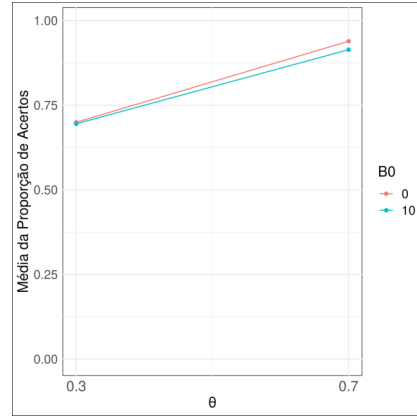


Figura 86: Interação entre θ e $B0$ para ARMA(0,1) para BIC

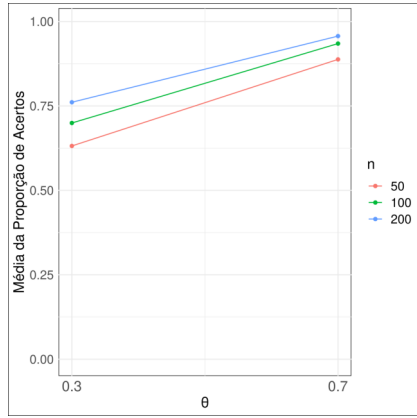


Figura 87: Interação entre θ e n para ARMA(0,1) para BIC

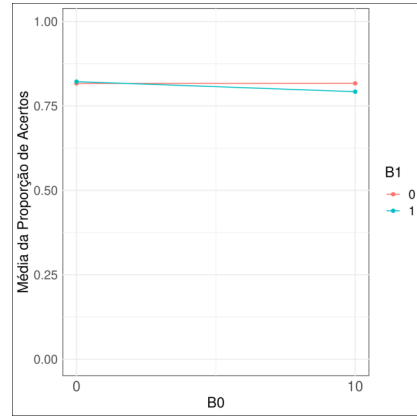


Figura 88: Interação entre $B1$ e $B0$ para ARMA(0,1) para BIC

No gráfico da Figura 85, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\theta=0.3$. Já para o gráfico da Figura 86, as diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta=0.3$. Para o gráfico da Figura 87, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 82 para ARMA(1,1), onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. E por fim, para o gráfico da Figura 88, a diferença entre as proporções de acertos para $B1=0$ e $B1=1$ é menor quando $B0=0$. Além disso, em $B0=0$, a maior proporção ocorreu para $B1=1$, enquanto que para $B0=10$ foi para $B1=0$.

c.4) Considerações: O critério BIC encontrado pelo modelo ARMA(1,1) apresenta comportamento mais robusto. O efeito principal de $B0$ não foi significativo no modelo ARMA(1,0), no entanto, ele interagiu com ϕ .

d) Critério T1: A Tabela 16 apresenta as razões de chance, erros padrão, valores de z e valores-p correspondentes aos coeficientes principais do modelo Beta para o critério T1.

Tabela 16: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T1.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-0.9975	0.1513	-6.592	<0.01	ARMA(0,1)
θ	3.9101	0.3378	11.5745	<0.01	
$B1$	0.3916	0.1271	3.0807	0.0021	
$B0$	0.0324	0.0127	2.553	0.0107	
n	0.0019	0.001	1.8471	0.0647	
$\theta:B1$	-1.0228	0.2677	-3.8209	<0.01	
$\theta:B0$	-0.0925	0.0268	-3.4554	<0.01	
$\theta:n$	0.0102	0.0024	4.2191	<0.01	
$B1:B0$	-0.0239	0.0096	-2.4972	0.0125	
(Intercept)	-0.8711	0.1088	-8.0094	<0.01	ARMA(1,0)
ϕ	3.7465	0.2461	15.2235	<0.01	
$B1$	0.0481	0.0913	0.5268	0.5983	
$B0$	0.0042	0.0091	0.4634	0.6431	
n	0.0011	7e-04	1.4253	0.1541	
$\phi:B1$	-0.4061	0.1951	-2.0819	0.0374	
$\phi:B0$	-0.0458	0.0195	-2.3461	0.019	
$\phi:n$	0.0112	0.0018	6.2777	<0.01	
$B1:B0$	-0.0133	0.0068	-1.9474	0.0515	
(Intercept)	-5.1128	0.3654	-13.9917	<0.01	ARMA(1,1)
ϕ	4.635	0.5435	8.5279	<0.01	
θ	4.657	0.5432	8.5728	<0.01	
$B1$	-0.1539	0.0604	-2.5506	0.0108	
$B0$	-0.0141	0.006	-2.3407	0.0192	
n	0.0132	0.0021	6.4344	<0.01	
$\phi:\theta$	1.5451	0.8322	1.8565	0.0634	
$\phi:n$	-0.0077	0.0026	-2.974	0.0029	
$\theta:n$	-0.0072	0.0026	-2.7558	0.0059	

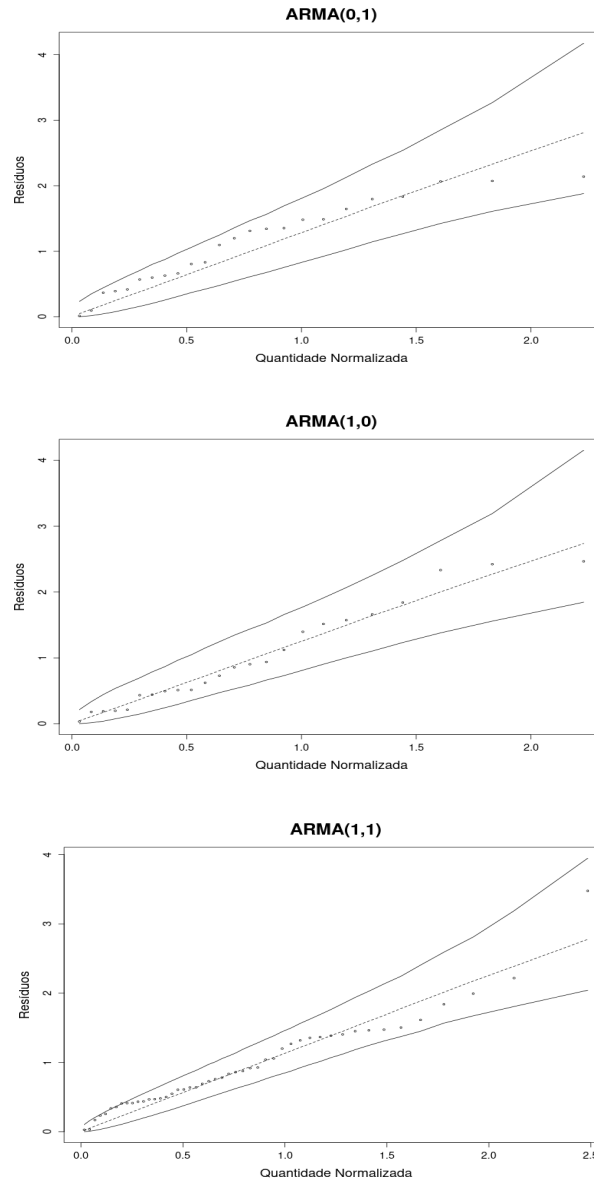


Figura 89: Gráficos envelope para modelos Beta de T1 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

d.1) ARMA(1,1): De acordo com a Tabela 16, para o modelo ARMA(1,1), todos os efeitos principais, com exceção de n , foram significativos. Existem evidências de que as proporções de acertos do critério T1 são maiores para $\theta=0.7$

(0,7506), $\phi=0.7$ (0,7465), $B0=0$ (0,5916), $B1=0$ (0,5929) e $n=200$ (0,6684). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T1 são afetados pela interação de ϕ com θ e com n e de θ com n , conforme figuras 90, 91 e 92.

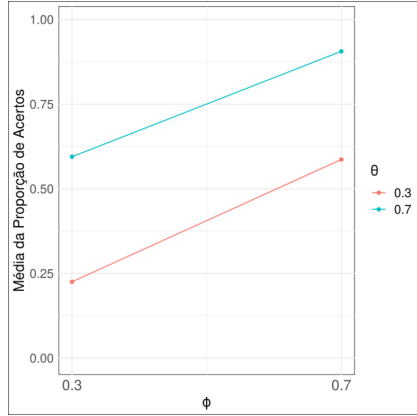


Figura 90: Interação entre ϕ e θ para ARMA(1,1) para T1

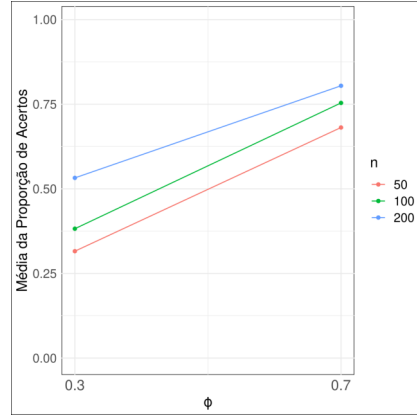


Figura 91: Interação entre ϕ e n para ARMA(1,1) para T1

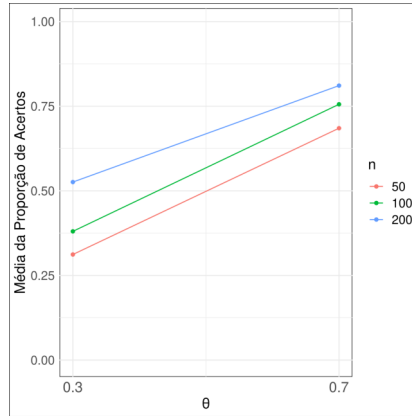


Figura 92: Interação entre θ e n para ARMA(1,1) para T1

No gráfico da Figura 90, a diferença entre as proporções de acertos para $\theta=0.3$ e $\theta=0.7$ é menor quando $\phi=0.7$. Para o gráfico da Figura 91, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\phi=0.7$. Já para o gráfico da Figura 92, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\theta=0.7$.

d.2) ARMA(1,0): De acordo com a Tabela 16, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), o efeito principal de ϕ é significativo. Existem evidências de que as proporções de acertos do critério T1 são maiores para $\phi=0.7$ (0.9133). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T1 são afetados pela interação de ϕ com $B1$, de ϕ com $B0$ e de ϕ com n conforme observado nas figuras 93, 94 e 95.

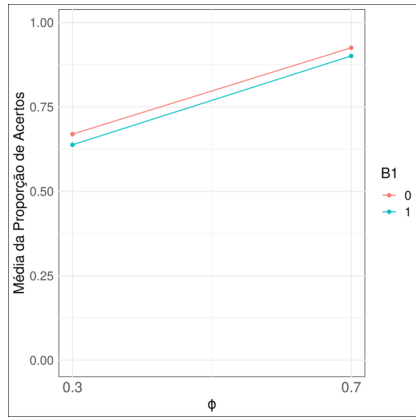


Figura 93: Interação entre ϕ e $B1$ para ARMA(1,0) para T1

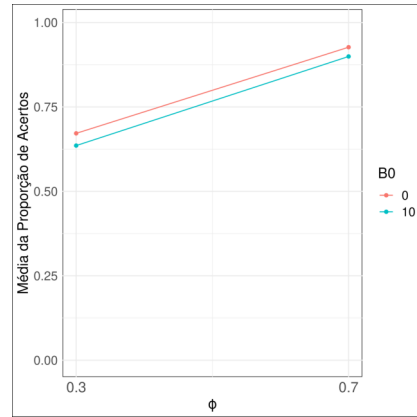


Figura 94: Interação entre ϕ e $B0$ para ARMA(1,0) para T1

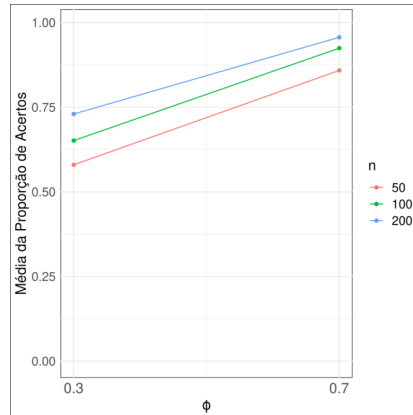


Figura 95: Interação entre ϕ e n para ARMA(1,0) para T1

No gráfico da Figura 93, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\phi=0.3$. Para o gráfico da Figura 94, as

diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta=0.3$. Enquanto que para o gráfico da Figura 95, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 91 para ARMA(1,1) com relação a T1, onde se conclui que estes fatores interagem em modelos ARMA(1,0) da mesma maneira.

d.3) ARMA(0,1): De acordo com a Tabela 16, se for retirado apenas o termo Autoregressivo, ficando o modelo do tipo ARMA(0,1), o efeito principal de n deixa de ser significativo. Existem evidências de que as proporções de acertos do critério T1 são maiores para $\theta=0.7$ (0,9045), $B1=0$ (0,8028) e $B0=0$ (0,8048). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T1 são afetados pela interação de θ com $B1$, de θ com $B0$, de θ com n e de $B1$ com $B0$, conforme observado nas figuras 96, 97, 98 e 99.

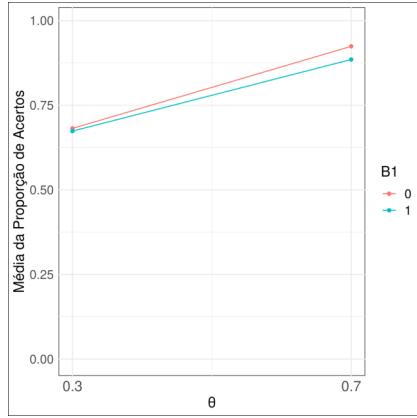


Figura 96: Interação entre θ e $B1$ para ARMA(0,1) para T1

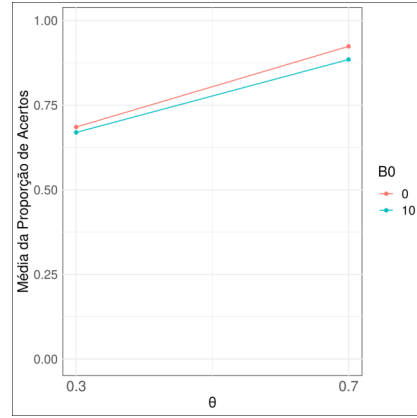


Figura 97: Interação entre θ e $B0$ para ARMA(0,1) para T1

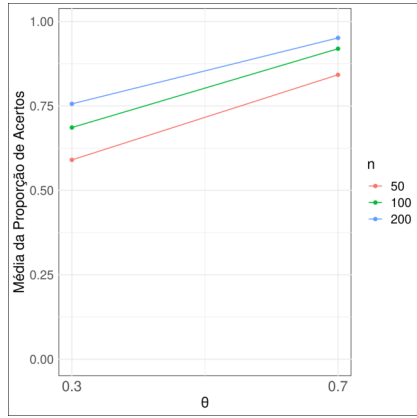


Figura 98: Interação entre θ e n para ARMA(0,1) para T1

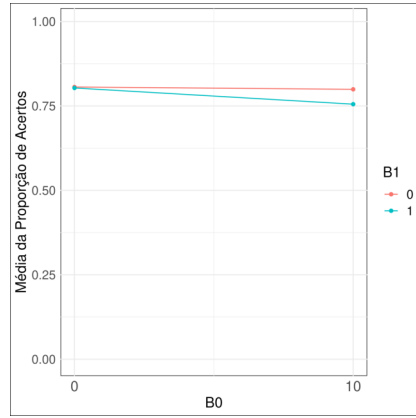


Figura 99: Interação entre $B1$ e $B0$ para ARMA(0,1) para T1

No gráfico da Figura 96, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\theta=0.3$. Já para o gráfico da Figura 97, as diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta=0.3$. Enquanto que para o gráfico da Figura 98, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 92 para ARMA(1,1) com relação a T1, onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. Por fim para o gráfico da Figura 99, a diferença entre as proporções de acertos para $B1=0$ e $B1=1$ é menor quando $B0=0$.

e) Critério T2: A Tabela 17 mostra os resultados finais obtidos para o critério T2, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 100 mostra os respectivos gráficos envelope.

Tabela 17: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T2.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-0.955	0.2062	-4.6316	<0.01	ARMA(0,1)
θ	3.471	0.4495	7.7225	<0.01	
$B1$	0.2328	0.1634	1.425	0.1541	
$B0$	0.021	0.0163	1.2836	0.1993	
n	0.0022	0.0014	1.5572	0.1194	
$\theta:B1$	-1.0717	0.353	-3.0356	0.0024	
$\theta:B0$	-0.1072	0.0353	-3.0324	0.0024	
$\theta:n$	0.0121	0.0033	3.7324	<0.01	
(Intercept)	-0.5695	0.1475	-3.8601	<0.01	ARMA(1,0)
ϕ	2.6557	0.3083	8.6132	<0.01	
$B1$	-0.2287	0.0555	-4.1216	<0.01	
$B0$	-0.0251	0.0055	-4.5154	<0.01	
n	0.0012	0.0012	0.9867	0.3238	
$\phi:n$	0.0131	0.0028	4.642	<0.01	
(Intercept)	-3.7268	0.2309	-16.1435	<0.01	ARMA(1,1)
ϕ	3.0126	0.3806	7.9156	<0.01	
θ	3.1197	0.381	8.1877	<0.01	
$B1$	0.4185	0.2178	1.9212	0.0547	
$B0$	-0.0157	0.0054	-2.9113	0.0036	
n	0.0059	0.0004	13.4872	<0.01	
$\phi:\theta$	2.3016	0.7124	3.2309	0.0012	
$\phi:B1$	-0.5606	0.2800	-2.0019	0.0453	
$\theta:B1$	-0.6638	0.2801	-2.3697	0.0178	

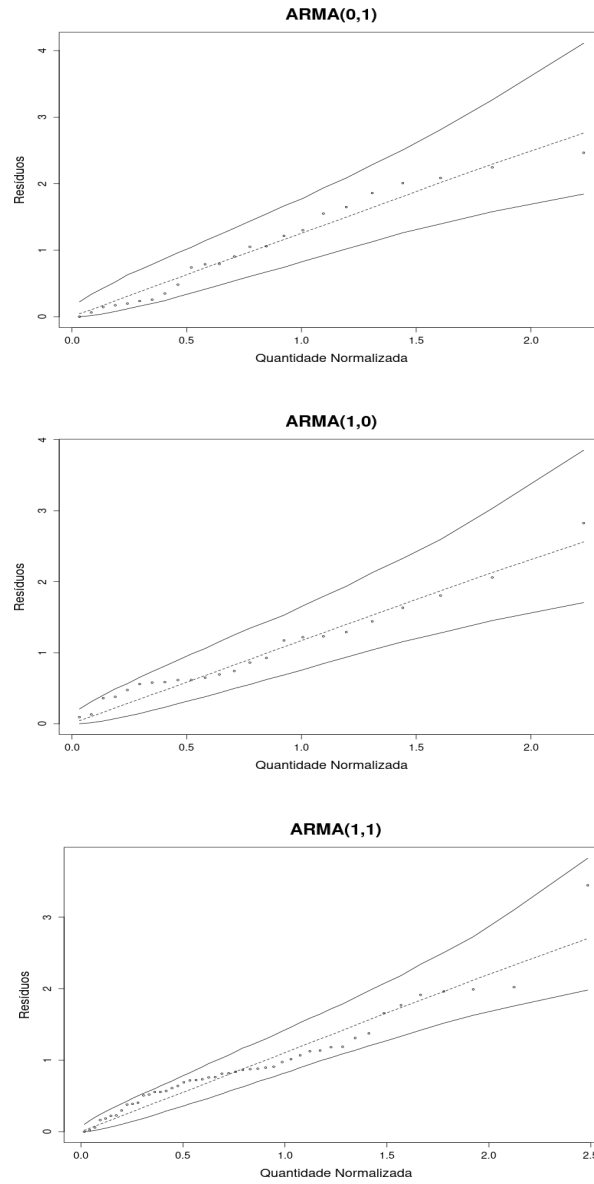


Figura 100: Gráficos envelope para modelos Beta de T2 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

e.1) ARMA(1,1): De acordo com a Tabela 17, para o modelo ARMA(1,1), o efeito principais de $B1$ não foi significativo. Existem evidências de que as proporções de acertos do critério T2 são maiores para $\theta=0.7$ (0,7322), $\phi=0.7$

(0,7295) e $n=200$ (0,6682). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T2 são afetados pela interação de ϕ com θ e com $B1$ e de θ com $B1$, conforme observado nas figuras 101, 102 e 103.

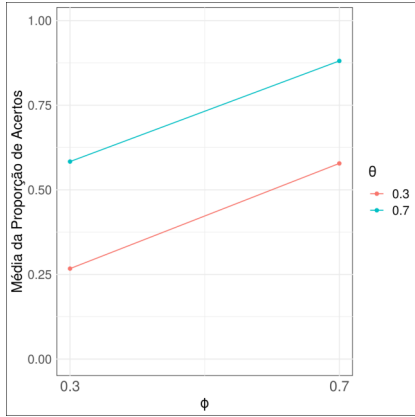


Figura 101: Interação entre ϕ e θ para ARMA(1,1) para T2

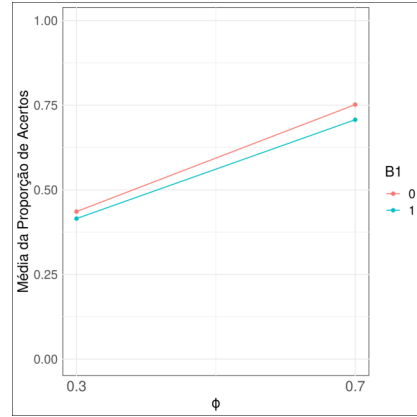


Figura 102: Interação entre ϕ e $B1$ para ARMA(1,1) para T2

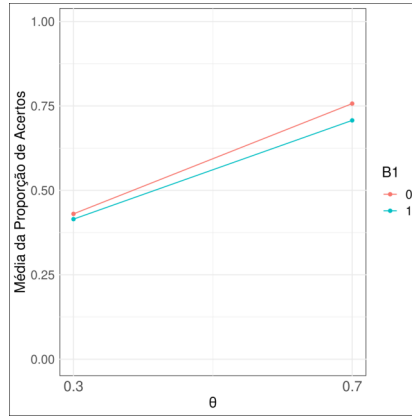


Figura 103: Interação entre θ e $B1$ para ARMA(1,1) para T2

No gráfico da Figura 101, a diferença entre as proporções de acertos para $\theta=0.3$ e $\theta=0.7$ é menor quando $\phi=0.3$. Já no gráfico da Figura 102, a diferença entre as proporções de acertos para $B1=0$ e $B1=1$ é menor quando $\phi=0.3$. Para o gráfico da Figura 103, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\theta=0.3$.

e.2) ARMA(1,0): De acordo com a Tabela 17, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), em relação aos efeitos principais somente o efeito para n não foi significativo e para este existem evidências de que a proporção de acerto do critério T2 é maior para $\phi=0.7$ (0.8906), $B0=0$ (0.7846) e $B1=0$ (0.7829). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T2 somente é afetado pela interação de ϕ com n , conforme observado na Figura 104.

