
Um procedimento para seleção de variáveis em
modelos lineares generalizados duplos

Lucas Leite Cavalaro

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

LUCAS LEITE CAVALARO

**UM PROCEDIMENTO PARA SELEÇÃO DE VARIÁVEIS EM MODELOS LINEARES
GENERALIZADOS DUPLOS**

Dissertação apresentada ao Departamento de Estatística – Des/UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Estatística - Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.

Orientador: Prof. Dr. Gustavo Henrique de Araujo Pereira

**São Carlos
Maio de 2019**

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

LUCAS LEITE CAVALARO

A PROCEDURE FOR VARIABLE SELECTION IN DOUBLE GENERALIZED LINEAR MODELS

Master dissertation submitted to the Department of Statistics – Des/UFSCar and Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Interagency Program Graduate in Statistics UFSCar-USP.

Advisor: Prof. Dr. Gustavo Henrique de Araujo Pereira

**São Carlos
May 2019**

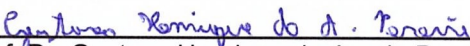


UNIVERSIDADE FEDERAL DE SÃO CARLOS

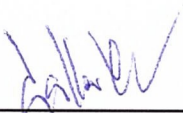
Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Lucas Leite Cavalaro, realizada em 01/04/2019:



Prof. Dr. Gustavo Henrique de Araujo Pereira
UFSCar



Prof. Dr. Diego Ignacio Gallardo Mateluna
UDA



Profa. Dra. Mônica Carneiro Sandoval
IME-USP

“É bom se divertir com os pequenos imprevistos. Ao máximo. Porque aí você encontrará o que é mais importante do que o que quer.”

Ging Freecss (Hunter x Hunter)

À João Reginaldo Cavalaro e Sandra
Aparecida Leite Cavalaro.

Agradecimentos

Agradeço primeiramente a Deus por ter me dado coragem, saúde, força e proteção nesta jornada de 2 anos.

Aos meus pais, João Reginaldo Cavalaro e Sandra Aparecida Leite Cavalaro, pelo amor, carinho, incentivo, apoio e compreensão nos momentos difíceis. Agradeço à minha mãe por sempre me incentivar a nunca desistir dos estudo e me aconselhar quando achei que não conseguiria completar este Mestrado em Estatística.

À minha irmã, Natália Leite Cavalaro, pelo apoio e pelas risadas ao longo desses anos.

Ao meu peixe de estimação, Bebezinho, por me fazer sempre companhia e me alegrar em momentos difíceis, e à calopsita Pituquinha, que adora aprontar e animar todos com sua cantoria.

Ao meu orientador, Prof. Dr. Gustavo Henrique de Araujo Pereira, pela excelente orientação, sem nunca restringir a minha criatividade, nesta Dissertação de Mestrado, no Trabalho de Graduação e nos dois projetos de Iniciação Científica. Muito obrigado Prof. Gustavo por todas as conversas, conselhos, risadas, ajudas, confiança e apoio. Acredito que, ao longo desses anos, além de um excelente orientador também encontrei um amigo.

À Prof^ª. Dr^ª. Alessandra Aparecida Verri (Departamento de Matemática - UFSCar) pela orientação na minha primeira Iniciação Científica e por me incentivar a estudar e aprender conteúdos que nós dois nunca havíamos estudado. Fazer este trabalho me influenciou a buscar outro projeto, mas desta vez no Departamento de Estatística.

A todos os professores que contribuíram na minha formação acadêmica. Em especial, agradeço ao Prof. Dr. Rafael Izbicki, ao Prof. Dr. Osvaldo Anacleto Junior, ao Prof. Dr. Diego Ignacio Gallardo Mateluna e à Prof^ª. Dr^ª. Mônica Carneiro Sandoval pelos ótimos questionamentos, sugestões e correções neste trabalho.

Aos amigos Paula Ianishi, André Arantes, Pedro Munari, Hêndria Moura, Victor Fernandes, Mateus Cabral, Victor Oliveira, Alana Salantim, Wellington Puerta, Guilherme Maia e João Carlos pelas conversas, risadas, conselhos e ajuda ao longo desses 2 anos.

À todos que foram alunos de pós-graduação comigo e tive uma ótima convivência.

À Universidade Federal de São Carlos, à Universidade de São Paulo e ao Programa Interinstitucional de Pós-Graduação em Estatística pela oportunidade de fazer o Mestrado em Estatística.

Aos programas PICMe (Programa de Iniciação Científica e Mestrado) e PIBIC (Programa Institucional de Bolsas de Iniciação Científica) pelos projetos de iniciação científica.

Ao CNPq pelo apoio financeiro no mestrado e nos três projetos de iniciação científica. Pesquisa desenvolvida com o auxílio dos recursos de HPC disponibilizados pela Superintendencia de Tecnologia da Informacao da Universidade de São Paulo.

Muito obrigado a todos que me ajudaram direta ou indiretamente nesta fase da minha vida.

Resumo

Cavalaro, L. L. **Um procedimento para seleção de variáveis em modelos lineares generalizados duplos.**

Os modelos lineares generalizados duplos (MLGD), diferentemente dos modelos lineares generalizados (MLG), permitem o ajuste do parâmetro de dispersão da variável resposta em função de variáveis preditoras, aperfeiçoando a forma de modelar fenômenos. Desse modo, os mesmos são uma possível solução quando a suposição de que o parâmetro de dispersão constante não é razoável e a variável resposta tem distribuição que pertence à família exponencial. Considerando nosso interesse em seleção de variáveis nesta classe de modelos, estudamos o esquema de seleção de variáveis em dois passos proposto por Bayer e Cribari-Neto (2015) e, com base neste método, desenvolvemos um esquema para seleção de variáveis em até “ k ” passos. Para verificar a performance do nosso procedimento, realizamos estudos de simulação de Monte Carlo em MLGD. Os resultados obtidos indicam que o nosso procedimento para seleção de variáveis apresenta, em geral, performance semelhante ou superior à das demais metodologias estudadas sem necessitar de um grande custo computacional. Também avaliamos o esquema para seleção de variáveis em até “ k ” passos em um conjunto de dados reais e o comparamos com diferentes métodos de regressão. Os resultados mostraram que o nosso procedimento pode ser também uma boa alternativa quando possui-se interesse em realizar previsões.

Palavras-chave: critérios de informação, modelos lineares generalizados duplos, seleção de variáveis, *Stepwise*.

Abstract

Cavalaro, L. L. **A procedure for variable selection in double generalized linear models.**

The double generalized linear models (DGLM), unlike the generalized linear model (GLM), allow the fit of the dispersion parameter of the response variable as a function of predictor variables, improving the way of modeling phenomena. Thus, they are a possible solution when the assumption that the constant dispersion parameter is unreasonable and the response variable has distribution belonging to the exponential family. Considering our interest in variable selection in this class of models, we studied the two-step variable selection scheme proposed by Bayer and Cribari-Neto (2015) and, based on this method, we developed a scheme to select variables in up to “ k ” steps. To check the performance of our procedure, we performed Monte Carlo simulation studies in DGLM. The results indicate that our procedure for variable selection presents, in general, similar or superior performance than the other studied methods without requiring a large computational cost. We also evaluated the scheme to select variables in up to “ k ” steps in a set of real data and compared it with different regression methods. The results showed that our procedure can also be a good alternative when the interest is in making predictions.

Keywords: double generalized linear models, information criteria, Stepwise, variable selection.

Sumário

Lista de Figuras	ix
Lista de Tabelas	x
1 Introdução	1
2 Modelos lineares generalizados duplos	3
2.1 Família exponencial e distribuições contínuas	3
2.2 Especificação do modelo	5
2.3 Interpretação dos parâmetros do modelo	6
2.3.1 Submodelo da média	6
2.3.2 Submodelo do parâmetro de dispersão	7
2.4 Inferência sobre os parâmetros	7
2.5 Análise de diagnóstico	8
3 Procedimentos para seleção de variáveis	9
3.1 Critérios de seleção de variáveis	10
3.1.1 Critério de Informação de Akaike (AIC)	11
3.1.2 Critério de Informação Bayesiano (BIC)	12
3.1.3 Critério de Informação de Hannan-Quinn (HQC)	13
3.1.4 Critério de Informação de Akaike corrigido (AICc)	13
3.2 Esquema de seleção de variáveis em dois passos	14
3.3 Esquema de seleção de variáveis em até “ <i>k</i> ” passos	14
3.4 Método <i>Stepwise</i>	17
4 Estudos de simulação	18
4.1 Resultados obtidos para a distribuição Gama	19
4.2 Resultados obtidos para a distribuição Gaussiana Inversa	25
4.3 Resultados obtidos para a distribuição Normal	27
4.4 Análise do número de passos para o esquema de seleção de variáveis em até “ <i>k</i> ” passos nos estudos de simulação	29
4.5 Comparação do número de modelos ajustados \times tempo entre o esquema de seleção de variáveis em até “ <i>k</i> ” passos e o método <i>Stepwise</i>	30

5	Aplicação	32
5.1	Descrição dos dados	32
5.2	Avaliação dos procedimentos para seleção de variáveis	34
5.3	Variáveis selecionadas no estudo de simulação da aplicação	37
5.4	Comparação do esquema de seleção de variáveis em até “ <i>k</i> ” passos com outros métodos de regressão	38
5.4.1	Regressão Ridge e Lasso	39
5.4.2	<i>Boosting</i> “tradicional”	40
5.4.3	<i>Boosting</i> para MLGD	40
5.4.4	Comparações entre as metodologias	41
6	Conclusão	44
6.1	Trabalhos futuros	45
A	Resultados completos dos estudos de simulação - Distribuição Gama	47
B	Resultados completos dos estudos de simulação - Distribuição Gaussiana Inversa	59
C	Resultados completos dos estudos de simulação - Distribuição Normal	70
	Referências Bibliográficas	82

Lista de Figuras

4.1	Gráfico de barras para o número de passos obtidos no estudo de simulação do cenário 5 para a distribuição Normal, com tamanho amostral $n = 50$ e o AIC como critério de informação.	29
5.1	Histograma dos ganhos totais dos golfistas (US\$).	34
5.2	<i>Box plot</i> dos riscos quadráticos para o melhor critério de informação em cada método.	36
5.3	<i>Box plot</i> dos riscos quadráticos para os métodos de regressão em estudo. . .	42

Lista de Tabelas

3.1	Cálculo do termo penalizador do AIC considerando as variáveis x_1 e x_3 no submodelo da média e $\mathbf{x} = (x_1, x_2, x_3)$ como possíveis covariáveis.	16
4.1	Coefficientes de correlação linear de Pearson utilizamos para gerar as variáveis preditoras nos Cenários 9.	20
4.2	Cenários - distribuição Gama.	20
4.3	Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Gama considerando o critério de informação com melhor desempenho.	24
4.4	Cenários - distribuição Gaussiana Inversa.	25
4.5	Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Gaussiana Inversa considerando o critério de informação com melhor desempenho.	26
4.6	Cenários - distribuição Normal.	27
4.7	Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Normal considerando o critério de informação com melhor desempenho.	28
4.8	Tempo, em minutos (número de passos - total de modelos ajustados), para encontrar um modelo candidato utilizando dados simulados com a distribuição Gama, o esquema de seleção de variáveis em até “k” passos e diferentes combinações de tamanhos amostrais (n) e números de parâmetros (p).	30
4.9	Tempo, em minutos (total de modelos ajustados), para obter um modelo candidato utilizando dados simulados com a distribuição Gama, o método <i>Stepwise</i> e diferentes combinações de tamanhos amostrais (n) e números de parâmetros (p).	31
5.1	Medidas descritivas da variável resposta - ganhos totais dos golfistas (US\$).	33

5.2	Coeficientes de correlação linear de Pearson calculados entre todas as variáveis em estudo.	34
5.3	Média e desvio padrão dos riscos quadráticos para cada método e critério de informação.	35
5.4	Mediana, média e desvio padrão dos riscos quadráticos para o melhor critério de informação em cada método.	36
5.5	Intervalos de confiança para a diferença das médias do risco quadrático (linha menos coluna) para os diferentes procedimentos de seleção de variáveis.	37
5.6	Contagem das variáveis selecionadas pelo esquema de seleção de variáveis em até “ k ” passos no estudo de simulação da aplicação para os dois submodelos ajustados.	37
5.7	Estimativas dos parâmetros do MLGD selecionado como modelo candidato a final.	38
5.8	Mediana, média e desvio padrão dos riscos quadráticos para os métodos de regressão em estudo.	42
5.9	Intervalos de confiança para a diferença das médias do risco quadrático (linha menos coluna) para os métodos de regressão em estudo.	42
A.1	Resultados completos dos estudos de simulação para o cenário 1 da distribuição Gama.	48
A.2	Resultados completos dos estudos de simulação para o cenário 2 da distribuição Gama.	49
A.3	Resultados completos dos estudos de simulação para o cenário 3 da distribuição Gama.	50
A.4	Resultados completos dos estudos de simulação para o cenário 4 da distribuição Gama.	51
A.5	Resultados completos dos estudos de simulação para o cenário 5 da distribuição Gama.	52
A.6	Resultados completos dos estudos de simulação para o cenário 6 da distribuição Gama.	53
A.7	Resultados completos dos estudos de simulação para o cenário 7 da distribuição Gama.	54
A.8	Resultados completos dos estudos de simulação para o cenário 8 da distribuição Gama.	55
A.9	Resultados completos dos estudos de simulação para o cenário 9 da distribuição Gama.	56
A.10	Resultados completos dos estudos de simulação para o cenário 10 da distribuição Gama.	57
A.11	Resultados completos dos estudos de simulação para o cenário 11 da distribuição Gama.	58

B.1	Resultados completos dos estudos de simulação para o cenário 1 da distribuição Gaussiana Inversa.	60
B.2	Resultados completos dos estudos de simulação para o cenário 2 da distribuição Gaussiana Inversa.	61
B.3	Resultados completos dos estudos de simulação para o cenário 3 da distribuição Gaussiana Inversa.	62
B.4	Resultados completos dos estudos de simulação para o cenário 4 da distribuição Gaussiana Inversa.	63
B.5	Resultados completos dos estudos de simulação para o cenário 5 da distribuição Gaussiana Inversa.	64
B.6	Resultados completos dos estudos de simulação para o cenário 6 da distribuição Gaussiana Inversa.	65
B.7	Resultados completos dos estudos de simulação para o cenário 7 da distribuição Gaussiana Inversa.	66
B.8	Resultados completos dos estudos de simulação para o cenário 8 da distribuição Gaussiana Inversa.	67
B.9	Resultados completos dos estudos de simulação para o cenário 9 da distribuição Gaussiana Inversa.	68
B.10	Resultados completos dos estudos de simulação para o cenário 10 da distribuição Gaussiana Inversa.	69
C.1	Resultados completos dos estudos de simulação para o cenário 1 da distribuição Normal.	71
C.2	Resultados completos dos estudos de simulação para o cenário 2 da distribuição Normal.	72
C.3	Resultados completos dos estudos de simulação para o cenário 3 da distribuição Normal.	73
C.4	Resultados completos dos estudos de simulação para o cenário 4 da distribuição Normal.	74
C.5	Resultados completos dos estudos de simulação para o cenário 5 da distribuição Normal.	75
C.6	Resultados completos dos estudos de simulação para o cenário 6 da distribuição Normal.	76
C.7	Resultados completos dos estudos de simulação para o cenário 7 da distribuição Normal.	77
C.8	Resultados completos dos estudos de simulação para o cenário 8 da distribuição Normal.	78
C.9	Resultados completos dos estudos de simulação para o cenário 9 da distribuição Normal.	79

C.10 Resultados completos dos estudos de simulação para o cenário 10 da distribuição Normal.	80
C.11 Resultados completos dos estudos de simulação para o cenário 11 da distribuição Normal.	81

Capítulo 1

Introdução

O modelo de regressão linear usual supõe que o erro aleatório (parte não explicada do modelo) é normalmente distribuído, com média zero e variância constante. Consequentemente, a distribuição da variável resposta, dadas as variáveis preditoras, possui distribuição normal com variância constante.

Quando as suposições de normalidade e/ou homocedasticidade não são razoáveis, a alternativa mais utilizada até o início da década de 1970 era a realização de transformações na variável resposta como, por exemplo, a proposta por [Box e Cox \(1964\)](#). Entretanto, com o desenvolvimento da teoria a respeito dos modelos lineares generalizados (MLG), por [Nelder e Wedderburn \(1972\)](#), permitiu-se a modelagem da variável resposta assumindo outras distribuições pertencentes à família exponencial ([Bolfarine e Sandoval, 2001](#)).

No contexto dos modelos lineares generalizados, [Smyth \(1989\)](#) introduziu a classe dos modelos lineares generalizados duplos (MLGD) para as situações em que a variável resposta tem distribuição que pertence à família exponencial e o parâmetro de dispersão não é constante.

Uma importante etapa do ajuste de um modelo de regressão é a seleção de variáveis, que consiste em escolher, dentre as variáveis preditoras disponíveis, aquelas que devem ser consideradas no ajuste da variável resposta. A partir da década de 1970 foram propostos procedimentos, medidas e algoritmos que não envolvem testes estatísticos para seleção de um modelo de regressão que seja parcimonioso, ou seja, que esteja bem ajustado e possua um número reduzido de parâmetros.

Resumidamente, como o valor do logaritmo da função de verossimilhança cresce com o aumento do número de parâmetros do modelo, a ideia dos critérios de informação é acrescentar a este valor uma quantidade, denominada termo penalizador, que aumenta conforme o número de parâmetros cresce. O primeiro destes critérios foi introduzido por [Akaike \(1974\)](#) e denominado AIC. Para resolver o problema de que o AIC muitas vezes inclui mais variáveis do que é necessário, diversas outras medidas foram propostas, como, por exemplo, o BIC ([Akaike, 1978](#); [Schwarz, 1978](#)), o HQC ([Hannan e Quinn, 1979](#)) e o AIC corrigido ([Hurvich e Tsai, 1989](#)). Outro conjunto de critérios são os baseados em bootstrap

(Efron e Tibshirani, 1997).

Utilizando como ferramenta a seleção de variáveis, desejamos encontrar um modelo suficientemente bom para explicar as relações presentes entre as variáveis preditoras e a resposta. Por outro lado, este modelo também deve ser simples o suficiente para ser entendido, explicado para outras pessoas e usado. Quando negociamos esse equilíbrio, os métodos de seleção de modelos entram em jogo. Eles fornecem um suporte formal para orientar os indivíduos na busca de modelos adequados para ajustar a variável resposta ou para determinar quais variáveis incluir quando o interesse está em realizar previsões e classificações (Claeskens e Hjort, 2008).

No contexto de MLGD, Bayer e Cribari-Neto (2015b) propuseram, dentre outras coisas, um esquema de seleção de variáveis em dois passos para o modelo de regressão beta duplo. Neste trabalho, os autores concluíram por meio de estudos de simulação de Monte Carlo que o método proposto é preciso e apresenta baixo custo computacional.

Baseando-se no esquema de seleção de variáveis proposto por Bayer e Cribari-Neto (2015b) desenvolvemos um procedimento para realizar seleção de variáveis em até “ k ” passos para modelos lineares generalizados duplos. Além disso, o implementamos computacionalmente no *software* R (R Development Core Team, 2008) e estudamos a sua performance. Com essa finalidade, desenvolvemos estudos de simulação de Monte Carlo considerando as distribuições Gama, Gaussiana Inversa e Normal, e comparamos nosso procedimento com o esquema de seleção em dois passos, a seleção dentre todos os possíveis modelos ajustados e um método *Stepwise* para MLGD (descrito em Stasinopoulos et al. (2017)).

Este trabalho está organizado da seguinte forma. O Capítulo 2 apresenta as principais características dos modelos lineares generalizados duplos, como especificação, interpretação, inferências sobre os parâmetros e análise de diagnóstico. O Capítulo 3 discute os procedimentos para seleção de variáveis, como os critérios de informação e procedimentos para seleção automática de variáveis. Neste capítulo também descrevemos detalhadamente o procedimento que desenvolvemos para seleção de variáveis em MLGD e as demais metodologias que utilizamos nos estudos de simulação. O Capítulo 4 discute como foram desenvolvidos os estudos de simulação, os resultados obtidos para as três distribuições estudadas, a avaliação geral do número de passos do nosso procedimento para a seleção de um modelo candidato e um estudo a respeito do número de passos em relação ao tempo necessário para esta seleção, comparando o nosso esquema com o método *Stepwise*. O Capítulo 5 compreende a aplicação que realizamos a um conjunto de dados reais e comparamos o poder preditivo dos métodos utilizados nos estudos de simulação. Além disso, também comparamos neste capítulo o poder preditivo do nosso esquema para seleção de variáveis com outros métodos de regressão. O Capítulo 6 apresenta a conclusão deste trabalho e sugestões para estudos posteriores. Os Apêndices A, B e C destinam-se aos resultados completos dos estudos de simulação para as distribuições Gama, Gaussiana Inversa e Normal, respectivamente.

Capítulo 2

Modelos lineares generalizados duplos

Neste capítulo discorreremos brevemente sobre as principais características dos modelos lineares generalizados duplos. Na Seção 2.1 apresentamos a família exponencial e as distribuições utilizadas no trabalho. Na Seção 2.2 definimos os MLGD. Na Seção 2.3 discutimos sobre interpretação dos parâmetros para os submodelos da média e do parâmetro de dispersão. Na Seção 2.4 comentamos a respeito da inferência sobre os parâmetros e, na última seção, explicitamos as principais ideias relacionadas com a análise de diagnóstico. Maiores detalhes a respeito desses modelos podem ser encontrados em [Paula \(2004\)](#) e [Smyth \(1989\)](#).

2.1 Família exponencial e distribuições contínuas

Inicialmente, considere a seguinte definição sobre a família exponencial:

Definição 2.1 Dizemos que uma variável aleatória Y pertence à família exponencial caso possamos reescrever a sua função densidade de probabilidade da seguinte forma

$$f(y; \theta, \phi) = \exp \left[\frac{y\theta - b(\theta)}{\phi^2} + C(y, \phi) \right] I_A(y), \quad (2.1)$$

em que a esperança de Y é dada por $E(Y) = \mu = b'(\theta)$, a função de variância da média é $V(\mu) = b''(\theta(\mu))$, a variância de Y pode ser obtida através de $Var(Y) = \phi^2 V(\mu)$ e A é o suporte de Y , ou seja, um conjunto que não depende nem de θ e nem de ϕ .

A parametrização (2.1) da família não é usualmente utilizada em livros que tratam de MLG e MLGD. Porém a utilizamos para facilitar a especificação dos termos presentes nos modelos lineares generalizados duplos que discutiremos a seguir. Além disso, também seguimos a forma utilizada no pacote *Generalized Additive Models for Location, Scale and Shape (GAMLSS)*, do *software R*, desenvolvido por [Stasinopoulos e Rigby \(2007\)](#), que engloba a classe de modelos que estamos interessados e usamos nos estudos de simulação de Monte Carlo.

Neste trabalho, utilizamos as seguintes distribuições contínuas:

- **Distribuição Normal**

Considere Y uma variável aleatória que assume distribuição Normal com média μ e variância ϕ^2 , denotada por $Y \sim N(\mu; \phi)$. A função densidade de probabilidade de Y é definida como

$$f(y|\mu, \phi) = \frac{1}{\sqrt{2\pi\phi^2}} \exp\left[-\frac{(y - \mu)^2}{2\phi^2}\right], \quad (2.2)$$

em que, μ e $y \in (-\infty, \infty)$ e $\phi > 0$.

Para a distribuição Normal pode-se demonstrar que

$$\theta = \mu, \quad b(\theta) = \frac{\theta^2}{2}, \quad c(y, \phi) = -\frac{1}{2} \left[\log(2\pi\phi^2) + \frac{y^2}{\phi^2} \right].$$

Podemos verificar também que a função de variância é dada por $V(\mu) = 1$ e, desse modo, $Var(Y) = \phi^2$.

- **Distribuição Gama**

Considere Y uma variável aleatória com distribuição Gama de média μ e parâmetro de dispersão ϕ , denotada por $Y \sim GA(\mu; \phi)$. A função densidade de probabilidade de Y é definida como

$$f(y|\mu, \phi) = \frac{1}{(\phi^2\mu)^{\frac{1}{\phi^2}} \Gamma\left(\frac{1}{\phi^2}\right)} y^{\frac{1}{\phi^2}-1} \exp\left[-\frac{y}{\phi^2\mu}\right], \quad (2.3)$$

em que $y > 0$, $\mu > 0$ e $\phi > 0$.

Para a distribuição Gama podemos demonstrar que

$$\theta = -\frac{1}{\mu}, \quad b(\theta) = \log\left(-\frac{1}{\theta}\right), \quad c(y, \phi) = -\frac{1}{\phi^2} \left[\log(\phi^2 y) \right] - \log\left(\Gamma\left(\frac{1}{\phi^2}\right)\right) - \log(y).$$

Pode-se verificar que, para a distribuição Gama, a função de variância é dada por $V(\mu) = \mu^2$ e, assim, $Var(Y) = \mu^2\phi^2$.

- **Distribuição Gaussiana Inversa**

Considere Y uma variável aleatória com distribuição Gaussiana Inversa de média μ e parâmetro de dispersão ϕ , denotada por $Y \sim GI(\mu; \phi)$. A função densidade de probabilidade de Y é definida como

$$f(y|\mu, \phi) = \frac{1}{\sqrt{2\pi\phi^2y^3}} \exp\left[-\frac{(y-\mu)^2}{2\mu^2\phi^2y}\right], \quad (2.4)$$

em que $y > 0$, $\mu > 0$ e $\phi > 0$.

Para a distribuição Gaussiana Inversa pode-se demonstrar que

$$\theta = -\frac{1}{2\mu^2}, \quad b(\theta) = -(-2\theta)^{\frac{1}{2}}, \quad c(y, \phi) = -\frac{1}{2} \left[\frac{1}{\phi^2y} + \log(2\pi\phi^2y^3) \right].$$

Podemos verificar que, para a distribuição Gaussiana Inversa, a função de variância é dada por $V(\mu) = \mu^3$ e, desse modo, $Var(Y) = \mu^3\phi^2$.

2.2 Especificação do modelo

Os modelos lineares generalizados duplos (MLGD), que foram introduzidos por Smyth (1989), são compostos por dois submodelos. O primeiro submodelo é destinado à modelagem da média, enquanto que o segundo é utilizado para ajustar o parâmetro de dispersão de uma determinada variável de interesse. Eles podem ser definidos da seguinte forma:

Definição 2.2 Sejam y_1, y_2, \dots, y_n variáveis aleatórias independentes que pertencem à família exponencial conforme definido na Equação (2.1). Os modelos lineares generalizados duplos (MLGD) são definidos por $f(y_i; \mu_i, \phi_i)$ e pelos seguintes componentes sistemáticos

$$\begin{cases} g(\mu_i) = \eta_i, \\ h(\phi_i) = \zeta_i, \end{cases} \quad (2.5)$$

em que $\eta_i = \mathbf{x}_{i1}^T \boldsymbol{\beta}$ e $\zeta_i = \mathbf{x}_{i2}^T \boldsymbol{\gamma}$ são os preditores lineares, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$ e $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_q)^T$ são os vetores de parâmetros desconhecidos, $\mathbf{x}_{i1} = (x_{i11}, x_{i21}, \dots, x_{ip1})^T$ e $\mathbf{x}_{i2} = (x_{i12}, x_{i22}, \dots, x_{iq2})^T$ são constantes que representam os valores das variáveis preditoras para os dois submodelos e, $g(\cdot)$ e $h(\cdot)$, são as funções de ligação estritamente monótonas e duplamente diferenciáveis para os submodelos de μ e ϕ , respectivamente.