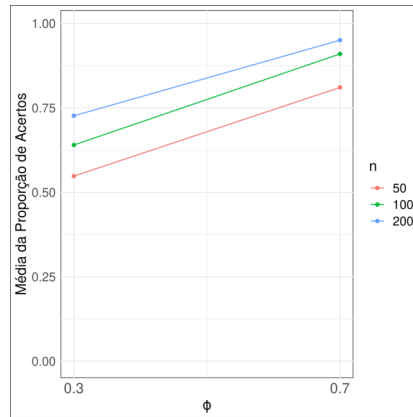


Figura 104: Interação entre ϕ e n para ARMA(1,0) para T2

No gráfico da Figura 104, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\phi=0.7$.

e.3) ARMA(0,1): De acordo com a Tabela 17, se for retirado apenas o termo Autorregressivo, ficando o modelo do tipo ARMA(0,1), somente o efeito de θ foi significativo. Existem evidências de que as proporções de acertos do critério T2 são maiores para $\theta=0.7$. Com relação às interações, existem evidências de que os resultados das proporções de acertos do T2 são afetados pela interação de θ com $B1$, com $B0$ e com n , conforme observado nas figuras 105, 106 e 107.

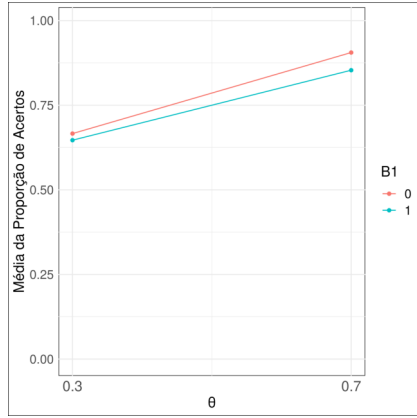


Figura 105: Interação entre θ e $B1$ para ARMA(0,1) para T2

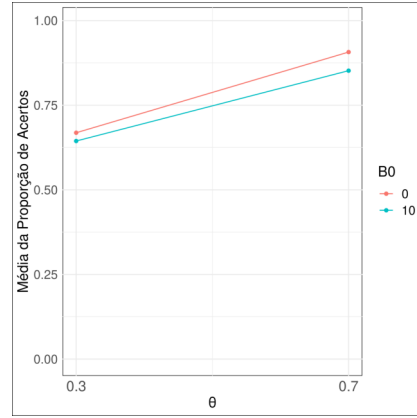


Figura 106: Interação entre θ e $B0$ para ARMA(0,1) para T2

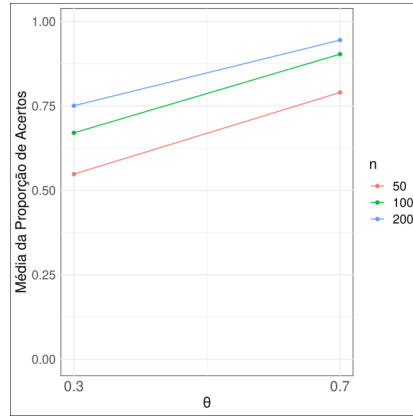


Figura 107: Interação entre θ e n para ARMA(0,1) para T2

O gráfico da Figura 105, que representa a interação entre θ e $B1$ é muito semelhante ao gráfico da Figura 103 para ARMA(1,1) com relação a T2, onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. Para o gráfico da Figura 106, as diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta=0.3$. Enquanto que para o gráfico da Figura 107, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\theta=0.7$.

e.4) Considerações: O critério T2 encontrado pelo modelo ARMA(1,0) apresenta comportamento mais robusto. O efeito principal de n não foi significativo

em dois modelos, mas em ARMA(0,1) ele interagiu com θ e em ARMA(1,0) interagiu com ϕ . O coeficiente θ , quando presente, interagiu com $B1$, sendo a proporção de acertos maior quando $B1=0$. Essa diferença aumenta para maiores valores de θ .

f) Critério T3: A Tabela 18 mostra os resultados finais obtidos para o critério T3, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 108 mostra os respectivos gráficos envelope.

Tabela 18: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T3.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-1.3165	0.0445	-29.5699	<0.01	ARMA(0,1)
θ	6.0562	0.0932	64.9741	<0.01	
$B1$	0.4666	0.0527	8.8532	<0.01	
$B0$	0.0399	0.0053	7.5770	<0.01	
n	0.0032	2e-04	14.6761	<0.01	
$\theta:B1$	-0.571	0.1052	-5.4268	<0.01	
$\theta:B0$	-0.0599	0.0105	-5.6892	<0.01	
$B1:B0$	-0.0148	0.0031	-4.7543	<0.01	
$B1:n$	-9e-04	3e-04	-3.4934	<0.01	
$B0:n$	-1e-04	2e-05	-2.5855	0.0097	
(Intercept)	-1.4102	0.0455	-31.0221	<0.01	ARMA(1,0)
ϕ	6.4526	0.0883	73.0810	<0.01	
$B1$	-0.0117	0.0342	-0.3435	0.7312	
$B0$	-0.0046	0.0034	-1.3498	0.1771	
n	0.0028	2e-04	14.2855	<0.01	
$B1:B0$	-0.0112	0.0048	-2.3451	0.0190	
(Intercept)	-12.2989	0.6715	-18.3151	<0.01	ARMA(1,1)
ϕ	12.1502	0.7110	17.0898	<0.01	
θ	12.4186	0.7113	17.4594	<0.01	
$B1$	-0.1473	0.0799	-1.8436	0.0652	
$B0$	-0.0127	0.0080	-1.5865	0.1126	
n	0.0448	0.0038	11.6648	<0.01	
$\phi:n$	-0.0391	0.0043	-9.075	<0.01	
$\theta:n$	-0.0398	0.0043	-9.2293	<0.01	

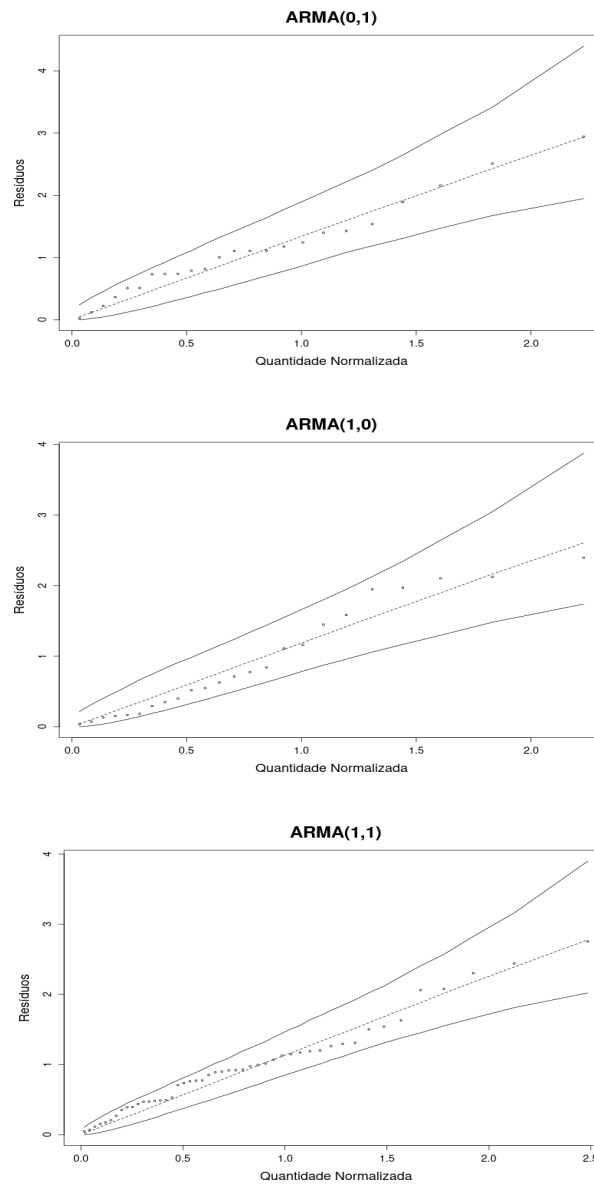


Figura 108: Gráficos envelope para modelos Beta de T3 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

f.1) ARMA(1,1): De acordo com a Tabela 18, para o modelo ARMA(1,1), apenas os efeitos principais de B_0 e B_1 não foram significativos. Existem evidências de que as proporções de acertos do critério T3 são maiores para

$\theta = 0.7$ (0,7925) e $\phi = 0.7$ (0,7836) e $n = 200$ (0,6623). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T3 são afetados pela interação de ϕ com n e de θ com n , conforme observado nas figuras 109 e 110.

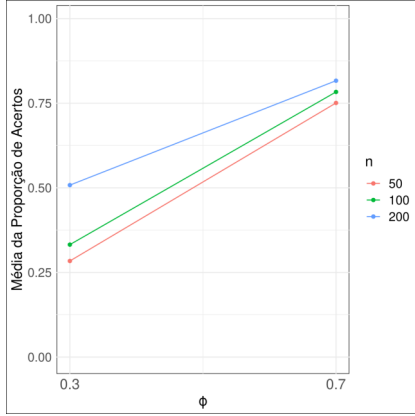


Figura 109: Interação entre ϕ e n para ARMA(1,1) para T3

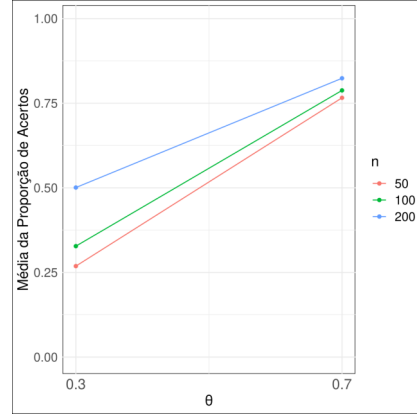


Figura 110: Interação entre θ e n para ARMA(1,1) para T3

No gráfico da Figura 109, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\phi=0.7$. Já para o gráfico da Figura 110, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $\theta=0.7$.

f.2) ARMA(1,0): De acordo com a Tabela 18, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), somente os fatores ϕ e n tiveram efeitos significativos. Existem evidências de que as proporções de acertos do critério T3 são maiores para $\phi=0.7$ (0,9663) e $n=200$ (0,8555). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T3 são afetados pela interação de $B1$ com $B0$ somente, no entanto, como as interações dos efeitos principais de $B1$ e de $B0$ não são significativos, decidiu-se por não realizar-se a análise dessa interação.

f.3) ARMA(0,1): De acordo com a Tabela 18, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(0,1), os efeitos principais deram significativos. Existem evidências de que as proporções de acertos do critério T3 são maiores para $\theta=0.7$ (0,9599), $B0=10$ (0,8473), $B1=1$ (0,8503) e $n=200$ (0,8684). Com relação às interações, existem evidências de que os

resultados das proporções de acertos do T3 são afetados pela interação de θ com $B1$ e com $B0$, de $B1$ com $B0$ e com n , e de $B0$ com n , conforme observado nas figuras 111, 112, 113, 114 e 115.

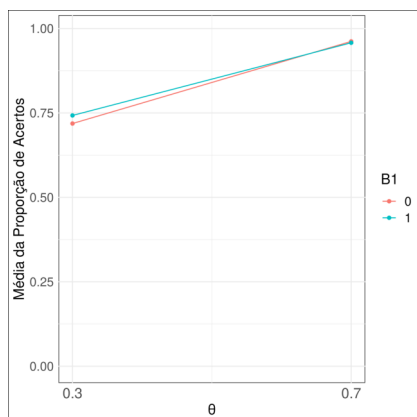


Figura 111: Interação entre θ e $B1$ para ARMA(0,1) para T3

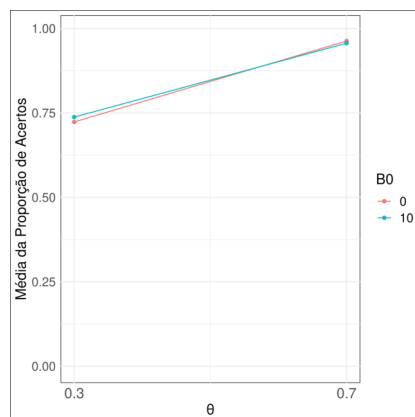


Figura 112: Interação entre θ e $B0$ para ARMA(0,1) para T3

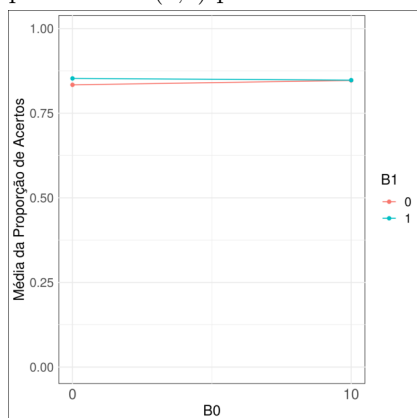


Figura 113: Interação entre $B1$ e $B0$ para ARMA(0,1) para T3

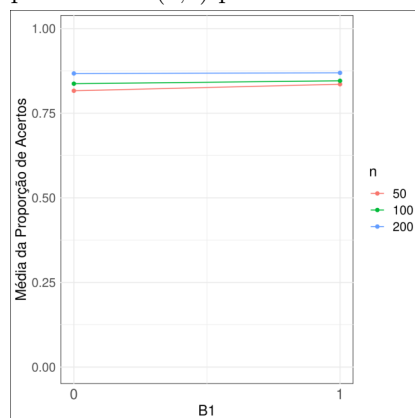


Figura 114: Interação entre $B1$ e n para ARMA(0,1) para T3

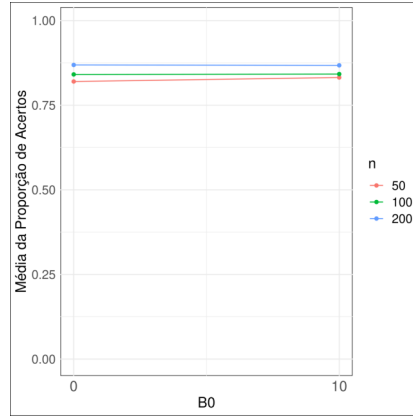


Figura 115: Interação entre $B0$ e n para ARMA(0,1) para T3

No gráfico da Figura 111, as diferenças entre as proporções de acertos para $B1=0$ e $B1=1$ são menores quando $\theta=0.7$. Para o gráfico da Figura 112, as diferenças entre as proporções de acertos para $B0=0$ e $B0=10$ são menores quando $\theta=0.7$. Já para o gráfico da Figura 113, a diferença entre as proporções de acertos para $B1=0$ e $B1=1$ é menor quando $B0=10$. E para o gráfico da Figura 114, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $B1=1$. Por fim, para o gráfico da Figura 115, as diferenças entre as proporções de acertos para $n=50$, $n=100$ e $n=200$ são menores quando $B0=10$.

f.4) Considerações: O critério T4 encontrado pelo modelo ARMA(1,1) apresenta comportamento mais robusto, enquanto o menos robusto foi encontrado no ARMA(0,1). O coeficiente ϕ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n=200$. Essa diferença diminui para maiores valores de ϕ . O coeficiente θ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n=200$. Essa diferença diminui para maiores valores de θ . Esse critério encontrado pelo modelo ARMA(0,1) apresentou comportamento diferenciado em relação às interações que envolveram $B1$ e $B0$, pois os resultados, embora próximos, se inverteram em relação à magnitude.

g) Critério T4: A Tabela 19 mostra os resultados finais obtidos para o critério T4, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 116 mostra os respectivos gráficos envelope.

Tabela 19: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T4.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-1.1206	0.0668	-16.7832	<0.01	ARMA(0,1)
θ	5.2876	0.1616	32.7211	<0.01	
$B1$	0.4143	0.0558	7.4222	<0.01	
$B0$	0.0358	0.0056	6.4181	<0.01	
n	0.002	4e-04	4.4545	<0.01	
$\theta:B1$	-0.8383	0.1279	-6.5556	<0.01	
$\theta:B0$	-0.0798	0.0128	-6.2313	<0.01	
$\theta:n$	0.004	0.0011	3.6915	<0.01	
$B1:B0$	-0.0177	0.0041	-4.3094	<0.01	
(Intercept)	-1.1374	0.0536	-21.2193	<0.01	ARMA(1,0)
ϕ	5.3011	0.1298	40.8345	<0.01	
$B1$	0.0124	0.0485	0.255	0.7987	
$B0$	-0.0063	0.0025	-2.5557	0.0106	
n	0.0021	4e-04	5.4895	<0.01	
$\phi:B1$	-0.1343	0.1136	-1.1818	0.2373	
$\phi:n$	0.0039	0.001	3.9641	<0.01	
$B1:B0$	-0.0135	0.0034	-3.9017	<0.01	
(Intercept)	-9.7583	0.5246	-18.6	<0.01	ARMA(1,1)
ϕ	9.6348	0.5746	16.767	<0.01	
θ	9.7636	0.5748	16.9851	<0.01	
$B1$	-0.1393	0.0786	-1.7722	0.0764	
$B0$	-0.0129	0.0079	-1.6454	0.0999	
n	0.0333	0.0032	10.4323	<0.01	
$\phi:n$	-0.0277	0.0037	-7.4107	<0.01	
$\theta:n$	-0.0277	0.0037	-7.3908	<0.01	

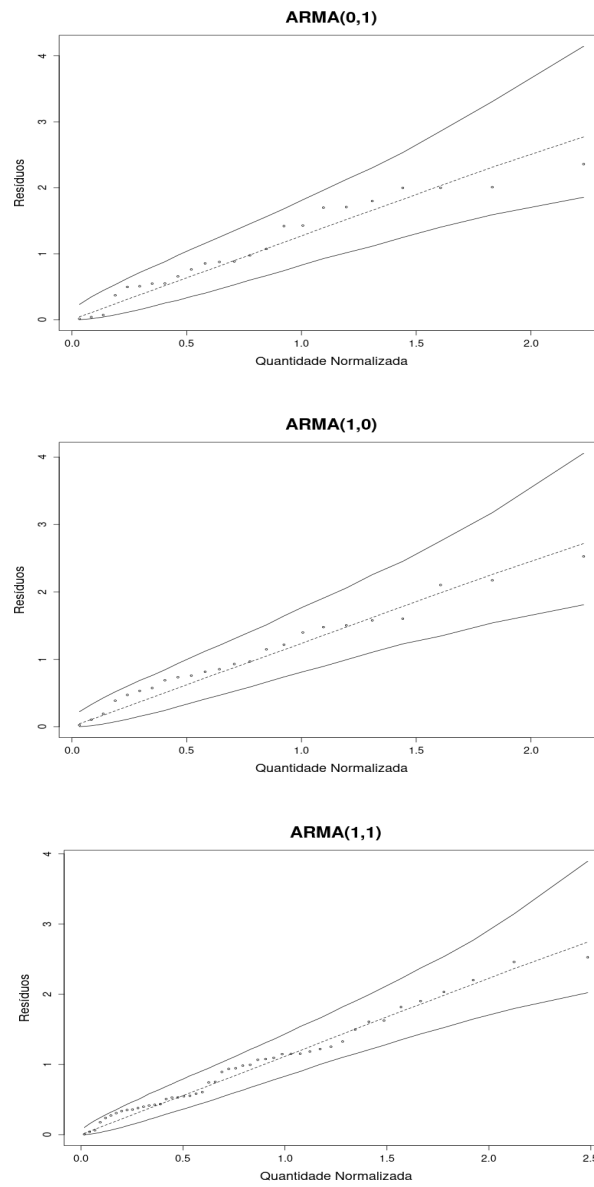


Figura 116: Gráficos envelope para modelos Beta de T4 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

g.1) ARMA(1,1): De acordo com a Tabela 19, para o modelo ARMA(1,1), apenas os efeitos principais de B_1 e B_0 não foram significativos. Existem evidências de que as proporções de acertos do critério T4 são maiores para

$\theta = 0.7$ (0,7818), $\theta = 0.7$ (0,7755) e $n = 200$ (0,6659). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T4 são afetados pela interação de ϕ com n e de θ com n , conforme observado nas figuras 117 e 118.

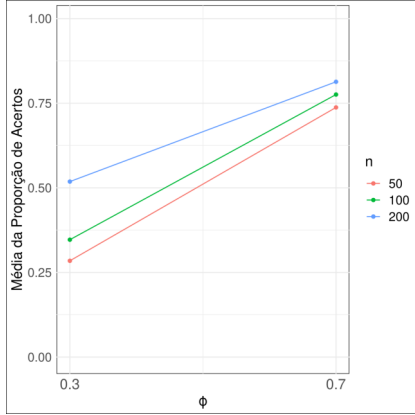


Figura 117: Interação entre ϕ e n para ARMA(1,1) para T4

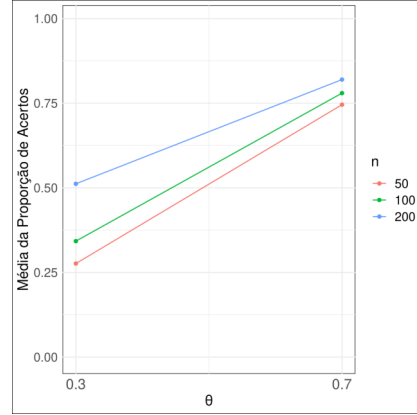


Figura 118: Interação entre θ e n para ARMA(1,1) para T4

No gráfico da Figura 117, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi = 0.7$. Já para o gráfico da Figura 118, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\theta = 0.7$.

g.2) ARMA(1,0): De acordo com a Tabela 19, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), somente os efeitos principais de ϕ , $B0$ e n são significativos. Existem evidências de que as proporções de acertos do critério T4 são maiores para $\phi = 0.7$ (0.9518), $B0 = 0$ (0.8236) e $n = 200$ (0.8514). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T4 são afetados pela interação de ϕ com n e de $B1$ com $B0$, conforme observado nas figuras 119 e 120.

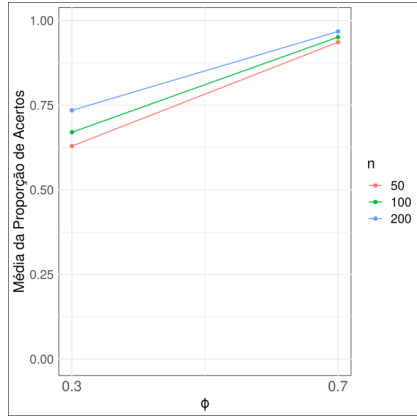


Figura 119: Interação entre ϕ e n para ARMA(1,0) para T4

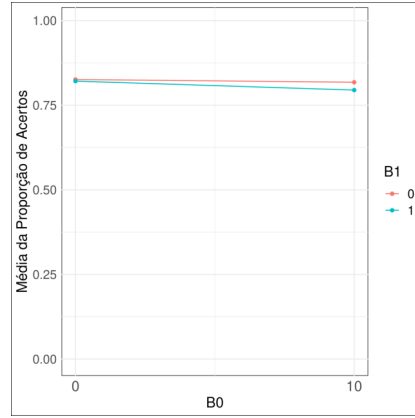


Figura 120: Interação entre $B1$ e $B0$ para ARMA(1,0) para T4

O gráfico da Figura 119, que representa a interação entre θ e n é muito semelhante ao gráfico da Figura 117 para ARMA(1,1) com relação a T4, onde se conclui que estes fatores interagem em modelos ARMA(1,0) da mesma maneira.