Os MLGs são um caso particular dos MLGDs. Com isso, quando consideramos em um MLGD que $h(\phi_i) = \zeta_i = \gamma_1$ (com γ_1 representando o intercepto do preditor linear), estamos ajustando um modelo nulo para o submodelo do parâmetro de dispersão e, desse modo, obtemos um MLG.

Na prática, um exemplo em que podemos observar o parâmetro de dispersão não constante pode ser obtido ao se medir determinada característica em pessoas, como a pressão arterial diastólica, e a mesma apresentar parâmetro de dispersão que aumenta

ou diminui conforme a idade, o sexo ou o peso das pessoas. Neste caso, o ajuste de um MLG não seria adequado porque o parâmetro de dispersão da variável de interesse altera-se de acordo com algumas covariáveis. Assim, para que a inferência relacionada com os parâmetros da média seja conduzida de forma correta, precisamos incorporar esta informação no ajuste. Neste caso, ajustar um MLGD pode ser uma possível solução para este problema.

2.3 Interpretação dos parâmetros do modelo

A interpretação dos parâmetros de um modelo estatístico é importante do ponto de vista inferencial e aplicado, pois por meio dela entendemos as relações existentes entre a variável resposta e as variáveis preditoras utilizadas na construção do modelo. Nesta seção, explicitaremos as interpretações relacionadas com as funções de ligação que utilizamos nos estudos de simulação.

2.3.1 Submodelo da média

Considerando a distribuição Normal, para relacionar a média μ com o componente linear do modelo, garantindo que a variável resposta estimada seja predita no intervalo contínuo $(-\infty, +\infty)$, a função de ligação comumente utilizada é a identidade. Desse modo,

$$g(\mu_i) = \mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip1}. \quad (2.6)$$

Para interpretarmos os parâmetros do submodelo da média com função de ligação identidade, ao considerarmos $x_{ij1} = l$, fixando-se as outras variáveis, temos que

$$\mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + l \times \beta_j + \cdots + \beta_p x_{ip1}. \quad (2.7)$$

Por outro lado, se $x_{ij1} = l + 1$, obteremos

$$\mu_i = \beta_0 + \beta_1 x_{i1} + \cdots + (l + 1) \times \beta_j + \cdots + \beta_p x_{ip1}. \quad (2.8)$$

Calculando-se (2.8) – (2.7) resulta-se em

$$\beta_0 + \cdots + (l + 1) \times \beta_j + \cdots + \beta_p x_{ip1} - (\beta_0 + \cdots + l \times \beta_j + \cdots + \beta_p x_{ip1}) = \beta_j. \quad (2.9)$$

Utilizando a Equação (2.9) podemos afirmar que β_j é o valor pelo qual é acrescida a média da variável resposta quando x_{ij1} tem acréscimo de uma unidade, enquanto que as demais variáveis preditoras mantêm-se constantes.

Em relação às distribuições Gama e Gaussiana Inversa, para que o valor de μ seja

estimado no intervalo contínuo $(0, +\infty)$, normalmente utiliza-se a função de ligação logarítmica. Com isso,

$$h(\mu_i) = \log(\mu_i) = \beta_0 + \beta_1 x_{i11} + \cdots + \beta_p x_{ip1}. \quad (2.10)$$

Para interpretarmos os parâmetros do submodelo da média com função de ligação logarítmica, considere $x_{ij1} = l$. Quando as demais variáveis mantêm-se constantes, temos

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i11} + \cdots + l \times \beta_j + \cdots + \beta_p x_{ip1}). \quad (2.11)$$

Entretanto, se $x_{ij1} = l + 1$, encontra-se

$$\mu_i = \exp(\beta_0 + \beta_1 x_{i11} + \cdots + (l + 1) \times \beta_j + \cdots + \beta_p x_{ip1}). \quad (2.12)$$

Fazendo-se $\frac{(2.12)}{(2.11)}$, obtemos

$$\frac{\exp(\beta_0 + \beta_1 x_{i11} + \cdots + (l + 1) \times \beta_j + \cdots + \beta_p x_{ip1})}{\exp(\beta_0 + \beta_1 x_{i11} + \cdots + l \times \beta_j + \cdots + \beta_p x_{ip1})} = \exp(\beta_j). \quad (2.13)$$

Através da Equação (2.13), podemos afirmar que $\exp(\beta_j)$ é o valor pelo qual é multiplicada a média da variável resposta quando x_{ij1} tem acréscimo de uma unidade, enquanto que as demais variáveis preditoras mantêm-se constantes.

2.3.2 Submodelo do parâmetro de dispersão

Em relação ao parâmetro de dispersão ϕ , para que seu valor seja estimado no intervalo contínuo $(0, +\infty)$, para as três distribuições que usamos neste trabalho normalmente utiliza-se a função de ligação logarítmica. Com isso, a sua interpretação é semelhante à descrita anteriormente para o submodelo da média com esta função de ligação, ou seja, $\exp(\gamma_k)$ é o valor pelo qual é multiplicado o parâmetro de dispersão da variável resposta quando x_{ik2} tem acréscimo de uma unidade, mantendo-se as demais variáveis preditoras constantes.

2.4 Inferência sobre os parâmetros

Segundo Paula (2004), o método comumente empregado para a estimação dos parâmetros desconhecidos em modelos lineares generalizados é o método da Máxima Verossimilhança. Esse método produz estimadores para os parâmetros desconhecidos que maximizam a função de verossimilhança, a qual é construída através da densidade de probabilidade conjunta dos dados coletados ao considerar como desconhecidos os parâmetros. Com isso, os estimadores resultantes são aqueles que melhor representam o

conjunto de dados observados.

Considerando que a maximização da função de verossimilhança na maioria dos casos em MLGD não possui solução algébrica, podemos concluir que a mesma não pode ser maximizada analiticamente. Entretanto, esse problema pode ser resolvido numericamente. Um dos primeiros métodos numéricos proposto para realizar a estimação dos parâmetros desconhecidos da função de verossimilhança é conhecido como “Mínimos Quadrados Reponderados Iterativamente” (*Iteratively Reweighted Least Squares - IRLS*), o qual é descrito detalhadamente em [McCullagh e Nelder \(1989\)](#). Este método consiste, resumidamente, em maximizar o logaritmo da função de verossimilhança resolvendo “equações de verossimilhança” por meio de métodos numéricos.

Atualmente, outros algoritmos também são bastante usados para maximizar a função de verossimilhança. No contexto de modelos lineares generalizados duplos, podemos destacar os algoritmos CG ([Cole e Green, 1992](#)) e RS ([Rigby e Stasinopoulos, 1996](#)). Os dois algoritmos são descritos em [Rigby e Stasinopoulos \(2005\)](#).

Os intervalos de confiança para os parâmetros podem ser obtidos por meio de propriedades assintóticas do estimador de máxima verossimilhança ([Sen et al., 2010](#)). Testes de hipóteses para os parâmetros podem ser feitos a partir dos testes da razão de verossimilhanças, score, Wald ([Paula, 2004](#)) e gradiente ([Terrell, 2002](#)).

2.5 Análise de diagnóstico

A análise de diagnóstico é fundamental no ajuste de um modelo, pois através dela é possível verificar se os pressupostos para a construção do modelo são satisfeitos e investigar a existência de observações alavancas, discrepantes e influentes ([Paula, 2004](#)). Para a identificação de observações discrepantes, utilizamos os resíduos, que quantificam a discrepância entre os valores ajustados e os observados. Na prática, dentre os resíduos mais utilizados podemos destacar o resíduo deviance ([Davison et al., 1989](#)), o resíduo de Pearson ([Paula, 2004](#)) e o resíduo quantílico ([Dunn e Smyth, 1996](#)).

Representações gráficas dos resíduos são extremamente importantes, pois visualmente conseguimos identificar a presença de observações discrepantes e possíveis violações nas suposições para que seja possível a construção do modelo. Geralmente construímos gráficos de resíduos contra valores observados, valores preditos e variáveis predictoras para realizar estas investigações. Além disso, também podemos obter bandas de confiança para os resíduos por meio do gráfico de probabilidade normal com envelope simulado. Esse método foi desenvolvido por [Atkinson \(1985\)](#) e, caso muitos pontos estejam fora do envelope simulado, podemos concluir que existem evidências de violação às suposições do modelo. Outros aspectos de análise de diagnóstico em MLGD, como a identificação de pontos de alavanca e influentes, podem ser vistos em [Paula \(2013\)](#).

Capítulo 3

Procedimentos para seleção de variáveis

Segundo Kutner et al. (2004), a identificação de “bons” subconjuntos de variáveis preditoras potencialmente úteis para serem incluídas no modelo de regressão final constitui um dos problemas mais difíceis na Análise de Regressão.

Em contrapartida, se todas as variáveis preditoras estão incluídas no modelo linear generalizado final, o mesmo poderá ficar super-ajustado aos dados. Desse modo, este modelo dificilmente produzirá boas previsões e a variância dos estimadores dos parâmetros será maior do que a de modelos mais simples. Nesse sentido, é fundamental utilizar uma medida e/ou um algoritmo para seleção de variáveis que possuam um bom desempenho e realizem essa tarefa em um tempo relativamente baixo em comparação com a seleção de um modelo ajustando-se todos os possíveis.

Como mencionamos anteriormente, a partir da década de 1970 foram propostos procedimentos, medidas e algoritmos que não envolvem testes estatísticos para seleção de um modelo de regressão que esteja bem ajustado e possua um número reduzido de parâmetros. Dentre as medidas, podemos citar os critérios de informação, que foram desenvolvidos com o objetivo de comparar modelos não encaixados.

Trabalhos que comparam critérios de informação consideram que será selecionado, na maioria dos casos, o melhor modelo segundo o critério escolhido dentre todos os modelos possíveis. Entretanto, devido ao alto tempo computacional necessário, é inviável utilizar este procedimento se o número de variáveis preditoras disponíveis não for pequeno. Nesse sentido, foram desenvolvidos algoritmos que procuram identificar o melhor modelo sem que seja necessário o ajuste de todos os possíveis.

Na Seção 3.1 discutimos a respeito de alguns trabalhos desenvolvidos sobre seleção de variáveis em MLGD e apresentamos os critérios de informação que utilizamos neste trabalho. Nas seções seguintes apresentamos os algoritmos para seleção de variáveis que usamos, o esquema de seleção de variáveis em dois passos e o método *Stepwise* para MLGD, e introduzimos o esquema de seleção de variáveis em até “ k ” passos.

3.1 Critérios de seleção de variáveis

Inúmeros trabalhos já compararam diversos critérios para a seleção de variáveis em modelos de regressão. Porém, apenas nos últimos anos foram produzidos trabalhos relacionados à seleção de variáveis em modelos de regressão com dispersão variável. [Cottet et al. \(2008\)](#) propuseram um critério de seleção de variáveis em modelos semi-paramétricos da família exponencial. Neste trabalho, os autores desenvolveram um quadro Bayesiano geral para a seleção de variáveis em modelos lineares generalizados que permite a sobre-dispersão e a subdispersão.

[Xu et al. \(2014\)](#) estudaram a seleção de variáveis e estimação dos parâmetros em modelos lineares generalizados duplos na presença de grande número de variáveis. No seu trabalho, os autores propuseram um método de seleção de modelos baseado na máxima penalização da pseudo-verossimilhança quando o número de parâmetros diverge com o tamanho da amostra. Além disso, eles realizaram estudos de simulação e uma análise de dados reais para avaliar o desempenho do método proposto em amostras finitas, mostrando que o mesmo funciona de forma satisfatória.

Critérios de seleção de variáveis em modelos com dispersão variável também foram estudados para a regressão beta com foco principal em amostras pequenas por [Bayer e Cribari-Neto \(2015a\)](#), os quais propuseram dois novos critérios de seleção: o “*bootstrapped likelihood quasi-CV*” e seu variante “*632QCV*”. Através de simulações de Monte Carlo, os autores compararam a performance dos dois critérios propostos com o AIC e alguns de seus variantes em amostras finitas para a regressão beta com dispersão variável. As evidências mostraram que os critérios de informação propostos tiveram boas performances em amostras pequenas.

[Wu et al. \(2012\)](#) propuseram um método unificado de verossimilhança penalizada que pode selecionar simultaneamente variáveis significativas nos modelos de média e dispersão para a regressão lognormal com dispersão variável. Além disso, o método de seleção de variáveis proposto pode executar simultaneamente estimativa de parâmetros e seleção de variáveis nos modelos de média e dispersão. [Wu e Li \(2012\)](#) fizeram um estudo semelhante para a regressão gaussiana inversa com dispersão variável.

Além dos trabalhos acima mencionados, podemos destacar o estudo de [Cavalaro \(2016\)](#) que, a partir de estudos de simulação de Monte Carlo, concluiu que penalizar de forma diferente as variáveis preditoras presentes no submodelo da média e do parâmetro de dispersão pode levar a um melhor critério de seleção para modelos lineares generalizados duplos (MLGD) com variável resposta assumindo distribuição Normal.

Neste trabalho, considerando que nosso principal objetivo é estudar a performance do procedimento que desenvolvemos para seleção de variáveis em modelos lineares generalizados duplos, utilizamos alguns dos critérios para seleção de variáveis mais conhecidos e amplamente utilizados:

- Critério de Informação de Akaike (AIC);

- Critério de Informação Bayesiano (BIC);
- Critério de Informação de Hannan-Quinn (HQC);
- Critério de Informação de Akaike corrigido (AICc).

3.1.1 Critério de Informação de Akaike (AIC)

Considere modelos paramétricos com parâmetro p -dimensional θ . Sejam y_1, y_2, \dots, y_n variáveis aleatórias independentes e $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$, para $i = 1, \dots, n$, as constantes que representam os valores das variáveis preditoras.

Akaike (1973) propôs o uso da Informação de Kullback-Leibler como a base fundamental para realizar seleção de modelos. Entretanto, a Informação de Kullback-Leibler não pode ser completamente calculada porque não possuímos completo conhecimento da verdadeira função de verossimilhança (f_0) e do verdadeiro valor de seus parâmetros (θ), para os utilizarmos nos modelos candidatos $g_i(x|\theta)$. Devido a este fato, Akaike propôs uma forma de estimar a Informação de Kullback-Leibler, baseado na função de log-verossimilhança avaliada em seu ponto de máximo. A Informação de Kullback-Leibler entre os modelos f_0 e g é definida, para funções contínuas, pela integral

$$I(f_0, g) = \int_{-\infty}^{\infty} f_0(x) \ln \left(\frac{f_0(x)}{g(x|\theta)} \right) dx, \quad (3.1)$$

em que $I(f_0, g)$ refere-se à “informação perdida quando g é utilizada para aproximar f_0 ”. Como uma interpretação heurística, $I(f_0, g)$ mede a distância de g para f_0 (Burnham e Anderson, 2003).

Akaike (1973, 1974) mostrou que um estimador aproximadamente não viesado para a $I(f_0, g)$ definida na Equação (3.1) para grandes amostras e “bons” modelos é

$$\log(L(\hat{\theta}|dados)) - k, \quad (3.2)$$

em que $\log(L(\hat{\theta}|dados))$ é o valor do logaritmo da função de verossimilhança aplicado no estimador de máxima verossimilhança dos parâmetros para o modelo de regressão construído e k é o número de parâmetros deste modelo (Burnham e Anderson, 2003).

Akaike (1974) definiu um critério de informação multiplicando $\log(L(\hat{\theta}|dados)) - k$ por -2 (devido à função Deviance) e, desse modo, obteve

$$AIC = -2 \log(L(\hat{\theta}|dados)) + 2k, \quad (3.3)$$

o qual ficou conhecido como “Critério de Informação de Akaike” ou AIC. Assim, ao invés de obtermos uma medida simples da distância direcionada entre dois modelos (isto é, a

Informação de Kullback-Leibler), temos uma estimativa de uma medida de proximidade (já que multiplicamos por -2) relativa entre o modelo ajustado e o mecanismo verdadeiro desconhecido (talvez de dimensão infinita) que realmente gerou os dados observados (Burnham e Anderson, 2003). Dessa forma, busca-se selecionar o modelo com o menor AIC.

Na Equação 3.3, o termo “ $2k$ ” é denominado termo penalizador. Com isso, a penalização para a inclusão de uma variável no modelo é igual 2 e, assim, de acordo com esse critério uma variável será inserida no modelo apenas se o logaritmo da função de verossimilhança aumentar mais de uma unidade com o acréscimo da mesma.

3.1.2 Critério de Informação Bayesiano (BIC)

O Critério de Informação Bayesiano (BIC), desenvolvido por Schwarz (1978) e Akaike (1978), foi proposto como uma abordagem alternativa para o problema de seleção de variáveis. Além disso, este critério procura resolver o problema de que o AIC algumas vezes inclui mais variáveis no modelo final do que realmente é necessário.

Em um modelo de regressão, as dimensões dos estimadores de máxima verossimilhança podem ser obtidas como limites de grandes amostras do estimador de Bayes para distribuições a priori arbitrárias (Schwarz, 1978).

Para construir o BIC, Schwarz (1978) buscou por uma apropriada modificação da máxima verossimilhança estudando o comportamento assintótico do estimador de Bayes em uma classe especial de prioris. Essas prioris não são absolutamente contínuas, uma vez que consideram probabilidade positiva em alguns subespaços de baixa dimensão do espaço paramétrico, ou seja, os subespaços correspondentes aos modelos concorrentes.

Nos limites de grandes amostras, o termo principal do estimador de Bayes é o estimador de máxima verossimilhança. Apenas no próximo termo algo novo é obtido. Isso é esperado pois o segundo termo que reflete as singularidades da distribuição a priori. Desenvolvendo esta ideia e utilizando alguns argumentos assintóticos, Schwarz (1978) definiu a escolha de um modelo quando o mesmo apresentar o maior valor para a seguinte equação

$$\log(L(\hat{\theta}|dados)) - \frac{1}{2} \log(n)k, \quad (3.4)$$

em que n refere-se ao número de observações presentes na amostra.

Multiplicando-se a equação acima por -2 , obtemos o Critério de Informação Bayesiano (BIC), que tem como objetivo minimizar a seguinte quantidade

$$BIC = -2 \log(L(\hat{\theta}|dados)) + \log(n)k. \quad (3.5)$$

A solução do estimador de Bayes consiste em selecionar o modelo mais provável a

posteriori (Schwarz, 1978). Desse modo, podemos entender que o BIC considera que o melhor modelo para os dados pode ser encontrado utilizando as covariáveis disponíveis. Além disso, a penalização para a inclusão de uma covariável no BIC é maior do que a penalização para o AIC quando utilizamos amostras com mais de 7 observações, ou seja, o BIC seleciona modelos com número de variáveis menor ou igual ao do modelo selecionado pelo AIC em amostras com mais de 7 observações.

3.1.3 Critério de Informação de Hannan-Quinn (HQC)

Hannan e Quinn (1979) introduziram o Critério de Informação de Hannan-Quinn (HQC) no contexto de modelos autorregressivos de séries temporais. Neste contexto, Shibata (1976) investigou as propriedades assintóticas do AIC e mostrou que o mesmo não é consistente, pois superestima assintoticamente o grau da autocorrelação estimado. Por outro lado, Hannan e Quinn (1979) mostraram que o BIC subestima, em grandes amostras, a ordem da autocorrelação estimada. Com isso, os autores construíram um critério de informação com o objetivo de garantir que o mesmo seja eficiente assintoticamente para estimar o grau da autocorrelação em modelos de séries temporais.

O HQC pode ser utilizado para seleção de variáveis em modelos de regressão através da minimização da seguinte função

$$HQC = -2 \log(L(\hat{\theta}|dados)) + 2(\log(\log(n)))k. \quad (3.6)$$

O termo de penalização para o HQC é intermediário entre o AIC e o BIC se o tamanho da amostra for superior a 15 observações.

3.1.4 Critério de Informação de Akaike corrigido (AICc)

Assim como o Critério de Informação de Hannan-Quinn, o Critério de Informação de Akaike corrigido (AICc), desenvolvido por Hurvich e Tsai (1989), também foi introduzido inicialmente no contexto de modelos autorregressivos de séries temporais. Neste caso, os autores propuseram uma correção para o viés do Critério de Informação de Akaike para ser utilizada quando a amostra não possui muitas observações, ou o número de parâmetros estimados é uma fração moderada ou grande do tamanho amostral.

Os autores também afirmam que o AICc é eficiente assintoticamente caso o modelo verdadeiro seja dimensionalmente infinito. Além disso, eles argumentam que, nos casos em que o modelo verdadeiro possui dimensão finita, o AICc apresenta melhor performance para escolher a ordem do modelo que qualquer outro método eficiente assintoticamente. O objetivo deste critério, no contexto de modelos de regressão, é minimizar a seguinte função

$$AICc = -2 \log(L(\hat{\theta}|dados)) + 2k \left(\frac{n}{n - k - 1} \right).$$

Podemos observar que a penalização para a inclusão de uma variável para o AICc é superior ao do AIC. Entretanto, quando $n \rightarrow \infty$, o termo de penalização do AICc converge para $2k$.

3.2 Esquema de seleção de variáveis em dois passos

Bayer e Cribari-Neto (2015b) propuseram o seguinte esquema de seleção de variáveis em dois passos para o modelo de regressão beta duplo:

1. Assumindo que o parâmetro de dispersão é constante, ajustam-se todos os possíveis modelos para a média e seleciona-se o modelo de acordo com algum critério definido (menor AIC, por exemplo);
2. Assumindo que o modelo selecionado na etapa (1) é correto para o submodelo da média, ajustam-se todos os possíveis modelos para o parâmetro de dispersão e seleciona-se o modelo de acordo com algum critério definido.

Podemos verificar a grande vantagem deste esquema quando comparado com a seleção conjunta dos dois submodelos em relação ao menor custo computacional. Considerando m covariáveis candidatas para os submodelos da média e do parâmetro de dispersão, na seleção conjunta temos um total de $2^m \times 2^m = 2^{2m}$ modelos candidatos a serem estimados, enquanto que pelo esquema de seleção de variáveis em dois passos precisamos ajustar apenas 2^{m+1} modelos. Por exemplo, se considerarmos $m = 10$, temos um total de $2^{20} = 1.048.576$ possíveis modelos, enquanto que pelo esquema em dois passos precisamos ajustar apenas $2^{11} = 2.048$ modelos, ou seja, aproximadamente 0,2% do total de modelos.

Através de estudos de simulação de Monte Carlo, Bayer e Cribari-Neto (2015b) concluíram que o esquema de seleção de variáveis em dois passos obteve um desempenho semelhante ou superior, em comparação com a seleção dentre todos os modelos, na proporção de escolha do modelo verdadeiro para o modelo de regressão beta duplo. Além disso, os autores mostraram que este método possui custo computacional muito menor do que a seleção ajustando-se todos os possíveis modelos. Em contrapartida, a performance desse procedimento ainda não foi avaliada em MLGD.

3.3 Esquema de seleção de variáveis em até “ k ” passos

Baseando-se no esquema de seleção de variáveis em dois passos, criado por Bayer e Cribari-Neto (2015b), introduzimos um esquema de seleção de variáveis para MLGD em até “ k ” passos. O intuito do procedimento proposto é aumentar a proporção de vezes que o esquema de seleção de variáveis consegue encontrar um bom modelo (ou o “melhor

modelo”) para explicar as relações presentes entre as covariáveis e a variável resposta, sem aumentar de forma excessiva o custo computacional. Buscamos realizar essa melhora no desempenho corrigindo o possível problema de má especificação do modelo presente no passo 1 do esquema de seleção de variáveis em dois passos, uma vez que no mesmo seleciona-se um MLG, mas o verdadeiro interesse do método está em obter um MLGD como modelo candidato.

Para o nosso esquema, aumentamos o número de passos que o algoritmo pode executar e, considerando que na prática a dispersão é um parâmetro de perturbação ajustado para que a inferência associada aos parâmetros da média seja conduzida de forma adequada penalizamos, com peso 2 nos passos pares, as variáveis não selecionadas para o submodelo da média no passo anterior. O peso é aplicado no cálculo do termo penalizador do critério de informação quando usamos variáveis não presentes no submodelo da média obtido no passo anterior. O objetivo deste procedimento é evitar que variáveis não importantes para o modelo final sejam incluídas no submodelo do parâmetro de dispersão. O nosso procedimento possui o seguinte algoritmo:

1. Assumindo que o parâmetro de dispersão é constante, ajustam-se todos os possíveis modelos para a média e seleciona-se o modelo de acordo com algum critério definido (menor AIC, por exemplo);
2. Assumindo que o modelo selecionado no passo (1) é correto para o submodelo da média, ajustam-se todos os possíveis modelos para o parâmetro de dispersão, considerando peso 2 para as variáveis não presentes no submodelo da média encontrado no passo anterior, e seleciona-se o modelo de acordo com algum critério definido;
3. Assumindo que o modelo selecionado na etapa anterior é correto para o submodelo do parâmetro de dispersão, seleciona-se novamente o submodelo da média dentre todos os possíveis. Caso os submodelos obtidos neste passo e à dois passos atrás forem iguais, o algoritmo termina;
4. Assumindo que o modelo selecionado na etapa anterior é correto para o submodelo da média, seleciona-se novamente o submodelo do parâmetro de dispersão dentre todos os possíveis, considerando peso 2 para as variáveis não presentes no submodelo da média selecionado no passo anterior. Caso os submodelos obtidos neste passo e à dois passos atrás forem iguais, o algoritmo termina; caso contrário, repete-se os passos 3 e 4 até que o procedimento atinja a convergência.

Para ilustrarmos o uso da penalização, considere um conjunto de dados com 3 variáveis preditoras, ou seja, $\mathbf{x} = (x_1, x_2, x_3)$, e o AIC como critério de informação. Além disso, suponha que no passo 1 o modelo que apresentou o menor AIC possui as variáveis x_1 e x_3 . Considerando que a variável x_2 não foi selecionada no passo 1 para o submodelo da média, temos uma evidência de que a mesma talvez não seja muito importante para o

modelo final e, devido a este fato, penalizamos a mesma com peso 2 no cálculo do AIC durante o passo 2. Na Tabela 3.1 exemplificamos o uso da penalização no cálculo do AIC para este caso e evidenciamos que, quando a variável x_2 está presente no submodelo do parâmetro de dispersão, a mesma possui sempre peso 2 no termo penalizador (“ $2k$ ” da Equação (3.3)).

Tabela 3.1: Cálculo do termo penalizador do AIC considerando as variáveis x_1 e x_3 no submodelo da média e $\mathbf{x} = (x_1, x_2, x_3)$ como possíveis covariáveis.