No gráfico da Figura 120, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$.

g.3) ARMA(0,1): De acordo com a Tabela 19, se for retirado apenas o termo Autoregressivo, ficando o modelo do tipo ARMA(0,1), todos os efeitos principais são significativos. Existem evidências de que as proporções de acertos do critério T4 são maiores para $\theta = 0.7$ (0,9454), $B0 = 0$ (0.8327), $B1 = 1$ (0.8308) e $n = 200$ (0,864). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T4 são afetados pela interação de θ com $B1$, de θ com $B0$, de θ com n e de $B1$ com B , conforme observado nas figuras 121, 122, 123 e 124.

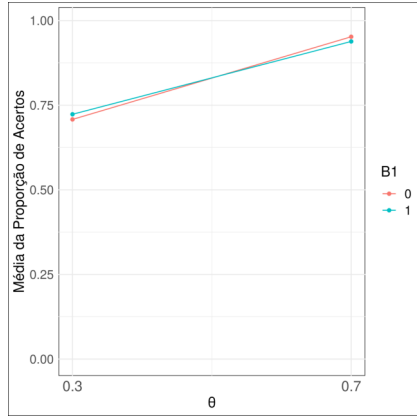


Figura 121: Interação entre θ e $B1$ para ARMA(0,1) para T4

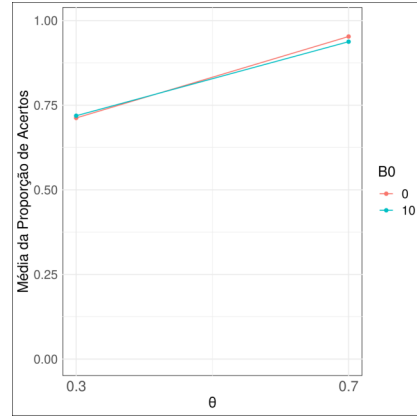


Figura 122: Interação entre θ e $B0$ para ARMA(0,1) para T4

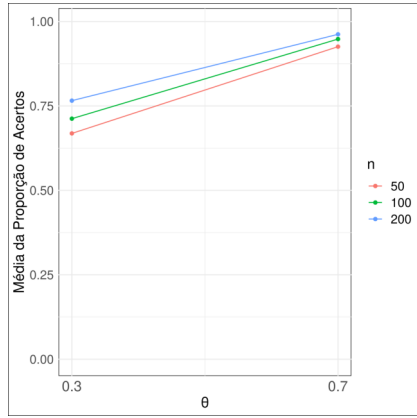


Figura 123: Interação entre θ e n para ARMA(0,1) para T4

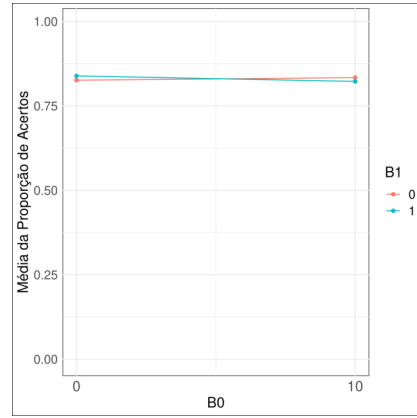


Figura 124: Interação entre $B1$ e $B0$ para ARMA(0,1) para T4

No gráfico da Figura 121, as diferenças entre as proporções de acertos para $B1=0$ e $B1 = 1$ são menores quando $\theta = 0.7$. Para o gráfico da Figura 122, as diferenças entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ são menores quando $\theta = 0.3$. Já para o gráfico da Figura 123, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 118 para ARMA(1,1) com relação a T4, onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. Por fim, para o gráfico da Figura 124, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$.

g.4) Considerações: O critério T4 encontrado pelo modelo ARMA(1,1) apresenta comportamento mais robusto e o menos robusto foi encontrado no ARMA(0,1). O coeficiente ϕ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de ϕ . O coeficiente θ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de θ . Esse critério encontrado pelo modelo ARMA(0,1) apresentou comportamento diferenciado em relação às interações que envolveram $B1$ e $B0$, pois os resultados, embora próximos, se inverteram em relação à magnitude.

h) Critério T5: A Tabela 20 mostra os resultados finais obtidos para o critério T5, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 125 mostra os respectivos gráficos envelope.

Tabela 20: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T5.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-0.9917	0.1274	-7.7863	<0.01	ARMA(0,1)
θ	4.1728	0.2891	14.4343	<0.01	
$B1$	0.3869	0.1069	3.6195	<0.01	
$B0$	0.0328	0.0107	3.069	0.0021	
n	0.0018	9e-04	2.1181	0.0342	
$\theta:B1$	-0.9499	0.2290	-4.1476	<0.01	
$\theta:B0$	-0.0886	0.0229	-3.8662	<0.01	
$\theta:n$	0.0088	0.0020	4.2924	<0.01	
$B1:B0$	-0.0221	0.008	-2.7491	0.0060	
(Intercept)	-0.9165	0.0899	-10.1901	<0.01	ARMA(1,0)
ϕ	4.0894	0.2074	19.7209	<0.01	
$B1$	0.0347	0.0754	0.4606	0.6451	
$B0$	0.0065	0.0076	0.8537	0.3933	
n	0.0011	6e-04	1.8816	0.0599	
$\phi:B1$	-0.3101	0.1645	-1.8853	0.0594	
$\phi:B0$	-0.0463	0.0165	-2.8094	0.0050	
$\phi:n$	0.0096	0.0015	6.5056	<0.01	
$B1:B0$	-0.014	0.0056	-2.4907	0.0128	
(Intercept)	-6.2303	0.4012	-15.528	<0.01	ARMA(1,1)
ϕ	5.7847	0.5733	10.0906	<0.01	
θ	5.8431	0.5731	10.1954	<0.01	
$B1$	0.2535	0.2677	0.9468	0.3438	
$B0$	-0.0088	0.0086	-1.0236	0.3060	
n	0.0172	0.0021	8.1807	<0.01	
$\phi:\theta$	1.3259	0.848	1.5636	0.1179	
$\phi:B1$	-0.354	0.3239	-1.0927	0.2745	
$\phi:n$	-0.0117	0.0026	-4.4841	<0.01	
$\theta:B1$	-0.3765	0.3242	-1.1614	0.2455	
$\theta:n$	-0.0112	0.0026	-4.2987	<0.01	
$B1:B0$	-0.0094	0.0120	-0.7829	0.4337	

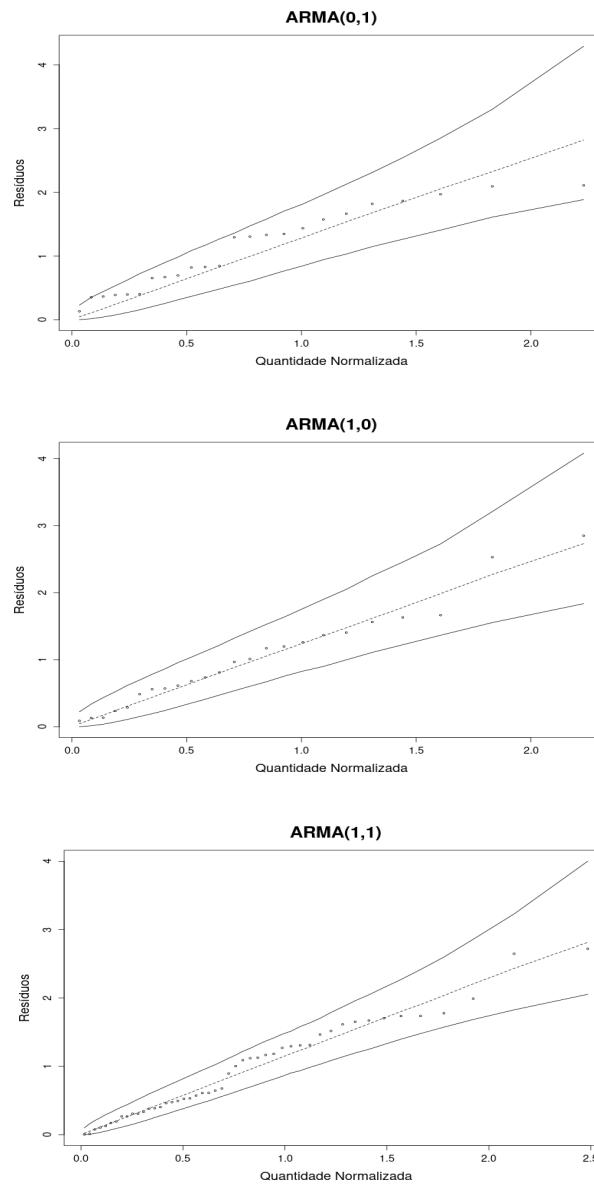


Figura 125: Gráficos envelope para modelos Beta de T5 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

h.1) ARMA(1,1): De acordo com a Tabela 20, para o modelo ARMA(1,1), apenas os efeitos principais de B_0 e B_1 não foram significativos. Existem evidências de que as proporções de acertos do critério T5 são maiores para

$\theta = 0.7$ (0,7598) e $\phi = 0.7$. Com relação às interações, existem evidências de que os resultados das proporções de acertos do T5 são afetados pela interação de ϕ com n e de θ com n , conforme observado nas figuras 126 e 127.

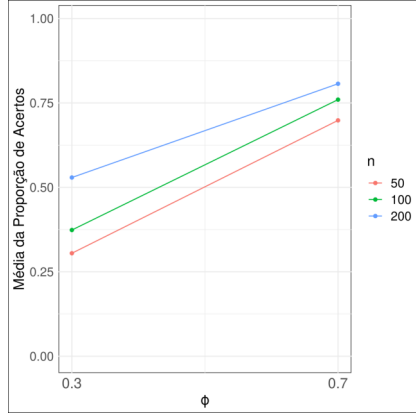


Figura 126: Interação entre ϕ e n para ARMA(1,1) para T5

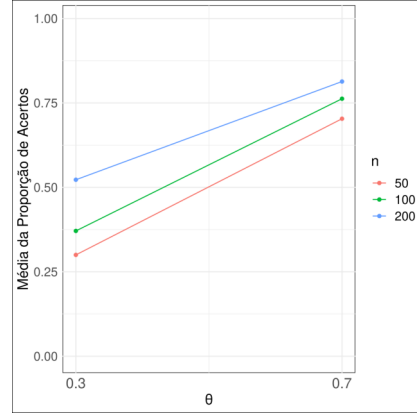


Figura 127: Interação entre θ e n para ARMA(1,1) para T5

No gráfico da Figura 126, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi = 0.7$.

No gráfico da Figura 127, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\theta = 0.7$.

h.2) ARMA(1,0): De acordo com a Tabela 20, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), somente o efeito ϕ é considerado significativo e existem evidências de que as proporções de acertos do critério T5 são maiores para $\phi = 0.7$ (0,9237). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T5 são afetados somente pela interação de ϕ com $B0$ e com n , conforme observado nas figuras 128 e 129.

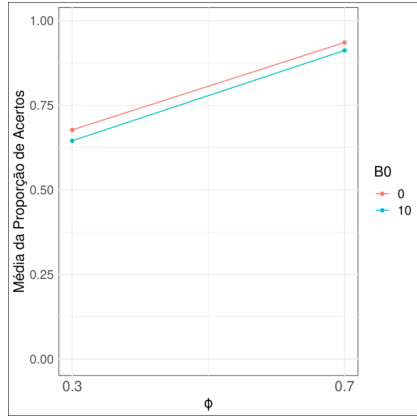


Figura 128: Interação entre ϕ e $B0$ para ARMA(1,0) para T5

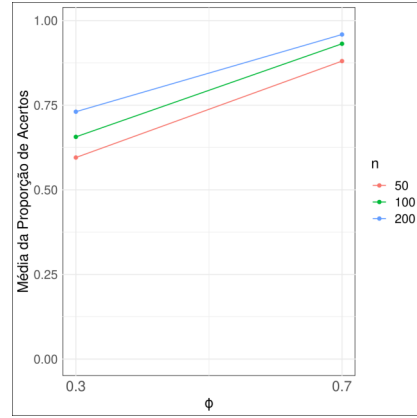


Figura 129: Interação entre ϕ e n para ARMA(1,0) para T5

No gráfico da Figura 128, a diferença entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ é menor quando $\phi = 0.7$. Para o gráfico da Figura 129, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 126 para ARMA(1,1) com relação a T5, onde se conclui que estes fatores interagem em modelos ARMA(1,0) da mesma maneira.

h.3) ARMA(0,1): De acordo com a Tabela 20, se for retirado apenas o termo Autoregressivo, ficando o modelo do tipo ARMA(0,1), os efeitos principais de todas as componentes são significativas. Existem evidências de que as proporções de acertos do critério T5 são maiores para $\theta = 0.7$ (0,9157), $B0 = 0$ (0,8124), $B1 = 0$ (0,8101) e $n = 200$ (0,856). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T5 são afetados pela interação de θ com $B1$, de θ com $B0$, de θ com n e de $B1$ com $B0$, conforme observado nas figuras 130, 131, 132 e 133.

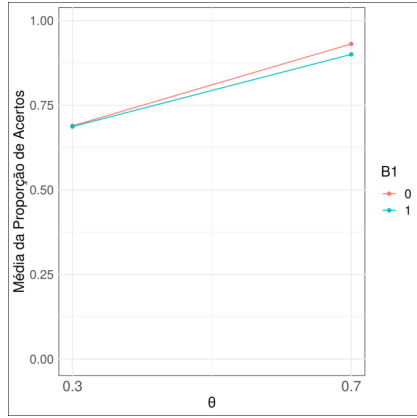


Figura 130: Interação entre θ e $B1$ para ARMA(0,1) para T5

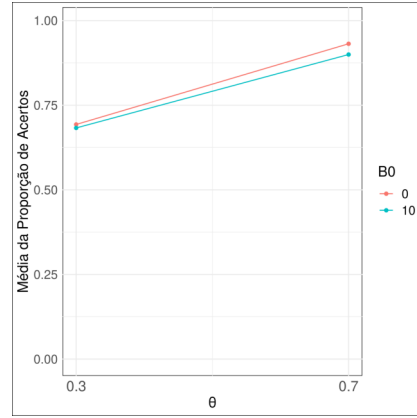


Figura 131: Interação entre θ e $B0$ para ARMA(0,1) para T5

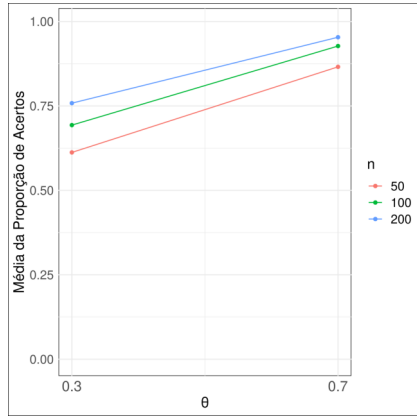


Figura 132: Interação entre θ e n para ARMA(0,1) para T5

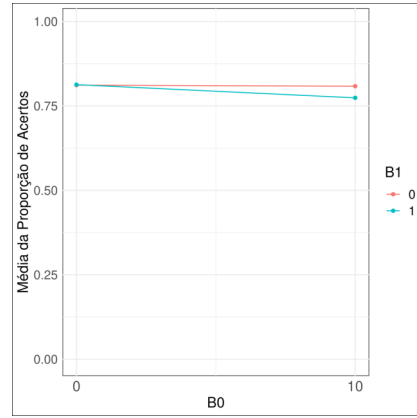


Figura 133: Interação entre $B0$ e $B1$ para ARMA(0,1) para T5

No gráfico da Figura 130, as diferenças entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ são menores quando $\theta = 0.3$. Para o gráfico da Figura 131, as diferenças entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ são menores quando $\theta = 0.3$. Já para o gráfico da Figura 132, que representa a interação entre θ e n é muito semelhante ao gráfico da Figura 127 para ARMA(1,1) com relação a T5, onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. Por fim, para o gráfico da Figura 133, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$.

h.4) Considerações: Os modelos mais parcimoniosos para descrever o comportamento do critério T5 no ARMA(1,1) e no ARMA(1,0), normalmente identificados como mais robustos, utilizaram muitos termos com efeitos não significativos. O coeficiente ϕ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de ϕ . O coeficiente θ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de θ . O efeito da interação de $B1$ com Bo foi significativo nos três modelos, mas só foi considerado no ARMA(0,1) porque nos demais o efeito principal desses fatores não foi significativo.

i) Critério T6: A Tabela 21 mostra os resultados finais obtidos para o critério T6, para cada um dos modelos ARMA(0,1), ARMA(1,0) e ARMA(1,1), enquanto a Figura 134 mostra os respectivos gráficos envelope.

Tabela 21: Coeficientes e significância obtidos pela regressão Beta para o Critério de Informação T6.

Componente	Coeficiente	Erro Padrão	Estatística de Teste	Valor-p	ARMA(p,q)
(Intercept)	-1.0622	0.0863	-12.3156	<0.01	ARMA(0,1)
θ	4.8567	0.204	23.8039	<0.01	
$B1$	0.3970	0.0723	5.4940	<0.01	
$B0$	0.0340	0.0072	4.7053	<0.01	
n	0.0018	6e-04	3.1970	<0.01	
$\theta:B1$	-0.8387	0.1617	-5.1868	<0.01	
$\theta:B0$	-0.0809	0.0162	-4.9932	<0.01	
$\theta:n$	0.0058	0.0014	4.158	<0.01	
$B1:B0$	-0.0182	0.0054	-3.3938	<0.01	
(Intercept)	-1.0305	0.0576	-17.903	<0.01	ARMA(1,0)
ϕ	4.8151	0.1356	35.5125	<0.01	
$B1$	-0.049	0.0271	-1.8086	<0.01	
$B0$	4e-04	0.0052	0.0746	0.9405	
n	0.0018	4e-04	4.4549	<0.01	
$\phi:B0$	-0.0237	0.0119	-1.9976	0.0458	
$\phi:n$	0.0055	0.001	5.3598	<0.01	
$B1:B0$	-0.0133	0.0038	-3.5456	<0.01	
(Intercept)	-8.4694	0.4334	-19.5409	<0.01	ARMA(1,1)
ϕ	8.3263	0.4855	17.1515	<0.01	
θ	8.4516	0.4856	17.4028	<0.01	
$B1$	-0.1375	0.0725	-1.8972	0.0578	
$B0$	-0.013	0.0072	-1.7959	0.0725	
n	0.0273	0.0027	9.9558	<0.01	
$\phi:n$	-0.0217	0.0033	-6.5704	<0.01	
$\theta:n$	-0.0216	0.0033	-6.5381	<0.01	

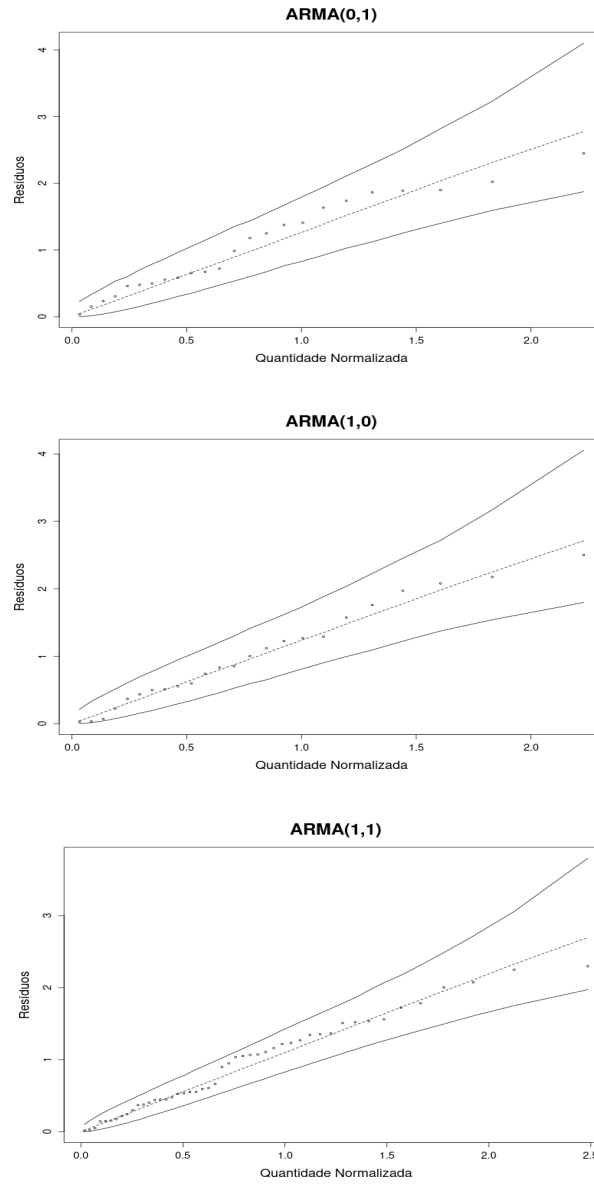


Figura 134: Gráficos envelope para modelos Beta de T6 de cima para baixo ARMA(0,1), ARMA(1,0) e ARMA(1,1)

i.1) ARMA(1,1): De acordo com a Tabela 21, para o modelo ARMA(1,1) somente θ e ϕ foram significativos. Existem evidências de que as proporções de acertos do critério T6 são maiores $\theta = 0.7$ (0,7754) e $\phi = 0.7$. Com relação

às interações, existem evidências de que os resultados das proporções de acertos do T6 são afetados pela interação de ϕ com n e de θ com n , conforme observado nas figuras 135 e 136.

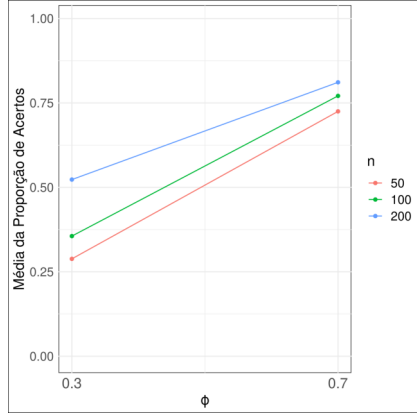


Figura 135: Interação entre ϕ e n para ARMA(1,1) para T6

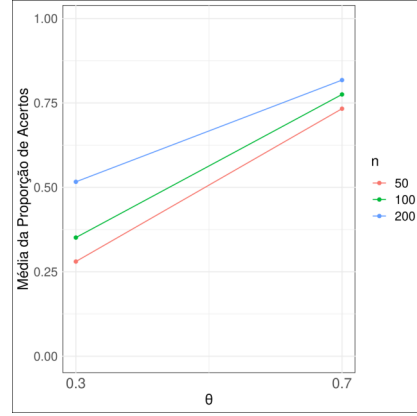


Figura 136: Interação entre θ e n para ARMA(1,1) para T6

No gráfico da Figura 135, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\phi = 0.7$. Enquanto para o gráfico da Figura 136, as diferenças entre as proporções de acertos para $n = 50$, $n = 100$ e $n = 200$ são menores quando $\theta = 0.7$.

i.2) ARMA(1,0): De acordo com a Tabela 21, se for retirado apenas o termo de Médias Móveis, ficando o modelo do tipo ARMA(1,0), o efeito principal de B_0 deixa de ser significativo. Existem evidências de que as proporções de acertos do critério T6 são maiores para, $\theta = 0.7$ (0.9368), $B_1 = 0$ (0.8232) e $n = 200$ (0.8615). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T6 são afetados pela interação de ϕ com B_0 , de ϕ com n e de B_1 com B_0 , conforme observado nas figura 137, 138 e 139.

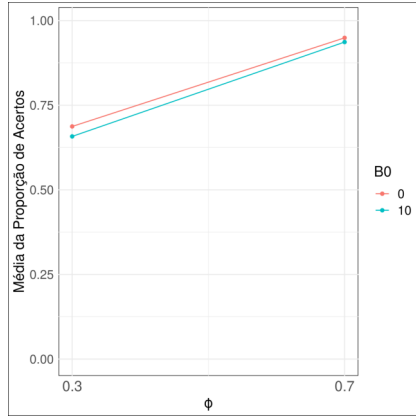


Figura 137: Interação entre ϕ e $B0$ para ARMA(0,1) para T6

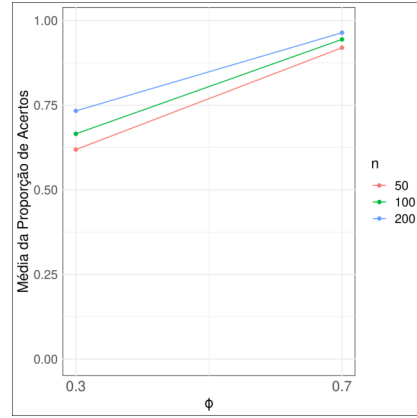


Figura 138: Interação entre ϕ e n para ARMA(0,1) para T6

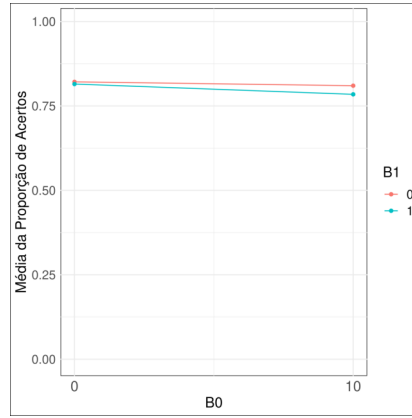


Figura 139: Interação entre $B1$ e $B0$ para ARMA(0,1) para T6

No gráfico da Figura 137, a diferença entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ é menor quando $\phi = 0.3$. Para o gráfico da Figura 138, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 135 para ARMA(1,1) com relação a T6, onde se conclui que estes fatores interagem em modelos ARMA(1,0) da mesma maneira. Já para o gráfico da Figura 139, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$.

i.3) ARMA(0,1): De acordo com a Tabela 21, se for retirado apenas o termo Autorregressivo, ficando o modelo do tipo ARMA(0,1), os efeitos principais de

θ , $B0$, $B1$ e n são significativos. Existem evidências de que as proporções de acertos do critério T6 são maiores para $\theta = 0.7$ (0,9368), $B0 = 0$ (0,826), $B1 = 0$ (0,8232) e $n = 200$ (0,8615). Com relação às interações, existem evidências de que os resultados das proporções de acertos do T6 são afetados pela interação de θ com $B1$, de θ com $B0$, de θ com n e de $B1$ com $B0$, conforme observado nas figura 140, 141, 142 e 143.

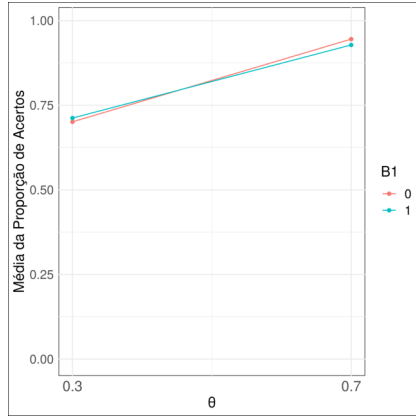


Figura 140: Interação entre θ e $B1$ para ARMA(0,1) para T6

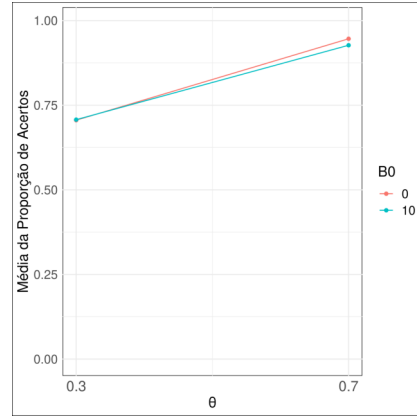


Figura 141: Interação entre θ e $B0$ para ARMA(0,1) para T6

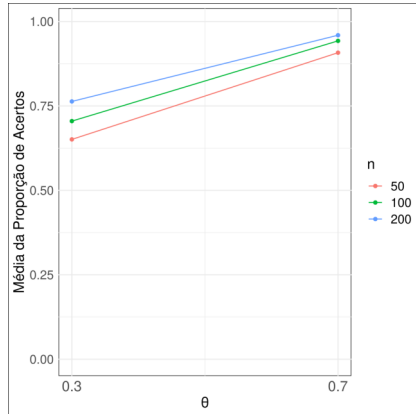


Figura 142: Interação entre θ e n para ARMA(0,1) para T6

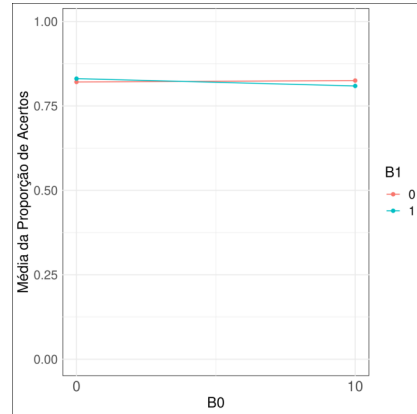


Figura 143: Interação entre $B1$ e $B0$ para ARMA(0,1) para T6

No gráfico da Figura 140, as diferenças entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ são menores quando $\theta = 0.3$. Para o gráfico da Figura 141, as diferenças entre as proporções de acertos para $B0 = 0$ e $B0 = 10$ são

menores quando $\theta = 0.3$. Já para o gráfico da Figura 142, que representa a interação entre θ e n , a mesma é muito semelhante ao gráfico da Figura 136 para ARMA(1,1) com relação a T6, onde se conclui que estes fatores interagem em modelos ARMA(0,1) da mesma maneira. Por fim, para o gráfico da Figura 143, a diferença entre as proporções de acertos para $B1 = 0$ e $B1 = 1$ é menor quando $B0 = 0$.

i.4) Considerações: O critério T6 encontrado pelo modelo ARMA(1,1) apresenta comportamento mais robusto enquanto o menos robusto foi encontrado no ARMA(0,1). O coeficiente ϕ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de ϕ . O coeficiente θ , quando presente, interagiu com n , sendo a proporção de acertos maior quando $n = 200$. Essa diferença diminui para maiores valores de θ .

A seguir, apresenta-se a Tabela 22 que sintetiza as interações entre os componentes da modelagem Beta nas configurações ARMA(0,1), ARMA(1,0) e ARMA(1,1). A análise focou nos critérios T3, T4 e T6, juntamente com o critério BIC, a fim de avaliar a significância das relações propostas, já que os resultados indicam que os critérios T3, T4 e T6 demonstram um desempenho superior em relação aos demais propostos, esta mesma tabela evidencia a relevância das interações significativas (marcadas com um “+”), não significativas (marcadas com um “-”) e as interações que não se enquadram nos modelos ARMA (indicadas com “x”).

Tabela 22: Interações significativas para os critérios BIC, T3, T4 e T6.

Componente	ARMA(0,1)	ARMA(1,0)	ARMA(1,1)	Critério
θ	+	x	+	BIC
ϕ	x	+	+	
$B1$	+	-	-	
$B0$	+	+	-	
n	+	+	+	
$\theta:\phi$	x	x	x	
$\theta:B1$	-	x	x	
$\theta:B0$	-	x	x	
$\theta:n$	+	x	-	
$\phi:B1$	x	x	x	
$\phi:B0$	x	-	x	
$\phi:n$	x	+	-	
$B1:B0$	-	-	x	
$B1:n$	x	x	x	
$B0:n$	x	x	x	
θ	+	x	+	T3
ϕ	x	+	+	
$B1$	+	-	-	
$B0$	+	-	-	
n	+	+	+	
$\theta:\phi$	x	x	x	
$\theta:B1$	-	x	x	
$\theta:B0$	-	x	x	
$\theta:n$	x	x	-	
$\phi:B1$	x	x	x	
$\phi:B0$	x	x	x	
$\phi:n$	x	x	-	
$B1:B0$	-	-	x	
$B1:n$	-	x	x	
$B0:n$	-	x	x	
θ	+	x	+	T4
ϕ	x	+	+	
$B1$	+	+	-	
$B0$	+	-	-	
n	+	+	+	
$\theta:\phi$	x	x	x	
$\theta:B1$	-	x	x	
$\theta:B0$	-	x	x	
$\theta:n$	+	x	-	
$\phi:B1$	x	-	x	
$\phi:B0$	x	x	x	
$\phi:n$	x	+	-	
$B1:B0$	-	-	x	
$B1:n$	x	x	x	
$B0:n$	x	x	x	
θ	+	x	+	T6
ϕ	x	+	+	
$B1$	+	-	-	
$B0$	+	+	-	
n	+	+	+	
$\theta:\phi$	x	x	x	
$\theta:B1$	-	x	x	
$\theta:B0$	-	x	x	
$\theta:n$	+	x	-	
$\phi:B1$	x	x	x	
$\phi:B0$	x	-	x	
$\phi:n$	x	+	-	
$B1:B0$	-	-	x	
$B1:n$	x	x	x	
$B0:n$	x	x	x	

Os resultados da Tabela 22 mostram que o parâmetro θ apresenta interações positivas (valores de coeficiente maior que zero) para os modelos ARMA(0,1) e ARMA(1,1). Isto indica que, aparentemente, um maior valor de θ resulta em um aumento da proporção de acertos. O parâmetro ϕ também demonstrou interações positivas em ARMA(1,0) e ARMA(1,1). As interações dos parâmetros $B1$ e $B0$ variaram conforme a complexidade do modelo. O $B1$ teve interações positivas em ARMA(0,1) e ARMA(1,0), mas tornou-se negativo em ARMA(1,1) para a maioria dos casos. O tamanho da amostra n foi consistentemente positivo em todas as configurações, indicando que um maior número de observações tende a uma melhora do desempenho dos critérios.

Considerações Gerais: Para a maioria dos Critérios de Informação analisados, embora os efeitos dessas interações tenham sido significativos, grande parte desses gráficos sugerem que não sejam grandes. Observa-se que a diferença entre as proporções associadas aos níveis de um fator apenas ficam um pouco menores em um dos níveis do outro fator considerado na interação. Os resultados obtidos também revelam uma tendência clara de que a proporção média de acertos é consistentemente menor quando os parâmetros $\phi = 0.3$ e $\theta = 0.3$ são utilizados em comparação com seus valores mais altos ($\phi = 0.7$ e $\theta = 0.7$). Essa relação se mantém estável em diferentes tamanhos de amostra sejam elas 50, 100 ou 200 e também entre os componentes analisados, como $B0$ e $B1$. Além disso, observou-se que os componentes $B1 = 0$ resultaram em uma taxa de acertos superior em relação a $B1 = 1$. Esses achados sugerem que parâmetros mais baixos para ϕ e θ não apenas reduzem a eficácia dos acertos, mas também indicam a necessidade de uma consideração cuidadosa desses valores em futuras investigações.

5 CONCLUSÃO

A presente pesquisa teve como hipótese a ideia de que, por meio de uma abordagem baseada em uma equação quadrática, seria possível desenvolver um Critério de Informação personalizado, adaptável a modelos de séries temporais.

Em suma, o objetivo geral foi alcançado, pois foi proposta a criação de um Critério de Informação abrangente e adaptável, que utilizou penalidades estruturadas a partir de uma equação Quadrática, sendo investigada a sua adaptabilidade a determinados tipos de modelo, mais especificamente, a séries temporais que pudessem ser descritas por modelos Autoregressivos e de Médias Móveis.

A análise detalhada da penalidade associada aos critérios AIC, AICc e BIC mostrou que estes critérios podem apresentar melhor ou pior desempenho de acordo com o mecanismo gerador dos dados que compõe a série temporal analisada. Também foi demonstrado que suas penalidades seguem a estrutura de uma Superfície Quádrica, permitindo a exploração de novas formas de adaptação dessas equações para diferentes tipos de modelos. Os objetivos específicos, portanto, também foram atingidos com êxito, incluindo a demonstração dessa estrutura geométrica e a adaptação dos critérios para identificar corretamente os modelos ARMA sugeridos.

A proposta de um critério generalizado, com penalidade baseada em dois coeficientes E e F , foi testada em 6 diferentes configurações, identificadas como T1, T2, T3, T4, T5 e T6. Um estudo de simulação permitiu concluir que algumas adaptações do Critério de Informação proposto (T3, T4 e T6) apresentaram melhor desempenho nos modelos ARMA(0,1) e ARMA(1,0), do que os tradicionais critérios AIC, AICc e BIC, embora a adaptação não tenha se mostrado satisfatória no modelo ARMA(1,1). Considerando que os critérios AIC, AICc e BIC são amplamente reconhecidos na literatura, pode-se concluir, de forma preliminar, que a adaptação T3, a melhor entre as três adaptações melhores do critério proposto, pode ser uma alternativa promissora na análise e seleção de modelos, principalmente em séries temporais cujo mecanismo gerador de seus dados seja um modelo Autoregressivo ou um modelo de Médias Móveis.

Por fim, também foi realizada uma simulação complementar, considerando coeficientes ARMA com efeito aleatório, encontrando-se resultados similares. (Apêndice A.3).

O estudo realizado oferece uma base sólida para futuras investigações, mas várias questões permanecem em aberto e podem ser exploradas em pesquisas subsequentes. Algumas direções possíveis incluem:

- **Ampliação dos valores de E e F para o critério proposto:** O Critério de Informação generalizado, com penalidade adaptável baseada em dois coeficientes de uma equação Quádrica, apresentou bons resultados, mas uma análise mais profunda sobre a variação desses coeficientes e sua sensibilidade a diferentes modelos pode contribuir ainda mais no aprimoramento desses resultados e consequente desempenho do critério proposto;
- **Ampliação da quantidade de coeficientes da equação Quádrica:** Uma equação Quádrica possui 9 coeficientes, aumentar a quantidade de coeficientes considerados na formulação da penalidade pode aumentar a

sensibilidade do critério de informação, tornando-o mais sensível ao tipo de modelo;

- **Aprofundamento na Análise dos Coeficientes da Superfície Quádrica:**

A penalidade w , modelada por meio de uma Superfície Quádrica, depende de múltiplos coeficientes, que determinam a forma e a sensibilidade da penalização. Um estudo mais profundo sobre como esses coeficientes interagem entre si e como sua variação impacta o desempenho dos Critérios de Informação pode ajudar a otimizar a adaptação do critério a diferentes tipos de modelos;

- **Exploração de conceitos geométricos mais avançados:** Explorar a integração de conceitos geométricos mais avançados, como Curva Geodésica e Torção, com a análise dos Critérios de Informação, e metodologias complementares de construção, conforme discutido preliminarmente nos Apêndices (A.4)-(A.8).

REFERÊNCIAS BIBLIOGRÁFICAS

- ACQUAH, H. D.-G. Weighted average information criterion for selection of an asymmetric price relationship. *Alanya Akademik Bakış*, v. 2, n. 2, p. 147–155, 2018.
- AKAIKE, H. A new look at the statistical model identification. *IEEE transactions on automatic control*, Ieee, v. 19, n. 6, p. 716–723, 1974.
- ANDERSON, D.; BURNHAM, K.; WHITE, G. Aic model selection in overdispersed capture-recapture data. *Ecology*, Wiley Online Library, v. 75, n. 6, p. 1780–1793, 1994.
- ANTON, H.; BIVENS, I.; DAVIS, S. *Cálculo-Volume II-10*. [S.l.]: Bookman Editora, 2014.
- AZEEZ, A. et al. Seasonality and trend forecasting of tuberculosis prevalence data in eastern cape, south africa, using a hybrid model. *International journal of environmental research and public health*, MDPI, v. 13, n. 8, p. 757, 2016.
- BOX, G. Box and jenkins: time series analysis, forecasting and control. In: *A Very British Affair*. [S.l.]: Springer, 2013. p. 161–215.
- BOZDOGAN, H. Akaike's information criterion and recent developments in information complexity. *Journal of mathematical psychology*, Elsevier, v. 44, n. 1, p. 62–91, 2000.
- BREWER, M. J.; BUTLER, A.; COOKSLEY, S. L. The relative performance of aic, aicc and bic in the presence of unobserved heterogeneity. *Methods in Ecology and Evolution*, Wiley Online Library, v. 7, n. 6, p. 679–692, 2016.
- BUENO, R. D. L. d. S. *Econometria de séries temporais*. 2018.
- BURNHAM, K. P. Model selection and multimodel inference. *A practical information-theoretic approach*, Springer, 1998.
- CARMO, M. P. D. *Geometria diferencial de curvas e superfícies*. [S.l.]: Sociedade Brasileira de Matemática, 2010.
- CHEN, J.; CHEN, Z. Extended bayesian information criteria for model selection with large model spaces. *Biometrika*, Oxford University Press, v. 95, n. 3, p. 759–771, 2008.
- CHEN, P.; WU, T.-J.; YANG, J. A comparative study of model selection criteria for the number of signals. *IET Radar, Sonar & Navigation*, IET, v. 2, n. 3, p. 180–188, 2008.
- DICKEY, D. A.; FULLER, W. A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: journal of the Econometric Society*, JSTOR, p. 1057–1072, 1981.

- DZIAK, J. J. et al. Sensitivity and specificity of information criteria. *Briefings in bioinformatics*, Oxford University Press, v. 21, n. 2, p. 553–565, 2020.
- EMILIANO, P. C. et al. Critérios de informação de akaike versus bayesiano: análise comparativa. *19^o Simpósio Nacional de Probabilidade e Estatística*, 2010.
- ESTEVES, L. G.; STERN, R. Inferência bayesiana. *Universidade Federal de São Carlos, São Carlos, SP-Brasil, Tech. Rep*, 2017.
- FITZGIBBON, L. J.; DOWE, D. L.; VAHID, F. Minimum message length autoregressive model order selection. In: IEEE. *International Conference on Intelligent Sensing and Information Processing, 2004. Proceedings of*. [S.l.], 2004. p. 439–444.
- GRÜN, B.; KOSMIDIS, I.; ZEILEIS, A. *Extended beta regression in R: Shaken, stirred, mixed, and partitioned*. [S.l.], 2011.
- GUJARATI, D. N.; PORTER, D. C. *Econometria básica-5*. [S.l.]: Amgh Editora, 2011.
- HARBECKE, J.; GRUNAU, J.; SAMANEK, P. Are the bayesian information criterion (bic) and the akaike information criterion (aic) applicable in determining the optimal fit and simplicity of mechanistic models? *International Studies in the Philosophy of Science*, Taylor & Francis, p. 1–20, 2024.
- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, Oxford University Press, v. 76, n. 2, p. 297–307, 1989.
- HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for r. *Journal of statistical software*, v. 27, p. 1–22, 2008.
- JAMES, G. et al. *An introduction to statistical learning*. [S.l.]: Springer, 2013. v. 112.
- KIM, H. et al. Appropriate model selection methods for nonstationary generalized extreme value models. *Journal of Hydrology*, Elsevier, v. 547, p. 557–574, 2017.
- KONISHI, S.; KITAGAWA, G. Information criteria and statistical modeling. Springer, 2008.
- KUNDU, D.; MITRA, A. Estimating the number of signals of the damped exponential models. *Computational statistics & data analysis*, Elsevier, v. 36, n. 2, p. 245–256, 2001.
- LEHMAN, C. Geometria analitica, edit. *LIMUSA SA, 2da. edic*, 1980.
- LU, D. et al. Effects of error covariance structure on estimation of model averaging weights and predictive performance. *Water Resources Research*, Wiley Online Library, v. 49, n. 9, p. 6029–6047, 2013.

- LUDDEN, T. M.; BEAL, S. L.; SHEINER, L. B. Comparison of the akaike information criterion, the schwarz criterion and the f test as guides to model selection. *Journal of pharmacokinetics and biopharmaceutics*, Springer, v. 22, n. 5, p. 431–445, 1994.
- MICHEAUX, P. L. de; DROUILHET, R.; LIQUET, B. *The R software*. [S.l.]: Springer, 2013.
- MORETTIN, P. A.; TOLOI, C. M. *Análise de séries temporais: modelos lineares univariados*. [S.l.]: Editora Blucher, 2018.
- MURTAUGH, P. A. Performance of several variable-selection methods applied to real ecological data. *Ecology letters*, Wiley Online Library, v. 12, n. 10, p. 1061–1068, 2009.
- PAPPAS, S. S. et al. Electricity demand load forecasting of the hellenic power system using an arma model. *Electric Power Systems Research*, Elsevier, v. 80, n. 3, p. 256–264, 2010.
- PFAFF, B. et al. Package ‘urca’. *Unit root and cointegration tests for time series data. R package version*, p. 1–2, 2016.
- PHILLIPS, P. C.; PERRON, P. Testing for a unit root in time series regression. *Biometrika*, Oxford University Press, v. 75, n. 2, p. 335–346, 1988.
- RICA, A. L. de la; CUENCA, A. de la V. *Geometría diferencial*. [S.l.]: Clagsa, 1997.
- SCHWARZ, G. Estimating the dimension of a model. *The annals of statistics*, JSTOR, p. 461–464, 1978.
- SILVEIRA, A. G. *Estratégia para definição do número de defasagens na aplicação do teste*. Tese (Doutorado) — Universidade Federal do Rio Grande – FURG, Programa de Pós-Graduação em Modelagem Computacional, Rio Grande/RS, 2023. Tese de Doutorado.
- SOHIL, F.; SOHALI, M. U.; SHABBIR, J. *An introduction to statistical learning with applications in R: by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, New York, Springer Science and Business Media, 2013, 41.98, eISBN : 978 – 1 – 4614 – 7137 – 7.*[S.l.] : Taylor&Francis, 2022.
- STEINBRUCH, A.; WINTERLE, P. *Álgebra linear*. [S.l.: s.n.], 1987.
- STEWART, J. Cálculo, volume 2, 5a edição, ed. thomson. *Learning*, 2013.
- SUGIURA, N. Further analysis of the data by akaike’s information criterion and the finite corrections: Further analysis of the data by akaike’s. *Communications in Statistics-theory and Methods*, Taylor & Francis, v. 7, n. 1, p. 13–26, 1978.
- TARSITANO, A.; AMERISE, I. L. Short-term load forecasting using a two-stage sarimax model. *Energy*, Elsevier, v. 133, p. 108–114, 2017.

VISMARA, E. d. S. *Mensuração da biomassa e construção de modelos para construção de equações de biomassa*. Tese (Doutorado) — Universidade de São Paulo, 2009.

WHITTAKER, T. A.; FURLOW, C. F. The comparison of model selection criteria when selecting among competing hierarchical linear models. *Journal of Modern Applied Statistical Methods*, v. 8, n. 1, p. 15, 2009.

WU, S. A failure process model with the exponential smoothing of intensity functions. *European Journal of Operational Research*, Elsevier, v. 275, n. 2, p. 502–513, 2019.

XAVIER, A. C. F. et al. Selecting “the best” nonstationary generalized extreme value (gev) distribution: on the influence of different numbers of gev-models. *Bragantia*, SciELO Brasil, v. 78, p. 606–621, 2019.