Submodelo - média	Submodelo - par. de dispersão	Termo penalizador do AIC
$x_1 + x_3$	modelo nulo	$2 \times (3 + 1) = 8$
$x_1 + x_3$	x_1	$2 \times (3 + 2) = 10$
$x_1 + x_3$	x_2	$2 \times (3 + 1 + 2 \times 1) = 12$
$x_1 + x_3$	x_3	$2 \times (3 + 2) = 10$
$x_1 + x_3$	$x_1 + x_2$	$2 \times (3 + 2 + 2 \times 1) = 14$
$x_1 + x_3$	$x_1 + x_3$	$2 \times (3 + 3) = 12$
$x_1 + x_3$	$x_2 + x_3$	$2 \times (3 + 2 + 2 \times 1) = 14$
$x_1 + x_3$	$x_1 + x_2 + x_3$	$2 \times (3 + 3 + 2 \times 1) = 16$

Conforme é possível observar na Tabela 3.1, o termo penalizador referente ao modelo com x_1 e x_3 no submodelo da média e x_2 no submodelo do parâmetro de dispersão será duas unidades maior do que o termo penalizador referente ao modelo com x_1 e x_3 no submodelo da média e x_1 ou x_3 no submodelo do parâmetro de dispersão. Neste sentido, o termo penalizador adicionando-se somente x_2 no submodelo do parâmetro de dispersão é igual ao termo penalizador com x_1 e x_3 neste submodelo (pois as mesmas possuem peso igual a 1, dado que foram incluídas no submodelo da média durante o passo anterior). Desta forma, garantimos que x_2 somente será incluída no submodelo do parâmetro de dispersão se for muito importante, uma vez que ela terá o dobro do peso no cálculo do AIC em comparação com as demais covariáveis. Posteriormente, supondo que no passo 3 as variáveis x_1 e x_2 passaram a compor o submodelo da média, durante o passo 4 esse procedimento penalizará a variável x_3 com peso 2.

Escolhemos o valor 2 para o peso porque este valor é bastante intuitivo e, em estudos preliminares, observamos que a performance do esquema de seleção de variáveis em até “ k ” passos melhora quando adotamos esse tipo de penalização em relação a sua não adoção. De acordo com o que verificamos, isso acontece porque é comum que o valor calculado para o critério de informação do modelo que consideramos como verdadeiro (nos estudos de simulação) seja ligeiramente superior ao de um modelo que tem variáveis preditoras que estão no submodelo do parâmetro de dispersão, mas não estão no submodelo da média. Com a penalização anteriormente descrita, dificultamos que estes tipos de modelos sejam selecionados. Também observamos alguns casos em que adotar a penalização pode levar o método a não escolher o modelo verdadeiro mas, por outro lado, essa proporção é baixa em comparação com as seleções incorretas que são corrigidas com a mesma. Além disso, nos estudos de simulação desenvolvemos um cenário em que adicionamos uma variável no submodelo do parâmetro de dispersão que não estava

presente no submodelo da média com o objetivo de analisar se a penalização poderia prejudicar o desempenho do método, o qual discutiremos no próximo capítulo (Cenário 10 para as três distribuições que estudamos).

3.4 Método *Stepwise*

O método *Stepwise* (Kutner et al., 2004) é uma das heurísticas (algoritmo que encontra boas soluções a maioria das vezes, mas não tem garantias de que sempre as encontrará) mais utilizadas em análise de regressão. O método *Stepwise* para modelos lineares generalizados duplos que utilizamos neste trabalho está descrito em Stasinopoulos et al. (2017) e tem o seguinte algoritmo:

1. Assumindo que o parâmetro de dispersão é constante, construa um modelo para a média utilizando o método *Forward* (Kutner et al., 2004) de acordo com algum critério definido (menor AIC, por exemplo);
2. Dado o submodelo obtido para a média no passo anterior, construa um modelo para o parâmetro de dispersão usando o método *Forward*;
3. Dados os submodelos para a média e para o parâmetro de dispersão, verifica-se se os termos presentes no submodelo do parâmetro de dispersão são realmente necessários através do método *Backward* (Kutner et al., 2004);
4. Dados os submodelos para a média e para o parâmetro de dispersão, verifica-se se os termos presentes no submodelo da média são realmente necessários utilizando o método *Backward*.

Através da nossa revisão bibliográfica não encontramos nenhum trabalho que compara a performance desse método com o de outros métodos de seleção em MLGD.

Capítulo 4

Estudos de simulação

No capítulo anterior apresentamos os critérios de informação e os procedimentos para seleção automática de variáveis que usamos no decorrer do trabalho. Neste capítulo, discutimos os estudos de simulação desenvolvidos para comparar o procedimento proposto para a seleção automática de variáveis com os demais procedimentos abordados no Capítulo 3 e com o procedimento que consiste em ajustar todos os possíveis modelos.

Usamos o *software* R para o desenvolvimento dos estudos de simulação de Monte Carlo. Em especial, utilizamos o pacote *Generalized Additive Models for Location, Scale and Shape* (*GAMLSS*) deste *software* desenvolvido por [Stasinopoulos e Rigby \(2007\)](#) que, dentre outros modelos, possibilita o ajuste dos modelos lineares generalizados duplos. Este pacote foi elaborado baseando-se em [Rigby e Stasinopoulos \(2005\)](#). O algoritmo utilizado para a obtenção das estimativas de máxima verossimilhança dos parâmetros do modelo é denominado RS ([Rigby e Stasinopoulos, 2005](#)), o qual não necessita de valores iniciais precisos para μ e ϕ para obtenção da convergência. Além disso, este método é rápido para fazer o ajuste em grandes conjuntos de dados.

Não usamos códigos de simulação prontos. Com isso, desenvolvemos todos os códigos no decorrer do trabalho e usamos o *software* R para elaborá-los. Apesar do método *Stepwise* estar implementado no pacote *GAMLSS*, como o mesmo apresentou erros em algumas situações durante o processo de seleção de variáveis, escolhendo o modelo nulo para os dois submodelos em casos que este não deveria ser o modelo escolhido, programamos o algoritmo sem utilizar o presente no pacote.

Para o desenvolvimento do trabalho, com o objetivo de obtermos estimativas da proporção de vezes que cada procedimento identifica o modelo que definimos como verdadeiro, utilizamos simulação de Monte Carlo e desenvolvemos diferentes cenários alterando alguma característica do cenário que consideramos como base para cada uma das três distribuições que explicitamos na Seção 2.1.

Utilizando o método de Monte Carlo realizamos 3000 repetições do processo de geração da variável resposta para obter uma estimativa com baixa variabilidade da proporção de vezes que cada procedimento identifica o modelo verdadeiro. O desvio padrão máximo

nesse caso é 0,00913, considerando proporção igual a $0,5 \left(\sqrt{0,5 \times 0,5 / 3000} = 0,00913 \right)$. Para o método de seleção de um modelo dentre todos os possíveis, devido ao tempo computacional, utilizamos apenas 1000 réplicas. Além disso, usamos amostras com 50, 100, 200 e 400 observações, e fixamos uma semente inicial para realizarmos as repetições do processo de geração da variável resposta.

No nosso estudo usamos seis variáveis preditoras candidatas e as mantivemos fixas em todas as repetições de cada cenário. Desse modo, temos um total de 4.096 (2^{6+6}) modelos diferentes. Para o esquema de seleção em dois passos, ajustamos $2 \times 2^6 = 128$ modelos diferentes. Além disso, utilizamos em todos os cenários para as distribuições Gama, Gaussiana Inversa e Normal função de ligação logarítmica para ϕ .

Durante as próximas três seções discutimos sobre os resultados dos estudos de simulação para as distribuições Gama, Gaussiana Inversa e Normal, respectivamente, na Seção 4.4, realizamos um estudo especial do número de passos para o nosso esquema de seleção de variáveis em MLGD e, na Seção 4.5, conduzimos um estudo a respeito no número de passos em relação ao tempo necessário para selecionar um modelo candidato, comparando o nosso esquema com o método *Stepwise*.

4.1 Resultados obtidos para a distribuição Gama

O cenário que utilizamos como base para o estudo das metodologias e critérios de informação utilizando a distribuição Gama possui como parâmetros do modelo verdadeiro $\beta = (2, 1, -1, 1, -1, 0, 0)$ e $\gamma = (-1, 1, -1, 0, 0, 0, 0)$. Para cada observação obtemos o verdadeiro valor de μ_i e ϕ_i , em que $i = 1, 2, \dots, n$, e geramos uma observação que segue distribuição $GA(\mu_i; \phi_i)$, com o objetivo de construir a variável resposta ($Y(i)$). Neste cenário, $\mu \in (1,00; 54,59)$, $\phi \in (0,14; 1,00)$ e $\sqrt{Var(Y)} \in (0,14; 54,59)$. Além disso, as seis variáveis preditoras foram geradas independentemente com a distribuição $U(0; 1)$ (Uniforme(0; 1)) e mantidas constantes em todas as replicações. Também adotamos função de ligação logarítmica para μ .

Na Tabela 4.2 apresentamos os cenários utilizados para a distribuição Gama. Considerando que $\beta_5, \beta_6, \gamma_3, \gamma_4$ e γ_5 foram sempre mantidos iguais a zero e $h(\phi) = \log(\phi)$ em todos os cenários, não os explicitamos nesta tabela. Além disso, nos cenários 2 a 11 modificamos o cenário base (cenário 1) em relação a algum aspecto.

Para o Cenário 9, nas três distribuições estudadas, usamos os coeficientes de correlação linear de Pearson apresentados na Tabela 4.1 para gerar as variáveis preditoras com o auxílio do pacote *MultiRNG: Multivariate Pseudo-Random Number Generation* (Demirtas e Allozi, 2018) do *software* R. Desse modo, definimos dois grupos correlacionados, sendo o primeiro formado pelas variáveis preditoras 1, 5 e 6, e o segundo composto pelas covariáveis 2, 3 e 4. Organizamos os grupos para que, dentro de cada um, existam variáveis que estão presentes no modelo verdadeiro e covariáveis que não incluímos no mesmo.

Na Tabela 4.3 mostramos a porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para esta distribuição considerando o critério de informação com melhor desempenho. Detalhamos em negrito alguns casos em que o nosso esquema apresentou bom desempenho em comparação com pelo menos dois dos outros procedimentos.

Tabela 4.1: Coeficientes de correlação linear de Pearson utilizamos para gerar as variáveis preditoras nos Cenários 9.

Variáveis	V_1	V_2	V_3	V_4	V_5	V_6
V_1	1,0					
V_2	0,2	1,0				
V_3	0,2	0,7	1,0			
V_4	0,2	0,7	0,7	1,0		
V_5	0,8	0,2	0,2	0,2	1,0	
V_6	0,8	0,2	0,2	0,2	0,8	1,0

Os resultados obtidos para cada cenário são exibidos detalhadamente nas Tabelas A.1 a A.11, presentes no Apêndice A. Além das informações apresentadas na Tabela 4.3, nestas tabelas também mostramos os resultados para os demais critérios de informação, a porcentagem de seleção correta para o número de parâmetros ($\# = k_0$), e as porcentagens com menos parâmetros ($\# < k_0$) e com mais parâmetros ($\# > k_0$) que o modelo verdadeiro para os parâmetros β e γ .

Tabela 4.2: Cenários - distribuição Gama.

Cenário	D1, D3, D5	D2, D4, D6	β_0	β_1	β_2	β_3	β_4	γ_0	γ_1	γ_2	γ_6	$g(\mu)$	Cov. Cor.*
1	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	1	-1	0	$\log(\mu)$	Não
2	$U(0;1)$	$U(0;1)$	2	0,5	-0,5	0,5	-0,5	-1	1	-1	0	$\log(\mu)$	Não
3	$U(0;1)$	$U(0;1)$	2	2	-2	2	-2	-1	1	-1	0	$\log(\mu)$	Não
4	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	0,5	-0,5	0	$\log(\mu)$	Não
5	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	2	-2	0	$\log(\mu)$	Não
6	$U(0;1)$	$U(0;1)$	2	2	-0,5	1	-1	-1	2	-0,5	0	$\log(\mu)$	Não
7	$U(0;1)$	$U(0;1)$	2	1	-1	0	0	-1	1	0	0	$\log(\mu)$	Não
8	$N(0,5; 0,25^2)$	$GA(0,4; 1)$	2	1	-1	1	-1	-1	1	-1	0	$\log(\mu)$	Não
9	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	1	-1	0	$\log(\mu)$	Sim
10	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	1	-1	1	$\log(\mu)$	Não
11	$U(0;1)$	$U(0;1)$	5	1	-1	1	-1	-1	1	-1	0	$\sqrt{\mu}$	Não

*: Covariáveis correlacionadas.

Observando o cenário base (cenário 1), apresentado na Tabela 4.3, podemos verificar que o esquema de seleção de variáveis em até “ k ” passos apresentou ótima performance, a qual foi semelhante à do método *Stepwise*. Analisando os demais métodos, é possível concluir que a seleção dentre todos os modelos possíveis e o esquema de seleção de variáveis em dois passos apresentaram performances levemente inferiores aos dois métodos citados anteriormente.

Dentre os critérios de informação, observando a Tabela A.1, podemos afirmar que o BIC obteve, para amostras com 100 ou mais observações, melhores performances quando

comparado com os demais critérios. Considerando a média de modelos ajustados, o *Stepwise* necessitou do ajuste de um número baixo de modelos para encontrar um modelo candidato. Entretanto, o nosso método também não necessitou de um número elevado de modelos ajustados. Com $n = 50$, o maior número médio de modelos ajustados foi de 245, menos de 6% do total de modelos possíveis. É interessante ressaltar ainda que esse número médio de modelos decresce com o aumento do tamanho amostral.

Em relação ao tamanho amostral, conforme aumentamos o número de observações eleva-se, como esperado, a porcentagem de vezes que os procedimentos encontram o modelo verdadeiro. Além disso, para amostras de 400 observações, a porcentagem de acertos para o BIC é superior a 90% para os 4 procedimentos, valor consideravelmente maior do que os observados para os demais critérios de informação.

Utilizando a Tabela A.1, também verificamos que os tipos de erros tendem a se concentrarem em alguns casos mais específicos. Para amostras menores, os critérios AIC e HQC apresentam concentrações com maiores porcentagens de erro, em geral, para os casos em que superestima-se o número de parâmetros para γ , enquanto que, para o BIC, verificamos o oposto. Somente o AICc, para amostras com 50 observações, possui maior equilíbrio em relação a este tipo de erro. Para amostras maiores, nos 4 critérios a maior concentração de erros acontece quando estima-se corretamente o número de parâmetros presentes no modelo verdadeiro para β e superestima-se o número de parâmetros para γ . Também é interessante observarmos que o percentual de modelos candidatos que selecionam menos variáveis preditoras do que é necessário para o submodelo da média ($\#\beta < 5$) é pequeno com $n = 50$ e quase nulo para $n = 100$, enquanto que, para os casos em que acerta-se o número de covariáveis ($\#\beta = 5$), verificamos porcentagem de acerto superior a 50%.

Um fato contraintuitivo observado na Tabela A.1 é o melhor desempenho dos métodos *Stepwise* e do esquema de seleção de variáveis em até “ k ” passos em relação à seleção dentre todos os modelos. Através da porcentagem de erros, podemos concluir que este método, em geral, superestima com maior frequência a quantidade de parâmetros para os dois submodelos do que os outros dois métodos citados. Antes de introduzirmos a penalização discutida na Seção 3.3 para o e esquema de seleção de variáveis em até “ k ” passos, o mesmo possuía performance semelhante com a seleção dentre todos os modelos. Após a introdução da mesma, observamos uma melhora do método porque reduziu-se a porcentagens de vezes que o procedimento identificava incorretamente modelos com mais parâmetros para γ .

Quando reduzimos os valores dos parâmetros do modelo verdadeiro para β , no cenário 2, aumentamos a dificuldade de identificação dos mesmos. Podemos verificar na Tabela 4.3 que o esquema de seleção de variáveis em até “ k ” passos obteve um excelente desempenho comparando-se com os demais métodos. Ele apresentou performance igual ou superior ao método *Stepwise* e ao esquema de seleção de variáveis em dois passos, em todos os casos, com diferenças grandes para amostras com 50 e 100 observações. Por outro lado, também

é possível concluir que o nosso procedimento apresenta, em geral, média de modelos ajustados superior ao exibido para o cenário 1. Além disso, por meio da Tabela A.2 observamos que os métodos *Stepwise* e Dois passos subestimam bastante, para pequenas amostras, o número de parâmetros presentes no modelo verdadeiro para β .

Em oposição ao cenário 2, no cenário 3 aumentamos os valores dos parâmetros do modelo verdadeiro para β , reduzindo a dificuldade para sua identificação. Neste caso, é possível afirmar, utilizando a Tabela 4.3 e a Tabela A.3, que os resultados obtidos são semelhantes aos encontrados para o cenário 1.

No cenário 4 reduzimos os valores dos parâmetros do modelo verdadeiro para γ . Desse modo, diminuímos a variabilidade do valor do parâmetro de dispersão entre as observações. Usando a Tabela 4.3 podemos afirmar que, em comparação com o cenário 1 e considerando principalmente amostras menores, a porcentagem de acertos do modelo verdadeiro diminuiu. Por outro lado, as conclusões em relação às metodologias mantêm-se as mesmas, com o método *Stepwise* e o esquema de seleção de variáveis em até “ k ” passos apresentando desempenhos semelhantes e, em geral, superiores ao dos demais métodos. Também é interessante verificarmos, utilizando a Tabela A.4, que neste cenário os métodos tendem a selecionar menos variáveis preditoras para o submodelo da dispersão do que realmente é necessário, ou seja, temos aumento da porcentagem para $\#\gamma < 3$.

Para o cenário 5 aumentamos os valores dos parâmetros do modelo verdadeiro para γ , ou seja, aumentamos a variabilidade do valor do parâmetro de dispersão entre as observações. Neste cenário, podemos verificar na Tabela 4.3 que, principalmente para pequenas amostras, o esquema de seleção de variáveis em até “ k ” passos obteve, em geral, desempenho consideravelmente superior ao dos métodos *Stepwise* e Dois passos. Considerando amostras com 50 observações e o AICc como critério de informação, nosso método encontrou o modelo verdadeiro em 46,5% das réplicas de Monte Carlo, enquanto que o método *Stepwise* apenas obteve esse modelo como candidato em 22,4% das réplicas e o método de seleção de variáveis em dois passos em 19,6% das réplicas. Além disso, diferentemente do que visualizamos no cenário anterior, usando a Tabela A.6 verificamos o comportamento oposto dos métodos de seleção de variáveis e, conseqüentemente, o aumento da porcentagem para $\#\gamma > 3$.

No cenário 6, mesclamos os valores dos parâmetros do modelo verdadeiro usando as ideias dos 4 cenários anteriores. Os resultados mostraram que, nos casos com valores dos parâmetros relativamente diferentes (em módulo), os procedimentos possuem mais dificuldade em identificar o modelo verdadeiro. Entretanto, o esquema de seleção em dois passos e o método *Stepwise* foram os métodos com os piores desempenhos.

Para o Cenário 7 reduzimos o número de parâmetros diferentes de zero no modelo verdadeiro. Neste caso, podemos observar na Tabela 4.3 que o esquema de seleção de variáveis em até “ k ” passos apresentou a melhor performance para todos os tamanhos amostrais, sendo a mesma bem superior para amostras com até 100 observações. Por outro lado, no cenário 8 alteramos as distribuições das variáveis preditoras, utilizando

para x_1, x_3 e x_5 distribuição $N(0,5; 0,25^2)$ e, para x_2, x_4 e x_6 , distribuição $GA(0,4; 1)$. Neste caso não verificamos grande diferença nas performances dos métodos para seleção de variáveis em comparação com o cenário base.

Com a adição de diferentes correlações entre as variáveis preditoras (Tabela 4.1), no cenário 9, conseguimos avaliar o desempenho do nosso procedimento para seleção de variáveis em dados que apresentam comportamento mais próximo ao observado na prática. Os resultados presentes na Tabela 4.3 indicam que o esquema de seleção de variáveis em até “ k ” passos obteve um ótimo desempenho, superior à das demais metodologias estudadas para todos os tamanhos amostrais considerando o critério de informação com melhor desempenho. Este cenário também nos mostra que o método *Stepwise* possui a pior performance para amostra menores (até 100 observações), evidenciando que, nestes casos, o mesmo pode não ser uma boa alternativa para utilizarmos em bancos de dados com correlações consideráveis entre as variáveis preditoras.

No cenário 10, adicionamos uma variável no submodelo do parâmetro de dispersão que não estava presente no submodelo da média. O objetivo deste cenário foi analisar se o peso, que atribuímos para as variáveis não selecionadas no submodelo da média, poderia prejudicar o desempenho do método. Os resultados, presentes na Tabela 4.3, indicam que houve uma perda no desempenho em relação à seleção de um modelo dentre todos os possíveis para amostras de até 100 observações mas, em comparação com os demais métodos, nosso esquema teve performance relativamente superior. Desse modo, podemos concluir que a adição do peso melhora o nosso procedimento ao evitar que variáveis não importantes sejam adicionadas no submodelo do parâmetro de dispersão, ao mesmo tempo que não o prejudica a ponto do mesmo apresentar desempenho inferior ao dos métodos *Stepwise* e Dois passos.

Para o cenário 11, alteramos a função de ligação do submodelo da média para a raiz quadrada. Neste cenário, buscamos aproximar o alcance dos parâmetros do cenário 1 e, assim, $\mu \in (9; 49)$, $\phi \in (0,14; 1,00)$ e $\sqrt{Var(Y)} \in (1,22; 49,00)$. Os resultados para este cenário, exibidos na Tabela 4.3, mostram a maior porcentagem de acertos do modelo definido como verdadeiro para o nosso esquema, principalmente para amostras menores, em comparação com os demais procedimentos.

Analisando os resultados obtidos para os 11 cenários com a distribuição Gama podemos afirmar que, dentre as metodologias, o esquema para seleção de variáveis em até “ k ” passos demonstrou ser o método mais estável, possuindo melhor desempenho em comparação com os métodos *Stepwise* e Dois passos, principalmente para amostras menores. Na comparação com a seleção de um modelo dentre todos os possíveis, o nosso esquema teve um desempenho consideravelmente inferior apenas nos cenários 6 e 10, e substancialmente superior nos cenários 7 e 9. Em relação aos critérios de informação, o BIC foi, em geral, o que apresentou melhores resultados, fato que pode ser justificado porque o modelo que gerou os dados está presente entre os modelos analisados e o BIC considera que o melhor modelo para os dados pode ser encontrado utilizando as covariáveis disponíveis,

conforme discutimos na Subseção 3.1.2.

Tabela 4.3: Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Gama considerando o critério de informação com melhor desempenho.

n	Método	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos
n = 50	Todos modelos	BIC	29,0	4.096	AICc	11,6	4.096	BIC	29,3	4.096	AICc	6,5	4.096
	Stepwise	AICc	30,9	43	AIC	5,8	40	BIC	35,8	41	AICc	7,7	39
	K passos	AICc	32,3	222	AICc	12,1	262	BIC	35,7	211	AICc	8,2	211
	Dois passos	BIC	28,2	128	AICc	6,0	128	BIC	33,6	128	AICc	7,1	128
n = 100	Todos modelos	BIC	69,8	4.096	BIC	59,0	4.096	BIC	69,8	4.096	HQC	16,2	4.096
	Stepwise	BIC	73,4	42	HQC	37,9	42	BIC	73,5	42	HQC	18,6	40
	K passos	BIC	74,6	202	BIC	55,4	241	BIC	74,7	202	HQC	20,5	203
	Dois passos	BIC	69,9	128	HQC	35,1	128	BIC	70,0	128	HQC	17,5	128
n = 200	Todos modelos	BIC	84,9	4.096	BIC	84,9	4.096	BIC	84,9	4.096	HQC	47,0	4.096
	Stepwise	BIC	89,1	43	BIC	84,6	42	BIC	89,1	43	HQC	48,5	42
	K passos	BIC	90,2	196	BIC	90,0	199	BIC	90,2	196	HQC	51,7	208
	Dois passos	BIC	87,1	128	BIC	82,8	128	BIC	87,1	128	HQC	46,3	128
n = 400	Todos modelos	BIC	92,0	4.096	BIC	92,0	4.096	BIC	92,0	4.096	BIC	84,3	4.096
	Stepwise	BIC	93,8	42	BIC	93,8	42	BIC	93,8	42	BIC	83,9	42
	K passos	BIC	93,8	195	BIC	93,8	195	BIC	93,8	195	BIC	85,2	194
	Dois passos	BIC	92,2	128	BIC	92,2	128	BIC	92,2	128	BIC	82,7	128
n	Método	Cenário 5			Cenário 6			Cenário 7			Cenário 8		
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos
n = 50	Todos modelos	BIC	56,4	4.096	AIC	5,3	4.096	BIC	33,9	4.096	AICc	19,7	4.096
	Stepwise	AICc	22,4	40	AIC	1,1	36	BIC	28,8	29	AICc	18,0	40
	K passos	AICc	46,5	275	AIC	4,9	307	BIC	39,7	223	AICc	21,8	234
	Dois passos	AICc	19,6	128	AIC	1,0	128	BIC	27,9	128	AICc	18,0	128
n = 100	Todos modelos	BIC	77,0	4.096	HQC	22,0	4.096	BIC	68,4	4.096	BIC	63,0	4.096
	Stepwise	BIC	57,5	41	AIC	7,5	42	BIC	69,8	31	BIC	66,9	42
	K passos	BIC	79,9	227	AIC	19,0	275	BIC	77,3	211	BIC	67,5	207
	Dois passos	BIC	56,1	128	AIC	6,6	128	BIC	65,7	128	BIC	64,6	128
n = 200	Todos modelos	BIC	85,2	4.096	BIC	63,7	4.096	BIC	81,2	4.096	BIC	83,0	4.096
	Stepwise	BIC	89,8	43	HQC	31,1	41	BIC	85,1	31	BIC	86,9	43
	K passos	BIC	89,9	197	HQC	53,6	243	BIC	88,4	198	BIC	87,9	200
	Dois passos	BIC	88,2	128	HQC	28,9	128	BIC	81,7	128	BIC	81,8	128
n = 400	Todos modelos	BIC	91,7	4.096	BIC	89,4	4.096	BIC	87,9	4.096	BIC	88,4	4.096
	Stepwise	BIC	94,3	42	HQC	60,4	43	BIC	90,8	31	BIC	92,6	42
	K passos	BIC	93,1	195	BIC	85,4	225	BIC	92,4	196	BIC	92,7	195
	Dois passos	BIC	92,9	128	HQC	56,1	128	BIC	88,1	128	BIC	90,7	128
n	Método	Cenário 9			Cenário 10			Cenário 11					
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos			
n = 50	Todos modelos	BIC	19,6	4.096	AICc	28,1	4.096	HQC	5,5	4.096			
	Stepwise	AIC	12,8	41	AIC	9,9	43	AIC	2,6	37			
	K passos	AICc	24,0	256	AIC	15,8	296	AIC	7,0	286			
	Dois passos	AICc	15,7	128	AIC	9,0	128	AIC	2,7	128			
n = 100	Todos modelos	BIC	57,1	4.096	BIC	73,8	4.096	BIC	42,5	4.096			
	Stepwise	BIC	52,6	42	HQC	47,4	46	AICc	22,3	43			
	K passos	BIC	66,3	210	HQC	56,5	239	HQC	43,8	256			
	Dois passos	BIC	55,7	128	HQC	43,2	128	AICc	20,7	128			
n = 200	Todos modelos	BIC	81,4	4.096	BIC	88,3	4.096	BIC	83,5	4.096			
	Stepwise	BIC	85,9	42	BIC	90,5	46	BIC	66,4	41			
	K passos	BIC	87,5	198	BIC	89,6	198	BIC	88,3	213			
	Dois passos	BIC	84,3	128	BIC	88,3	128	BIC	65,7	128			
n = 400	Todos modelos	BIC	89,4	4.096	BIC	92,6	4.096	BIC	92,0	4.096			
	Stepwise	BIC	92,0	43	BIC	95,4	46	BIC	93,3	42			
	K passos	BIC	92,9	205	BIC	93,8	195	BIC	93,7	195			
	Dois passos	BIC	90,1	128	BIC	93,2	128	BIC	92,3	128			

4.2 Resultados obtidos para a distribuição Gaussiana Inversa

Para estudar as metodologias e critérios de informação utilizando a distribuição Gaussiana Inversa escolhemos como parâmetros do modelo verdadeiro para o cenário de base $\beta = (2, 1, -1, 1, -1, 0, 0)$ e $\gamma = (-1, 8, 1, -1, 0, 0, 0, 0)$. Para cada observação obtemos o verdadeiro valor de μ_i e ϕ_i , em que $i = 1, 2, \dots, n$, e geramos uma observação que segue distribuição $GI(\mu_i; \phi_i)$, com o objetivo de construir a variável resposta ($Y(i)$). Neste cenário, $\mu \in (1, 00; 54, 59)$, $\phi \in (0, 06; 0, 45)$ e $\sqrt{Var(Y)} \in (0, 06; 181, 50)$.