ZIVOT, E.; ANDREWS, D. W. K. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of business & economic statistics*, Taylor & Francis, v. 20, n. 1, p. 25–44, 2002.

A APÊNDICE

A.1 Pseudo-código para as simulações

```
#comentarios começam e terminam com sustenido #
Inicio{#inicio do pseudo código#
  criterios = [T1,T2,T3,T4,T5,T6,AIC,AICc,BIC]
  modelos(p,q) = [ARMA(0,1),ARMA(0,2),ARMA(1,0),ARMA(1,1)]#modelos
  com as condicoes especificadas#
  n = [50,100,200] #vetor de tamanho das amostras#
  para( n1 em n) faça {
    para(c em criterios){
      ac[c] = 0 #vetor de acertos para cada criterio de informação
    }

  para(model em modelos){
    para(s em 1:300){# 3 mil simulações
      gera[s,n,model] = modelo(model) #gera modelo de tamanho n para cada modelo ARMA(p,q)#
      para(crit em criterios){
        para(p em 0:2){
          para(q em 0:2){
            se( p e q diferentes de 0){
              candidatos[crit,p,q,n] = modelagem(gera[s,n,model]) #gera modelos candidatos
              se (menorcrit(candidatos[crit,p,q,n]) = ordem(model)){#vefifica se o menor valor
                de criterio corresponde a ordem do modelo gerado#
                ac[crit] = ac[crit] + 1 #acumula acerto para o criterio
              }}}}
            }
          }
        }
      }
    }
  }
}
```

A.2 Tabelas de resultados

Tabela 23: Proporção de acertos em séries simuladas com $n = 50$ para coeficientes ARMA não aleatórios

modelo	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC	ϕ	θ	B1	B0	$n = 50$
1 arma01	0,5977	0,5753	0,6563	0,6383	0,6097	0,6277	0,4583	0,4970	0,6180	0	0.3	0	0	50
2 arma01	0,6023	0,5630	0,6987	0,6723	0,6237	0,6570	0,4490	0,5023	0,6410	0	0.3	0	10	50
3 arma01	0,6187	0,5723	0,7153	0,6873	0,6370	0,6713	0,4463	0,5073	0,6547	0	0.3	1	0	50
4 arma01	0,5423	0,4820	0,7217	0,6770	0,5793	0,6480	0,3803	0,4547	0,6127	0	0.3	1	10	50
5 arma01	0,8993	0,8777	0,9537	0,9403	0,9087	0,9310	0,7087	0,7657	0,9213	0	0.7	0	0	50
6 arma01	0,8543	0,7993	0,9567	0,9303	0,8773	0,9127	0,6350	0,7160	0,8980	0	0.7	0	10	50
7 arma01	0,8530	0,8030	0,9550	0,9280	0,8747	0,9107	0,6500	0,7203	0,8910	0	0.7	1	0	50
8 arma01	0,7630	0,6803	0,9507	0,9040	0,8030	0,8770	0,5353	0,6400	0,8417	0	0.7	1	10	50
9 arma10	0,6270	0,6067	0,6700	0,6570	0,6343	0,6483	0,4877	0,5303	0,6420	0.3	0	0	0	50
10 arma10	0,5830	0,5553	0,6410	0,6283	0,5990	0,6193	0,4467	0,5003	0,6107	0.3	0	0	10	50
11 arma10	0,5807	0,5507	0,6483	0,6307	0,5940	0,6193	0,4460	0,5007	0,6057	0.3	0	1	0	50
12 arma10	0,5290	0,4803	0,6267	0,6010	0,5543	0,5887	0,3943	0,4577	0,5700	0.3	0	1	10	50
13 arma11	0,1003	0,1537	0,0057	0,0213	0,0727	0,0310	0,2973	0,2973	0,0477	0.3	0.3	0	0	50
14 arma11	0,1060	0,1750	0,0000	0,0077	0,0730	0,0213	0,2823	0,2750	0,0440	0.3	0.3	0	10	50
15 arma11	0,1087	0,1827	0,0000	0,0113	0,0790	0,0260	0,2747	0,2677	0,0457	0.3	0.3	1	0	50
16 arma11	0,1120	0,1887	0,0000	0,0060	0,0780	0,0217	0,2523	0,2440	0,0450	0.3	0.3	1	10	50
17 arma11	0,5620	0,5500	0,5833	0,5763	0,5663	0,5740	0,4943	0,5190	0,5693	0.3	0.7	0	0	50
18 arma11	0,5440	0,5220	0,5790	0,5717	0,5497	0,5650	0,4637	0,5053	0,5557	0.3	0.7	0	10	50
19 arma11	0,5190	0,5003	0,5590	0,5510	0,5300	0,5470	0,4473	0,4850	0,5397	0.3	0.7	1	0	50
20 arma11	0,4737	0,4343	0,5397	0,5297	0,4903	0,5200	0,3883	0,4377	0,5073	0.3	0.7	1	10	50
21 arma10	0,8990	0,8713	0,9580	0,9387	0,9077	0,9290	0,6920	0,7537	0,9163	0.7	0	0	0	50
22 arma10	0,8647	0,8187	0,9683	0,9443	0,8853	0,9247	0,6720	0,7440	0,9050	0.7	0	0	10	50
23 arma10	0,8657	0,8187	0,9617	0,9360	0,8900	0,9237	0,6647	0,7430	0,9067	0.7	0	1	0	50
24 arma10	0,8057	0,7350	0,9647	0,9260	0,8387	0,9023	0,5997	0,7020	0,8713	0.7	0	1	10	50
25 arma11	0,5587	0,5447	0,5767	0,5763	0,5627	0,5717	0,4913	0,5147	0,5667	0.7	0.3	0	0	50
26 arma11	0,5180	0,4993	0,5367	0,5413	0,5233	0,5367	0,4530	0,4877	0,5307	0.7	0.3	0	10	50
27 arma11	0,5197	0,5043	0,5380	0,5397	0,5277	0,5367	0,4617	0,4917	0,5317	0.7	0.3	1	0	50
28 arma11	0,4717	0,4527	0,4883	0,5063	0,4837	0,4973	0,4213	0,4550	0,4903	0.7	0.3	1	10	50
29 arma11	0,9043	0,8783	0,9683	0,9507	0,9170	0,9407	0,7510	0,8120	0,9297	0.7	0.7	0	0	50
30 arma11	0,8637	0,8137	0,9723	0,9437	0,8867	0,9293	0,6983	0,7793	0,9113	0.7	0.7	0	10	50
31 arma11	0,8560	0,8063	0,9647	0,9330	0,8817	0,9143	0,6983	0,7727	0,8953	0.7	0.7	1	0	50
32 arma11	0,7577	0,6850	0,9617	0,9097	0,8050	0,8740	0,5897	0,6937	0,8417	0.7	0.7	1	10	50

Tabela 24: Proporção de acertos em séries simuladas com $n = 100$ para coeficientes ARMA não aleatórios

Modelo	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC	ϕ	θ	B1	B0	$n = 100$
1 arma01	0,6767	0,6687	0,6947	0,6897	0,6813	0,6860	0,5050	0,5283	0,6843	0	0.3	0	0	100
2 arma01	0,6957	0,6793	0,7297	0,7207	0,7013	0,7113	0,4833	0,5137	0,7050	0	0.3	0	10	100
3 arma01	0,7060	0,6897	0,7417	0,7313	0,7147	0,7260	0,4733	0,5020	0,7203	0	0.3	1	0	100
4 arma01	0,6663	0,6440	0,7287	0,7080	0,6760	0,6973	0,4313	0,4673	0,6887	0	0.3	1	10	100
5 arma01	0,9400	0,9327	0,9630	0,9580	0,9450	0,9527	0,6973	0,7210	0,9493	0	0.7	0	0	100
6 arma01	0,9313	0,9170	0,9617	0,9523	0,9340	0,9467	0,6667	0,7073	0,9400	0	0.7	0	10	100
7 arma01	0,9313	0,9170	0,9647	0,9537	0,9370	0,9503	0,6687	0,7087	0,9440	0	0.7	1	0	100
8 arma01	0,8757	0,8470	0,9483	0,9287	0,8940	0,9217	0,5627	0,6143	0,9060	0	0.7	1	10	100
9 arma10	0,6673	0,6607	0,6850	0,6813	0,6717	0,6787	0,4973	0,5197	0,6753	0.3	0	0	0	100
10 arma10	0,6637	0,6537	0,6867	0,6803	0,6690	0,6747	0,4923	0,5183	0,6723	0.3	0	0	10	100
11 arma10	0,6640	0,6540	0,6900	0,6837	0,6690	0,6793	0,4877	0,5180	0,6750	0.3	0	1	0	100
12 arma10	0,6107	0,5930	0,6480	0,6350	0,6163	0,6290	0,4277	0,4573	0,6233	0.3	0	1	10	100
13 arma11	0,1737	0,2143	0,0743	0,1010	0,1563	0,1197	0,3237	0,3340	0,1363	0.3	0.3	0	0	100
14 arma11	0,1817	0,2220	0,0510	0,0880	0,1590	0,1133	0,3297	0,3410	0,1340	0.3	0.3	0	10	100
15 arma11	0,1700	0,2180	0,0530	0,0877	0,1470	0,1053	0,3157	0,3250	0,1273	0.3	0.3	1	0	100
16 arma11	0,1597	0,2140	0,0300	0,0663	0,1357	0,0853	0,2697	0,2843	0,1093	0.3	0.3	1	10	100
17 arma11	0,6190	0,6150	0,6307	0,6277	0,6220	0,6260	0,5220	0,5357	0,6233	0.3	0.7	0	0	100
18 arma11	0,5933	0,5887	0,6103	0,6073	0,5970	0,6057	0,4927	0,5130	0,6017	0.3	0.7	0	10	100
19 arma11	0,5997	0,5913	0,6130	0,6097	0,6023	0,6077	0,4937	0,5150	0,6050	0.3	0.7	1	0	100
20 arma11	0,5600	0,5447	0,5950	0,5860	0,5700	0,5823	0,4323	0,4587	0,5760	0.3	0.7	1	10	100
21 arma10	0,9367	0,9293	0,9603	0,9523	0,9410	0,9483	0,6973	0,7273	0,9437	0.7	0	0	0	100
22 arma10	0,9303	0,9153	0,9673	0,9547	0,9363	0,9483	0,6813	0,7207	0,9427	0.7	0	0	10	100
23 arma10	0,9330	0,9203	0,9687	0,9583	0,9407	0,9527	0,6827	0,7187	0,9463	0.7	0	1	0	100
24 arma10	0,8977	0,8753	0,9563	0,9397	0,9083	0,9290	0,6193	0,6643	0,9200	0.7	0	1	10	100
25 arma11	0,6290	0,6243	0,6423	0,6377	0,6323	0,6360	0,5330	0,5473	0,6343	0.7	0.3	0	0	100
26 arma11	0,5943	0,5893	0,6077	0,6043	0,5977	0,6013	0,5097	0,5280	0,6003	0.7	0.3	0	10	100
27 arma11	0,5933	0,5870	0,6073	0,6037	0,5957	0,6017	0,4990	0,5217	0,5987	0.7	0.3	1	0	100
28 arma11	0,5403	0,5303	0,5557	0,5513	0,5443	0,5480	0,4463	0,4673	0,5453	0.7	0.3	1	10	100
29 arma11	0,9370	0,9267	0,9603	0,9553	0,9423	0,9510	0,7577	0,7820	0,9460	0.7	0.7	0	0	100
30 arma11	0,9250	0,9097	0,9673	0,9560	0,9327	0,9487	0,7273	0,7687	0,9407	0.7	0.7	0	10	100
31 arma11	0,9293	0,9103	0,9677	0,9583	0,9373	0,9530	0,7247	0,7583	0,9437	0.7	0.7	1	0	100
32 arma11	0,8823	0,8483	0,9587	0,9387	0,8970	0,9280	0,6397	0,6877	0,9133	0.7	0.7	1	10	100

Tabela 25: Proporção de acertos em séries simuladas com $n = 200$ para coeficientes ARMA não aleatórios

modelo	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC	ϕ	θ	B1	B0	$n = 200$
1 arma01	0,7550	0,7527	0,7607	0,7583	0,7557	0,7570	0,5577	0,5670	0,7560	0	0.3	0	0	200
2 arma01	0,7620	0,7583	0,7723	0,7690	0,7630	0,7667	0,5350	0,5523	0,7637	0	0.3	0	10	200
3 arma01	0,7590	0,7533	0,7723	0,7693	0,7607	0,7663	0,5340	0,5467	0,7643	0	0.3	1	0	200
4 arma01	0,7490	0,7387	0,7763	0,7663	0,7540	0,7637	0,4900	0,5063	0,7600	0	0.3	1	10	200
5 arma01	0,9687	0,9663	0,9750	0,9737	0,9697	0,9723	0,7087	0,7247	0,9713	0	0.7	0	0	200
6 arma01	0,9503	0,9410	0,9620	0,9590	0,9520	0,9570	0,6600	0,6770	0,9543	0	0.7	0	10	200
7 arma01	0,9517	0,9457	0,9680	0,9643	0,9543	0,9610	0,6643	0,6860	0,9577	0	0.7	1	0	200
8 arma01	0,9357	0,9277	0,9607	0,9520	0,9383	0,9480	0,6050	0,6287	0,9447	0	0.7	1	10	200
9 arma10	0,7497	0,7467	0,7537	0,7530	0,7503	0,7527	0,5670	0,5787	0,7513	0.3	0	0	0	200
10 arma10	0,7263	0,7250	0,7343	0,7310	0,7270	0,7297	0,5300	0,5413	0,7283	0.3	0	0	10	200
11 arma10	0,7427	0,7387	0,7500	0,7463	0,7427	0,7453	0,5383	0,5503	0,7437	0.3	0	1	0	200
12 arma10	0,7010	0,6970	0,7153	0,7090	0,7033	0,7057	0,4860	0,5000	0,7047	0.3	0	1	10	200
13 arma11	0,3983	0,4097	0,3627	0,3760	0,3910	0,3817	0,3787	0,3853	0,3867	0.3	0.3	0	0	200
14 arma11	0,4007	0,4113	0,3403	0,3640	0,3917	0,3737	0,3640	0,3713	0,3827	0.3	0.3	0	10	200
15 arma11	0,4080	0,4173	0,3483	0,3733	0,4007	0,3860	0,3743	0,3843	0,3947	0.3	0.3	1	0	200
16 arma11	0,3843	0,3997	0,3100	0,3433	0,3753	0,3567	0,3550	0,3627	0,3667	0.3	0.3	1	10	200
17 arma11	0,7020	0,7007	0,7077	0,7063	0,7030	0,7057	0,5977	0,6060	0,7037	0.3	0.7	0	0	200
18 arma11	0,6680	0,6660	0,6750	0,6730	0,6703	0,6723	0,5507	0,5607	0,6717	0.3	0.7	0	10	200
19 arma11	0,6627	0,6597	0,6673	0,6660	0,6627	0,6650	0,5497	0,5567	0,6633	0.3	0.7	1	0	200
20 arma11	0,6347	0,6293	0,6530	0,6457	0,6390	0,6437	0,4803	0,4923	0,6413	0.3	0.7	1	10	200
21 arma10	0,9653	0,9623	0,9763	0,9733	0,9680	0,9717	0,7087	0,7210	0,9703	0.7	0	0	0	200
22 arma10	0,9563	0,9500	0,9723	0,9687	0,9583	0,9637	0,6780	0,6980	0,9593	0.7	0	0	10	200
23 arma10	0,9623	0,9563	0,9770	0,9720	0,9650	0,9693	0,6837	0,7037	0,9673	0.7	0	1	0	200
24 arma10	0,9427	0,9347	0,9647	0,9577	0,9447	0,9520	0,6430	0,6653	0,9493	0.7	0	1	10	200
25 arma11	0,6630	0,6607	0,6670	0,6663	0,6637	0,6657	0,5503	0,5563	0,6647	0.7	0.3	0	0	200
26 arma11	0,6630	0,6610	0,6703	0,6690	0,6650	0,6677	0,5670	0,5750	0,6667	0.7	0.3	0	10	200
27 arma11	0,6583	0,6550	0,6650	0,6630	0,6597	0,6620	0,5557	0,5620	0,6607	0.7	0.3	1	0	200
28 arma11	0,6317	0,6280	0,6430	0,6403	0,6337	0,6383	0,5190	0,5317	0,6360	0.7	0.3	1	10	200
29 arma11	0,9637	0,9587	0,9740	0,9713	0,9647	0,9700	0,7620	0,7750	0,9683	0.7	0.7	0	0	200
30 arma11	0,9603	0,9553	0,9773	0,9727	0,9623	0,9677	0,7293	0,7487	0,9650	0.7	0.7	0	10	200
31 arma11	0,9617	0,9540	0,9743	0,9703	0,9650	0,9693	0,7317	0,7490	0,9680	0.7	0.7	1	0	200
32 arma11	0,9343	0,9243	0,9613	0,9537	0,9413	0,9487	0,6610	0,6787	0,9447	0.7	0.7	1	10	200

A.3 Resultados da simulação de experimento com coeficientes ARMA não determinísticos

Inicialmente foram determinadas as proporções de acertos de cada um dos critérios AIC, AICc e BIC, além dos seis propostos nesse estudo.

É importante destacar que, embora os valores atribuídos a θ e ϕ tenham valores de 0,3 e 0,7 (como no caso do experimento descrito na Seção 4.1), respectivamente, esses valores foram ajustados com um peso aleatório em cada rodada da simulação, conforme descrito na Tabela 26.

Tabela 26: Valores atribuídos aos coeficientes utilizados na geração de modelos $ARMA(p, q)$

Fatores	Níveis
ϕ_1	0,3 e 0,7 com desvio 0,05
θ_1	0,3 e 0,7 com desvio 0,05
β_0	0 e 10
β_1	0 e 1

As medidas descritivas dos resultados encontrados que incluem Mínimo, Primeiro Quartil, Mediana, Média, Terceiro Quartil, Máximo, Desvio Padrão, Coeficiente de Curtose e Coeficiente de Assimetria estão detalhadas na Tabela 27.

Tabela 27: Medidas descritivas da proporção de acertos de cada Critério de Informação para as séries simuladas.

Medidas	T1	T2	T3	T4	T5	T6	AIC	AICc	BIC
Mínimo	0,0833	0,0927	0,0167	0,0517	0,0727	0,0753	0,1393	0,0850	0,0913
1º Quartil	0,4819	0,4703	0,4204	0,4372	0,4704	0,4496	0,4220	0,4338	0,4660
Mediana	0,5871	0,5840	0,5863	0,5850	0,5851	0,5861	0,4897	0,5019	0,5866
Média	0,5753	0,5641	0,5480	0,5573	0,5701	0,5648	0,5146	0,5115	0,5709
3º Quartil	0,7366	0,6960	0,7191	0,7211	0,7338	0,7254	0,5736	0,5881	0,7256
Máximo	0,9610	0,9583	0,9750	0,9677	0,9620	0,9653	0,7020	0,7153	0,9630
Desvio-padrão	0,3006	0,2913	0,3294	0,3155	0,3052	0,3063	0,2712	0,2819	0,2984
Curtose	0,8614	1,0113	0,5662	0,5641	0,7839	0,5556	1,7932	1,6234	0,6847
Assimetria	-0,6013	-0,4973	-0,5947	-0,5334	-0,5884	-0,5038	0,5907	0,1101	-0,5071

Obs: Maiores resultados = negrito itálico; Menores resultados = negrito.

Na Tabela 27 observa-se que, em geral, os maiores valores para a maior parte das medidas estão associados ao critério T1. O critério AIC continua tendo a maior parte de valores menores.

Em relação ao menor valor mínimo, o critério T3 apresentou o menor valor (0,0167), indicando que possui a menor proporção de acertos mais baixos em comparação aos demais critérios. O critério AICc mostrou o maior valor mínimo (0,1393), sugerindo uma proporção mais alta de acertos.

Quanto à curtose, todas as distribuições foram classificadas como leptocúrticas, o que significa que elas tem mais caudas pesadas, com um pico mais acentuado em torno da média do que uma distribuição normal. Em contraste, o critério AIC apresentou a menor curtose, sugerindo uma distribuição com menor pico e menos concentrada em torno da tendência central.

No que diz respeito a assimetria, com excessão dos critérios AIC e AICc, todas as outras distribuições foram identificadas como assimétricas negativas.

Na Figura 144 é possível ver os gráficos Boxplots para as proporções de acertos de cada critério estudado, enquanto que na Figura 145, os gráficos Boxplots desses mesmos critérios por modelo ARIMA.

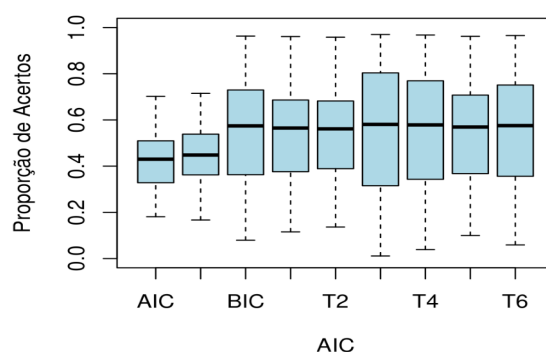


Figura 144: Boxplot da proporção de acertos para cada Critério de Informação

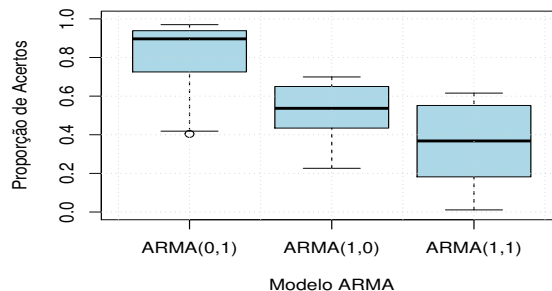


Figura 145: Boxplot da proporção de acertos para cada Critério de Informação

Avaliando a Figura 144, é possível notar que os critérios BIC, T1, T2, T3, T4, T5 e T6 tem comportamentos semelhantes, e além de que nenhum desses exibiu valores que se situam fora do padrão. Já na Figura 145, é possível notar que dentre os modelos, os critérios se saem melhor quando o modelo considerado é o ARMA(0,1).

A Tabela 28 mostra os valores de proporção média para os critérios utilizados, subdividindo-os em relação ao tipo de modelo e ao tamanho da série.