Na Tabela 4.4 mostramos os cenários utilizados para a distribuição Gaussiana Inversa e, considerando que $\beta_5, \beta_6, \gamma_3, \gamma_4$ e γ_5 foram sempre mantidos iguais a zero e, $g(\mu) = \log(\mu)$ e $h(\phi) = \log(\phi)$ em todos os cenários, não os explicitamos nesta tabela. Na Tabela 4.5 apresentamos a porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para esta distribuição considerando o critério de informação com melhor desempenho. Também detalhamos em negrito alguns casos em que o nosso esquema apresentou bom desempenho em comparação com pelo menos dois dos outros procedimentos. Os resultados obtidos para cada cenário são detalhados nas Tabelas B.1 a B.10, presentes no Apêndice B.

Tabela 4.4: Cenários - distribuição Gaussiana Inversa.

Cenário	D1, D3, D5	D2, D4, D6	β_0	β_1	β_2	β_3	β_4	γ_0	γ_1	γ_2	γ_6	Cov. Cor.*
1	$U(0; 1)$	$U(0; 1)$	2	1	-1	1	-1	-1,8	1	-1	0	Não
2	$U(0; 1)$	$U(0; 1)$	2	0,5	-0,5	0,5	-0,5	-1,8	1	-1	0	Não
3	$U(0; 1)$	$U(0; 1)$	2	2	-2	2	-2	-1,8	1	-1	0	Não
4	$U(0; 1)$	$U(0; 1)$	2	1	-1	1	-1	-1,8	0,5	-0,5	0	Não
5	$U(0; 1)$	$U(0; 1)$	2	1	-1	1	-1	-1,8	2	-2	0	Não
6	$U(0; 1)$	$U(0; 1)$	2	2	-0,5	1	-1	-1,8	2	-0,5	0	Não
7	$U(0; 1)$	$U(0; 1)$	2	1	-1	0	0	-1,8	1	0	0	Não
8	$N(0,5; 0,25^2)$	$GA(0,4; 1)$	2	1	-1	1	-1	-1,8	1	-1	0	Não
9	$U(0; 1)$	$U(0; 1)$	2	1	-1	1	-1	-1,8	1	-1	0	Sim
10	$U(0; 1)$	$U(0; 1)$	2	1	-1	1	-1	-1,8	1	-1	1	Não

*: Covariáveis correlacionadas.

Os cenários construídos para as distribuições Gaussiana Inversa e Normal (discutidos na próxima seção) seguiram a mesma lógica explicada para a distribuição Gama na seção anterior. Entretanto, para a distribuição Gaussiana Inversa não conseguimos montar o cenário 11, pois alguns problemas com a convergência do algoritmo numérico utilizado na estimação dos parâmetros pelo pacote que usamos não permitiram que alterássemos a função de ligação do submodelo da média.

Considerando a Tabela 4.5, podemos afirmar que as conclusões gerais obtidas para a distribuição Gaussiana Inversa são semelhantes às discutidas utilizando a distribuição Gama.

Tabela 4.5: Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Gaussiana Inversa considerando o critério de informação com melhor desempenho.

n	Método	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos
n = 50	Todos modelos	BIC	27,0	4.096	AICc	5,4	4.096	BIC	30,1	4.096	AICc	6,1	4.096
	Stepwise	AICc	24,3	40	AIC	1,9	36	BIC	36,9	40	AICc	7,5	39
	K passos	AICc	26,8	234	AIC	5,9	281	BIC	36,6	208	AICc	7,4	221
	Dois passos	BIC	24,0	128	AIC	2,1	128	BIC	38,2	128	AICc	7,3	128
n = 100	Todos modelos	BIC	68,2	4.096	BIC	43,0	4.096	BIC	68,1	4.096	AICc	15,9	4.096
	Stepwise	BIC	73,1	42	AICc	22,4	43	BIC	75,0	42	AICc	19,9	42
	K passos	BIC	73,0	202	HQC	42,4	257	BIC	73,2	201	AICc	20,5	216
	Dois passos	BIC	71,9	128	AICc	21,3	128	BIC	74,7	128	AICc	18,6	128
n = 200	Todos modelos	BIC	84,8	4.096	BIC	83,6	4.096	BIC	84,5	4.096	HQC	46,5	4.096
	Stepwise	BIC	89,3	42	BIC	67,4	41	BIC	89,5	42	HQC	49,0	42
	K passos	BIC	89,7	197	BIC	87,8	211	BIC	89,8	195	HQC	51,7	205
	Dois passos	BIC	88,5	128	BIC	67,4	128	BIC	89,3	128	HQC	48,1	128
n = 400	Todos modelos	BIC	92,0	4.096	BIC	91,9	4.096	BIC	91,2	4.096	BIC	83,8	4.096
	Stepwise	BIC	94,0	42	BIC	93,4	42	BIC	94,0	42	BIC	83,1	42
	K passos	BIC	93,2	194	BIC	93,4	195	BIC	93,0	194	BIC	83,5	193
	Dois passos	BIC	93,7	128	BIC	92,5	128	BIC	93,9	128	BIC	82,8	128
n	Método	Cenário 5			Cenário 6			Cenário 7					
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos			
n = 50	Todos modelos	BIC	53,6	4.096	AIC	1,1	4.096	BIC	21,7	4.096			
	Stepwise	AIC	7,6	38	AIC	0,0	31	AICc	14,6	31			
	K passos	AICc	34,8	300	AIC	0,9	305	BIC	22,2	223			
	Dois passos	AICc	9,9	128	AIC	0,1	128	AICc	14,0	128			
n = 100	Todos modelos	BIC	74,3	4.096	AICc	13,7	4.096	BIC	61,8	4.096			
	Stepwise	HQC	48,8	43	AIC	0,7	36	BIC	56,1	30			
	K passos	BIC	70,3	239	AICc	10,6	304	BIC	69,5	212			
	Dois passos	HQC	49,8	128	AIC	0,6	128	BIC	55,7	128			
n = 200	Todos modelos	BIC	84,7	4.096	BIC	45,3	4.096	BIC	80,6	4.096			
	Stepwise	BIC	87,9	42	AIC	11,8	41	BIC	84,9	31			
	K passos	BIC	89,7	198	AICc	33,5	279	BIC	87,8	199			
	Dois passos	BIC	88,5	128	AICc	11,6	128	BIC	83,0	128			
n = 400	Todos modelos	BIC	91,0	4.096	BIC	82,2	4.096	BIC	88,0	4.096			
	Stepwise	BIC	94,3	42	AICc	33,1	43	BIC	91,5	31			
	K passos	BIC	93,1	195	HQC	69,5	245	BIC	92,0	196			
	Dois passos	BIC	94,2	128	HQC	32,1	128	BIC	90,0	128			
n	Método	Cenário 8			Cenário 9			Cenário 10					
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos			
n = 50	Todos modelos	AICc	9,0	4.096	BIC	17,5	4.096	AICc	14,8	4.096			
	Stepwise	AIC	6,1	40	AIC	8,5	40	AIC	2,1	38			
	K passos	AIC	9,2	255	AICc	18,4	271	AIC	7,3	308			
	Dois passos	AICc	7,0	128	AICc	9,6	128	AIC	3,0	128			
n = 100	Todos modelos	BIC	60,7	4.096	BIC	54,4	4.096	BIC	64,9	4.096			
	Stepwise	BIC	55,6	41	HQC	44,3	43	AICc	35,6	46			
	K passos	BIC	57,3	216	BIC	57,8	226	HQC	43,0	263			
	Dois passos	BIC	55,7	128	BIC	46,6	128	AICc	34,1	128			
n = 200	Todos modelos	BIC	82,4	4.096	BIC	81,4	4.096	BIC	87,9	4.096			
	Stepwise	BIC	86,8	43	BIC	85,2	42	BIC	82,9	46			
	K passos	BIC	85,8	198	BIC	87,4	197	BIC	86,8	205			
	Dois passos	BIC	84,4	128	BIC	85,3	128	BIC	83,9	128			
n = 400	Todos modelos	BIC	89,4	4.096	BIC	89,4	4.096	BIC	92,5	4.096			
	Stepwise	BIC	92,8	42	BIC	92,4	42	BIC	95,2	46			
	K passos	BIC	92,4	195	BIC	93,3	194	BIC	93,5	195			
	Dois passos	BIC	92,3	128	BIC	91,9	128	BIC	94,5	128			

4.3 Resultados obtidos para a distribuição Normal

O cenário que utilizamos como base para o estudo das metodologias e critérios de informação utilizando a distribuição Normal possui como parâmetros do modelo verdadeiro $\beta = (30, 4, -4, 4, -4, 0, 0)$ e $\gamma = (1, 1, -1, 0, 0, 0, 0)$. Para cada observação obtemos o verdadeiro valor de μ_i e ϕ_i , em que $i = 1, 2, \dots, n$, e geramos uma observação que segue distribuição $N(\mu_i; \phi_i)$, com o objetivo de construir a variável resposta ($Y(i)$). Neste cenário, $\mu \in (22,00; 38,00)$ e $\sqrt{Var(Y)} = \phi \in (1,00; 7,39)$.

Na Tabela 4.6 exibimos os cenários utilizados para a distribuição Normal. Considerando que $\beta_5, \beta_6, \gamma_3, \gamma_4$ e γ_5 foram sempre mantidos iguais a zero e, $h(\phi) = \log(\phi)$ em todos os cenários, não os mostramos nesta tabela. Detalhamos cada cenário na Tabela 4.7, em que mostramos a porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para esta distribuição considerando o critério de informação com melhor desempenho. Nesta tabela também detalhamos em negrito alguns casos em que o nosso procedimento apresentou bom desempenho em comparação com pelo menos dois dos outros métodos. Mostramos os resultados obtidos para cada cenário detalhadamente nas Tabelas C.1 a C.11, presentes no Apêndice C.

Tabela 4.6: Cenários - distribuição Normal.

Cenário	D1, D3, D5	D2, D4, D6	β_0	β_1	β_2	β_3	β_4	γ_0	γ_1	γ_2	γ_6	$g(\mu)$	Cov. Cor.*
1	$U(0;1)$	$U(0;1)$	30	4	-4	4	-4	1	1	-1	0	μ	Não
2	$U(0;1)$	$U(0;1)$	30	2	-2	2	-2	1	1	-1	0	μ	Não
3	$U(0;1)$	$U(0;1)$	30	6	-6	6	-6	1	1	-1	0	μ	Não
4	$U(0;1)$	$U(0;1)$	30	4	-4	4	-4	1	0,5	-0,5	0	μ	Não
5	$U(0;1)$	$U(0;1)$	30	4	-4	4	-4	1	2	-2	0	μ	Não
6	$U(0;1)$	$U(0;1)$	30	6	-2	4	-4	1	2	-0,5	0	μ	Não
7	$U(0;1)$	$U(0;1)$	30	4	-4	0	0	1	1	0	0	μ	Não
8	$N(0,5; 0,25^2)$	$GA(0,4; 1)$	30	4	-4	4	-4	1	1	-1	0	μ	Não
9	$U(0;1)$	$U(0;1)$	30	4	-4	4	-4	1	1	-1	0	μ	Sim
10	$U(0;1)$	$U(0;1)$	30	4	-4	4	-4	1	1	-1	1	μ	Não
11	$U(0;1)$	$U(0;1)$	2	1	-1	1	-1	-1	1	-1	0	$\log(\mu)$	Não

*: Covariáveis correlacionadas.

Utilizando a Tabela 4.7, podemos verificar que as conclusões gerais obtidas para a distribuição Normal são semelhantes com as discutidas anteriormente para as distribuições Gama e Gaussiana Inversa. Entretanto, para a distribuição Normal podemos observar, em alguns cenários (como o 5) e principalmente para amostras menores, um desempenho ainda melhor do esquema de seleção de variáveis em até “ k ” passos em comparação com o método *Stepwise* e com o esquema de seleção de variáveis em dois passos. Além disso, nosso procedimento tem desempenho próximo ou superior ao da seleção de um modelo através do ajuste de todos os possíveis, com tempo computacional bastante inferior.

No cenário 11, em que alteramos a função de ligação do submodelo da média para a logarítmica, com exceção do método Dois passos que apresentou desempenho inferior,

não observamos grandes diferenças entre a performance dos demais métodos através dos resultados presentes na Tabela 4.7. Entretanto, o nosso procedimento para seleção de variáveis apresentou performance levemente superior em comparação com os demais para todos os tamanhos amostrais considerando o critério de informação com melhor desempenho. Neste cenário, $\mu \in (1,00; 54,59)$ e $\sqrt{Var(Y)} = \phi \in (0,14; 1,00)$.

Tabela 4.7: Porcentagem de seleção do modelo verdadeiro e média do número de modelos ajustados para cada cenário, tamanho amostral e método estudado para a distribuição Normal considerando o critério de informação com melhor desempenho.

n	Método	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos
n = 50	Todos modelos	AICc	13,6	4.096	AIC	1,0	4.096	BIC	26,5	4.096	AIC	3,2	4.096
	Stepwise	AICc	7,9	38	AIC	0,4	35	AICc	24,3	42	AIC	2,9	40
	K passos	AICc	13,0	269	AIC	1,8	296	AICc	26,4	234	AIC	4,1	235
	Dois passos	AICc	8,3	128	AIC	0,8	128	AICc	20,5	128	AIC	2,7	128
n = 100	Todos modelos	BIC	62,7	4.096	AIC	9,6	4.096	BIC	68,2	4.096	AICc	16,0	4.096
	Stepwise	HQC	44,5	43	AIC	3,6	41	BIC	71,1	42	AICc	17,0	42
	K passos	BIC	60,7	230	AIC	13,7	275	BIC	72,2	205	AICc	19,3	218
	Dois passos	BIC	41,8	128	AICc	4,1	128	BIC	68,1	128	AICc	15,4	128
n = 200	Todos modelos	BIC	88,0	4.096	HQC	45,4	4.096	BIC	88,0	4.096	BIC	42,4	4.096
	Stepwise	BIC	87,8	43	AICc	24,6	44	BIC	89,1	43	HQC	45,8	42
	K passos	BIC	90,0	197	HQC	54,0	241	BIC	90,0	196	HQC	48,4	203
	Dois passos	BIC	85,2	128	HQC	22,7	128	BIC	86,5	128	HQC	43,7	128
n = 400	Todos modelos	BIC	92,4	4.096	BIC	88,8	4.096	BIC	92,4	4.096	BIC	82,2	4.096
	Stepwise	BIC	93,9	42	HQC	62,5	43	BIC	93,9	42	BIC	83,0	42
	K passos	BIC	94,7	195	BIC	90,8	217	BIC	94,7	195	BIC	84,3	195
	Dois passos	BIC	91,4	128	HQC	58,0	128	BIC	91,4	128	BIC	81,9	128
n	Método	Cenário 5			Cenário 6			Cenário 7			Cenário 8		
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos
n = 50	Todos modelos	BIC	34,2	4.096	AIC	1,4	4.096	AIC	7,6	4.096	AICc	6,0	4.096
	Stepwise	AICc	6,4	36	HQC	0,2	29	AICc	6,1	30	AIC	4,2	40
	K passos	AICc	38,4	292	AIC	0,9	305	AICc	11,5	270	AIC	6,9	262
	Dois passos	AICc	5,7	128	AIC	0,4	128	AICc	5,6	128	AIC	4,5	128
n = 100	Todos modelos	BIC	74,6	4.096	AICc	6,0	4.096	BIC	31,1	4.096	BIC	58,6	4.096
	Stepwise	AICc	16,6	42	AIC	0,8	38	HQC	23,6	31	HQC	32,2	42
	K passos	BIC	78,9	266	AIC	7,3	312	HQC	36,0	244	BIC	49,0	245
	Dois passos	AICc	12,8	128	AIC	0,9	128	HQC	22,0	128	HQC	29,4	128
n = 200	Todos modelos	BIC	87,4	4.096	HQC	26,5	4.096	BIC	71,9	4.096	BIC	82,9	4.096
	Stepwise	HQC	55,8	43	AIC	8,4	42	BIC	60,8	29	BIC	75,3	42
	K passos	BIC	90,2	223	AICc	25,9	275	BIC	75,9	209	BIC	86,5	208
	Dois passos	BIC	51,0	128	AICc	7,7	128	BIC	59,6	128	BIC	71,4	128
n = 400	Todos modelos	BIC	92,2	4.096	BIC	59,2	4.096	BIC	86,0	4.096	BIC	91,1	4.096
	Stepwise	BIC	89,1	42	AICc	22,4	44	BIC	88,9	31	BIC	93,1	42
	K passos	BIC	94,9	200	HQC	60,4	243	BIC	91,6	197	BIC	93,9	197
	Dois passos	BIC	86,2	128	HQC	19,2	128	BIC	85,6	128	BIC	91,2	128
n	Método	Cenário 9			Cenário 10			Cenário 11					
		Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos	Critério	Modelo Correto(%)	Média de Modelos			
n = 50	Todos modelos	AICc	5,7	4.096	AICc	3,2	4.096	BIC	23,5	4.096			
	Stepwise	AIC	1,2	37	AIC	1,0	39	BIC	24,4	40			
	K passos	AIC	7,4	292	AIC	3,7	299	BIC	25,9	209			
	Dois passos	AIC	1,9	128	AIC	1,4	128	BIC	22,1	128			
n = 100	Todos modelos	HQC	32,5	4.096	HQC	30,5	4.096	BIC	66,5	4.096			
	Stepwise	AICc	16,6	42	AIC	7,1	45	BIC	66,8	43			
	K passos	HQC	36,1	271	AICc	30,3	284	BIC	69,9	206			
	Dois passos	AICc	18,0	128	AIC	6,6	128	BIC	56,3	128			
n = 200	Todos modelos	BIC	75,8	4.096	BIC	76,8	4.096	BIC	86,7	4.096			
	Stepwise	HQC	55,4	42	HQC	38,9	45	BIC	86,8	43			
	K passos	BIC	80,0	220	BIC	75,3	244	BIC	89,5	201			
	Dois passos	HQC	55,0	128	HQC	35,7	128	BIC	77,1	128			
n = 400	Todos modelos	BIC	91,8	4.096	BIC	93,3	4.096	BIC	91,5	4.096			
	Stepwise	BIC	81,4	42	BIC	75,5	46	BIC	92,8	43			
	K passos	BIC	93,6	199	BIC	94,6	208	BIC	94,1	200			
	Dois passos	BIC	87,1	128	BIC	73,3	128	BIC	83,7	128			

4.4 Análise do número de passos para o esquema de seleção de variáveis em até “ k ” passos nos estudos de simulação

Quando programamos o procedimento para seleção de variáveis em até “ k ” passos, delimitamos o número máximo de passos com o objetivo de evitar a possibilidade de que a execução do algoritmo não finalizasse durante os estudos de simulação. Desse modo, utilizamos como número máximo de passos possíveis 64, ou seja, o método pode ajustar até 4.096 modelos, valor igual ao número máximo de modelos possíveis.

Para a grande maioria dos critérios de informação e tamanhos amostrais estudados, em mais de 95% das repetições foram necessários 6 ou menos passos para o esquema de seleção de variáveis em até k passos encontrar um modelo candidato. Com isso, ajustam-se nesses casos até 384 modelos (apenas 9,4% do total) e, assim, além das vantagens em desempenho também conseguimos um grande ganho em relação ao tempo computacional comparando-se com a seleção de um modelo através do ajuste de todos os possíveis.

Na Figura 4.1 apresentamos graficamente o número de passos obtidos no estudo de simulação para o cenário 5 com a distribuição Normal, considerando amostras com 50 observações e o AIC como critério de informação. Em geral, o comportamento observado para os demais casos foi semelhante a este, em que podemos observar as maiores concentrações entre 3 e 6 passos.

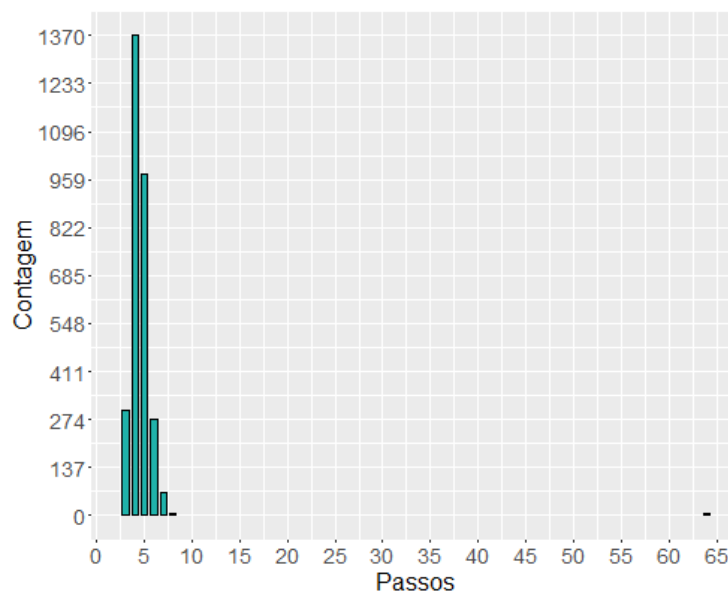


Figura 4.1: Gráfico de barras para o número de passos obtidos no estudo de simulação do cenário 5 para a distribuição Normal, com tamanho amostral $n = 50$ e o AIC como critério de informação.

Observando a Figura 4.1 podemos verificar também que uma pequena parcela das réplicas de Monte Carlo não convergiu. Esse comportamento foi observado na maioria dos cenários, independentemente da distribuição, tamanho amostral e critério de informação. Entretanto, a porcentagem de não convergência foi inferior a 1,4% do total de réplicas em todos os casos. Desse modo, não consideramos este comportamento como um fator

que possa comprometer o desempenho do método. Além disso, estudando uma das amostras em que a seleção utilizando o nosso procedimento não converge, verificamos que o modelo candidato permanece alternando-se sempre entre os mesmos dois modelos após um curto número de passos. Em especial para esta amostra, um desses dois modelos candidatos era o modelo que definimos como verdadeiro.

4.5 Comparação do número de modelos ajustados \times tempo entre o esquema de seleção de variáveis em até “ k ” passos e o método *Stepwise*

Com o objetivo de estudarmos a viabilidade do nosso método na prática, realizamos um breve estudo de simulação comparando o esquema de seleção de variáveis em até “ k ” passos e o método *Stepwise* em relação ao tempo e ao número de modelos necessários para a obtenção de um modelo candidato. Para isso, simulamos um único conjunto de dados utilizando a distribuição Gama para diferentes combinações de números de parâmetros (p) e tamanhos amostrais (n). Para construirmos o modelo verdadeiro, seguimos uma ideia semelhante à apresentada para o cenário base utilizando esta distribuição. Usamos os mesmos valores para os interceptos e alteramos as variáveis preditoras entre 1 e -1 até $\frac{p}{2}$ para o submodelo da média, e até $\frac{p}{2} - 1$, para o submodelo do parâmetro de dispersão. As demais covariáveis foram mantidas iguais a zero. Na Tabela 4.8 apresentamos os resultados para o nosso procedimento e, na Tabela 4.9, os resultados para o método *Stepwise*.

Utilizando a Tabela 4.8, podemos observar que, no esquema de seleção de variáveis em até “ k ” passos, o tempo computacional necessário para que um modelo candidato seja encontrado cresce rapidamente com o aumento do número de variáveis presentes no conjunto de dados, pois é necessário um número maior de modelos a serem ajustados em cada passo. Por outro lado, o aumento no tamanho amostral também produz um crescimento no tempo necessário, mas o mesmo é significativamente inferior em comparação com o aumento no número de covariáveis. Além disso, também é interessante observar que o número de passos não é superior a quatro na grande maioria dos casos.

Tabela 4.8: Tempo, em minutos (número de passos - total de modelos ajustados), para encontrar um modelo candidato utilizando dados simulados com a distribuição Gama, o esquema de seleção de variáveis em até “ k ” passos e diferentes combinações de tamanhos amostrais (n) e números de parâmetros (p).

$p \backslash n$	50	100	200	500	1.000	5.000	10.000
5	0,05 (3 - 96)	0,05 (3 - 96)	0,07 (3 - 96)	0,12 (3 - 96)	0,22 (3 - 96)	1,26 (3 - 96)	2,58 (3 - 96)
8	0,38 (3 - 768)	0,44 (3 - 768)	0,76 (4 - 1.024)	1,12 (4 - 1.024)	1,53 (3 - 768)	6,67 (3 - 768)	14,96 (3 - 768)
10	1,48 (3 - 3.072)	2,03 (3 - 3.072)	3,91 (4 - 4.096)	5,95 (4 - 4.096)	11,07 (3 - 3.072)	45,53 (3 - 3.072)	109,04 (3 - 3.072)
12	5,78 (3 - 12.288)	14,56 (5 - 20.480)	15,24 (4 - 16.384)	21,07 (3 - 12.288)	40,92 (3 - 12.288)	193,96 (3 - 12.288)	458,75 (3 - 12.288)
15	107,73 (5 - 163.840)	187,74 (6 - 196.608)	149,62 (4 - 131.072)	197,68 (3 - 98.304)	310,1 (3 - 98.304)	1.890,92 (3 - 98.304)	3.891,97 (3 - 98.304)

Através dos resultados discutidos acerca da Tabela 4.8 sugerimos que, para realizar seleção de variáveis em MLGD para um conjunto de dados com mais de 15 variáveis preditoras, é interessante fazer uma pré-seleção de variáveis com o intuito de reduzir o número das mesmas e o tempo computacional para o esquema de seleção de variáveis em até “ k ” passos encontrar um modelo candidato. Esta pré-seleção pode ser realizada, por exemplo, através de uma análise descritiva, retirando do conjunto de dados covariáveis correlacionadas, e/ou que aparentemente não possuem muita relação com a variável resposta de interesse. Outra possível alternativa para conduzir esta pré-seleção de variáveis consiste em utilizar, primeiramente, algum método de regressão capaz de classificar as covariáveis em nível de importância, como florestas aleatórias (Breiman, 2001), e utilizar no nosso procedimento até 15 das variáveis preditoras mais importantes para o método inicialmente aplicado.

Em relação ao método *Stepwise*, analisando a Tabela 4.9 verificamos que o número de modelos ajustados e o tempo computacional necessário são baixos em comparação com o nosso procedimento. Desse modo, considerando que nos estudos de simulação, para a maioria dos cenários, a porcentagem de vezes que o modelo verdadeiro é selecionado com amostras de 400 observações está próxima entre estes dois métodos, o método *Stepwise* pode ser utilizado como uma alternativa caso o tamanho amostral e o número de variáveis sejam grandes e não se deseje realizar uma pré-seleção de variáveis. Entretanto, como mostramos em alguns cenários com dados simulados que apresentam comportamento mais próximo ao observado na prática (como o cenário 9), o método *Stepwise* obteve performance relativamente inferior à do nosso esquema.