Tabela 28: Proporção média de acertos de cada Critério de Informação por tipo de modelo e ao tamanho da série.

n	ARMA(p,q)	Critérios de Informação								
		AIC	AICc	BIC	T1	T2	T3	T4	T5	T6
50	ARMA(0,1)	0.5465	0.6342	0.8325	0.7773	0.7179	0.9189	0.8803	0.8056	0.8574
	ARMA(1,0)	0.3026	0.3531	0.4468	0.4228	0.395	0.4748	0.4653	0.4351	0.4579
	ARMA(1,1)	0,259	0,2419	0,1013	0,1406	0,1736	0,0332	0,0632	0,1211	0,0822
100	ARMA(0,1)	0.5911	0.6342	0.9107	0.8908	0.8702	0.9412	0.9265	0.9014	0.9189
	ARMA(1,0)	0.3853	0.4096	0.575	0.5647	0.5525	0.5915	0.5843	0.5701	0.5795
	ARMA(1,1)	0.3929	0.4073	0.3626	0.3792	0.3913	0.3146	0.3414	0.3711	0.3528
200	ARMA(0,1)	0,6351	0,6552	0,9527	0,9473	0,9400	0,9637	0,9584	0,9500	0,9558
	ARMA(1,0)	0.4731	0.4866	0.6775	0.674	0.6702	0.6846	0.681	0.676	0.679
	ARMA(1,1)	0.4859	0.4949	0.5848	0.5852	0.5854	0.5820	0.5837	0.5853	0.5841
Todos	ARMA(0,1)	0.5909	0.6412	0.8986	0.8718	0.8427	0.9412	0.9217	0.8857	0.9107
	ARMA(1,0)	0.3870	0.4164	0.5664	0.5538	0.5392	0.5837	0.5769	0.5604	0.5721
	ARMA(1,1)	0.3793	0.3814	0.3496	0.3683	0.3834	0.3099	0.3294	0.3592	0.3397

Obs: Maiores resultados = negrito itálico; Menores resultados = negrito.

Na Tabela 28 se observa que a proporção média de acertos dos Critérios de Informação varia entre os diferentes modelos ARMA e tamanhos de série. Notavelmente, o modelo ARMA(0,1) consistentemente apresentou as maiores proporções médias de acertos, destacando-se especialmente em séries de tamanho 200, onde os resultados são notavelmente superiores. Em contraste, o modelo ARMA(1,1) apresentou as menores proporções de acertos, especialmente nos critérios AIC e T3. O critério T3 destacou-se com uma ligeira maior proporção média de acertos, especialmente no modelo ARMA(0,1) para séries de tamanho 200, alcançando 0,9637.

A.3.1 Análise de Variância

Utilizando o teste de Friedman, se obteve $\chi^2 = 4740,9$, com 8 graus de liberdade, e $valorp = 2,10^{-16}$. Esse resultado indica que há evidência estatística para concluir que há pelo menos uma diferença significativa entre as proporções de acertos obtidas pelos critérios estudados.

Para analisar mais detalhadamente estas diferenças detectadas foram realizadas comparações múltiplas utilizando o teste de Wilcoxon. Os resultados podem ser observados na Tabela 29.

Tabela 29: Resultados do *valor – p* obtidos pelos testes de Wilcoxon

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	<0,01	-	-	-	-	-	-	-
BIC	<0,01	<0,01	-	-	-	-	-	-
T1	<0,01	<0,01	>0,99	-	-	-	-	-
T2	<0,01	<0,01	>0,99	>0,99	-	-	-	-
T3	<0,01	<0,01	>0,99	>0,99	>0,99	-	-	-
T4	<0,01	<0,01	>0,99	>0,99	>0,99	>0,99	-	-
T5	<0,01	<0,01	>0,99	>0,99	>0,99	>0,99	>0,99	-
T6	<0,01	<0,01	>0,99	>0,99	>0,99	>0,99	>0,99	>0,99

A Tabela 29 indica que há evidência de diferença entre todas as proporções de acertos obtidas entre os critérios AIC, AICc e o restante. Para os critérios BIC e T1 a T6 parece não haver indicativo de diferença.

No entanto, parece crucial considerar que essas proporções podem estar sendo influenciadas por alguns dos fatores considerados no experimento, especialmente o tipo de modelo ARMA, objeto de estudo desta análise. Os Boxplots apresentados nas Figuras 146, 147, 149, 150, 151, 152, 153 e 154, parecem corroborar com essa observação.

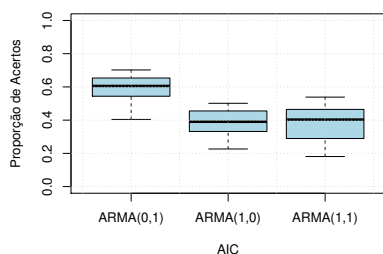


Figura 146: Gráfico Boxplot com os modelos considerados para o critério AIC

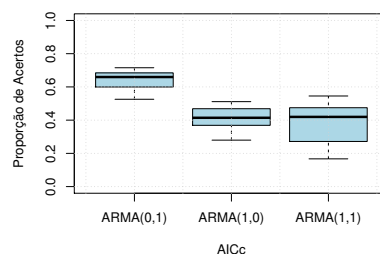


Figura 147: Gráfico Boxplot com os modelos considerados para o critério AICc

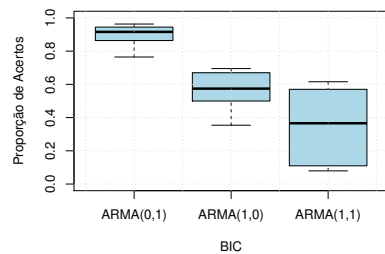


Figura 148: Gráfico Boxplot com os modelos considerados para o critério BIC

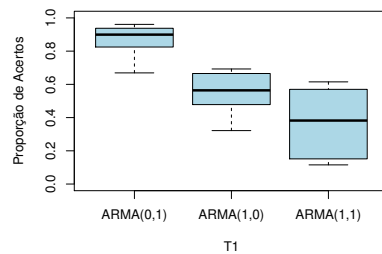


Figura 149: Gráfico Boxplot com os modelos considerados para o critério T1

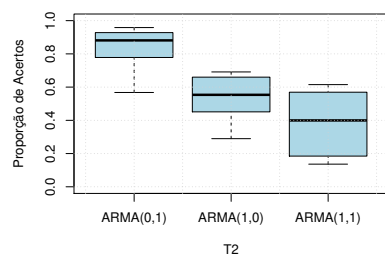


Figura 150: Gráfico Boxplot com os modelos considerados para o critério T2

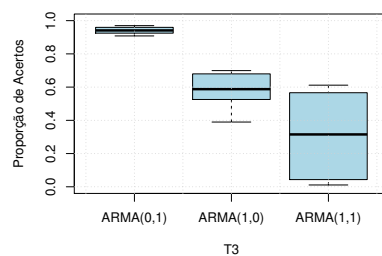


Figura 151: Gráfico Boxplot com os modelos considerados para o critério T3

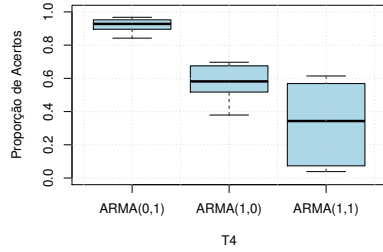


Figura 152: Gráfico Boxplot com os modelos considerados para o critério T4

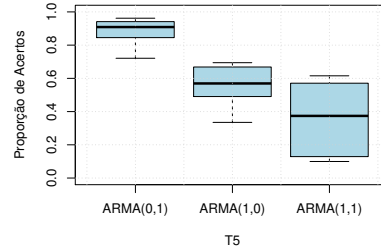


Figura 153: Gráfico Boxplot com os modelos considerados para o critério T5

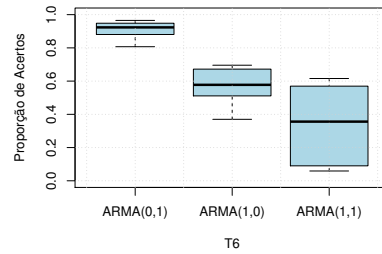


Figura 154: Gráfico Boxplot com os modelos considerados para o critério T6

Nos gráficos de boxplot, observa-se uma maior variabilidade nos modelos ARMA(1,1) para todos os critérios, enquanto o modelo ARMA(0,1) apresenta a menor variabilidade, sendo a menor de todas para o critério T3. Os critérios T1, T2 e T5 mostraram valores fora do padrão.

Novos testes de Friedman foram realizados, para cada um dos tipos de modelo (ARMA(0,1), ARMA(1,0), ARMA(1,1)), encontrando-se *valor - p* = $2,10^{-16}$ com 7 graus de liberdade nas três aplicações. Os resultados das comparações múltiplas utilizando o teste de Wilcoxon podem ser observados nas Tabelas 30, 31 e 32.

Tabela 30: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(0,1)

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	<0,01	-	-	-	-	-	-	-
BIC	<0,01	<0,01	-	-	-	-	-	-
T1	<0,01	<0,01	<0,01	-	-	-	-	-
T2	<0,01	<0,01	<0,01	<0,01	-	-	-	-
T3	<0,01	<0,01	<0,01	<0,01	<0,01	-	-	-
T4	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	-
T5	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-
T6	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01

Tabela 31: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,0)

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	0,0108	-	-	-	-	-	-	-
BIC	0,0108	<0,01	-	-	-	-	-	-
T1	<0,01	<0,01	<0,01	-	-	-	-	-
T2	0,0108	<0,01	<0,01	<0,01	-	-	-	-
T3	<0,01	<0,01	<0,01	<0,01	<0,01	-	-	-
T4	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-	-
T5	0,0108	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	-
T6	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01

Tabela 32: Resultados do *valor - p* para os testes de Wilcoxon para todos os tamanhos considerados para o modelo ARMA(1,1)

	AIC	AICc	BIC	T1	T2	T3	T4	T5
AICc	>0,99	-	-	-	-	-	-	-
BIC	0,8486	>0,99	-	-	-	-	-	-
T1	>0,99	>0,99	<0,01	-	-	-	-	-
T2	>0,99	>0,99	<0,01	<0,01	-	-	-	-
T3	0,4101	0,0843	<0,01	1e-04	1e-04	-	-	-
T4	0,5136	0,4276	<0,01	<0,01	<0,01	<0,01	-	-
T5	0,8486	>0,99	<0,01	<0,01	<0,01	1e-04	<0,01	-
T6	0,5578	0,8486	<0,01	<0,01	<0,01	<0,01	<0,01	<0,01

Os resultados apresentados nas Tabelas 30 e 31 indicam que há evidência de diferença entre as proporções de acertos obtidos por cada um dos Critérios de Informação para séries descritas por modelos do tipo ARMA(0,1) e ARMA(1,0), em função dos baixos valores de probabilidade (*valor - p* < 0.05), o mesmo não acontece para séries descritas por modelos do tipo ARMA(1,1). Neste caso, os testes de Wilcoxon, apresentados na Tabela 32, indicam que não existem diferenças significativas entre alguns Critérios de Informação, especialmente, AICc e AIC em relação aos demais, considerando nível de significância de 0.05, entre BIC, T1, T2, T3, T4, T5 e T6 foram encontradas evidências de diferença.

Os resultados obtidos pela regressão Beta para o experimento na condição não determinística revelaram padrões que se alinham com os observados no caso determinístico em grande parte das situações. No entanto, em alguns casos, especialmente aqueles que envolveram o modelo ARMA(1,1), não foi possível alcançar um ajuste abaixo de 5% para o gráfico envelope em quase todos os critérios. Por outro lado, as figuras de interação relacionadas à proporção exibiram um padrão semelhante aos obtidos para o caso determinístico. Para evitar repetições excessivas na descrição dos resultados, optou-se por focar e registrar apenas as situações mais específicas deste experimento. Nesse contexto, observou-se uma tendência de que a proporção média de acertos é consistentemente menor quando são utilizados os parâmetros $\phi=0.3 \pm 0,05$ e $\theta=0.3 \pm 0,05$, em comparação com os valores mais elevados ($\phi = 0.7 \pm 0,05$ e $\theta = 0.7 \pm 0,05$). Essa relação se mostrou relativamente estável nos diferentes tamanhos de amostra (50, 100 e 200) e também entre os componentes analisados, como B0 e B1..

A.4 Resultados encontrados do Gradiente e da Derivada Direcional para os critérios AIC, AICc e BIC

Uma função de várias variáveis em $x_1, x_2, \dots, x_m$ é uma função $w = w(x_1, x_2, \dots, x_m)$ que associa a cada conjunto de valores de $x_1, x_2, \dots, x_m$ um valor w . Um campo vetorial, no entanto, é uma estrutura matemática fundamental na qual cada ponto em um espaço está associado a um vetor. Essencialmente, um campo vetorial descreve a distribuição de vetores em uma região do espaço, atribuindo a cada ponto um vetor específico que indica a direção e a magnitude. Se um campo vetorial estiver restrito a uma função de duas variáveis k e n , matematicamente, o campo em um sistema bidimensional será descrito conforme Equação (106) (Anton et al., 2014):

$$\vec{F}(k, n) = A(k, n)\hat{i} + B(k, n)\hat{j}, \quad (106)$$

onde $A(k, n)$ e $B(k, n)$ são funções escalares de duas variáveis nas variáveis k e n e $\hat{i}$ e $\hat{j}$ são vetores unitários geralmente associados a um sistema retangular de eixos bidimensional.

O vetor gradiente é um conceito fundamental em Cálculo, principalmente em tópicos de Cálculo Vetorial, e é usado para descrever a taxa de variação de uma função em uma determinada direção. Ele é definido como um vetor que aponta

na direção de maior crescimento da função e cuja magnitude representa a taxa de variação em relação a essa direção (Anton et al., 2014). Se $w = w(k, n)$ for uma função de duas variáveis (k e n), então a equação para o vetor gradiente nessas condições é dada pela Equação (107), ou seja:

$$\nabla w = \frac{\partial w}{\partial k} \hat{i} + \frac{\partial w}{\partial n} \hat{j}, \quad (107)$$

onde $w = w(k, n)$ é uma função das variáveis k e n .

Sendo as derivadas parciais da função w funções das variáveis k e n , pode-se dizer intrinsecamente que o gradiente forma um campo vetorial nessas variáveis, sendo que este vetor ∇w aponta na direção de maior crescimento da função w e cuja magnitude representa a taxa de variação em relação a essas variáveis (Anton et al., 2014).

A derivada direcional, no entanto, pode ser entendida como a medida da taxa de variação de uma determinada função em uma dada direção específica, que pode ser qualquer direção no espaço. Essa função é definida como o produto escalar do vetor gradiente da função w por um vetor unitário que aponta na direção desejada (Anton et al., 2014). Sendo então $w = w(k, n)$ uma função de duas variáveis e $\vec{u} = (u_1, u_2)$ um vetor unitário, a derivada direcional pode ser definida a partir do produto escalar (*) dado pela Equação (108), ou seja:

$$D_{\vec{u}}w = \nabla w * \vec{u}. \quad (108)$$

Utilizando os conceitos de produto escalar e lembrando que o módulo de $\vec{u} = 1$, uma outra forma de avaliar esta derivada direcional é por meio da Equação (109) (Anton et al., 2014):

$$D_{\vec{u}}w = |\nabla w| \cos(\psi), \quad (109)$$

onde ψ representa o ângulo entre o vetor gradiente e o vetor unitário.

Sendo as penalidades dos critérios AIC, AICc e BIC funções de duas variáveis, podemos determinar essas grandezas para encontrarmos relações que podem ser interessantes.

Os vetores do campo Gradiente, determinados a partir da Equação (107), associados a penalidade de AIC, podem ser observado a partir da Equação (110), sendo ele:

$$\nabla w_{AIC} = 2\hat{i}, \quad (110)$$

onde: $\hat{i} = (1, 0)$.

O campo Gradiente, definido a partir da Equação (110), em um espaço bidimensional, tem o seu comportamento representado na Figura 155.

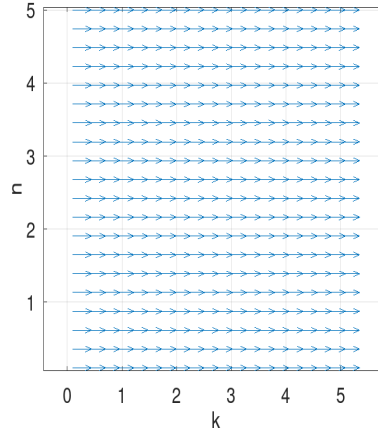


Figura 155: Comportamento do Campo Gradiente associado à função da penalidade de AIC

No gráfico da Figura 155, é possível observar que a direção dos vetores desse campo Gradiente para a penalidade utilizada pelo AIC está na direção de k , enquanto o módulo desse gradiente, que representa a taxa de crescimento nessa direção, é igual a dois.

O campo gradiente mostrado na Figura 155, determinado em relação a penalidade do AIC, tende a apontar exclusivamente para a direção de k (direita) à medida que n cresce. Este achado pode ser demonstrado matematicamente. Para isso considere que um vetor $\vec{v} = (1, 1)$, que também pode ser escrito como $\vec{v} = \hat{i} + \hat{j}$. É possível identificar um ponto (k, n) no plano como $(k, n) = k\hat{i} + n\hat{j}$. Considere também um vetor $\vec{u}$ unitário, que tem a mesma direção de $\vec{v}$, sendo definido como $\vec{u} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$.

Para a penalidade de AIC, se for determinada a derivada direcional na direção desse vetor, encontramos a expressão dada pela Equação (111):

$$D_{\vec{u}}w = \sqrt{2}, \quad (111)$$

mostrando que a derivada direcional para esse caso independe de n e de k . Usando a Equação (109), é possível determinar o valor de $\cos(\psi)$, que assume

o valor apresentado na Equação (112):

$$\cos(\psi) = \frac{\sqrt{2}}{2}, \quad (112)$$

sendo possível concluir que $\psi = \frac{\pi}{4} rad$ independente de k e n .

Os vetores Gradiente, determinados a partir da Equação (107), associados à penalidade de AICc está apresentado na Equação 113, ou seja:

$$\nabla w_{AICc} = \frac{2n(n-1)}{(n-k-1)^2} \hat{i} - \frac{2k(k+1)}{(n-k-1)^2} \hat{j}, \quad (113)$$

onde: $\hat{i} = (1, 0)$ e $\hat{j} = (0, 1)$.

O campo Gradiente, definido a partir da Equação (113) em um espaço bidimensional, tem o seu comportamento representado na Figura 156.

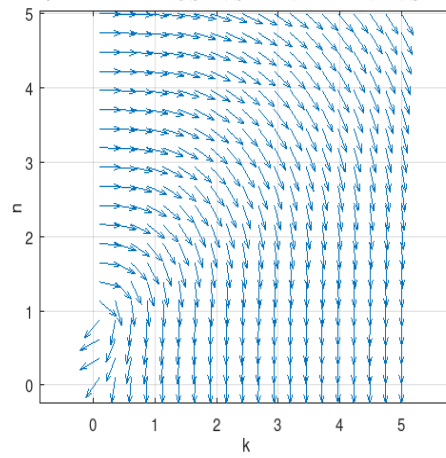


Figura 156: Comportamento do Campo Gradiente associado à função da penalidade de AICc

A Figura 156 revela que os vetores do campo Gradiente associados a penalidade do critério AICc tendem a ter uma curvatura para baixo a medida que k aumenta em relação a n . Esse padrão se revela ao longo de todo n . Também é possível observar que a medida que n se torna maior que k este campo tende a ter um comportamento idêntico ao do critério AIC. Também é possível observar que a medida que n cresce, o vetor gradiente da penalidade deste critério tende a ter o mesmo comportamento do vetor gradiente relativo a penalidade do critério AIC.

Para demonstrarmos essas características interessantes, utilizemos o mesmo vetor $\vec{u} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, nessa situação a derivada direcional para o caso toma a forma dada pela Equação 114:

$$D_{\vec{u}}w_{AICc} = \frac{\sqrt{2}n(n-1)}{(n-k-1)^2} - \frac{\sqrt{2}k(k+1)}{(n-k-1)^2}, \quad (114)$$

mostrando que seu resultado depende de k e de n . Se tomarmos a condição de n muito maior que um k fixo em um limite onde n tende ao infinito é possível mostrar que a derivada direcional de AICc tende a ficar com a mesma característica da penalidade de AIC, ou seja:

$$\lim_{n \rightarrow \infty} D_{\vec{u}}w_{AICc} = \lim_{n \rightarrow \infty} \frac{\sqrt{2}n(n-1)}{(n-k-1)^2} - \frac{\sqrt{2}k(k+1)}{(n-k-1)^2} = \sqrt{2}. \quad (115)$$

Para entendermos a curvatura desse campo tomemos o caso extremo em que $n = k$, nesse caso o valor de $\cos(\psi)$ é dado pela Equação 116:

$$\cos(\psi) = \frac{-1}{\sqrt{n^2 + 1}}. \quad (116)$$

Nesse caso, o comportamento dessa medida para a penalidade de AICc em relação ao $\cos(\psi)$ pode ser observado na Figura 157.

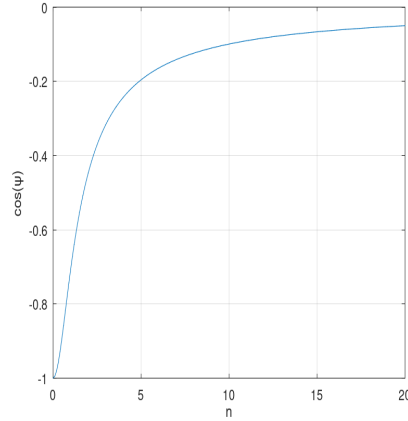


Figura 157: Comportamento do $\cos(\psi)$ quando $k = n$ para a penalidade de AICc

Pode-se observar que para n grande, a Equação (157) mostra que o campo Gradiente tende a formar um ângulo de $\frac{\pi}{2} rad$ com o vetor $\vec{u}$. Demonstrando assim o comportamento relacionado.

Generalizando, para qualquer par de valores de k e n , iguais ou não, o $\cos(\psi)$ para a penalidade AICc será dada pela Equação (117), ou seja:

$$\cos(\psi) = \frac{\sqrt{2}}{2} \frac{n(n-1) - k(k+1)}{\sqrt{n^2(n-1)^2 + k^2(k+1)^2}}. \quad (117)$$

É possível observar nessa equação que quando n é bem grande o cosseno da penalidade de AICc tende a ficar próximo ao da penalidade de AIC. No entanto, vale ressaltar que quando $n = k$, conforme Equação (116), não temos o mesmo comportamento que a Equação (112), já que essa equação para $n = k$ e n grande fica sempre igual a $\sqrt{2}/2$. Isso ocorre porque as penalidades dos critérios AIC e AICc, apesar de terem comportamento semelhante quando n é muito grande, não possuem o mesmo comportamento quando $n = k$, já que para AIC o valor da penalidade é $w_{AIC} = 2n$, enquanto que para AICc o valor da penalidade é $w_{AICc} = -2n^2$, conforme pode ser verificado utilizando as Equações (18) e (19).

O vetor Gradiente, determinado a partir da Equação (107), associado à penalidade de BIC, está apresentado na Equação (118), ou seja:

$$\nabla w_{BIC} = (\ln(n)) \hat{i} + \left(\frac{k}{n}\right) \hat{j}, \quad (118)$$

onde: $\hat{i} = (1, 0)$ e $\hat{j} = (0, 1)$.

O campo Gradiente, definido a partir das Equação (118) em um espaço bidimensional, tem o seu comportamento representado na Figura 158.

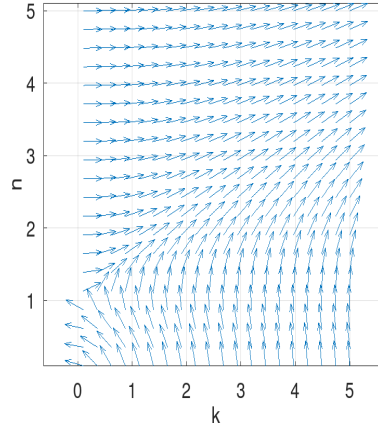


Figura 158: Comportamento do Campo Gradiente associado à função da penalidade de BIC

A Figura 158 mostra que o vetor Gradiente associado à penalidade do BIC tende a realizar uma curva para cima a medida que k aumenta em relação a n . O gráfico também mostra que ele tende a apontar para uma dada direção conforme a relação entre k e n vai aumentando. Da mesma maneira que foi feito com os critérios AIC e AICc, tomemos a derivada direcional no sentido do vetor $\vec{u} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, onde encontramos a Equação (119):

$$D_{\vec{u}} w_{BIC} = \frac{\sqrt{2}}{2} \ln(n) + \frac{\sqrt{2}k}{2n}. \quad (119)$$

Já o ângulo ψ , formado pelo vetor $\vec{u}$ acima definido, e o vetor Gradiente determinado pela Equação (120), para $n = k$, é determinado pela Equação (109):

$$\cos(\psi) = \frac{\sqrt{2}}{2} \left(\frac{\ln(n) + 1}{\sqrt{\ln^2(n) + 1}} \right), \quad (120)$$

mostrando que seu valor depende n e de k . Nesse caso, o comportamento deste $\cos(\psi)$, para esse caso, pode ser observado pela Figura 159,

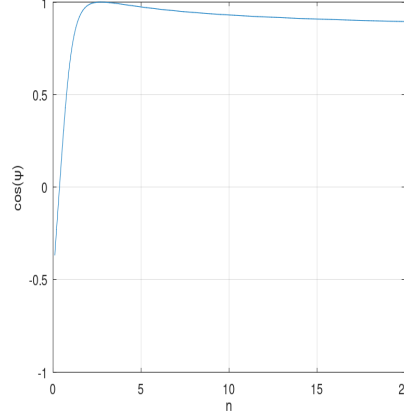


Figura 159: Comportamento do $\cos(\psi)$ para a penalidade de BIC quando $k = n$

Pode-se observar que a medida que n aumenta, o $\cos(\psi)$ tende a se aproximar do valor $\frac{\sqrt{2}}{2}$, o que indica que para n grande e uma modelagem na qual o número de parâmetros k seja igual ao tamanho da amostra n , o vetor Gradiente forma um ângulo de $\frac{\pi}{4} rad$ em relação ao vetor $\vec{u}$.

Para a penalidade de BIC, a generalização será dada pela Equação (121), ou seja:

$$\cos(\psi) = \frac{\sqrt{2} \left(\ln(n) + \frac{k}{n} \right)}{2 \left(\ln^2(n) + \frac{k^2}{n^2} \right)}. \quad (121)$$

As expressões dadas pelas Equações (112), (117) e (121) conectam as relações dadas pelas taxas de variação em k e em n em uma única equação, ou mais precisamente ao movimento relacionado aos modelos candidatos (variando k) e ao movimento relacionado ao tamanho da amostra para esses Critérios de Informação (variando n).

Baseado nessas ideias, pode-se dizer que o Gradiente funciona como um tipo de transportador da informação à medida que os parâmetros n e k dos modelos variam. Enquanto isso, a derivada direcional especifica como esse transporte ocorre, e o $\cos(\psi)$ pode ser considerado como o indicador detalhado desse processo, especialmente em relação à penalidade. E como pode ser percebido, por mais que dois critérios tenham comportamentos semelhantes em relação a esse

cosseno, esse valor é uma espécie de assinatura de cada tipo de penalidade atribuída ao critério de informação. Para tanto, definiremos o $\cos(\psi)$ como sendo o grau de complexidade da penalidade do critério de informação.

Para os casos extremos em que $n = k$, conforme observado pelas Equações (116) e (120), e que descrevem o cosseno relacionado a penalidade dos critérios de informação AICc e BIC, respectivamente, definiremos a medida do cosseno para esses casos como sendo o grau de complexidade máxima. Dado que o $\cos(\psi)$ do critério AIC é sempre o mesmo valor, se considerarmos que para n relativamente grande os cossenos para AICc e BIC também se estabilizam, conforme observado, diremos nesse caso que o grau de complexidade máxima atingiu seu máximo valor. Aqui é interessante notar que essas características peculiares são sempre garantidas devido ao fato de o cosseno ser restrito a valores dentro do intervalo $[-1, 1]$. Podemos generalizar essa característica para toda penalidade de um Critério de Informação que possui dependências com as variáveis k e n . Podemos nesse momento generalizar para uma penalidade de um Critério de Informação em termos de uma coleção de variáveis $x_1, x_2, \dots, x_n$. Nesse contexto, dado que a penalidade pode ser escrita como $w = w(x_1, x_2, \dots, x_n)$ dado que o gradiente, a derivada direcional serão determinadas em termos dessas variáveis, em situações onde o gradiente seja não nulo, podemos garantir que essas características também serão garantidas para o cosseno, generalizando assim que o grau de complexidade é uma medida intrínseca de um Critério de Informação e não necessariamente somente dos critérios AIC, AICc e BIC.

Dito isso, considera-se que essas medidas podem ser utilizadas para complementar o conceito de complexidade de um modelo, pois conectam simultaneamente as grandezas relevantes k e n , com relação a penalidade, em termos dos seus movimentos relativos (derivadas parciais) em relação a essas variáveis.

A.5 Diferencial e equação diferencial para penalidades

Uma das características dos campos gradientes é que os mesmos são conservativos pelo teorema de Stokes (Stewart, 2013), e uma das características disso é que a integral sobre uma curva fechada é nula independente do caminho $\vec{r}$ escolhido, matematicamente expresso a partir da Equação (122):

$$\oint \nabla w * d\vec{r} = 0. \quad (122)$$

Uma outra característica dos campos conservativos do tipo $\vec{F} = \vec{F}(\vec{r})$ é que

é possível expressar diretamente uma grandeza W simplesmente se conhecendo os pontos inicial a e final b usando a Equação (123), ou seja:

$$W = \int_a^b \vec{F}(\vec{r}) \cdot d\vec{r} = w(b) - w(a), \quad (123)$$

O que permite calcular W pelo campo conservativo, de um ponto a outro, apenas a partir da função w (em alguns contextos chamado de campo potencial) em cada um desses pontos a partir da Equação (124).

$$W = w(b) - w(a). \quad (124)$$

Sendo $\vec{F}$ conservativo, então o campo $\vec{F}$ pode ser escrito como o Gradiente de um campo potencial w , ou mais precisamente $\vec{F} = \nabla w$.

Tomemos a diferencial sobre $w = w(k, n)$, assim:

$$dw = \frac{\partial w(n, k)}{\partial k} dk + \frac{\partial w(n, k)}{\partial n} dn = \nabla w * d\vec{r}, \quad (125)$$

onde $d\vec{r} = dk\hat{i} + dn\hat{j}$.

Vamos impor que a superfície $w(k, n)$ seja uma superfície equipotencial (calcular W independe do caminho escolhido). Assim, a Equação (125) toma a forma dada pela Equação (126):

$$\frac{\partial w(n, k)}{\partial k} dk + \frac{\partial w(n, k)}{\partial n} dn = 0. \quad (126)$$

Uma forma mais generalizada, dado que nem sempre sabemos quem é $w(k, n)$, é dada sob a forma da Equação (127):

$$A(k, n)dk + B(k, n)dn = 0. \quad (127)$$

Que é uma equação diferencial de primeira ordem em k e n , sendo sua interpretação em cima das taxas dk e dn . Mais precisamente, podemos dizer que a taxa dk muda a partir de um termo $A(k, n)$, enquanto dn muda a partir de um termo $B(k, n)$. Isso é interessante, pois como visto anteriormente no Apêndice A.4, a penalidade de um critério está intimamente relacionada a sua taxa de variação.

Um exemplo de construção poderia envolver a definição de uma penalidade que responda de maneira diferenciada a variações em dois parâmetros, dk e dn . Especificamente, a penalidade deve aumentar proporcionalmente conforme dk aumenta, seguindo uma função $c(k, n)$ e também aumentar proporcionalmente

conforme dn aumenta, mas de acordo com uma função constante α que deve aumentar com k e diminuir com n . Nesse cenário, uma equação que poderia ser utilizada para descrever essa penalidade é:

$$c(k, n)dk + \alpha \frac{k}{n} dn = 0. \quad (128)$$

Por outro lado, se o objetivo é que a estrutura da penalidade se assemelhe a uma função específica, como as taxas de penalização do BIC, a formulação anterior pode ser ajustada de maneira diferente, sem seguir necessariamente a hipótese conservativa. Nesse contexto, a penalidade pode ser modelada para refletir o comportamento característico do BIC. Uma possível equação para essa penalidade, inspirada na estrutura do BIC, poderia ser expressa como:

$$c(k, n)dk + \alpha \frac{n}{k} dn = \ln(n) + \frac{k}{n}. \quad (129)$$

Onde $c(k, n)$ e α são termos que determinam a intensidade dessa penalidade.

É possível observar que a solução da Equação (127) pode ser vista como uma alternativa viável na construção das penalidades dos Critérios de Informação que são escritos na forma da Equação (12).

A solução de uma equação diferencial gera uma constante de integração que pode ser determinada a partir de condições iniciais. Partindo de uma análise em cima dos critérios conhecidos, as Equações (130), (131) e (132), que são equações diferenciais ditas exatas, representam as penalidades de AIC, AICc e BIC, respectivamente.

$$2dk + 0dn = 0, \quad (130)$$

$$\frac{2n(n-1)}{(n-k-1)^2} dk + \frac{-2k(k+1)}{(n-k-1)^2} dn = 0, \quad (131)$$

$$\ln(n)dk + \frac{k}{n} dn = 0. \quad (132)$$

As Equações (130), (131) e (132) tem como soluções gerais, respectivamente, as Equações (133), (134) e (135).

$$w(k, n) = 2k + c_1(k, n), \quad (133)$$

$$w(k, n) = \frac{2kn}{n - k - 1} + c_2(k, n), \quad (134)$$

$$w(k, n) = k \ln(n) + c_3(k, n). \quad (135)$$

Uma condição para as soluções dessas equações (para o caso dos critérios originais), dado que $c_1(k, n)$, $c_2(k, n)$ e $c_3(k, n)$ são nulas para essas penalidades, é que $w(0, n) = 0$ (há outros casos possíveis como $w(k, 0) = 0$ ou $w(k, 1)$ nas Equações (134) ou (135), respectivamente).

Pensando em termos de Equações Diferenciais, uma aplicação interessante da constante de integração é que ela possibilita o ajuste dos Critérios de Informação de forma a equilibrá-los em um ponto específico (k, n) . Esse ponto que chamaremos de ponto de equilíbrio entre Critérios de Informação atua como um parâmetro de ajuste, permitindo sintonizar e alinhar os diferentes Critérios de Informação, assegurando que, em (k, n) , eles sejam idênticos. Assim, ela pode desempenhar um papel crucial na harmonização das penalizações e na análise de modelos, ajustando os critérios para refletir de maneira mais precisa as características e os requisitos a partir de um ponto (k, n) de interesse específico.

Por exemplo, os autores Emiliano et al. (2010) sugerem que para um n grande o critério BIC se sai relativamente melhor que os critérios AIC e AICc. Podemos impor, por exemplo, que para este n grande, os critérios AIC e AICc tenham o mesmo valor de BIC neste n em $k = 1$, criando-se assim um valor comum para ambos critérios, nesse caso teria-se:

$$\begin{aligned} w(1, n) &= 2 + c_1(k, n) = \ln(n), \\ w(1, n) &= \frac{2n}{n - 2} + c_2(k, n) = \ln(n). \end{aligned}$$

E a correção para a penalidade w dos critérios AIC e AICc ficariam respectivamente:

$$\begin{aligned} w(AIC) &= 2k + \ln(n) - 2, \\ w(AICc) &= \frac{2kn}{n - k - 1} + \ln(n) - \frac{2n}{n - 2}. \end{aligned}$$

A Tabela 33 mostra como seriam as penalidades desses critérios corrigidos para os valores de $n = 50$, $n = 100$ e $n = 200$, para os primeiros 5 valores de k .

Naturalmente, não sabe-se ao certo se esta correção seria convenientemente

Tabela 33: Valores das penalidades dos Critérios de Informação $w(\text{AIC})$ e $w(\text{AICc})$ e BIC para $n = 50$, $n = 100$ e $n = 200$

	50			100			200		
k	w(AIC)	w(AICc)	BIC	w(AIC)	w(AICc)	BIC	w(AIC)	w(AICc)	BIC
1	3,9120	3,9120	3,9120	4,6052	4,6052	4,6052	5,2983	5,2983	5,2983
2	5,9120	6,0840	7,8240	6,6052	6,6881	9,2103	7,2983	7,3390	10,5966
3	7,9120	8,3504	11,7361	8,6052	8,8144	13,8155	9,2983	9,4006	15,8950
4	9,9120	10,7176	15,6481	10,6052	10,9854	18,4207	11,2983	11,4832	21,1933
5	11,9120	13,1923	19,5601	12,6052	13,2027	23,0259	13,2983	13,5874	26,4916

*Obs: em negrito os valores do ponto de equilíbrio em $k=1$ para cada tamanho.

útil para os Critérios de Informação AIC e AICc, mas se poderia possivelmente, a partir de simulações, comprovar ou não essa hipótese.

Outro fato interessante é que a Equação (132) revela que para BIC tem-se

$$\frac{dn}{dk} = \frac{-n \ln(n)}{k}. \quad (136)$$

Se tomarmos a solução da Equação diferencial (136) tem-se:

$$\frac{dn}{n \ln(n)} = -\frac{dk}{k}.$$

Utilizando os procedimentos do cálculo integral para essa última equação e uma escolha conveniente para a constante de integração p , chegamos a:

$$n = e^{\frac{p}{k}}.$$

Sendo p uma constante, que para o caso representa a penalidade do critério de informação BIC, pois aplicando $\ln$ em ambos os lados tem-se:

$$\ln(n) = \ln\left(e^{\frac{p}{k}}\right),$$

impondo que $p = k \ln(n)$ como demonstrado.

Dado o que foi exposto nessa seção, é possível construir a penalidade dos Critérios de Informação a partir da Equação (127). No entanto, é interessante que para levantar hipóteses sobre $A(n, k)$ e $B(n, k)$, pode-se levar em conta algumas características que foram observadas no Gradiente de critérios conhecidos, como foi feito na seção anterior.

A.6 Construção e análise de Penalidades Integrando Hipóteses a Equações Diferenciais

A intenção deste capítulo é apresentar uma abordagem argumentativa para a construção de uma penalidade, utilizando equações diferenciais com base em argumentos hipotéticos. Nesse contexto, as ideias que podem ser expressas como hipóteses na formulação de uma equação diferencial para descrever uma penalidade w em termos de k e n podem ser desenvolvidas ao estabelecer um conjunto de argumentos e suposições sobre o comportamento desses parâmetros em relação à penalidade.

Essa construção pode ser realizada a partir de um conjunto de argumentos e hipóteses sobre a penalidade. Para isso, utilizaremos um conjunto de parâmetros adicionais α_1 , α_2 , α_3 e α_4 , que possuem dependências com as variáveis k , n e w , que ajudariam a descrever a equação.

Nesse caso, tomemos os seguintes postulados, sendo eles:

- O conjunto de coeficientes α_1 , α_2 , α_3 e α_4 variam com os parâmetros k , n e w . Isso implicaria em uma relação entre as mudanças nos parâmetros e as mudanças na penalidade.
- O conjunto de coeficientes α_1 , α_2 , α_3 e α_4 são independentes um do outro, por exemplo, mudanças em α_1 não podem afetar de forma direta α_2 e vice-versa.
- Os coeficientes α_1 , α_2 , α_3 e α_4 variam de acordo com a complexidade do modelo representada por k , o tamanho da amostra representado por n e a própria penalidade representada por w . Por exemplo, α_1 e α_3 podem aumentar à medida que k aumenta, enquanto α_2 e α_4 diminuem. Isso refletiria uma penalidade maior para modelos mais complexos ou com conjuntos de dados menores.
- Uma hipótese adicional pode ser que os coeficientes dependem da penalidade atual w . Por exemplo, pode-se considerar que α_1 e α_2 aumentam à medida que w aumenta, indicando um aumento na penalidade à medida que o modelo se torna mais complexo ou a amostra se torne menor.

Por hipótese, podemos explorar a influência da quantidade de parâmetros sobre a penalidade do modelo. Nesse cenário podemos argumentar que a taxa de variação da penalidade w em relação à quantidade de parâmetros k , representada por $\frac{dw}{dk}$, deve ser proporcional ao produto entre k e uma combinação de termos dependentes dos parâmetros k , n e w , denotados como α_1 e α_2 . Essa formulação sugere que o aumento do número de parâmetros k , ou seja, a complexidade do modelo, resulta em um crescimento não linear da penalidade dos Critérios

de Informação, afetando a magnitude da penalidade. Essa relação pode ser representada pela Equação (137), supondo que os coeficientes α_1 e α_2 possuem dependências complexas em relação a k , n e w

$$\frac{dw}{dk} = k(\alpha_1 + \alpha_2 n). \quad (137)$$

Para o caso BIC por exemplo, tem-se:

$$\frac{dw}{dk} = \ln(n), \quad (138)$$

e para o caso $\alpha_1 = 0$ enquanto $\alpha_2 = \ln(n)/n$.

Uma outra hipótese que pode ser levantada é a respeito da influência do tamanho da amostra na penalidade. Nesse contexto, podemos argumentar que a taxa de variação da penalidade em relação ao tamanho da amostra, representada por $\frac{dw}{dn}$, deve ser proporcional ao produto entre n e uma combinação de termos dependentes dos parâmetros k , n e w , denotados como α_3 e α_4 . Isso sugere que o aumento da amostra n resulta em um crescimento não linear da penalidade dos Critérios de Informação, afetando a magnitude da penalidade. Essa relação pode ser representada pela Equação (139), supondo que os coeficientes α_3 e α_4 possuem dependências complexas em relação a k , n e w

$$\frac{dw}{dn} = n(\alpha_3 + \alpha_4 k). \quad (139)$$

Para o caso BIC por exemplo, tem-se:

$$\frac{dw}{dn} = \frac{k}{n}, \quad (140)$$

e para o caso $\alpha_3 = 0$, enquanto $\alpha_4 = 1/n$.

Uma terceira hipótese pode surgir pela combinação das hipóteses anteriores. Nesse caso, ao considerarmos o sistema conjunto das Equações (137) e (139), observamos que a variação da penalidade dos Critérios de Informação em relação a k e n não é independente. Isso sugere uma interação complexa entre o número de parâmetros e o tamanho da amostra na determinação da penalidade. Podemos interpretar que mudanças no número de parâmetros podem influenciar não apenas a taxa de variação da penalidade em relação a k , mas também em relação a n , e vice-versa. Para esse caso temos a Equação (141):

$$\frac{\partial w}{\partial k} + \frac{\partial w}{\partial n} = k(\alpha_1 + \alpha_2 n) + n(\alpha_3 + \alpha_4 k). \quad (141)$$

Se desejarmos incorporar a possibilidade de variações nas próprias taxas, podemos introduzir os coeficientes $A = A(w, k, n)$ e $B = B(w, k, n)$, exigindo que eles sejam independentes um do outro. Assim, a Equação (141) se torna mais generalizada a partir da Equação (142):

$$A \frac{\partial w}{\partial k} + B \frac{\partial w}{\partial n} = k(\alpha_1 + \alpha_2 n) + n(\alpha_3 + \alpha_4 k). \quad (142)$$

No caso do critério BIC, dado que

$$\frac{\partial w}{\partial k} + \frac{\partial w}{\partial n} = \ln(n) + \frac{k}{n},$$

pode-se dizer que $\alpha_1 = 0$, $\alpha_2 = \ln(n)/n$, $\alpha_3 = 1/n$, $\alpha_4 = 0$ $A(w, k, n) = 1$ e $B(w, k, n) = 1$ seriam os mais corretos.

Essas hipóteses estruturadas em forma de Equações Diferenciais podem fornecer uma estrutura teórica para entender como o processo de ajustes de parâmetros do modelo podem influenciar a penalidade dos Critérios de Informação, podendo vir a contribuir para uma compreensão mais profunda do processo de modelagem estatística. Naturalmente, para muitas situações, a solução dessas equações seria numérica, podendo ser necessário recorrer a métodos computacionais para sua resolução.

A.7 Análise do Laplaciano para a penalidade dos Critérios de Informação e metodologia complementar para a criação de um Critério de Informação a partir de uma Equação Diferencial Parcial

Uma outra grandeza que pode ser determinada a partir dos Critérios de Informação conhecidos é o Laplaciano, que nada mais é do que o divergente do gradiente. Segundo Anton et al. (2014), é dada pela seguinte expressão:

$$\nabla^2 w(x, y) = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}. \quad (143)$$

Se fizermos o Laplaciano para as penalidades de AIC, AICc e BIC, adaptando para as variáveis k e n , teremos, respectivamente:

$$\nabla^2 w = 0, \quad (144)$$

$$\nabla^2 w = \frac{2 + 2k - 4n}{(n - k - 1)^3}, \quad (145)$$

$$\nabla^2 w = -\frac{k}{n^2}. \quad (146)$$

As Equações (145) e (146) revelam que a segunda derivada parcial de w em relação a k é diferente de zero (a não ser em casos onde $n = (1+k)/2$ na Equação (145) e $k = 0$ na Equação (146) respectivamente). Se fizermos o gráfico desses Laplacianos fixando o valor de k se obtém a interpretação da curva relativa ao se variar n . A interpretação gráfica para as Equações (145) e (146) para $k = 1$ podem ser vistas nas figuras 160 e 161, que mostram que a curva para ambos Laplacianos é uma hipérbole (ambas com comportamentos muito semelhantes).

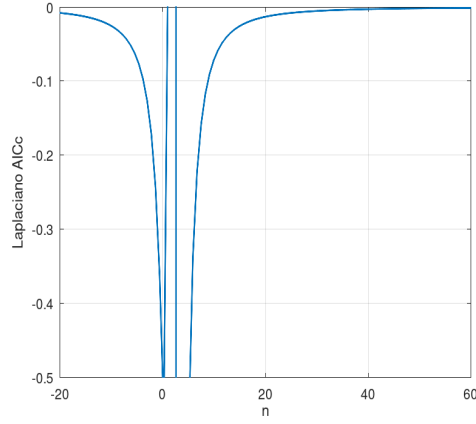


Figura 160: Comportamento da curva relativa ao Laplaciano de AICc

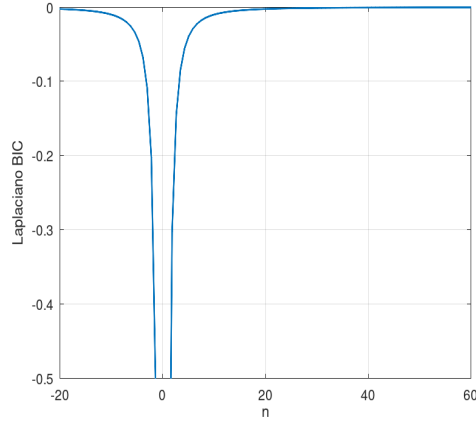


Figura 161: Comportamento da curva relativa ao Laplaciano de BIC

Essas interpretações da penalidade dos Critérios de Informação conhecidos podem revelar uma segunda alternativa para a construção de novos Critérios de Informação. Essa alternativa estaria relacionada ao Laplaciano, que envolve a derivada segunda, e aparentemente leva em conta Laplacianos que decaem conforme n aumenta. Laplacianos que poderiam ser testados seriam superfícies do tipo hiperbólicas na forma

$$\nabla^2 w = \frac{\partial^2 w}{\partial k^2} + \frac{\partial^2 w}{\partial n^2} = \frac{ak^v}{n^c}, \quad (147)$$

sendo a , v e c números reais.