Tabela 4.9: Tempo, em minutos (total de modelos ajustados), para obter um modelo candidato utilizando dados simulados com a distribuição Gama, o método *Stepwise* e diferentes combinações de tamanhos amostrais (n) e números de parâmetros (p).

$p \backslash n$	50	100	200	500	1.000	5.000	10.000
5	0,02 (28)	0,02 (32)	0,02 (32)	0,03 (32)	0,05 (32)	0,23 (32)	0,46 (32)
8	0,02 (40)	0,02 (40)	0,04 (64)	0,08 (72)	0,10 (64)	0,35 (64)	0,73 (64)
10	0,04 (58)	0,06 (95)	0,07 (95)	0,12 (105)	0,16 (95)	0,56 (95)	1,17 (95)
12	0,05 (60)	0,07 (91)	0,12 (125)	0,16 (132)	0,23 (132)	0,76 (132)	1,56 (132)
15	0,32 (88)	0,36 (114)	0,47 (199)	0,55 (207)	0,68 (207)	1,85 (207)	3,36 (207)

Capítulo 5

Aplicação

Neste capítulo apresentamos uma aplicação que realizamos a um conjunto de dados reais. Na Seção 5.1 apresentamos os dados e realizamos uma breve análise descritiva dos mesmos. Durante a Seção 5.2 comparamos o poder preditivo dos métodos utilizados nos estudos de simulação. Na Seção 5.3 analisamos as variáveis selecionadas pelo estudo apresentado na seção anterior e mostramos um dos modelos candidatos a final e, na última seção, também comparamos o poder preditivo do nosso esquema para seleção de variáveis com outros métodos de regressão (regressão Ridge, Lasso, *Boosting* “tradicional” e um método *Boosting* para MLGD).

5.1 Descrição dos dados

O golfe é um esporte que possui origem incerta, sendo que a mais plausível é a sua criação pelos escoceses, que possuem registros da prática deste esporte por volta do ano de 1400. O golfe chegou ao Brasil no final do século XIX, quando monges beneditinos cederam parte do terreno do Mosteiro de São Bento para que engenheiros ingleses e escoceses, que construíam a Estrada de Ferro Santos - Jundiaí, fizessem o primeiro campo de golfe do país. Na atualidade, esta região está compreendida entre a Estação da Luz e o rio Tietê ([Golfe.tur, 2019](#)).

Resumidamente, o jogo de golfe baseia-se em jogar uma bola, usando um taco, desde a área do *tee* (lugar onde se inicia o jogo de um buraco) até o buraco, executando uma ou mais tacadas em conformidade com as regras. Há muitas modalidades de jogo de golfe. As mais utilizadas são *Stroke Play*, em que o vencedor é definido pelo menor número na somatória das tacadas, e *Match Play*, em que os pontos são conseguidos a cada buraco, e quem somar o maior número de pontos é o campeão ([SportRegras, 2019](#)).

O golfe pode ser jogado individualmente ou em grupos de dois a quatro jogadores. Nesse esporte, não existe um “adversário” propriamente dito, dado que não há nada que um golfista possa fazer para atrapalhar o desempenho de outros jogadores. O resultado depende de seu esforço individual e sorte, com cada golfista lutando para baixar a sua

pontuação total no campo (*Golfe.tur, 2019*). Este campo é composto por 9 ou 18 buracos, incluindo vários obstáculos como bancos de areia, árvores, lagos, diferentes relevos, entre outros. Desse modo, todos os campos são diferentes uns dos outros e, por isso, o golfe é um jogo que requer muita técnica e leitura da situação (*SportRegras, 2019*).

A base de dados reais que utilizamos neste trabalho traz informações referentes às estatísticas de desempenho e ganhos para 196 golfistas participantes do PGA (*Professional Golfers Association*) durante a temporada de 2004. Esta base de dados está disponível em *sportsillustrated.cnn (2004)* e, na mesma, há onze variáveis com informações sobre os golfistas. Selecionamos como variável resposta para o nosso estudo os ganhos totais dos jogadores, em dólares (US\$). Além disso, retiramos da análise três variáveis que não seriam razoáveis serem consideradas nos ajustes dos modelos: nome, ganho médio dos golfistas e posição em que os mesmos ocupavam no *ranking* de ganhos totais. Desse modo, usamos as seguintes variáveis preditoras no nosso estudo:

- **Idade:** idade do golfista, em anos;
- **Média_drive:** distância, em média, da primeira tacada a partir do *tee* (em jardas);
- **Precisão_drive:** precisão da primeira tacada, em porcentagem, para atingir o *green* (zona próxima ao buraco que está sendo jogado, onde a grama é fina, compacta e aparada rente ao solo);
- **GIR:** *greens on regulation* - estatística utilizada pelos golfistas para avaliarem suas rodadas (em porcentagem);
- **Média_putt:** número médio de tacadas no *green* para atingir o buraco;
- **Sand_save:** percentual de tacadas que atingiram o *green* seguidas por apenas uma tacada para acertar o buraco;
- **Eventos:** número de eventos que o golfista participou.

Inicialmente, realizamos a análise descritiva da variável resposta - ganhos totais dos golfistas (US\$). Na Tabela 5.1 apresentamos as medidas descritivas e na Figura 5.1 o histograma da mesma. Observando a Tabela 5.1 e a Figura 5.1 podemos afirmar que os ganhos totais dos golfistas apresenta uma assimetria à direita. Além disso, metade dos jogadores ganhou até US\$ 815.000 e o ganho máximo observado é US\$ 10.910.000.

Tabela 5.1: Medidas descritivas da variável resposta - ganhos totais dos golfistas (US\$).

Variável	Mínimo	Q1	Mediana	Média	Q3	Máximo	Desv. padrão
Ganhos_totais	21.250	436.600	815.000	1.135.000	1.408.000	10.910.000	1.214.783

Posteriormente, para avaliarmos as relações entre todas as variáveis presentes no nosso estudo calculamos o coeficiente de correlação linear de Pearson entre as mesmas e

mostramos os resultados encontrados na Tabela 5.2. Através destes resultados, não observamos nenhuma correlação alta ou muito alta entre as variáveis. Entretanto, verificamos que existem correlações moderadas entre as a variável resposta (*Ganhos_totais*) e as variáveis *GIR* e *Média_putt* (iguais a 0,44, em módulo), e entre as variáveis *Precisão_drive* e *Média_drive* (-0,61).

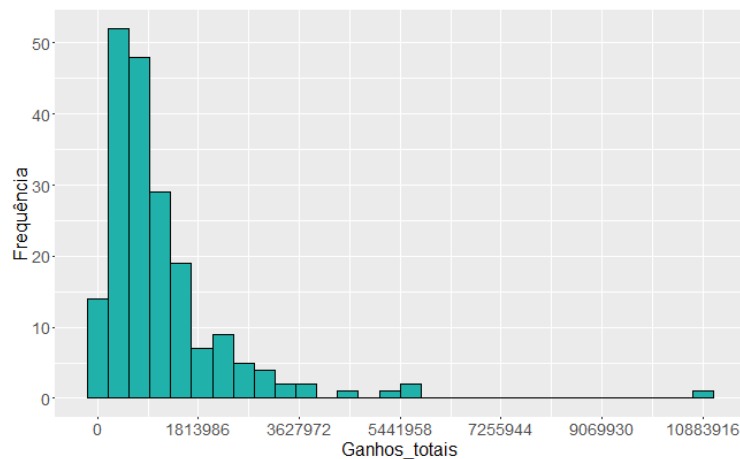


Figura 5.1: Histograma dos ganhos totais dos golfistas (US\$).

Tabela 5.2: Coeficientes de correlação linear de Pearson calculados entre todas as variáveis em estudo.

Variáveis	Ganhos_totais	Idade	Média_drive	Precisão_drive	GIR	Média_putt	Sand_save	Eventos
Ganhos_totais	1,00							
Idade	-0,06	1,00						
Média_drive	0,20	-0,40	1,00					
Precisão_drive	-0,09	0,24	-0,61	1,00				
GIR	0,44	0,05	0,19	0,29	1,00			
Média_putt	-0,44	0,02	0,00	0,07	-0,11	1,00		
Sand_save	0,20	0,06	-0,16	0,05	-0,09	-0,34	1,00	
Eventos	-0,15	-0,16	0,06	0,02	-0,04	0,09	-0,10	1,00

5.2 Avaliação dos procedimentos para seleção de variáveis

Propomos inicialmente um MLGD com distribuição Gama para a variável resposta e função de ligação logarítmica para os dois submodelos. Através da análise de resíduos de alguns dos MLGD selecionados pelos métodos descritos no Capítulo 3, verificamos que os mesmos estavam bem ajustados e, desse modo, prosseguimos com o estudo utilizando esta distribuição.

Para realizar a avaliação dos métodos dividimos a base de dados em base de treinamento (70% da amostra – 137 observações) e de teste (30% restante – 59 observações). A partir disso, utilizamos os métodos de seleção de variáveis em MLGD na base de treinamento e usamos o estimador do risco quadrático na base de teste. O estimador do risco quadrático neste caso é definido da seguinte maneira:

$$R(g) = \frac{1}{59} \sum_{i=1}^{59} (y_i - \hat{y}_i)^2, \quad (5.1)$$

em que g é o modelo candidato, y_i é o valor observado para o ganho do jogador e \hat{y}_i é a previsão relacionada com o i -ésimo valor observado pelo modelo g .

O melhor modelo é aquele com menor estimador do risco quadrático, pois o mesmo quantifica uma média das distâncias ao quadrado entre os valores preditos para a variável de interesse e os valores realmente observados. Com isso, este pode ser um critério para escolher qual metodologia está melhor prevendo o valor da variável resposta na prática.

Para não viesarmos o processo de cálculo da estimativa do risco quadrático, replicamos 100 vezes esse procedimento selecionando diferentes bases de treinamento e respectivas bases de teste. Além disso, mantivemos fixas as mesmas amostras para o cálculo dessa medida nos quatro métodos analisados. Na Tabela 5.3 apresentamos a média e o desvio padrão das estimativas dos riscos quadráticos para cada método e critério de informação. Por outro lado, na Tabela 5.4, considerando que na prática escolhemos o modelo com menor estimativa do erro quadrático, para cada uma das 100 réplicas e método selecionamos, dentre os quatro critérios de informação, a menor estimativa do risco quadrático e calculamos a mediana, média e desvio padrão das mesmas para cada um dos métodos. Também analisamos de forma gráfica essas estimativas através da Figura 5.2, que apresenta o *box plot* dos riscos quadráticos para o melhor critério de informação em cada método.

Tabela 5.3: Média e desvio padrão dos riscos quadráticos para cada método e critério de informação.

Método	Critério de informação							
	AIC		AICc		BIC		HQC	
	Média	DP	Média	DP	Média	DP	Média	DP
Todos modelos	$72,40 \times 10^{10}$	$32,16 \times 10^{10}$	$72,87 \times 10^{10}$	$32,39 \times 10^{10}$	$72,99 \times 10^{10}$	$30,20 \times 10^{10}$	$73,23 \times 10^{10}$	$31,52 \times 10^{10}$
<i>Stepwise</i>	$70,95 \times 10^{10}$	$30,79 \times 10^{10}$	$70,08 \times 10^{10}$	$30,36 \times 10^{10}$	$73,68 \times 10^{10}$	$32,05 \times 10^{10}$	$71,21 \times 10^{10}$	$30,78 \times 10^{10}$
K passos	$71,06 \times 10^{10}$	$30,84 \times 10^{10}$	$70,39 \times 10^{10}$	$30,16 \times 10^{10}$	$71,59 \times 10^{10}$	$29,34 \times 10^{10}$	$69,95 \times 10^{10}$	$29,99 \times 10^{10}$
Dois passos	$71,04 \times 10^{10}$	$31,20 \times 10^{10}$	$69,99 \times 10^{10}$	$30,19 \times 10^{10}$	$72,18 \times 10^{10}$	$30,63 \times 10^{10}$	$70,05 \times 10^{10}$	$30,76 \times 10^{10}$

Observando a Tabela 5.3, é possível verificar que, para dois critérios de informação o nosso procedimento apresentou a melhor performance (BIC e HQC), enquanto que o método *Stepwise* obteve o menor risco quadrático médio para o AIC e a seleção em dois passos para o AICc. Por outro lado, através da Tabela 5.4 e da Figura 5.2, podemos afirmar que as menores estimativas médias e medianas dos riscos para cada procedimento são, em geral, próximas. Entretanto, o esquema de seleção de variáveis em até “ k ” passos apresentou a menor estimativa do risco quadrático médio e mediano dentre todos os métodos, com o menor desvio padrão, considerando as menores estimativas do risco quadrático dentre os quatro critérios de informação para cada uma das 100 réplicas.

Para verificarmos a existência de diferença estatística entre as médias do risco quadrático para as metodologias estudadas, com as 100 réplicas usadas na Tabela 5.4 em cada método construímos intervalos de confiança para a diferença de médias em amostras

pareadas comparando, duas a duas, as metodologias. Assumimos aqui que a diferença entre a estimativa do risco quadrático observada em quaisquer duas metodologias em análise tem distribuição normal. Esse procedimento é plausível de ser realizado porque as amostras são iguais em cada réplica para todos os métodos. Caso o zero esteja incluído no intervalo, podemos afirmar que não existe diferença estatística entre os dois procedimentos em análise para a base de dados em estudo. Na Tabela 5.5 apresentamos os intervalos de confiança para a diferença das médias do risco quadrático para os diferentes procedimentos em estudo.

Tabela 5.4: Mediana, média e desvio padrão dos riscos quadráticos para o melhor critério de informação em cada método.

Método	Mediana	Média	DP
Todos modelos	$60,56 \times 10^{10}$	$69,55 \times 10^{10}$	$29,39 \times 10^{10}$
Stepwise	$60,56 \times 10^{10}$	$68,23 \times 10^{10}$	$28,62 \times 10^{10}$
K passos	$59,71 \times 10^{10}$	$67,41 \times 10^{10}$	$27,68 \times 10^{10}$
Dois passos	$60,56 \times 10^{10}$	$67,46 \times 10^{10}$	$28,46 \times 10^{10}$

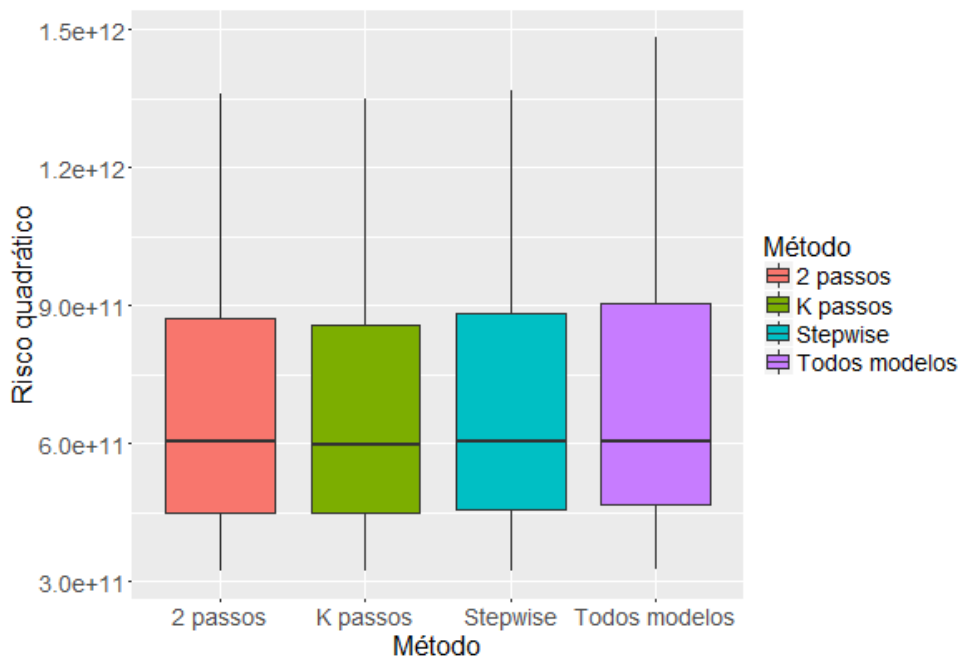


Figura 5.2: Box plot dos riscos quadráticos para o melhor critério de informação em cada método.

Através da Tabela 5.5, podemos verificar que o esquema de seleção de variáveis em até “k” passos apresenta médias do risco quadrático estatisticamente inferiores às conseguidas pela seleção dentre todos os modelos e pelo método *Stepwise* (os limites dos intervalos de confiança são negativos). Por outro lado, é possível afirmar que o nosso procedimento apresentou desempenho estatisticamente igual ao esquema de seleção de variáveis em dois passos (intervalo de confiança compreende o zero).

Tabela 5.5: Intervalos de confiança para a diferença das médias do risco quadrático (linha menos coluna) para os diferentes procedimentos de seleção de variáveis.

Método	Todos Modelos	Stepwise	K passos	2 passos
Todos Modelos	-	-	-	-
Stepwise	$(-24,8 \times 10^8; -15,6 \times 10^8)$	-	-	-
K passos	$(-33,5 \times 10^9; -92,3 \times 10^8)$	$(-15,0 \times 10^9; -13,0 \times 10^8)$	-	-
2 passos	$(-33,1 \times 10^9; -86,3 \times 10^8)$	$(-12,8 \times 10^9; -25,3 \times 10^8)$	$(-51,9 \times 10^8; 61,4 \times 10^8)$	-

5.3 Variáveis selecionadas no estudo de simulação da aplicação

Para analisar quais foram as variáveis preditoras mais importantes no estudo de simulação da aplicação para explicar as relações presentes entre as variáveis preditoras e a resposta, selecionamos o esquema de seleção de variáveis em até “k” passos e contamos quantas vezes cada variável foi selecionada nos modelos finais, considerando o critério de informação com melhor desempenho em cada réplica. Na Tabela 5.6 apresentamos as contagens das variáveis selecionadas para os submodelos da média e do parâmetro de dispersão.

Tabela 5.6: Contagem das variáveis selecionadas pelo esquema de seleção de variáveis em até “k” passos no estudo de simulação da aplicação para os dois submodelos ajustados.

Submodelo da média							
Variável	Idade	Média_drive	Precisão_drive	GIR	Média_putt	Sand_save	Eventos
Número de seleções	10	8	84	100	100	92	7

Submodelo do parâmetro de dispersão							
Variável	Idade	Média_drive	Precisão_drive	GIR	Média_putt	Sand_save	Eventos
Número de seleções	1	7	57	5	53	4	21

Observando a Tabela 5.6, podemos verificar que as variáveis preditoras GIR e Média_putt foram incluídas em todas as repetições para o submodelo da média, o que é coerente com o observado na análise descritiva, já que essas variáveis são as que apresentam maior correlação, em módulo, com a variável resposta. Já as variáveis Precisão_drive e Sand_save estavam no modelo candidato a final em mais de 80% das repetições. Por outro lado, em relação ao submodelo do parâmetro de dispersão, as variáveis Precisão_drive e Média_putt foram selecionadas em mais de 50% das repetições, sendo as mais escolhidas para este submodelo.

Para estudar o nível de significância das variáveis preditoras presentes em um possível modelo candidato a final, utilizando o nosso procedimento e o conjunto de dados completo, encontramos um modelo candidato para cada um dos quatro critérios de informação e selecionamos como modelo candidato a final o escolhido pela maioria dos critérios. Na Tabela 5.7 apresentamos as estimativa dos parâmetros deste modelo.

Tabela 5.7: Estimativas dos parâmetros do MLGD selecionado como modelo candidato a final.

Parâmetro	Variável	Estimativa	Erro Padrão	Valor t	p-valor	exp()
β	Intercepto	34,08	3,81	8,94	<0,001	$6,35 \times 10^{14}$
	Precisão_drive	-0,02	0,01	-3,19	0,002	$9,76 \times 10^{-1}$
	GIR	0,17	0,02	10,54	<0,001	1,18
	Média_putt	-17,36	1,89	-9,14	<0,001	$2,88 \times 10^{-8}$
	Sand_save	0,02	0,01	2,94	0,003	1,03
γ	Intercepto	-6,67	3,68	-1,81	0,071	0,001
	Precisão_drive	-0,02	0,01	-2,64	0,009	0,98
	Média_putt	4,28	2,06	2,08	0,039	71,99

Por meio da Tabela 5.7 observamos que, utilizando-se o teste de Wald, todas as variáveis preditoras presentes neste modelo são significativas ao nível de significância de 5%. Também podemos concluir, por exemplo, que estima-se que o aumento de uma unidade na variável Sand_save, ou seja, o aumento de um ponto percentual das tacadas que atingiram o *green* seguidas por apenas uma tacada para acertar o buraco, acarretam no aumento de 3% na média dos ganhos do golfista. Por outro lado, considerando o submodelo do parâmetro de dispersão e a variável Precisão_drive, estima-se que o aumento de uma unidade na mesma provoca, em média, redução de 2% no valor do parâmetro de dispersão. Além disso, todos os sinais dos parâmetros estimados para o submodelo da média ($\hat{\beta}$) estão coerentes com a análise descritiva (sinais são os mesmos das correlações apresentadas na Tabela 5.2).

Também é interessante destacar que as variáveis preditoras do modelo presentes na Tabela 5.7 também foram as mais selecionadas no estudo de simulação da aplicação que mostramos na Tabela 5.6. Além disso, chamamos o modelo apresentado na Tabela 5.7 de candidato a final porque consideramos que o modelo final deve ser discutido com o pesquisador ou indivíduo interessado no estudo. Desse modo, caso o mesmo deseje incluir alguma variável que o método não selecionou, ou retirar possíveis variáveis não significativas, garantimos que o modelo final atenda as necessidades esperadas.

5.4 Comparação do esquema de seleção de variáveis em até “k” passos com outros métodos de regressão

Considerando nosso interesse em estudar o desempenho preditivo do nosso procedimento, também o comparamos com a regressão Ridge, Lasso, *Boosting* “tradicional” e um método *Boosting* para MLGD. Nas próximas três subseções descrevemos brevemente estas metodologias.

5.4.1 Regressão Ridge e Lasso

A regressão Ridge, introduzida por [Hoerl e Kennard \(1970\)](#), e o Lasso, desenvolvido por [Tibshirani \(1996\)](#), tem como objetivo encontrar um estimador para uma regressão linear que possua risco menor (melhor poder preditivo) que o de mínimos quadrados. Com essa finalidade, nos dois métodos introduz-se um termo penalizador no processo de minimização da soma de quadrado de resíduos e a diferença entre os mesmos está no termo penalizador.

O Lasso utiliza a penalização $L1$, a qual pode ser definida como:

$$\lambda \sum_{j=1}^d |\beta_j|, \quad (5.2)$$

em que $\lambda \geq 0$ e d é o número de variáveis preditoras disponíveis.

Desse modo, no Lasso procuramos por:

$$\inf_{\beta} \left(\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^d \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^d |\beta_j| \right), \quad (5.3)$$

em que y refere-se à variável resposta, β é o vetor de parâmetros a serem estimados, $X_{n \times (d+1)}$ representa a matriz de regressão e n é o número de observações.

O parâmetro λ é chamado de *tuning parameter* (parâmetro de sintonização), pois ele possui a função de controlar o balanço entre viés e variância ([Friedman et al., 2001](#)). Note que, quando $\lambda = 0$, obtemos os estimadores usuais da regressão linear. Na prática, λ é escolhido minimizando-se alguma quantidade de interesse em um processo de validação cruzada, de modo a não produzir uma estimativa otimista para este *tuning parameter*.

Por outro, na regressão Ridge usamos a penalização $L2$, definida como:

$$\lambda \sum_{j=1}^d \beta_j^2, \quad (5.4)$$

em que $\lambda \geq 0$.

Desse modo, na Regressão Ridge buscamos por:

$$\inf_{\beta} \left(\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^d \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^d \beta_j^2 \right). \quad (5.5)$$

A penalização $L1$ reduz as estimativas dos parâmetros, tornando as mesmas iguais a zero em vários parâmetros e, assim, também realiza uma seleção de variáveis. Por outro lado, a penalização $L2$ apenas tende a reduzir os valores dos parâmetros. Além disso, a Regressão Ridge para o modelo linear possui solução analítica e, nos dois métodos,

normalmente o valor estimado para o intercepto não é penalizado e as variáveis preditoras são padronizadas, para que as estimativas dos parâmetros possuam a mesma escala.

Para ajustar esses métodos de regressão utilizamos o pacote do *software* R desenvolvido por [Friedman et al. \(2010\)](#), chamado “*glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models*”.

5.4.2 *Boosting* “tradicional”

O *Boosting* foi introduzido na área de *Machine Learning* como um algoritmo para classificação de variáveis binárias, chamado de *AdaBoost* ([Freund et al., 1996](#)). Posteriormente, mostrou-se que o *Boosting* pode ser interpretado como um algoritmo de gradiente descendente ([Friedman, 2001](#)).

No *Boosting*, o estimador, que denotaremos por $g(x)$, é construído através da agregação de diferentes estimadores da função de regressão, sendo construído incrementalmente. Considerando r para os resíduos, o algoritmo do *Boosting* pode ser descrito como:

1. Definimos $g(x) \equiv 0$ e $r_i = y_i \forall i$;
2. Para $b = 1, \dots, B$:
 - (a) Ajustamos uma regressão linear simples (também é possível ajustar árvore de regressão) para $(x_1, r_1), \dots, (x_n, r_n)$. Seja $g^b(x)$ sua função de predição;
 - (b) Atualizamos g e os resíduos: $g(x) \leftarrow g(x) + \lambda g^b(x)$ e $r_i \leftarrow y_i - g(x)$.
3. Retornamos o modelo final $g(x)$.

Observe que os *tuning parameters* do *Boosting* são B , d e λ . Tipicamente λ é pequeno (por exemplo, 0,0001), $B \approx 1000$ e d é da ordem de 2 ou 4. Maiores informações sobre o *Boosting* podem ser encontradas em [James et al. \(2013\)](#).

Para ajustar o *Boosting* usamos o pacote do *software* R desenvolvido por [Wang \(2019\)](#), chamado *bst: Gradient Boosting*.

5.4.3 *Boosting* para MLGD

O método *Boosting* para modelos lineares generalizados duplos que utilizamos neste trabalho foi desenvolvido por [Mayr et al. \(2012\)](#). Este método utiliza um algoritmo gradiente descendente que rotaciona entre os diferentes preditores a serem estimados, incluindo uma intrínseca seleção de variáveis (resultante do algoritmo gradiente descende). Além disso, ele consegue trabalhar com conjuntos de dados com alta dimensão.

Resumidamente, a estratégia usada pelos autores consiste em calcular, para cada iteração do algoritmo, o valor negativo das derivadas parciais da função de log-verossimilhança negativa com respeito a cada um dos preditores a serem estimados. Em outras palavras,

podemos afirmar que os autores elaboraram uma possível generalização do *Boosting* “tradicional”, pois o algoritmo do *Boosting* para MLGD é semelhante ao descrito na subseção anterior, com o acréscimo da possibilidade do ajuste de até quatro submodelos. Para realizar esse procedimento, na atualização de cada submodelo as estimativas obtidas para os demais são consideradas como *offset*, ou seja, supõe-se que os parâmetros têm valor conhecido para que seja possível atualizar o submodelo em análise.

Para o nosso caso, consideramos $\eta_{\theta_k}, k = 1, 2$. Estes dois preditores são atualizados sucessivamente em cada iteração, como explicitamos abaixo, no esquema ilustrado para a iteração $m + 1$

$$\begin{array}{l} (\hat{\mu}^{[m]}, \hat{\phi}^{[m]}) \xrightarrow{\text{atualização}} \hat{\eta}_{\mu}^{[m+1]} \rightarrow \hat{\mu}^{[m+1]}, \\ (\hat{\mu}^{[m+1]}, \hat{\phi}^{[m]}) \xrightarrow{\text{atualização}} \hat{\eta}_{\phi}^{[m+1]} \rightarrow \hat{\phi}^{[m+1]}. \end{array}$$

As funções de predição são atualizadas para cada preditor η_{θ_k} até que a iteração de parada m_{stop} seja atingida. Além disso, os estimadores base utilizados para a estimação dos parâmetros em cada iteração são modelos lineares. Para maiores detalhes sobre este método, consulte [Mayr et al. \(2012\)](#).

Para o ajuste utilizando o método *Boosting* para MLGD utilizamos o pacote *gamboostLSS: An R package for model building and variable selection in the GAMLSS framework*, desenvolvido por [Hofner et al. \(2014\)](#). Além disso, selecionamos a distribuição Normal para ser utilizada no algoritmo gradiente descendente, uma vez que o ajuste com a distribuição Gama foi impossibilitado devido a erros numéricos durante o processo de estimação do modelo.

5.4.4 Comparações entre as metodologias

Após obtermos as estimativas do risco quadrático para as 100 réplicas usando cada um dos métodos anteriormente explicitados, calculamos a mediana, média e desvio padrão dos mesmos e apresentamos os resultados na Tabela 5.8. Também acrescentamos nesta tabela os resultados obtidos com o nosso esquema para a menor estimativa do risco quadrático em cada réplica (Tabela 5.4). Além disso, na da Figura 5.3, analisamos de forma gráfica os riscos quadráticos através do *box plot* dos mesmos para cada um dos métodos presentes na Tabela 5.8.

Observando a Tabela 5.8, podemos afirmar que o esquema de seleção de variáveis em até “ k ” passos também obteve a menor estimativa do risco quadrático médio e mediano, com o menor desvio padrão, quando comparado com outros métodos de regressão. Dentre as demais metodologias, o Lasso e a regressão Ridge apresentaram as melhores performances, enquanto que o *Boosting* “tradicional” obteve as piores valores médio e mediano. Através da Figura 5.3 podemos visualizar graficamente a grande dispersão dos dois métodos *Boosting*, em relação aos demais, e o melhor desempenho do nosso procedimento. Além disso, os métodos de regressão que usamos nesta seção para compararmos com o

nosso esquema também tiveram desempenho inferior aos demais métodos considerados na seção anterior.

Tabela 5.8: Mediana, média e desvio padrão dos riscos quadráticos para os métodos de regressão em estudo.

Método	Mediana	Média	DP
Boosting MLGD	$99,73 \times 10^{10}$	$13,38 \times 10^{11}$	$81,77 \times 10^{10}$
Boosting trad.	$11,38 \times 10^{11}$	$14,27 \times 10^{11}$	$73,49 \times 10^{10}$
Lasso	$75,26 \times 10^{10}$	$10,19 \times 10^{11}$	$53,31 \times 10^{10}$
Ridge	$71,45 \times 10^{10}$	$10,15 \times 10^{11}$	$54,44 \times 10^{10}$
K passos	$59,71 \times 10^{10}$	$67,41 \times 10^{10}$	$27,68 \times 10^{10}$

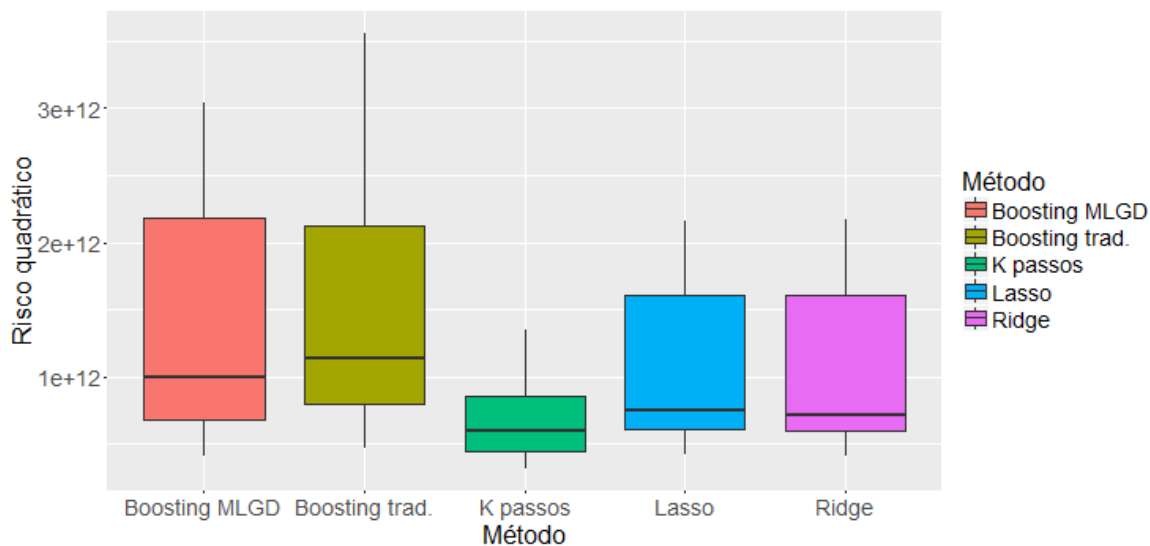


Figura 5.3: Box plot dos riscos quadráticos para os métodos de regressão em estudo.

Para verificarmos a existência de diferença estatística entre as médias do risco quadrático para os diferentes métodos, construímos intervalos de confiança para a diferença de médias em amostras pareadas comparando as metodologias, duas a duas, do mesmo modo que fizemos com os procedimentos para seleção de variáveis na Tabela 5.5. Caso o zero esteja incluído no intervalo, podemos afirmar que não existe diferença estatística entre os dois procedimentos em estudo para a base de dados utilizada. Na Tabela 5.9 apresentamos os intervalos de confiança para a diferença das médias do risco quadrático para as diferentes metodologias analisadas.

Tabela 5.9: Intervalos de confiança para a diferença das médias do risco quadrático (linha menos coluna) para os métodos de regressão em estudo.

Método	Boosting MLGD	Boosting trad.	Lasso	Ridge	K passos
Boosting MLGD	-				
Boosting trad.	$(-69,4 \times 10^9; 24,9 \times 10^{10})$	-			
Lasso	$(-38,1 \times 10^{10}; -25,7 \times 10^9)$	$(-54,4 \times 10^{10}; -27,3 \times 10^{10})$	-		
Ridge	$(-38,1 \times 10^{10}; -26,4 \times 10^{10})$	$(-54,9 \times 10^{10}; -27,6 \times 10^{10})$	$(-80,9 \times 10^8; 63,9 \times 10^7)$	-	
K passos	$(-77,9 \times 10^{10}; -54,8 \times 10^{10})$	$(-88,4 \times 10^{10}; -62,2 \times 10^{10})$	$(-40,4 \times 10^{10}; -28,5 \times 10^{10})$	$(-40,2 \times 10^{10}; -27,9 \times 10^{10})$	-

Através da Tabela 5.9, podemos verificar que o esquema de seleção de variáveis em até “ k ” passos apresenta médias do risco quadrático estatisticamente inferiores às conseguidas pelos demais métodos, ou seja, para este conjunto de dados temos evidências de que o nosso esquema possui performance preditiva superior aos outros métodos de regressão estudados. Por outro lado, os dois métodos *Boosting* tiveram os piores desempenhos, com médias do risco quadrático estatisticamente iguais entre si.

Capítulo 6

Conclusão

Neste trabalho, discutimos sobre os modelos lineares generalizados duplos (MLGD) que, diferentemente dos modelos lineares generalizados (MLG), permitem o ajuste do parâmetro de dispersão da variável resposta em função de variáveis preditoras. Desse modo, os mesmos são uma possível solução quando a suposição de que o parâmetro de dispersão constante não é razoável e a variável resposta tem distribuição que pertence à família exponencial. Considerando nosso interesse em seleção de variáveis nesta classe de modelos e baseando-se no esquema de seleção de variáveis proposto por [Bayer e Cribari-Neto \(2015b\)](#), desenvolvemos um procedimento para realizar seleção de variáveis em até “ k ” passos para modelos lineares generalizados duplos.

Implementamos computacionalmente no *software* R o nosso procedimento e estudamos a sua performance. Com essa finalidade, desenvolvemos estudos de simulação de Monte Carlo considerando as distribuições Gama, Gaussiana Inversa e Normal, e comparamos nosso procedimento com o esquema de seleção em dois passos ([Bayer e Cribari-Neto, 2015b](#)), a seleção dentre todos os possíveis modelos ajustados e um método *Stepwise* para MLGD ([Stasinopoulos et al. \(2017\)](#)). Além disso, utilizamos nos estudos de simulação como critérios de informação AIC (Critério de Informação de Akaike), AICc (Critério de Informação de Akaike corrigido), BIC (Critério de Informação Bayesiano) e HQC (Critério de Informação de Hannan-Quinn).

Utilizando das conclusões discutidas no Capítulo 3, podemos afirmar que, dentre as metodologias, o esquema de seleção de variáveis em até “ k ” passos demonstrou ser o método mais estável, possuindo bom desempenho em comparação com os demais, principalmente para amostras menores. Este desempenho foi semelhante para as três distribuições estudadas. Além disso, é possível afirmar que o nosso procedimento, na maioria dos casos, tem desempenho próximo ou superior ao da seleção de um modelo através do ajuste de todos os possíveis, com tempo computacional bastante inferior. Para a maioria dos critérios de informação e tamanhos amostrais estudados, em mais de 95% das repetições foram necessários 6 ou menos passos para o esquema de seleção de variáveis em até “ k ” passos encontrar um modelo candidato. Com isso, ajustam-se geralmente nesses

casos até 384 modelos (apenas 9,4% do total) se há 6 variáveis preditoras disponíveis.

O método *Stepwise* para modelos lineares generalizados duplos em geral funcionou bem com um custo computacional muito baixo, o que pode ser uma boa alternativa em casos com número muito grande de variáveis. Entretanto, como mostramos em alguns cenários com dados simulados que apresentam comportamento mais próximo ao observado na prática (como o cenário 9), o método *Stepwise* obteve performance relativamente inferior à do nosso esquema. Desse modo, para conjuntos de dados com mais de 15 variáveis preditoras, recomendamos fazer uma pré-seleção de variáveis com o intuito de reduzir o número das mesmas e o tempo computacional para o esquema de seleção de variáveis em até “ k ” passos. Esta pré-seleção pode ser realizada, por exemplo, através de uma análise descritiva ou com a utilização de algum método de regressão que consegue classificar as covariáveis em nível de importância, como florestas aleatórias e a regressão Ridge.

Analisando a aplicação que realizamos a um conjunto de dados reais, verificamos que, nesse banco de dados, o esquema de seleção de variáveis em até “ k ” passos apresenta média do risco quadrático estatisticamente inferior às conseguidas pela seleção dentre todos os modelos e pelo método *Stepwise*, e possui desempenho estatisticamente igual ao esquema de seleção de variáveis em dois passos. Comparando a performance preditiva do nosso método com a regressão Ridge, o Lasso, o *Boosting* “tradicional” e um método *Boosting* para MLGD, concluímos que o esquema de seleção de variáveis em até “ k ” passos apresenta média do risco quadrático estatisticamente inferior às destas últimas metodologias. Estes resultados indicam que o nosso procedimento também é uma boa alternativa quando possui-se interesse em realizar previsões.

6.1 Trabalhos futuros

Escolhemos o valor 2, por ser intuitivo, para o peso que usamos no esquema de seleção de variáveis em até “ k ” passos para o cálculo do termo penalizador do critério de informação dos passos pares quando utiliza-se variáveis não presentes no submodelo da média. Realizamos este procedimento para evitar que variáveis não importantes para o modelo final sejam incluídas no submodelo da dispersão. Em trabalhos futuros, pode-se realizar um estudo para avaliar se existe um outro valor para o peso, ou intervalo de valores, capaz de melhorar o desempenho do esquema de seleção de variáveis em até “ k ” passos. Também é interessante estudar formas de evitar a pequena porcentagem de não convergência como, por exemplo, comparar mais passos além do anterior a dois passos e definir um critério de parada.

Na literatura, há modelos em que se é possível ajustar mais de dois parâmetros da distribuição de uma variável resposta de interesse (Rigby e Stasinopoulos, 2005). Nesse sentido, uma proposta futura é estender o nosso método para estes modelos e comparar

sua performance com outros métodos.

Para conjuntos de dados com mais de 15 variáveis preditoras, futuramente pode-se estudar a performance de diferentes métodos, que consideram a importância das covariáveis (como florestas aleatórias e a regressão Ridge), para realizar a pré-seleção de variáveis que sugerimos anteriormente antes de aplicar o esquema de seleção de variáveis em até “ k ” passos.

Com o objetivo de viabilizar a utilização do nosso método, pretendemos desenvolver um pacote no *software* R para o esquema de seleção de variáveis em até “ k ” passos. O pacote permitirá o ajuste dos 4 métodos comparados nos estudos de simulação deste trabalho para o ajuste de MLGD.

Apêndice A

Resultados completos dos estudos de simulação - Distribuição Gama

Tabela A.1: Resultados completos dos estudos de simulação para o cenário 1 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,2	0,0	6,2	21,9	25,8	4,1	13,8	28,0	16,0	4.096
		Stepwise	1,0	1,0	1,1	7,1	31,2	36,7	3,7	8,0	10,1	22,6	45
		K passos	0,2	0,1	0,0	8,1	28,5	20,3	5,2	18,3	19,3	22,4	245
		Dois passos	0,7	0,8	0,9	5,8	25,2	31,3	4,4	12,8	18,1	17,9	128
	AICc	Todos modelos	0,3	0,4	0,2	12,6	35,9	21,3	6,6	13,3	9,4	25,3	4.096
		Stepwise	2,3	2,7	0,7	16,4	42,4	22,1	5,2	5,6	2,6	30,9	43
		K passos	1,2	0,2	0,0	17,0	40,9	12,2	7,7	15,0	5,7	32,3	222
		Dois passos	1,8	2,2	0,6	13,7	37,5	20,8	6,2	11,1	6,1	26,6	128
	BIC	Todos modelos	1,3	0,6	0,1	26,8	39,1	11,5	6,6	9,1	4,9	29,0	4.096
		Stepwise	10,5	3,0	0,8	29,0	38,7	10,8	3,9	2,5	0,9	29,1	39
		K passos	5,9	0,0	0,0	31,8	38,6	6,5	6,4	8,9	1,9	31,5	218
		Dois passos	8,3	2,9	0,8	26,3	38,0	11,9	4,9	5,0	1,9	28,2	128
HQC	Todos modelos	0,3	0,4	0,1	11,7	31,9	21,8	5,6	12,6	15,6	22,7	4.096	
	Stepwise	3,0	2,9	0,8	15,2	39,0	25,5	4,4	5,5	3,8	28,5	43	
	K passos	1,3	0,2	0,0	15,8	37,2	13,9	6,8	15,7	9,1	29,2	224	
	Dois passos	2,3	2,3	0,7	12,7	34,9	23,9	5,1	10,1	8,0	25,1	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,4	31,2	32,9	0,1	15,3	20,1	30,8	4.096
		Stepwise	0,0	0,0	0,0	0,5	39,6	42,7	0,2	7,2	9,7	38,6	46
		K passos	0,0	0,0	0,0	0,5	39,2	25,0	0,3	18,4	16,5	38,3	227
		Dois passos	0,0	0,0	0,0	0,4	32,7	36,1	0,3	13,5	17,0	31,8	128
	AICc	Todos modelos	0,0	0,0	0,0	0,6	40,7	30,7	0,1	14,9	13,0	39,8	4.096
		Stepwise	0,0	0,0	0,0	1,1	49,4	36,3	0,2	7,3	5,7	47,9	45
		K passos	0,0	0,0	0,0	1,1	49,3	21,1	0,5	17,8	10,2	48,1	225
		Dois passos	0,0	0,0	0,0	0,8	41,7	31,5	0,5	14,5	11,0	40,4	128
	BIC	Todos modelos	0,0	0,0	0,0	7,4	71,9	11,7	1,0	6,3	1,7	69,8	4.096
		Stepwise	0,1	0,0	0,0	7,9	76,0	13,1	0,7	1,9	0,3	73,4	42
		K passos	0,0	0,0	0,0	8,3	76,0	7,3	1,1	6,6	0,7	74,6	202
		Dois passos	0,1	0,0	0,0	7,6	72,3	12,7	0,9	5,5	1,0	69,9	128
HQC	Todos modelos	0,0	0,0	0,0	1,8	53,8	24,5	0,6	11,9	7,4	52,1	4.096	
	Stepwise	0,0	0,0	0,0	2,6	60,9	28,2	0,3	5,3	2,7	58,9	44	
	K passos	0,0	0,0	0,0	2,7	61,4	16,1	0,6	14,2	5,0	60,1	212	
	Dois passos	0,0	0,0	0,0	2,2	55,0	25,9	0,6	10,8	5,5	53,3	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,8	36,4	0,0	14,0	16,8	32,8	4.096
		Stepwise	0,0	0,0	0,0	0,0	41,3	44,0	0,0	6,9	7,9	41,3	46
		K passos	0,0	0,0	0,0	0,0	42,9	26,3	0,0	17,5	13,3	42,9	228
		Dois passos	0,0	0,0	0,0	0,0	34,1	37,8	0,0	13,8	14,2	34,1	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,1	34,3	0,0	14,4	13,2	38,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,5	40,5	0,0	6,7	6,2	46,5	45
		K passos	0,0	0,0	0,0	0,0	48,5	23,8	0,0	16,9	10,7	48,5	220
		Dois passos	0,0	0,0	0,0	0,0	38,7	35,7	0,0	14,4	11,2	38,7	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	84,9	9,9	0,0	4,9	0,3	84,9	4.096
		Stepwise	0,0	0,0	0,0	0,1	89,1	9,6	0,0	1,0	0,2	89,1	43
		K passos	0,0	0,0	0,0	0,1	90,2	4,8	0,0	4,5	0,4	90,2	196
		Dois passos	0,0	0,0	0,0	0,1	87,1	9,3	0,0	3,1	0,4	87,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,7	22,5	0,0	10,9	3,9	62,7	4.096	
	Stepwise	0,0	0,0	0,0	0,0	69,4	25,1	0,0	3,9	1,5	69,4	44	
	K passos	0,0	0,0	0,0	0,0	71,3	14,0	0,0	11,6	3,0	71,3	207	
	Dois passos	0,0	0,0	0,0	0,0	64,9	23,3	0,0	8,3	3,4	64,9	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,5	32,6	0,0	16,2	14,7	36,5	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,1	42,5	0,0	7,3	7,1	43,1	46
		K passos	0,0	0,0	0,0	0,0	44,4	24,2	0,0	18,3	13,1	44,4	230
		Dois passos	0,0	0,0	0,0	0,0	35,4	35,4	0,0	14,9	14,4	35,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,6	32,1	0,0	16,1	13,2	38,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	45,6	41,3	0,0	7,2	5,9	45,6	45
		K passos	0,0	0,0	0,0	0,0	47,1	23,5	0,0	17,9	11,5	47,1	228
		Dois passos	0,0	0,0	0,0	0,0	37,9	34,5	0,0	14,8	12,8	37,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,0	5,1	0,0	2,6	0,3	92,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	93,8	5,6	0,0	0,6	0,0	93,8	42
		K passos	0,0	0,0	0,0	0,0	93,8	3,0	0,0	3,1	0,1	93,8	195
		Dois passos	0,0	0,0	0,0	0,0	92,2	5,5	0,0	2,2	0,1	92,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,7	19,4	0,0	8,2	2,7	69,7	4.096	
	Stepwise	0,0	0,0	0,0	0,0	73,9	21,9	0,0	3,4	0,9	73,9	44	
	K passos	0,0	0,0	0,0	0,0	75,9	12,0	0,0	10,4	1,7	75,9	205	
	Dois passos	0,0	0,0	0,0	0,0	68,3	20,0	0,0	9,0	2,8	68,3	128	

Tabela A.2: Resultados completos dos estudos de simulação para o cenário 2 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	3,9	8,8	15,1	4,0	17,9	21,1	2,8	8,7	17,7	9,5	4,096
		Stepwise	12,5	28,6	24,3	4,0	11,1	12,2	1,6	2,8	2,8	5,8	40
		K passos	14,4	13,6	6,3	8,2	19,8	11,8	4,4	11,3	10,1	11,7	284
		Dois passos	10,8	24,8	21,1	4,0	11,8	14,1	2,0	5,1	6,3	5,3	128
	AICc	Todos modelos	10,9	19,2	16,1	6,4	20,5	12,3	3,3	6,5	4,8	11,6	4,096
		Stepwise	23,3	37,4	16,1	5,8	10,0	4,1	1,6	1,4	0,4	5,4	37
		K passos	29,6	18,8	3,6	12,1	19,2	3,9	4,4	6,9	1,6	12,1	262
		Dois passos	20,1	33,1	15,2	6,4	12,7	5,9	2,0	3,1	1,4	6,0	128
	BIC	Todos modelos	33,2	23,5	10,2	7,8	14,2	3,8	2,2	3,5	1,6	9,0	4,096
		Stepwise	53,1	30,9	5,8	4,9	3,6	0,8	0,6	0,3	0,0	2,1	30
		K passos	64,7	11,1	0,9	10,2	7,9	1,0	1,9	2,1	0,4	5,1	251
		Dois passos	49,4	29,9	6,5	5,7	5,1	1,6	0,9	0,7	0,2	2,8	128
HQC	Todos modelos	11,6	17,1	14,7	5,8	18,3	14,1	2,8	7,1	8,5	10,5	4,096	
	Stepwise	26,6	35,7	15,4	5,1	8,4	5,2	1,6	1,5	0,5	4,6	36	
	K passos	32,1	16,1	3,7	10,7	18,0	4,6	4,4	7,1	3,1	11,3	259	
	Dois passos	23,2	32,7	15,2	5,5	10,4	6,7	1,8	2,9	1,8	5,2	128	
n = 100	AIC	Todos modelos	0,1	0,9	1,0	0,4	30,8	32,7	0,1	14,6	19,4	30,0	4,096
		Stepwise	0,4	8,6	10,3	0,4	32,8	34,3	0,1	5,5	7,5	30,3	45
		K passos	0,2	1,2	0,9	0,9	38,5	24,5	0,5	17,7	15,6	36,9	240
		Dois passos	0,3	7,3	8,5	0,5	28,4	31,0	0,2	10,5	13,4	25,1	128
	AICc	Todos modelos	0,2	1,3	1,9	0,5	40,0	29,6	0,1	14,0	12,4	38,4	4,096
		Stepwise	0,6	12,5	10,1	0,9	38,5	27,9	0,1	5,3	4,0	35,3	44
		K passos	0,5	2,0	1,3	1,8	47,7	19,7	0,6	16,8	9,6	45,2	240
		Dois passos	0,5	10,6	8,4	0,8	34,8	25,9	0,3	10,5	8,2	30,6	128
	BIC	Todos modelos	3,4	9,8	1,2	6,0	62,0	10,3	0,7	5,2	1,4	59,0	4,096
		Stepwise	9,9	39,3	6,8	4,7	32,8	5,3	0,4	0,8	0,1	30,5	40
		K passos	9,0	9,3	0,7	11,9	57,9	4,7	1,2	4,8	0,5	55,4	241
		Dois passos	9,3	36,3	6,7	4,8	34,0	5,6	0,5	2,4	0,4	30,4	128
HQC	Todos modelos	0,4	2,9	2,0	1,6	51,5	23,4	0,4	11,0	6,8	48,7	4,096	
	Stepwise	2,3	20,9	10,5	2,1	41,5	17,5	0,3	3,1	1,8	37,9	42	
	K passos	1,6	3,8	0,9	5,1	56,6	14,0	0,9	12,8	4,3	53,6	237	
	Dois passos	2,0	18,5	9,5	1,9	40,1	17,1	0,6	6,7	3,6	35,1	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,8	36,4	0,0	14,0	16,8	32,8	4,096
		Stepwise	0,0	0,2	0,1	0,0	41,3	43,9	0,0	6,7	7,8	41,1	46
		K passos	0,0	0,0	0,0	0,0	42,9	26,2	0,0	17,5	13,3	42,9	228
		Dois passos	0,0	0,2	0,1	0,0	34,1	37,8	0,0	13,7	14,2	34,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,1	34,3	0,0	14,4	13,2	38,1	4,096
		Stepwise	0,0	0,2	0,2	0,0	46,5	40,4	0,0	6,6	6,1	46,3	45
		K passos	0,0	0,0	0,0	0,0	48,6	23,8	0,0	16,9	10,7	48,6	220
		Dois passos	0,0	0,2	0,1	0,0	38,6	35,6	0,0	14,2	11,1	38,5	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	84,9	9,9	0,0	4,9	0,3	84,9	4,096
		Stepwise	0,0	4,5	0,5	0,1	84,8	9,0	0,0	0,9	0,1	84,6	42
		K passos	0,0	0,0	0,0	0,3	90,0	4,8	0,0	4,5	0,4	90,0	199
		Dois passos	0,0	4,2	0,5	0,1	83,3	8,8	0,0	2,7	0,4	82,8	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,7	22,5	0,0	10,9	3,9	62,7	4,096	
	Stepwise	0,0	1,0	0,6	0,0	68,6	24,6	0,0	3,7	1,5	68,5	44	
	K passos	0,0	0,0	0,0	0,0	71,4	14,0	0,0	11,5	3,0	71,3	208	
	Dois passos	0,0	0,8	0,5	0,0	64,3	23,0	0,0	8,1	3,2	64,1	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,5	32,6	0,0	16,2	14,7	36,5	4,096
		Stepwise	0,0	0,0	0,0	0,0	43,1	42,5	0,0	7,3	7,1	43,1	46
		K passos	0,0	0,0	0,0	0,0	44,4	24,2	0,0	18,3	13,1	44,4	230
		Dois passos	0,0	0,0	0,0	0,0	35,4	35,4	0,0	14,9	14,4	35,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,6	32,1	0,0	16,1	13,2	38,6	4,096
		Stepwise	0,0	0,0	0,0	0,0	45,6	41,3	0,0	7,2	5,9	45,6	45
		K passos	0,0	0,0	0,0	0,0	47,1	23,5	0,0	17,9	11,5	47,1	228
		Dois passos	0,0	0,0	0,0	0,0	37,9	34,5	0,0	14,8	12,8	37,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,0	5,1	0,0	2,6	0,3	92,0	4,096
		Stepwise	0,0	0,0	0,0	0,0	93,8	5,6	0,0	0,6	0,0	93,8	42
		K passos	0,0	0,0	0,0	0,0	93,8	3,0	0,0	3,1	0,1	93,8	195
		Dois passos	0,0	0,0	0,0	0,0	92,2	5,5	0,0	2,2	0,1	92,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,7	19,4	0,0	8,2	2,7	69,7	4,096	
	Stepwise	0,0	0,0	0,0	0,0	73,9	21,9	0,0	3,4	0,9	73,9	44	
	K passos	0,0	0,0	0,0	0,0	75,9	12,0	0,0	10,4	1,7	75,9	205	
	Dois passos	0,0	0,0	0,0	0,0	68,3	20,0	0,0	9,0	2,8	68,3	128	