Uma interpretação para essa equação é que estamos agora interessados em avaliar a taxa de variação em relação a k e n quando as taxas $(\frac{\partial w}{\partial k}$ e $\frac{\partial w}{\partial n})$ variam, já que podemos reescrever a Equação (147) como:

$$\nabla^2 w = \frac{\partial}{\partial k} \left(\frac{\partial w}{\partial k} \right) + \frac{\partial}{\partial n} \left(\frac{\partial w}{\partial n} \right) = \frac{ak^v}{n^c}. \quad (148)$$

Mais precisamente, o termo do lado direito da equação, $\frac{ak^v}{n^c}$ indica que essa taxa de variação é influenciada pela relação entre k e n . Conforme k aumenta, a taxa de variação aumenta proporcionalmente conforme uma potência v , enquanto que em relação a n , a taxa de variação diminui inversamente proporcional a n por uma potência c (desde que v e c sejam positivos), enquanto o parâmetro a controla a magnitude desse efeito.

Uma solução particular dessa última equação (sob certas condições de con-

torno) seria para $c \neq 1$, $c \neq 2$ e $v = 1$:

$$w = \frac{akn^{(2-c)}}{(2-c)(1-c)}. \quad (149)$$

Uma segunda e terceira solução para $c = 1$, $c = 2$ e com $v = 1$ nos dão respectivamente :

$$w = ak(\ln(n) - n), \quad (150)$$

$$w = ak\ln(n). \quad (151)$$

Na equação (151) pode-se observar que a penalidade atribuída ao BIC é um caso particular dessa equação quando $a = 1$.

Formas de extensão do Critério de Informação Bayesiano já foram propostas como em Chen and Chen (2008), que propõem a forma de Critério de Informação, conforme Equação (152) a seguir:

$$bic_{\gamma_1}(s) = 2\ln L(\theta(s)) + \nu(s)\ln(n) + 2\gamma_1\ln(\tau(Sj)). \quad (152)$$

No entanto, o caso da Equação (151) recai ao caso da Equação (152) quando considera-se $\gamma_1 = 0$ e $\nu(s)$ como sendo a . O mesmo autor em seu artigo dispõe de alguns resultados simulados para alguns valores de γ_1 e algumas situações.

Situações mais generalizadas para a Equação (147) proposta envolveriam Equações Diferenciais Parciais de segunda ordem, onde algumas dessas soluções poderiam ser realizadas por algum método de solução conhecido, sendo que para algumas somente haveriam soluções numéricas. A validação para algum a (Equação (151)) que se adapte a algum modelo pode ser testada a partir de simulações.

Um ponto de destaque reside no fato de termos derivado duas equações distintas que podem definir a penalidade dos Critérios de Informação por meio de Equações Diferenciais. Uma delas é a Equação (127), uma Equação Diferencial Ordinária de Primeira Ordem, e a outra é a Equação (143), uma Equação Diferencial Parcial de Segunda Ordem. Como são Equações Diferenciais, e as mesmas representam taxas de variações em termos das variáveis k e n , essas equações generalizam as superfícies que representam a penalidade dos Critérios de Informação (onde pode-se incluir as Superfícies Quádricas). Além disso, essas equações nos dão a representação analítica do movimento relacionado a modelos candidatos (seja com movimentos somente na quantidade de parâmetros, como o movimento no tamanho das amostras). Fortalecendo assim o estudo

relacionado a penalidade e a sua integração com superfícies em geral.

Para uma análise complementar da Equação (149), e uma possível metodologia para encontrar a penalidade, tomemos o seu gradiente, assim:

$$\nabla w = a \left(\frac{n^{2-c}}{(2-c)(1-c)} \hat{i} + k \frac{n^{1-c}}{1-c} \hat{j} \right). \quad (153)$$

O comportamento, por exemplo, para $a = 1.8$, $c=0$ e $c=1.1$ desse gradiente podem ser vistos nas figuras 162 e 163.

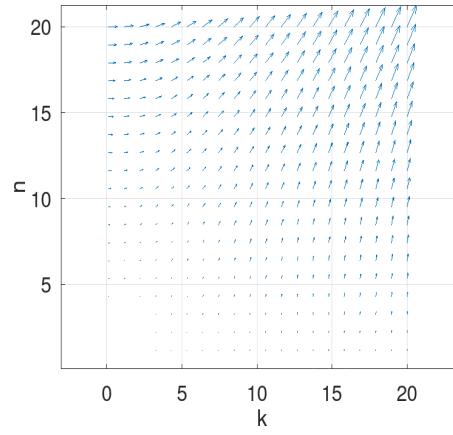


Figura 162: Comportamento do campo Gradiente para $c = 0$ e $a = 1.8$

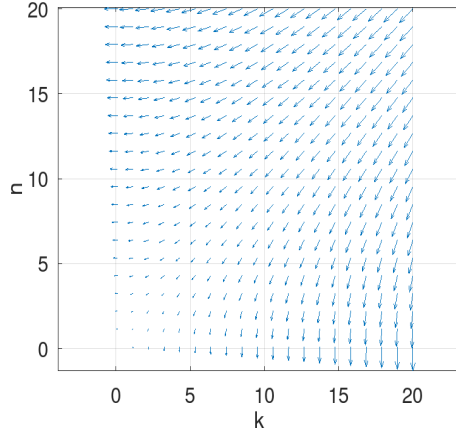


Figura 163: Comportamento do campo Gradiente para $c = 1.1$ e $a = 1.8$

O que se observaria se modificássemos o parâmetro c , mantendo a fixo, é que os valores de c determinariam o quanto a penalidade w enverga-se para cima, ou para baixo a medida que se aumenta ou diminui esse valor em relação a k e n . Já o valor de a vai determinar o valor da direção, caso seja positivo ou negativo, do módulo do gradiente e da derivada direcional. Essa estrutura de comportamento para o gradiente pode ser utilizada para decidir sobre o comportamento da penalidade em relação a sua complexidade. Se quisermos um comportamento conforme o de BIC, que penaliza modelos com muitos parâmetros de forma mais intensa, basta escolhermos valores de c de tal maneira que o campo se envergue para cima a medida que k aumenta (conforme observado na Figura 158). Contudo, se quisermos que o critério penalize modelos com características semelhantes ao do AICc, parece razoável que o gradiente tenha sua envergadura para baixo (conforme observado na Figura 156).

O valor do $\cos(\psi)$ para a Equação (153) pode ser utilizado para se obter a medida auxiliar da complexidade que se dá a partir da expressão:

$$\cos(\psi) = \frac{\sqrt{2} \left(\frac{n^{2-c}}{(2-c)(1-c)} + \frac{kn^{1-c}}{1-c} \right)}{2 \sqrt{\left(\frac{n^{2(2-c)}}{((2-c)(1-c))^2} + \frac{k^2 n^{(1-c)^2}}{(1-c)^2} \right)}}. \quad (154)$$

Utilizando as expressões da Equação (153), combinados com os comportamentos gráficos, e a expressão dada pela Equação (154), pode-se decidir sobre o valor de a e c . Após isso, usa-se então a Equação (149) para montar a pena-

lidade.

A.8 Análise da curvatura Gaussiana para a penalidade dos critérios AIC AICc e BIC e métrica para BIC(a)

Como a penalidade dos Critérios de Informação representa uma superfície, é interessante determinarmos o que ocorre com esta superfície em um determinado ponto (k, n) .

Uma das grandezas importantes de serem estudadas para esses casos e interessante de serem determinadas é a sua curvatura, que descreve como uma superfície se curva em relação a uma referência plana, sendo uma delas a curvatura Gaussiana. Essa curvatura, segundo o teorema Egregium de Gauss(de la Rica and de la Villa Cuenca (1997)), afirma que em uma superfície essa grandeza não é afetada por deformações que não envolvam cortes ou colagens. Porém, essa grandeza pode ser de grande valia ao se estudar as penalidades de alguns dos critérios estudados nesse trabalho.

Segundo Do Carmo (2010), é possível através das funções de curvatura encontrar possíveis características relacionadas a essas superfícies, o que também facilita a compreensão das características associadas aos Critérios de Informação. Em relação a curvatura Gaussiana, caso a curvatura seja zero, teríamos uma característica plana, caso ela seja positiva, um padrão elíptico, e caso negativa, um padrão hipérbolico.

Para determinarmos a curvatura Gaussiana para os critérios AIC, AICc e BIC, necessitamos escrever a forma vetorial de sua superfície a partir da seguinte equação:

$$\vec{R} = \vec{R}_0 + \vec{C}, \quad (155)$$

onde $\vec{R}_0 = (0, 0, -2\ln(L(\theta)))$ e $\vec{C} = (k, n, w(k, n))$

Sem perda de generalidade, já que queremos somente o comportamento da penalidade, tomaremos a forma mais simples, impondo $\vec{R}_0 = \vec{0}$, tendo assim para a superfície dessas penalidades as Equações (156), (157) e (158) para os critérios AIC AICc e BIC, respectivamente, como sendo:

$$\vec{R} = (k, n, 2k), \quad (156)$$

$$\vec{R} = (k, n, \frac{2kn}{n-k-1}), \quad (157)$$

$$\vec{R} = (k, n, k \ln(n)). \quad (158)$$

O cálculo da curvatura Gaussiana κ_G , segundo de la Rica and de la Villa Cuenca (1997), é dada pela equação

$$\kappa_G = \frac{h_{11}h_{22} - h_{12}^2}{g_{11}g_{22} - g_{12}^2}. \quad (159)$$

Os termos g_{ij} e h_{ij} (com i e $j = \{1, 2\}$) para ambas as equações são determinadas a partir das seguintes equações, ou seja:

$$g_{11} = \frac{\partial \vec{R}}{\partial k} * \frac{\partial \vec{R}}{\partial k}, \quad (160)$$

$$g_{12} = g_{21} = \frac{\partial \vec{R}}{\partial k} * \frac{\partial \vec{R}}{\partial n}, \quad (161)$$

$$g_{22} = \frac{\partial \vec{R}}{\partial n} * \frac{\partial \vec{R}}{\partial n}, \quad (162)$$

$$h_{11} = \frac{\partial^2 \vec{R}}{\partial k^2} * \vec{\nu}, \quad (163)$$

$$h_{12} = h_{21} = \frac{\partial^2 \vec{R}}{\partial k \partial n} * \vec{\nu}, \quad (164)$$

$$h_{22} = \frac{\partial^2 \vec{R}}{\partial n^2} * \vec{\nu}, \quad (165)$$

onde ν é o vetor dado pela equação

$$\vec{\nu} = \frac{\frac{\partial \vec{R}}{\partial k} \times \frac{\partial \vec{R}}{\partial n}}{|\frac{\partial \vec{R}}{\partial k} \times \frac{\partial \vec{R}}{\partial n}|}. \quad (166)$$

Calculando essa curvatura para a penalidade do critério AIC tem-se:

$$k_G^{AIC} = 0, \quad (167)$$

o que faz sentido pois um plano não possui curvatura (de la Rica and de la Villa Cuenca (1997)). Para a superfície da penalidade de AICc tem-se:

$$\kappa_G^{AICc} = \frac{s^6 16nk(n-1)(k+1) - 4s(k(2n-1) + n-1)^2}{((4n^2(n-1)^2 + s^4)(4k^2(k+1)^2 + s^2) - 16k^2n^2(n-1)^2(k+1)^2)u}, \quad (168)$$

onde:

$$s = n - k - 1 \text{ e } u = 2n(n-1)^2 + (2k(k+1))^2 + s^2.$$

Para a superfície da penalidade de BIC tem-se:

$$k_G^{BIC} = \frac{-1/n}{(\ln^2(n) + k^2/n^2 + 1)^2}. \quad (169)$$

Para qualquer ponto (k, n) o valor da curvatura de BIC é negativa, já que n deve ser positivo devido ao $\ln(n)$. Segundo Do Carmo (2010), pontos onde a curvatura Gaussiana é negativa é dita Hiperbólica.

Se calcularmos $k_G^{BIC(a)}$ para a expressão da Equação (151) generalizada do BIC a partir de a , tem-se:

$$k_G^{BIC(a)} = \frac{-1/n}{a^3 (\ln^2(n) + 2k^2/n^2 + 1)^2}. \quad (170)$$

Que recai ao caso particular de k_G^{BIC} quando $a = 1$ e naturalmente invalidando $a = 0$ (pelo menos no que diz respeito a geometria, não se discute curvatura de função nula).

No caso da curvatura dada pela Equação (170), a curvatura fica definida a partir do valor a , permitindo que a curvatura seja positiva quando $a < 0$.

Outra grandeza importante que pode ser explorada, aproveitando o conceito de curvatura é a forma quadrática, que diz respeito a forma na qual é determinada a distância entre dois pontos infinitesimalmente próximos da superfície. De forma geral, essa expressão é dada para o caso de duas variáveis por (de la Rica and de la Villa Cuenca, 1997):

$$ds^2 = \begin{pmatrix} dk & dn \end{pmatrix} \begin{bmatrix} g_{ij} \end{bmatrix} \begin{pmatrix} dk \\ dn \end{pmatrix}. \quad (171)$$

Sendo sua métrica dada por:

$$\begin{bmatrix} g_{ij} \end{bmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}. \quad (172)$$

Uma característica interessante da Equação (171) é sua invariância (não

mudar o valor independente da mudança de parâmetros) diante de alterações do tipo $k = k(\alpha, \beta)$ e $n = n(\alpha, \beta)$, conforme discutido por de la Rica and de la Villa Cuenca (1997).

Explorando essa Equação para a forma dada pela Equação (151), tem-se para a métrica:

$$[g_{ij}] = \begin{pmatrix} 1 + a^2 \ln^2(n) & \frac{a^2 k \ln(n)}{n} \\ \frac{a^2 k \ln(n)}{n} & 1 + \frac{a^2 k^2}{n^2} \end{pmatrix}, \quad (173)$$

sendo a distância infinitesimal dada por:

$$ds^2 = (1 + a^2 \ln^2(n)) dk^2 + 2 \left(\frac{a^2 k \ln(n)}{n} \right) dk dn + \left(1 + \frac{a^2 k^2}{n} \right) dn^2. \quad (174)$$

A Equação (174) possui dependências com a, k e n . Essa mesma equação possui a presença do termo g_{12} , indicando naturalmente a presença do termo cruzado $dkdn$, indicando correlação entre as duas variáveis k e n . Essa mesma equação pode ser utilizada para se determinar o ângulo entre duas curvas parametrizadas, cálculo de áreas, etc. Para o caso que estamos tratando, essa equação basicamente nos diz como devemos determinar a distância geométrica na superfície de $ak \ln(n)$ quando mudamos de um determinado ponto (k, n) para um outro ponto vizinho, ou mais precisamente, quando mudamos de um modelo para outro modificando o quantidade de parâmetros e o tamanho da amostra. Sendo que essa mudança ocorre a partir dos valores encontrados através da métrica dada pela Equação (173).

É interessante observar que quando $a = 0$ a Equação (174) recai a distância euclidiana convencional (Pitagórica), como esperado.

$$ds^2 = dk^2 + dn^2.$$

O vetor unitário ortogonal a superfície pode ser expresso por (de la Rica and de la Villa Cuenca, 1997):

$$\hat{\nu} = \frac{\frac{\partial \vec{R}}{\partial k} \times \frac{\partial \vec{R}}{\partial n}}{\sqrt{g_{11}g_{22} - (g_{12})^2}}. \quad (175)$$

Percebamos que para que $\hat{\nu}$ exista o valor $g_{11}g_{22} - (g_{12})^2$ deve ser positivo ($\det(g_{ij}) > 0$), assim para o caso BIC(a) tem-se:

$$1 + \frac{a^2 k^2}{n} + a^2 \ln^2(n) > 0.$$

Os valores de a são dependentes de k e n e não há restrição explícita para k , mas para n o mesmo deve ser positivo (se admitir-se que não se está interessado em valores fora do conjunto dos números reais).

Tomemos o caso mais simples sendo $k = 1$ e $n = 1$, então:

$$a^2 + 1 > 0,$$

o que fortifica a validade de a para todos números reais diferentes de zero, como esperado.

É interessante fazermos algumas observações a partir da Equação (173) que podem ser relevantes para o critério BIC(a), assim como para outros Critérios de Informação que possam se beneficiar deste tipo de análise (para o BIC, por exemplo, basta colocar $a = 1$). Em relação a Equação (173), o termo g_{11} representa como a distância ou a curvatura na direção do número de parâmetros k é afetada conforme aumenta-se k por dk . Essa situação ocorre, em termos de modelagem, quando se mantém n fixo e se varia somente a quantidade de parâmetros (caso de modelos candidatos). Já o termo $g_{12} = g_{21}$ representa como a distância ou curvatura na direção do número de parâmetros k e o tamanho da amostra n (ou vice e versa) são afetadas conforme aumenta-se k por dk e n por dn . Essa condição, em termos de modelagem, ocorre quando se faz diferentes modelos candidatos para diferentes tamanhos de amostras (por exemplo modelos candidatos para ajuste e para previsão). O termo g_{22} descreve como a curvatura ou a distância na direção do tamanho das amostras n são afetadas conforme aumenta-se n por dn , em termos de modelagem, essa situação ocorre quando se utiliza um mesmo modelo candidato em amostras de tamanhos diferentes.

Para o caso observado no critério BIC(a), o valor do termo $g_{11} = 1 + a^2 \ln^2(n)$ sugere que a curvatura, ou a métrica aumentam com o quadrado do logaritmo do tamanho da amostra n . Isso implica que o critério se torna mais sensível a variações no número de parâmetros k , representadas por dk conforme n aumenta (ou seja especialmente em modelos com amostras maiores). À medida que o tamanho da amostra aumenta, a penalidade no Critério de Informação, em decorrência da variação dk , também cresce de forma significativa. Essa dinâmica mostra que a penalidade se eleva mais rapidamente com alterações em k (dadas pelo termo dk) à medida que n aumenta, indicando que modelos com maior número de parâmetros enfrentam penalizações mais severas nesse caso.

O valor de $g_{12} = \frac{a^2 k \ln(n)}{n}$ sugere que a curvatura ou a métrica na direção da

taxa $dkdn$ se eleva rapidamente com o valor de k , no entanto, essa mesma taxa de aumento diminui proporcionalmente com o crescimento de n . Assim, g_{12} evidencia como o número de parâmetros k e o tamanho da amostra n interagem na penalização do Critério de Informação tanto na direção dk como na direção dn (devido a $dkdn$).

A expressão $g_{22} = \frac{a^2 k^2}{n^2}$ revela que a curvatura ou a métrica, à medida que o número de parâmetros k aumenta, diminuem com o aumento de n . Isso significa que o impacto do aumento de n no critério de informação se torna menos sensível em modelos mais complexos, ou seja, aqueles com maior k . Em outras palavras, a variação no tamanho da amostra exerce uma influência reduzida na penalidade aplicada a modelos mais complexos na variação de n .

Apesar de se ter analisado especificamente a equação para o BIC(a), é possível extrair considerações relevantes sobre os Critérios de Informação e sua penalidade a partir dos conceitos vistos, sendo eles:

- A métrica g_{ij} : Esta métrica ajuda a entender como a distância e a curvatura são afetadas conforme se aumentam k ou n , nas direções de k e n , através das variações dk e dn . Oferecendo uma visão sobre como os Critérios de Informação penalizam os modelos de acordo com a complexidade (número de parâmetros) e o tamanho da amostra.

- Curvatura Gaussiana: Reflete como a penalidade nas variáveis k (número de parâmetros) e n (tamanho da amostra) se comportam em torno de um ponto (k, n) , refletindo a forma local da penalidade. Ela revela se a penalização é mais acentuada ou suave, influenciando a maneira como diferentes valores de k e n afetam a penalização imposta pelo critério de informação.

Especificamente para a Curvatura Gaussiana tem-se as seguintes condições:

- Positiva: Indica que a penalização é mais rígida em todas as direções (k ou n), como uma superfície convexa (elíptica). Isso sugere que a penalidade aumenta de maneira mais pronunciada à medida que se varia de um modelo candidato para outro.

- Negativa: Indica uma superfície em forma de sela (hiperbólica), onde a penalização é suave em uma direção e acentuada em outra (em termos de k e n). Isso significa que a penalização pode ser menos severa em relação a k e mais severa em relação a n , ou mais severa em relação a n e menos severa em relação a k à medida que se varia de um modelo candidato para outro.

- zero: Indica que a penalização é plana em uma das direções (k ou n), sugerindo que variações em k ou n nessa direção não resultam em uma variação da penalização ao se variar de um modelo candidato para outro (naturalmente, a

penalidade em um ponto (k,n) pode ser maior ou menor dependendo da condição plana da penalidade).

É interessante observar que tanto a métrica g_{ij} quanto a curvatura se complementam em termos de análise. Enquanto a métrica fornece uma compreensão detalhada das distâncias e direções de variação em um espaço de parâmetros (para o caso k e n), a curvatura oferece perspectivas sobre como essas variações impactam a penalização do Critério de Informação.

Portanto, ao estudar a métrica, podemos extrair conclusões sobre como os Critérios de Informação penalizam modelos com base na complexidade e no tamanho da amostra, podendo ser uma ferramenta fundamental para a análise da penalidade e do critério como um todo.

A.9 Equações para os Critérios de Informação Baseados em Superfícies Quadráticas (QSIC)

- Critérios de Informação Baseado em Superfícies Quadráticas

$$QSIC = -2\ln(L) + \left(\frac{-Ak^2 - Bn^2 - 2Dkn - Mk - Nn - Q}{2Ek + 2Fn + P} \right). \quad (176)$$

- Critérios de Informação Baseados em γ

$$C_n^{\gamma, E, F} = -2\ln(L) + \frac{2F\gamma n^2 k}{e^2(Ek + Fn)}. \quad (177)$$

- Critérios de Informação baseados em γ com semelhança em BIC

$$C_n^{E, F} = -2\ln(L) + \frac{Fkn.\ln(n)}{Ek + Fn}. \quad (178)$$

- Critérios de Informação T1 a T6

$$T1 = -2\ln(L) + \left(\frac{-kn.\ln(n)}{-k - n} \right). \quad (179)$$

$$T2 = -2\ln(L) + \left(\frac{-0.5kn.\ln(n)}{-k - 0.5n} \right). \quad (180)$$

$$T3 = -2\ln(L) + \left(\frac{0.5kn.\ln(n)}{-k + 0.5n} \right). \quad (181)$$

$$T4 = -2\ln(L) + \left(\frac{kn.\ln(n)}{-k + n} \right). \quad (182)$$

$$T5 = -2\ln(L) + \left(\frac{-kn.\ln(n)}{-0.5k - n} \right). \quad (183)$$

$$T6 = -2\ln(L) + \left(\frac{kn.\ln(n)}{-0.5k + n} \right). \quad (184)$$