Tabela A.3: Resultados completos dos estudos de simulação para o cenário 3 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,0	0,0	6,1	21,9	25,6	4,1	13,9	28,4	16,0	4.096
		Stepwise	0,0	0,0	0,0	7,4	32,3	37,6	3,8	8,2	10,7	23,7	45
		K passos	0,0	0,0	0,0	7,6	28,6	20,4	5,2	18,4	19,8	22,8	243
		Dois passos	0,0	0,0	0,0	5,9	26,0	31,9	4,5	13,1	18,7	18,8	128
	AICc	Todos modelos	0,0	0,0	0,0	12,8	36,3	21,3	6,7	13,3	9,6	25,6	4.096
		Stepwise	0,0	0,0	0,0	16,7	45,8	23,2	5,4	6,1	2,9	33,8	43
		K passos	0,0	0,0	0,0	16,2	42,0	12,9	7,6	15,6	5,8	33,6	220
		Dois passos	0,0	0,0	0,0	13,8	39,7	21,6	6,3	12,1	6,5	28,7	128
	BIC	Todos modelos	0,0	0,0	0,0	27,8	39,6	11,6	6,8	9,2	5,0	29,3	4.096
		Stepwise	0,0	0,0	0,0	31,8	47,1	12,7	4,2	3,0	1,1	35,8	41
		K passos	0,0	0,0	0,0	30,7	43,3	7,4	6,5	10,0	2,0	35,7	211
		Dois passos	0,0	0,0	0,0	28,5	44,7	13,5	5,2	6,0	2,1	33,6	128
HQC	Todos modelos	0,0	0,0	0,0	11,9	32,4	21,6	5,7	12,6	15,8	23,1	4.096	
	Stepwise	0,0	0,0	0,0	15,6	42,4	27,1	4,7	6,0	4,3	31,5	43	
	K passos	0,0	0,0	0,0	14,9	38,4	14,7	6,7	16,1	9,2	30,6	221	
	Dois passos	0,0	0,0	0,0	12,9	37,1	25,1	5,3	10,9	8,6	27,2	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,4	31,2	32,9	0,1	15,3	20,1	30,8	4.096
		Stepwise	0,0	0,0	0,0	0,5	39,6	42,7	0,2	7,2	9,7	38,6	46
		K passos	0,0	0,0	0,0	0,5	39,2	25,0	0,3	18,4	16,5	38,3	227
		Dois passos	0,0	0,0	0,0	0,4	32,7	36,1	0,3	13,5	17,0	31,8	128
	AICc	Todos modelos	0,0	0,0	0,0	0,6	40,7	30,7	0,1	14,9	13,0	39,8	4.096
		Stepwise	0,0	0,0	0,0	1,1	49,4	36,3	0,2	7,3	5,7	47,9	45
		K passos	0,0	0,0	0,0	1,1	49,3	21,1	0,5	17,8	10,2	48,1	225
		Dois passos	0,0	0,0	0,0	0,8	41,7	31,5	0,5	14,5	11,0	40,4	128
	BIC	Todos modelos	0,0	0,0	0,0	7,4	71,9	11,7	1,0	6,3	1,7	69,8	4.096
		Stepwise	0,0	0,0	0,0	7,9	76,0	13,1	0,7	1,9	0,3	73,5	42
		K passos	0,0	0,0	0,0	8,2	76,1	7,3	1,1	6,6	0,7	74,7	202
		Dois passos	0,0	0,0	0,0	7,6	72,3	12,7	0,9	5,5	1,0	70,0	128
HQC	Todos modelos	0,0	0,0	0,0	1,8	53,8	24,5	0,6	11,9	7,4	52,1	4.096	
	Stepwise	0,0	0,0	0,0	2,6	60,9	28,2	0,3	5,3	2,7	58,9	44	
	K passos	0,0	0,0	0,0	2,7	61,4	16,1	0,6	14,2	5,0	60,1	212	
	Dois passos	0,0	0,0	0,0	2,2	55,0	25,9	0,6	10,8	5,5	53,3	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,8	36,4	0,0	14,0	16,8	32,8	4.096
		Stepwise	0,0	0,0	0,0	0,0	41,3	44,0	0,0	6,9	7,9	41,3	46
		K passos	0,0	0,0	0,0	0,0	42,9	26,3	0,0	17,5	13,3	42,9	228
		Dois passos	0,0	0,0	0,0	0,0	34,1	37,8	0,0	13,8	14,2	34,1	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,1	34,3	0,0	14,4	13,2	38,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,5	40,5	0,0	6,7	6,2	46,5	45
		K passos	0,0	0,0	0,0	0,0	48,5	23,8	0,0	16,9	10,7	48,5	220
		Dois passos	0,0	0,0	0,0	0,0	38,7	35,7	0,0	14,4	11,2	38,7	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	84,9	9,9	0,0	4,9	0,3	84,9	4.096
		Stepwise	0,0	0,0	0,0	0,1	89,1	9,6	0,0	1,0	0,2	89,1	43
		K passos	0,0	0,0	0,0	0,1	90,2	4,8	0,0	4,5	0,4	90,2	196
		Dois passos	0,0	0,0	0,0	0,1	87,1	9,3	0,0	3,1	0,4	87,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,7	22,5	0,0	10,9	3,9	62,7	4.096	
	Stepwise	0,0	0,0	0,0	0,0	69,4	25,1	0,0	3,9	1,5	69,4	44	
	K passos	0,0	0,0	0,0	0,0	71,3	14,0	0,0	11,6	3,0	71,3	207	
	Dois passos	0,0	0,0	0,0	0,0	64,9	23,3	0,0	8,3	3,4	64,9	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,5	32,6	0,0	16,2	14,7	36,5	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,1	42,5	0,0	7,3	7,1	43,1	46
		K passos	0,0	0,0	0,0	0,0	44,4	24,2	0,0	18,3	13,1	44,4	230
		Dois passos	0,0	0,0	0,0	0,0	35,4	35,4	0,0	14,9	14,4	35,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,6	32,1	0,0	16,1	13,2	38,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	45,6	41,3	0,0	7,2	5,9	45,6	45
		K passos	0,0	0,0	0,0	0,0	47,1	23,5	0,0	17,9	11,5	47,1	228
		Dois passos	0,0	0,0	0,0	0,0	37,9	34,5	0,0	14,8	12,8	37,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,0	5,1	0,0	2,6	0,3	92,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	93,8	5,6	0,0	0,6	0,0	93,8	42
		K passos	0,0	0,0	0,0	0,0	93,8	3,0	0,0	3,1	0,1	93,8	195
		Dois passos	0,0	0,0	0,0	0,0	92,2	5,5	0,0	2,2	0,1	92,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,7	19,4	0,0	8,2	2,7	69,7	4.096	
	Stepwise	0,0	0,0	0,0	0,0	73,9	21,9	0,0	3,4	0,9	73,9	44	
	K passos	0,0	0,0	0,0	0,0	75,9	12,0	0,0	10,4	1,7	75,9	205	
	Dois passos	0,0	0,0	0,0	0,0	68,3	20,0	0,0	9,0	2,8	68,3	128	

Tabela A.4: Resultados completos dos estudos de simulação para o cenário 4 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,1	0,0	23,2	20,9	13,6	12,4	13,4	16,4	5,3	4.096
		Stepwise	0,2	0,1	0,1	28,7	25,8	16,9	12,7	9,3	6,1	6,9	42
		K passos	0,2	0,1	0,1	29,8	19,4	9,4	16,8	14,9	9,3	6,8	216
		Dois passos	0,2	0,1	0,0	25,1	22,7	16,3	14,1	12,0	9,6	5,7	128
	AICc	Todos modelos	0,2	0,3	0,1	41,5	23,0	6,9	13,6	9,5	4,9	6,5	4.096
		Stepwise	0,4	0,3	0,0	47,2	25,0	6,9	13,5	5,0	1,4	7,7	39
		K passos	0,4	0,1	0,1	48,5	19,9	3,5	17,5	7,7	2,3	8,2	211
		Dois passos	0,3	0,2	0,0	43,5	23,7	7,6	14,9	7,1	2,7	7,1	128
	BIC	Todos modelos	1,5	0,2	0,1	60,0	18,1	2,9	11,3	4,5	1,4	5,9	4.096
		Stepwise	2,2	0,3	0,1	67,0	16,8	2,8	8,8	1,7	0,2	6,2	36
		K passos	1,7	0,2	0,0	67,1	13,9	1,5	11,7	3,4	0,5	6,7	203
		Dois passos	1,6	0,2	0,0	63,8	17,9	3,6	9,9	2,4	0,5	6,3	128
HQC	Todos modelos	0,4	0,2	0,0	39,3	21,2	8,6	12,7	10,1	7,5	5,9	4.096	
	Stepwise	0,6	0,3	0,0	45,5	24,5	9,0	12,3	5,4	2,4	7,4	39	
	K passos	0,5	0,1	0,1	46,0	19,5	4,7	16,7	8,8	3,7	8,1	213	
	Dois passos	0,5	0,1	0,1	41,9	23,1	9,8	13,2	7,2	4,0	6,8	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	20,2	25,4	19,8	8,3	13,9	12,4	13,3	4.096
		Stepwise	0,0	0,0	0,0	21,4	31,1	23,1	7,8	9,9	6,7	16,5	43
		K passos	0,0	0,0	0,0	23,3	28,1	13,8	10,4	15,1	9,2	18,2	224
		Dois passos	0,0	0,0	0,0	18,9	27,9	21,2	9,2	12,8	10,1	14,6	128
	AICc	Todos modelos	0,0	0,0	0,0	26,4	28,3	17,0	9,4	12,7	6,2	14,7	4.096
		Stepwise	0,0	0,0	0,0	29,1	33,8	17,3	8,0	8,1	3,6	18,1	42
		K passos	0,0	0,0	0,0	30,6	30,8	10,0	11,5	12,2	4,8	20,3	212
		Dois passos	0,0	0,0	0,0	26,0	30,9	16,5	9,9	11,0	5,7	16,0	128
	BIC	Todos modelos	0,0	0,0	0,0	66,3	22,0	2,2	6,5	2,6	0,4	12,6	4.096
		Stepwise	0,0	0,0	0,0	69,3	22,4	2,7	4,6	0,9	0,2	13,7	37
		K passos	0,0	0,0	0,0	70,9	19,0	1,8	6,2	1,8	0,2	14,4	199
		Dois passos	0,0	0,0	0,0	67,1	22,1	3,1	5,9	1,7	0,2	13,2	128
HQC	Todos modelos	0,0	0,0	0,0	40,2	29,3	10,3	8,5	9,1	2,6	16,2	4.096	
	Stepwise	0,0	0,0	0,0	43,7	32,6	10,5	7,4	4,3	1,6	18,6	40	
	K passos	0,0	0,0	0,0	46,0	28,9	5,9	10,2	6,8	2,3	20,5	203	
	Dois passos	0,0	0,0	0,0	41,3	30,8	10,4	8,7	6,3	2,5	17,5	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	4,0	33,2	32,0	2,4	13,7	14,7	28,3	4.096
		Stepwise	0,0	0,0	0,0	5,1	37,6	35,6	1,8	9,6	10,2	32,7	45
		K passos	0,0	0,0	0,0	6,4	40,4	22,2	2,3	16,0	12,7	37,1	211
		Dois passos	0,0	0,0	0,0	4,5	34,0	32,8	2,2	13,1	13,4	29,5	128
	AICc	Todos modelos	0,0	0,0	0,0	5,4	36,8	30,0	2,8	13,4	11,6	31,6	4.096
		Stepwise	0,0	0,0	0,0	6,6	41,2	32,9	2,2	9,5	7,6	36,2	44
		K passos	0,0	0,0	0,0	7,8	44,3	20,2	2,6	15,7	9,3	40,7	215
		Dois passos	0,0	0,0	0,0	5,9	37,2	30,6	2,6	13,5	10,3	32,5	128
	BIC	Todos modelos	0,0	0,0	0,0	41,5	48,3	5,3	2,7	2,1	0,1	44,1	4.096
		Stepwise	0,0	0,0	0,0	44,4	47,9	4,7	1,9	1,1	0,1	44,1	39
		K passos	0,0	0,0	0,0	44,9	47,9	2,4	2,3	2,4	0,1	45,8	194
		Dois passos	0,0	0,0	0,0	43,3	47,8	4,9	2,1	1,8	0,1	44,0	128
HQC	Todos modelos	0,0	0,0	0,0	15,7	52,9	16,5	2,8	9,6	2,5	47,0	4.096	
	Stepwise	0,0	0,0	0,0	18,4	54,8	17,5	2,6	5,1	1,6	48,5	42	
	K passos	0,0	0,0	0,0	19,6	55,7	10,3	3,4	9,0	2,1	51,7	208	
	Dois passos	0,0	0,0	0,0	17,7	52,4	17,0	2,9	7,5	2,5	46,3	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,3	38,3	32,6	0,0	14,3	14,5	37,9	4.096
		Stepwise	0,0	0,0	0,0	0,4	40,4	37,9	0,1	10,5	10,7	40,1	45
		K passos	0,0	0,0	0,0	0,4	45,5	23,7	0,1	17,6	12,8	45,3	218
		Dois passos	0,0	0,0	0,0	0,3	35,8	34,6	0,1	14,9	14,3	35,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,4	40,4	31,9	0,0	14,6	12,7	40,1	4.096
		Stepwise	0,0	0,0	0,0	0,4	42,4	37,0	0,1	10,6	9,4	42,1	45
		K passos	0,0	0,0	0,0	0,5	47,8	23,0	0,1	17,5	11,2	47,5	218
		Dois passos	0,0	0,0	0,0	0,4	38,0	33,4	0,1	14,9	13,1	37,8	128
	BIC	Todos modelos	0,0	0,0	0,0	8,1	85,0	4,9	0,0	2,0	0,0	84,3	4.096
		Stepwise	0,0	0,0	0,0	8,9	84,6	5,1	0,3	1,0	0,1	83,9	42
		K passos	0,0	0,0	0,0	9,1	85,5	2,8	0,4	2,2	0,1	85,2	194
		Dois passos	0,0	0,0	0,0	8,9	83,4	5,1	0,3	2,3	0,1	82,7	128
HQC	Todos modelos	0,0	0,0	0,0	2,0	68,3	18,4	0,3	7,9	3,1	68,1	4.096	
	Stepwise	0,0	0,0	0,0	2,1	69,8	20,6	0,3	5,7	1,5	69,3	43	
	K passos	0,0	0,0	0,0	2,2	74,5	11,2	0,4	9,9	1,8	74,1	201	
	Dois passos	0,0	0,0	0,0	1,9	66,6	19,6	0,5	8,9	2,5	66,1	128	

Tabela A.5: Resultados completos dos estudos de simulação para o cenário 5 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,0	0,0	0,0	19,6	29,3	0,0	15,8	35,3	19,6	4,096
		Stepwise	1,2	21,6	24,4	0,1	18,2	27,4	0,1	2,3	4,9	16,9	43
		K passos	0,2	0,1	0,0	1,5	29,7	22,3	1,4	22,1	22,6	28,3	281
		Dois passos	1,0	17,6	21,9	0,2	16,5	24,9	0,1	6,5	11,4	13,5	128
	AICc	Todos modelos	0,0	0,0	0,0	0,1	40,1	27,5	0,0	16,1	16,2	40,0	4,096
		Stepwise	3,3	33,8	19,8	0,2	23,6	16,5	0,1	1,6	1,2	22,4	40
		K passos	1,0	0,4	0,1	5,1	50,0	14,7	1,5	19,0	8,3	46,5	275
		Dois passos	2,8	29,9	18,5	0,2	22,8	15,8	0,2	5,7	4,0	19,6	128
	BIC	Todos modelos	0,0	0,0	0,1	0,3	56,5	20,4	0,3	13,2	9,2	56,4	4,096
		Stepwise	15,8	46,5	11,4	0,4	18,2	6,8	0,0	0,5	0,4	17,6	35
		K passos	8,2	0,4	0,1	16,8	48,8	6,8	2,2	14,1	2,6	46,4	278
		Dois passos	13,7	44,3	11,5	0,5	19,5	7,1	0,1	2,2	1,1	17,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,1	35,8	27,8	0,0	15,8	20,5	35,7	4,096	
	Stepwise	4,7	34,8	20,1	0,2	20,4	16,8	0,1	1,5	1,4	19,5	39	
	K passos	1,5	0,2	0,1	5,3	44,0	16,2	2,3	20,1	10,2	41,1	278	
	Dois passos	3,9	30,8	18,9	0,2	20,1	16,3	0,1	5,2	4,5	17,4	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,0	30,1	33,0	0,0	15,0	21,9	30,1	4,096
		Stepwise	0,0	3,7	5,1	0,0	40,0	44,5	0,0	2,4	4,4	39,5	45
		K passos	0,0	0,0	0,0	0,0	37,6	25,8	0,0	20,1	16,5	37,6	250
		Dois passos	0,0	3,2	4,5	0,0	33,2	38,2	0,0	9,1	11,8	32,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,2	31,1	0,0	16,1	13,6	39,2	4,096
		Stepwise	0,0	5,2	5,0	0,0	48,7	37,0	0,0	1,9	2,2	47,9	45
		K passos	0,0	0,0	0,0	0,0	48,8	22,6	0,0	18,7	9,9	48,8	241
		Dois passos	0,0	4,6	4,5	0,0	42,2	32,5	0,0	8,6	7,6	41,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	77,0	13,3	0,0	8,2	1,5	77,0	4,096
		Stepwise	0,3	24,5	7,1	0,0	58,0	9,7	0,0	0,4	0,0	57,5	41
		K passos	0,0	0,0	0,0	0,9	80,1	9,2	0,0	8,9	0,9	79,9	227
		Dois passos	0,2	23,6	6,8	0,0	56,9	9,6	0,0	2,5	0,3	56,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	52,9	24,6	0,0	14,3	8,2	52,9	4,096	
	Stepwise	0,0	10,8	6,7	0,0	54,9	25,5	0,0	1,0	1,1	54,1	43	
	K passos	0,0	0,0	0,0	0,1	61,5	17,3	0,0	15,9	5,3	61,5	230	
	Dois passos	0,0	10,1	6,5	0,0	50,5	23,6	0,0	5,9	3,5	49,2	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	33,4	34,0	0,0	14,6	18,0	33,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	46,0	48,5	0,0	2,9	2,6	46,0	46
		K passos	0,0	0,0	0,0	0,0	42,4	26,3	0,0	19,6	11,7	42,4	231
		Dois passos	0,0	0,0	0,0	0,0	37,6	40,1	0,0	11,0	11,3	37,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,8	33,4	0,0	14,3	13,5	38,8	4,096
		Stepwise	0,0	0,0	0,0	0,0	51,5	44,0	0,0	2,5	2,0	51,5	45
		K passos	0,0	0,0	0,0	0,0	48,4	24,4	0,0	18,3	8,9	48,4	226
		Dois passos	0,0	0,0	0,0	0,0	42,6	37,6	0,0	11,2	8,6	42,6	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	85,2	9,2	0,0	5,3	0,3	85,2	4,096
		Stepwise	0,0	0,3	0,0	0,0	89,8	9,5	0,0	0,4	0,0	89,8	43
		K passos	0,0	0,0	0,0	0,0	89,9	4,8	0,0	5,0	0,3	89,9	197
		Dois passos	0,0	0,3	0,0	0,0	88,2	9,3	0,0	2,0	0,2	88,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,2	23,3	0,0	10,5	4,0	62,2	4,096	
	Stepwise	0,0	0,0	0,0	0,0	72,1	26,5	0,0	1,0	0,3	72,1	44	
	K passos	0,0	0,0	0,0	0,0	71,5	14,4	0,0	11,4	2,7	71,5	215	
	Dois passos	0,0	0,0	0,0	0,0	66,8	24,8	0,0	6,2	2,2	66,8	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,1	34,1	0,0	15,1	14,7	36,1	4,096
		Stepwise	0,0	0,0	0,0	0,0	47,0	47,3	0,0	3,0	2,6	47,0	46
		K passos	0,0	0,0	0,0	0,0	44,0	25,4	0,0	19,0	11,6	44,0	233
		Dois passos	0,0	0,0	0,0	0,0	37,0	38,3	0,0	12,9	11,7	37,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,0	32,7	0,0	15,2	13,1	39,0	4,096
		Stepwise	0,0	0,0	0,0	0,0	49,5	45,4	0,0	2,9	2,2	49,5	46
		K passos	0,0	0,0	0,0	0,0	46,2	24,7	0,0	18,7	10,5	46,2	230
		Dois passos	0,0	0,0	0,0	0,0	39,5	37,1	0,0	12,7	10,7	39,5	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,7	5,0	0,0	3,1	0,2	91,7	4,096
		Stepwise	0,0	0,0	0,0	0,0	94,3	5,6	0,0	0,2	0,0	94,3	42
		K passos	0,0	0,0	0,0	0,0	93,1	3,0	0,0	3,8	0,1	93,1	195
		Dois passos	0,0	0,0	0,0	0,0	92,9	5,5	0,0	1,6	0,1	92,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	68,9	18,3	0,0	10,3	2,5	68,9	4,096	
	Stepwise	0,0	0,0	0,0	0,0	77,3	21,6	0,0	1,0	0,1	77,3	43	
	K passos	0,0	0,0	0,0	0,0	76,1	11,0	0,0	11,2	1,6	76,1	205	
	Dois passos	0,0	0,0	0,0	0,0	71,4	20,2	0,0	6,8	1,5	71,4	128	

Tabela A.6: Resultados completos dos estudos de simulação para o cenário 6 da distribuição Gama.

n	Critério	Método	$\beta < 5$			$\beta = 5$			$\beta > 5$			Modelo Correto	Média de Modelos
			$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$		
n = 50	AIC	Todos modelos	8,3	11,4	11,9	8,1	15,6	16,8	3,4	10,1	14,4	5,3	4.096
		Stepwise	25,3	39,4	26,1	1,6	3,5	3,3	0,2	0,3	0,3	1,1	36
		K passos	18,9	13,9	4,1	17,5	16,2	5,7	8,5	10,0	5,2	4,9	307
		Dois passos	22,9	35,9	25,4	2,6	5,5	4,6	0,8	1,2	1,1	1,0	128
	AICc	Todos modelos	17,7	22,6	10,5	12,1	16,1	8,1	3,1	6,0	3,8	4,6	4.096
		Stepwise	37,6	43,2	14,6	1,5	1,9	0,8	0,2	0,1	0,0	0,6	33
		K passos	34,7	15,5	2,0	23,1	10,9	1,9	6,3	4,5	1,0	3,5	298
		Dois passos	35,6	40,8	14,7	3,0	3,4	1,5	0,4	0,4	0,2	0,7	128
	BIC	Todos modelos	37,4	24,2	6,2	11,5	10,8	2,9	2,3	2,9	1,8	4,3	4.096
		Stepwise	63,9	30,3	4,7	0,6	0,3	0,1	0,0	0,0	0,0	0,1	28
		K passos	61,5	9,2	0,5	19,2	4,0	0,3	3,7	1,5	0,1	1,5	284
		Dois passos	62,3	30,1	5,2	1,4	0,6	0,1	0,1	0,1	0,0	0,2	128
HQC	Todos modelos	17,7	20,4	9,7	10,9	15,3	10,0	3,4	6,1	6,5	4,6	4.096	
	Stepwise	41,3	40,7	14,0	1,2	1,6	0,8	0,1	0,2	0,0	0,4	32	
	K passos	35,6	13,7	2,2	22,0	10,5	2,3	7,3	5,0	1,3	3,3	291	
	Dois passos	39,0	38,6	14,6	2,4	2,8	1,6	0,4	0,4	0,2	0,5	128	
n = 100	AIC	Todos modelos	1,7	5,3	4,4	6,9	27,3	23,9	3,8	11,5	15,2	17,3	4.096
		Stepwise	9,9	31,0	27,7	4,0	13,7	10,5	0,3	1,2	1,6	7,5	42
		K passos	4,0	6,2	2,6	14,4	28,5	14,5	6,6	14,1	9,0	19,0	275
		Dois passos	7,8	25,5	24,0	5,1	16,6	13,2	0,8	3,4	3,7	6,6	128
	AICc	Todos modelos	2,5	7,4	5,1	11,2	30,5	20,0	3,2	11,8	8,3	20,0	4.096
		Stepwise	13,4	37,3	23,0	5,0	12,3	7,1	0,4	0,7	0,8	7,0	41
		K passos	6,3	7,4	2,3	20,0	30,9	10,4	6,6	11,4	4,6	20,9	270
		Dois passos	11,2	32,0	20,5	6,1	15,7	9,2	1,1	2,4	1,9	6,6	128
	BIC	Todos modelos	17,9	17,6	1,9	25,4	26,6	4,7	2,1	3,5	0,3	21,1	4.096
		Stepwise	50,6	38,4	5,9	2,6	2,0	0,3	0,1	0,0	0,0	1,5	34
		K passos	31,9	7,6	0,5	40,4	13,4	1,1	3,3	1,6	0,1	10,0	278
		Dois passos	49,0	37,7	5,9	3,8	2,8	0,5	0,1	0,1	0,0	1,5	128
HQC	Todos modelos	5,0	10,8	4,6	17,9	30,5	15,2	3,7	8,5	3,8	22,0	4.096	
	Stepwise	25,0	42,7	16,2	4,6	7,8	3,3	0,2	0,1	0,1	4,6	38	
	K passos	12,3	8,5	1,4	30,3	27,3	5,2	6,7	6,4	1,8	18,8	278	
	Dois passos	23,0	39,2	15,8	5,8	10,4	4,0	0,7	0,9	0,3	4,6	128	
n = 200	AIC	Todos modelos	0,0	0,8	0,6	1,6	32,8	34,3	0,7	13,9	15,3	31,3	4.096
		Stepwise	0,8	15,6	14,8	1,2	30,3	30,7	0,2	3,0	3,5	27,1	44
		K passos	0,2	0,5	0,4	3,3	41,6	23,5	1,4	17,5	11,6	39,5	253
		Dois passos	0,6	12,7	12,4	1,1	27,5	28,7	0,4	8,5	8,1	22,3	128
	AICc	Todos modelos	0,0	0,9	0,6	2,0	37,7	32,0	0,7	13,9	12,2	36,1	4.096
		Stepwise	0,9	17,4	14,9	1,5	32,6	27,4	0,2	2,5	2,6	29,2	44
		K passos	0,2	0,7	0,4	4,1	45,5	21,7	1,5	16,4	9,5	43,2	254
		Dois passos	0,7	14,3	12,5	1,4	30,1	26,5	0,4	8,0	6,1	24,5	128
	BIC	Todos modelos	2,9	6,3	0,5	14,1	65,8	6,0	0,8	3,2	0,4	63,7	4.096
		Stepwise	18,6	49,0	4,7	5,6	20,0	1,9	0,0	0,2	0,1	19,0	39
		K passos	6,9	4,2	0,1	37,9	44,4	2,5	1,7	2,2	0,1	42,2	253
		Dois passos	18,0	47,8	4,7	5,7	20,9	2,0	0,2	0,6	0,1	19,0	128
HQC	Todos modelos	0,5	2,3	1,0	5,1	56,5	19,7	0,7	11,0	3,2	54,4	4.096	
	Stepwise	4,7	34,0	11,5	3,7	33,4	11,1	0,2	1,0	0,5	31,1	41	
	K passos	1,5	2,0	0,4	15,4	56,3	10,5	2,0	10,0	1,9	53,6	243	
	Dois passos	4,2	31,3	10,7	3,8	33,5	11,4	0,4	3,6	1,1	28,9	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	37,7	31,9	0,0	15,8	14,6	37,6	4.096
		Stepwise	0,0	4,4	3,7	0,0	42,3	41,6	0,0	4,3	3,7	41,8	45
		K passos	0,0	0,0	0,0	0,0	43,7	25,7	0,0	18,8	11,8	43,7	236
		Dois passos	0,0	3,4	2,9	0,0	35,0	35,1	0,0	12,5	11,2	33,6	128
	AICc	Todos modelos	0,0	0,1	0,0	0,0	39,9	31,8	0,0	15,4	12,8	39,8	4.096
		Stepwise	0,0	4,8	3,6	0,0	44,2	40,0	0,0	4,1	3,3	43,7	45
		K passos	0,0	0,0	0,0	0,0	46,9	24,2	0,0	18,4	10,4	46,9	238
		Dois passos	0,0	3,7	2,9	0,0	36,9	34,2	0,0	12,4	10,0	35,5	128
	BIC	Todos modelos	0,0	0,4	0,1	1,8	89,8	5,0	0,0	2,8	0,1	89,4	4.096
		Stepwise	1,3	37,8	2,0	1,1	54,6	3,1	0,0	0,1	0,0	54,2	41
		K passos	0,1	0,3	0,0	7,8	85,7	2,5	0,2	3,3	0,1	85,4	225
		Dois passos	1,2	37,1	1,9	1,1	54,5	3,3	0,0	0,9	0,0	53,4	128
HQC	Todos modelos	0,0	0,2	0,0	0,2	68,6	18,5	0,0	9,8	2,7	68,5	4.096	
	Stepwise	0,1	15,4	4,1	0,2	61,0	17,7	0,0	1,2	0,3	60,4	43	
	K passos	0,0	0,1	0,0	0,9	75,5	10,8	0,2	11,1	1,4	75,3	217	
	Dois passos	0,1	14,3	3,6	0,2	57,6	17,1	0,1	5,5	1,5	56,1	128	

Tabela A.7: Resultados completos dos estudos de simulação para o cenário 7 da distribuição Gama.

n	Critério	Método	$\#\beta < 3$			$\#\beta = 3$			$\#\beta > 3$			Modelo Correto	Média de Modelos
			$\#\gamma < 2$	$\#\gamma = 2$	$\#\gamma > 2$	$\#\gamma < 2$	$\#\gamma = 2$	$\#\gamma > 2$	$\#\gamma < 2$	$\#\gamma = 2$	$\#\gamma > 2$		
n = 50	AIC	Todos modelos	0,0	0,8	3,4	0,5	13,3	21,8	0,5	15,4	44,3	11,7	4.096
		Stepwise	0,2	5,3	8,2	0,8	22,5	32,1	1,0	10,6	19,3	18,3	36
		K passos	0,6	2,2	1,2	1,6	25,4	13,0	1,6	27,9	26,5	22,4	268
		Dois passos	0,2	3,5	5,8	0,8	18,9	27,7	0,9	15,0	27,3	14,4	128
	AICc	Todos modelos	0,2	2,6	4,6	1,4	24,5	22,1	1,0	17,1	26,5	21,8	4.096
		Stepwise	0,5	9,6	9,0	1,9	32,2	25,8	1,8	9,7	9,5	26,1	33
		K passos	1,7	4,7	1,4	3,4	37,5	10,5	2,4	27,5	11,1	32,6	246
		Dois passos	0,4	7,0	6,4	1,8	28,6	24,5	1,7	15,0	14,7	21,9	128
	BIC	Todos modelos	3,5	10,3	4,9	5,7	38,6	14,1	0,9	12,7	9,3	33,9	4.096
		Stepwise	6,3	24,6	8,8	6,2	34,9	11,8	1,2	4,7	1,6	28,8	29
		K passos	13,3	10,2	0,6	8,7	44,4	4,1	1,7	14,8	2,2	39,7	223
		Dois passos	6,0	21,5	7,9	6,0	35,2	12,2	1,2	7,1	2,9	27,9	128
HQC	Todos modelos	0,3	3,5	4,5	2,0	26,7	19,3	1,0	15,7	27,0	23,7	4.096	
	Stepwise	1,1	12,5	9,7	2,6	32,5	23,4	1,5	8,4	8,4	26,2	33	
	K passos	3,0	5,6	1,2	4,3	38,5	9,4	2,1	25,2	10,7	33,7	246	
	Dois passos	1,0	9,7	7,7	2,4	30,0	23,1	1,4	12,3	12,4	23,0	128	
n = 100	AIC	Todos modelos	0,0	0,1	0,1	0,0	19,9	23,9	0,2	20,6	35,2	19,7	4.096
		Stepwise	0,0	0,3	0,3	0,0	26,9	37,0	0,1	14,2	21,1	26,5	37
		K passos	0,0	0,0	0,0	0,0	31,4	14,4	0,2	30,4	23,5	31,4	242
		Dois passos	0,0	0,2	0,1	0,0	20,4	29,2	0,1	19,8	30,2	20,0	128
	AICc	Todos modelos	0,0	0,1	0,2	0,1	26,8	25,8	0,1	20,9	26,0	26,5	4.096
		Stepwise	0,0	0,4	0,4	0,1	34,1	34,6	0,1	14,5	15,6	33,5	36
		K passos	0,0	0,1	0,0	0,2	39,6	12,7	0,3	30,6	16,6	39,5	231
		Dois passos	0,0	0,3	0,2	0,1	26,6	28,3	0,1	21,3	23,1	25,9	128
	BIC	Todos modelos	0,3	0,9	0,3	1,9	68,9	11,9	0,3	12,0	3,5	68,4	4.096
		Stepwise	0,3	4,7	0,9	2,2	71,6	13,1	0,4	5,2	1,4	69,8	31
		K passos	0,8	0,8	0,0	2,8	77,7	3,6	0,5	12,2	1,5	77,3	211
		Dois passos	0,3	4,4	0,8	2,0	67,5	13,0	0,4	9,5	2,0	65,7	128
HQC	Todos modelos	0,1	0,2	0,2	0,2	42,3	23,4	0,2	18,7	14,7	41,9	4.096	
	Stepwise	0,1	1,3	0,8	0,4	51,3	27,6	0,3	11,0	7,3	50,2	33	
	K passos	0,2	0,2	0,0	0,7	57,2	9,1	0,4	24,9	7,4	56,8	224	
	Dois passos	0,1	1,0	0,5	0,4	44,5	24,9	0,3	17,4	10,8	43,5	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	19,5	29,5	0,0	20,1	30,9	19,5	4.096
		Stepwise	0,0	0,0	0,0	0,0	27,3	38,6	0,0	14,2	19,8	27,3	37
		K passos	0,0	0,0	0,0	0,0	32,9	16,0	0,0	31,3	19,8	32,9	242
		Dois passos	0,0	0,0	0,0	0,0	20,7	30,2	0,0	20,6	28,6	20,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	22,9	29,4	0,0	20,4	27,3	22,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	30,5	37,5	0,0	14,8	17,3	30,4	36
		K passos	0,0	0,0	0,0	0,0	36,5	15,1	0,0	31,5	16,8	36,5	241
		Dois passos	0,0	0,0	0,0	0,0	23,3	30,1	0,0	21,7	25,0	23,2	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	81,2	10,3	0,0	6,8	1,7	81,2	4.096
		Stepwise	0,0	0,0	0,0	0,0	85,1	10,4	0,0	3,6	0,8	85,1	31
		K passos	0,0	0,0	0,0	0,0	88,4	2,4	0,0	8,8	0,4	88,4	198
		Dois passos	0,0	0,0	0,0	0,0	81,7	10,0	0,0	6,9	1,3	81,7	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	50,4	22,7	0,0	18,3	8,6	50,4	4.096	
	Stepwise	0,0	0,0	0,0	0,0	59,6	25,0	0,0	10,0	5,4	59,6	33	
	K passos	0,0	0,0	0,0	0,0	64,3	8,5	0,0	22,9	4,4	64,3	216	
	Dois passos	0,0	0,0	0,0	0,0	53,4	23,0	0,0	16,1	7,6	53,3	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	22,8	26,5	0,0	22,8	27,9	22,8	4.096
		Stepwise	0,0	0,0	0,0	0,0	28,2	37,9	0,0	14,9	19,0	28,2	37
		K passos	0,0	0,0	0,0	0,0	32,3	16,3	0,0	32,3	19,0	32,3	239
		Dois passos	0,0	0,0	0,0	0,0	20,6	29,2	0,0	22,3	27,9	20,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	24,1	27,0	0,0	23,4	25,5	24,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	29,7	37,7	0,0	15,0	17,7	29,7	36
		K passos	0,0	0,0	0,0	0,0	34,3	16,2	0,0	31,9	17,6	34,3	231
		Dois passos	0,0	0,0	0,0	0,0	21,8	29,1	0,0	22,7	26,4	21,8	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	87,9	6,2	0,0	5,5	0,4	87,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	90,8	7,0	0,0	2,0	0,1	90,8	31
		K passos	0,0	0,0	0,0	0,0	92,4	1,9	0,0	5,6	0,1	92,4	196
		Dois passos	0,0	0,0	0,0	0,0	88,1	6,7	0,0	4,7	0,4	88,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	59,3	19,4	0,0	15,3	6,0	59,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	64,7	23,3	0,0	8,7	3,3	64,7	33	
	K passos	0,0	0,0	0,0	0,0	70,3	7,5	0,0	19,4	2,8	70,3	207	
	Dois passos	0,0	0,0	0,0	0,0	57,3	20,6	0,0	15,9	6,1	57,3	128	

Tabela A.8: Resultados completos dos estudos de simulação para o cenário 8 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,7	1,3	2,3	6,4	20,2	20,1	6,2	15,8	27,0	12,8	4,096
		Stepwise	5,5	7,8	6,4	10,9	25,3	26,1	4,0	7,4	6,7	15,0	43
		K passos	4,4	2,8	1,1	11,6	23,8	15,8	7,0	17,2	16,3	16,9	248
		Dois passos	4,2	6,0	4,9	9,2	23,0	26,0	4,4	9,7	12,4	13,5	128
	AICc	Todos modelos	2,9	4,0	3,0	15,4	28,6	13,5	8,2	13,5	10,9	19,7	4,096
		Stepwise	12,2	11,1	4,3	20,2	29,1	12,9	4,7	4,2	1,3	18,0	40
		K passos	10,5	3,9	1,1	22,3	30,0	8,6	7,8	11,8	4,0	21,8	234
		Dois passos	10,0	8,7	3,3	17,9	29,1	15,0	5,7	7,3	2,9	18,0	128
	BIC	Todos modelos	12,7	5,5	1,9	26,2	25,0	7,9	7,1	8,6	5,1	17,6	4,096
		Stepwise	31,3	10,8	2,5	27,1	19,0	4,4	3,3	1,3	0,3	12,1	36
		K passos	26,8	4,1	0,4	31,4	20,9	3,3	6,3	5,6	1,3	15,8	220
		Dois passos	26,1	10,0	2,6	27,2	21,0	5,9	3,6	2,6	0,8	13,7	128
HQC	Todos modelos	2,9	3,3	2,6	13,8	25,5	15,2	6,7	13,7	16,3	18,1	4,096	
	Stepwise	13,3	10,9	4,8	18,6	26,7	15,3	4,2	4,2	2,0	16,6	40	
	K passos	11,0	3,5	1,0	20,5	27,6	9,9	7,7	12,4	6,5	20,2	233	
	Dois passos	10,9	8,8	3,8	16,5	26,5	17,7	5,0	6,8	4,0	16,5	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	1,4	29,5	34,7	0,8	14,7	18,9	27,8	4,096
		Stepwise	0,0	0,1	0,1	1,8	38,2	43,6	0,7	6,9	8,6	36,3	46
		K passos	0,0	0,0	0,0	1,5	37,6	25,8	1,0	18,0	16,0	35,9	234
		Dois passos	0,0	0,1	0,1	1,2	30,6	36,4	1,0	13,7	16,9	28,9	128
	AICc	Todos modelos	0,0	0,0	0,0	2,6	38,0	31,4	0,7	15,0	12,3	35,6	4,096
		Stepwise	0,0	0,2	0,2	2,6	48,2	36,1	0,7	6,9	5,0	45,6	45
		K passos	0,0	0,0	0,0	2,7	47,1	21,8	1,0	17,7	9,6	45,3	221
		Dois passos	0,0	0,1	0,1	1,9	39,6	31,9	1,2	15,1	10,0	37,3	128
	BIC	Todos modelos	0,0	0,0	0,0	11,4	66,6	12,5	1,4	6,4	1,7	63,0	4,096
		Stepwise	1,9	1,1	0,2	11,1	70,8	11,7	0,7	2,0	0,4	66,9	42
		K passos	0,2	0,0	0,0	13,2	70,4	6,0	1,3	8,0	1,0	67,5	207
		Dois passos	1,7	1,0	0,1	10,3	68,5	11,5	1,4	4,7	0,8	64,6	128
HQC	Todos modelos	0,0	0,0	0,0	4,8	50,3	24,5	1,1	12,4	6,9	47,3	4,096	
	Stepwise	0,2	0,5	0,1	4,7	60,1	26,4	0,9	4,4	2,7	56,8	44	
	K passos	0,0	0,0	0,0	5,2	59,2	15,3	1,2	14,1	5,0	56,8	219	
	Dois passos	0,1	0,3	0,2	4,0	53,9	24,2	1,4	10,4	5,5	50,7	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	33,0	34,1	0,0	15,2	17,7	33,0	4,096
		Stepwise	0,0	0,0	0,0	0,1	41,0	43,1	0,0	7,6	8,3	40,9	46
		K passos	0,0	0,0	0,0	0,1	40,3	26,6	0,0	18,3	14,7	40,3	231
		Dois passos	0,0	0,0	0,0	0,1	30,7	34,1	0,0	17,6	17,5	30,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	37,9	32,7	0,0	14,3	15,1	37,9	4,096
		Stepwise	0,0	0,0	0,0	0,1	45,5	40,6	0,0	7,3	6,5	45,4	46
		K passos	0,0	0,0	0,0	0,1	45,2	24,8	0,0	18,2	11,7	45,1	230
		Dois passos	0,0	0,0	0,0	0,1	35,4	32,6	0,0	17,2	14,8	35,3	128
	BIC	Todos modelos	0,0	0,0	0,0	1,0	83,1	10,6	0,1	4,7	0,5	83,0	4,096
		Stepwise	0,0	0,0	0,0	1,1	87,0	10,1	0,1	1,6	0,2	86,9	43
		K passos	0,0	0,0	0,0	1,2	87,9	5,1	0,1	5,2	0,5	87,9	200
		Dois passos	0,0	0,0	0,0	1,0	81,9	9,8	0,1	6,6	0,6	81,8	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	61,3	24,0	0,0	10,9	3,8	61,3	4,096	
	Stepwise	0,0	0,0	0,0	0,1	68,3	26,0	0,1	4,1	1,4	68,2	44	
	K passos	0,0	0,0	0,0	0,1	69,4	14,9	0,1	12,5	3,0	69,3	212	
	Dois passos	0,0	0,0	0,0	0,1	59,5	23,4	0,1	12,6	4,4	59,3	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	31,8	35,6	0,0	15,9	16,7	31,8	4,096
		Stepwise	0,0	0,0	0,0	0,0	41,6	44,1	0,0	6,7	7,6	41,6	46
		K passos	0,0	0,0	0,0	0,0	41,2	27,3	0,0	18,2	13,3	41,2	225
		Dois passos	0,0	0,0	0,0	0,0	34,7	37,2	0,0	13,5	14,6	34,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	34,2	34,4	0,0	15,7	15,7	34,2	4,096
		Stepwise	0,0	0,0	0,0	0,0	43,8	42,4	0,0	6,6	7,2	43,8	46
		K passos	0,0	0,0	0,0	0,0	44,0	26,3	0,0	17,5	12,2	44,0	222
		Dois passos	0,0	0,0	0,0	0,0	36,9	35,9	0,0	13,4	13,8	36,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	88,4	7,5	0,0	3,7	0,4	88,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	92,6	6,7	0,0	0,6	0,1	92,6	42
		K passos	0,0	0,0	0,0	0,0	92,7	3,6	0,0	3,6	0,1	92,7	195
		Dois passos	0,0	0,0	0,0	0,0	90,7	6,5	0,0	2,5	0,3	90,7	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	65,4	20,5	0,0	10,5	3,6	65,4	4,096	
	Stepwise	0,0	0,0	0,0	0,0	74,3	20,9	0,0	3,6	1,2	74,3	43	
	K passos	0,0	0,0	0,0	0,0	73,5	13,0	0,0	11,4	2,0	73,5	205	
	Dois passos	0,0	0,0	0,0	0,0	69,0	19,6	0,0	8,8	2,6	69,0	128	

Tabela A.9: Resultados completos dos estudos de simulação para o cenário 9 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,9	2,1	1,7	5,6	27,1	21,3	2,5	18,1	20,7	12,6	4.096
		Stepwise	8,7	17,7	6,8	8,6	31,3	15,2	2,2	6,2	3,3	12,8	41
		K passos	2,2	1,8	0,9	9,5	34,7	15,0	4,9	19,9	11,1	19,1	254
		Dois passos	2,8	9,1	5,1	7,3	33,8	21,3	2,3	10,5	7,8	12,7	128
	AICc	Todos modelos	2,7	4,3	1,8	10,0	40,4	15,6	3,4	13,9	7,9	18,1	4.096
		Stepwise	18,5	23,5	4,6	12,5	28,8	7,0	1,8	2,6	0,6	12,1	39
		K passos	6,2	3,2	0,5	20,5	42,2	7,7	4,8	12,3	2,7	24,0	256
		Dois passos	8,2	14,3	3,8	12,8	39,3	11,1	2,8	5,4	2,2	15,7	128
	BIC	Todos modelos	8,9	7,0	1,3	16,2	42,3	10,5	3,4	6,9	3,5	19,6	4.096
		Stepwise	49,6	22,2	2,1	10,1	12,5	1,8	0,8	0,8	0,1	5,5	33
		K passos	21,4	3,8	0,1	29,2	31,8	3,3	3,9	5,6	0,8	19,8	246
		Dois passos	27,3	17,6	2,5	16,7	26,8	4,9	1,7	1,9	0,5	11,0	128
HQC	Todos modelos	2,8	4,1	1,5	9,1	37,3	15,7	3,2	15,8	10,5	16,4	4.096	
	Stepwise	21,8	24,2	5,1	10,7	25,6	7,8	1,6	2,7	0,6	10,2	38	
	K passos	6,9	2,9	0,5	19,3	38,9	8,3	4,7	14,1	4,5	21,8	259	
	Dois passos	9,8	14,7	4,3	11,6	36,6	12,7	2,4	5,3	2,5	14,1	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,4	39,3	25,7	0,2	19,8	14,6	28,2	4.096
		Stepwise	0,0	0,3	0,1	0,7	50,5	32,8	0,1	9,1	6,3	36,3	45
		K passos	0,0	0,0	0,0	0,6	45,5	19,7	0,2	21,4	12,6	36,1	233
		Dois passos	0,0	0,3	0,1	0,5	43,8	31,4	0,1	13,4	10,4	31,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,6	49,5	22,4	0,2	17,7	9,6	35,8	4.096
		Stepwise	0,0	0,5	0,2	0,9	60,6	25,6	0,1	8,3	3,6	43,6	45
		K passos	0,0	0,0	0,0	1,0	56,1	16,0	0,2	19,4	7,3	45,2	219
		Dois passos	0,0	0,3	0,1	0,7	54,0	25,6	0,1	13,1	6,1	39,2	128
	BIC	Todos modelos	0,0	0,1	0,0	4,6	80,2	7,2	0,9	6,0	1,0	57,1	4.096
		Stepwise	2,7	7,3	0,5	5,5	75,0	6,6	0,3	1,9	0,2	52,6	42
		K passos	0,5	0,1	0,0	6,4	80,6	4,4	0,7	6,9	0,4	66,3	210
		Dois passos	1,1	4,4	0,4	4,8	78,0	7,6	0,4	3,0	0,3	55,7	128
HQC	Todos modelos	0,0	0,1	0,0	0,9	63,1	16,2	0,3	13,7	5,7	44,8	4.096	
	Stepwise	0,1	1,7	0,4	1,3	71,0	18,2	0,2	5,7	1,5	50,7	44	
	K passos	0,0	0,0	0,0	1,6	68,6	10,7	0,3	15,3	3,6	56,1	214	
	Dois passos	0,1	1,1	0,3	1,0	67,2	18,7	0,1	8,9	2,5	48,6	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,2	31,9	0,0	14,5	17,4	35,3	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,5	38,6	0,0	7,4	7,4	45,0	45
		K passos	0,0	0,0	0,0	0,0	45,5	23,6	0,0	18,2	12,7	44,4	227
		Dois passos	0,0	0,0	0,0	0,0	39,1	35,5	0,0	13,2	12,2	37,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	40,8	31,6	0,0	14,0	13,6	39,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	51,4	35,8	0,0	7,3	5,5	49,5	45
		K passos	0,0	0,0	0,0	0,0	51,2	21,7	0,0	17,6	9,5	49,7	224
		Dois passos	0,0	0,0	0,0	0,0	43,8	32,9	0,0	13,3	10,0	42,4	128
	BIC	Todos modelos	0,0	0,0	0,0	0,4	87,7	6,0	0,0	5,6	0,3	81,4	4.096
		Stepwise	0,0	0,0	0,0	0,2	92,8	5,5	0,0	1,5	0,1	85,9	42
		K passos	0,0	0,0	0,0	0,2	91,2	3,2	0,0	5,4	0,1	87,5	198
		Dois passos	0,0	0,0	0,0	0,2	90,7	5,7	0,0	3,1	0,3	84,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	64,9	20,4	0,0	10,9	3,8	61,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	73,9	20,5	0,0	4,4	1,2	69,8	43	
	K passos	0,0	0,0	0,0	0,0	73,5	12,3	0,0	11,9	2,4	70,8	206	
	Dois passos	0,0	0,0	0,0	0,0	69,0	19,8	0,0	8,2	3,0	65,5	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,0	35,5	0,0	14,2	14,3	36,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,1	41,2	0,0	8,6	7,1	43,1	45
		K passos	0,0	0,0	0,0	0,0	44,2	26,2	0,0	18,3	11,4	44,2	223
		Dois passos	0,0	0,0	0,0	0,0	36,0	36,3	0,0	14,1	13,6	36,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,4	34,5	0,0	14,1	13,0	38,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,0	39,5	0,0	8,2	6,3	46,0	45
		K passos	0,0	0,0	0,0	0,0	46,8	24,7	0,0	18,3	10,2	46,8	218
		Dois passos	0,0	0,0	0,0	0,0	38,7	34,9	0,0	14,1	12,2	38,7	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	90,4	5,7	0,0	3,5	0,4	89,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	93,1	5,8	0,0	0,9	0,1	92,0	42
		K passos	0,0	0,0	0,0	0,0	93,0	3,5	0,0	3,3	0,2	92,9	195
		Dois passos	0,0	0,0	0,0	0,0	91,2	5,8	0,0	2,8	0,2	90,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,7	18,1	0,0	9,1	3,1	69,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	75,2	19,6	0,0	4,1	1,1	74,7	43	
	K passos	0,0	0,0	0,0	0,0	76,0	12,0	0,0	9,9	2,1	75,9	205	
	Dois passos	0,0	0,0	0,0	0,0	69,8	19,0	0,0	8,9	2,3	69,4	128	

Tabela A.10: Resultados completos dos estudos de simulação para o cenário 10 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$		
n = 50	AIC	Todos modelos	2,2	3,3	4,1	6,4	25,9	18,8	3,8	14,3	21,2	19,2	4.096
		Stepwise	18,3	27,4	15,1	5,6	15,5	11,3	1,8	2,6	2,5	9,9	43
		K passos	7,8	4,3	1,4	16,8	23,5	9,8	9,3	16,3	10,8	15,8	296
		Dois passos	15,6	23,1	14,2	6,2	16,6	11,2	2,4	5,6	5,1	9,0	128
	AICc	Todos modelos	6,1	7,6	4,7	15,2	36,1	10,2	5,2	9,4	5,5	28,1	4.096
		Stepwise	34,0	30,9	7,2	8,7	12,6	3,8	1,2	1,2	0,5	8,8	40
		K passos	21,3	5,7	0,5	30,4	21,1	3,1	8,2	8,1	1,5	15,1	296
		Dois passos	30,0	27,7	7,1	9,3	15,1	4,6	2,6	2,7	0,9	9,0	128
	BIC	Todos modelos	20,5	9,1	3,0	20,3	28,8	5,5	3,7	6,8	2,3	22,7	4.096
		Stepwise	64,0	20,6	3,0	6,5	4,4	0,6	0,7	0,2	0,0	3,1	33
		K passos	49,3	3,3	0,2	29,1	8,7	0,8	5,0	3,2	0,4	6,4	284
		Dois passos	58,7	21,2	3,6	7,9	5,9	1,0	1,1	0,6	0,1	3,7	128
HQC	Todos modelos	5,3	6,4	4,7	11,9	33,1	13,0	4,7	10,5	10,4	26,2	4.096	
	Stepwise	36,0	30,2	8,1	7,1	10,9	4,8	1,2	1,1	0,6	7,5	39	
	K passos	20,5	5,3	0,7	26,1	20,5	4,8	9,0	9,9	3,2	14,4	290	
	Dois passos	31,5	28,0	8,3	7,3	13,1	5,7	2,3	2,5	1,3	7,8	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,2	38,5	26,6	0,1	19,0	15,6	37,8	4.096
		Stepwise	0,4	8,8	7,2	0,5	42,1	30,4	0,1	5,8	4,8	39,9	48
		K passos	0,0	0,0	0,0	1,2	41,0	22,6	0,5	22,0	12,6	39,7	237
		Dois passos	0,3	7,3	6,2	0,4	36,2	26,7	0,1	12,8	10,1	33,1	128
	AICc	Todos modelos	0,0	0,1	0,1	0,4	48,3	24,2	0,3	17,9	8,7	47,4	4.096
		Stepwise	1,1	11,6	6,7	1,1	48,5	23,7	0,1	4,7	2,5	46,0	47
		K passos	0,0	0,1	0,0	3,4	51,5	17,7	0,9	19,3	7,0	49,9	239
		Dois passos	0,9	10,2	5,6	1,0	42,9	21,9	0,2	11,4	5,8	39,5	128
	BIC	Todos modelos	0,3	0,3	0,1	5,6	75,2	9,5	0,5	7,5	1,0	73,8	4.096
		Stepwise	14,7	31,5	4,2	5,0	39,9	4,0	0,2	0,5	0,1	38,2	43
		K passos	1,8	0,6	0,1	30,9	54,5	3,7	1,8	6,2	0,4	51,8	256
		Dois passos	13,5	30,5	4,3	5,5	39,6	4,1	0,3	1,9	0,2	37,3	128
HQC	Todos modelos	0,1	0,1	0,1	1,1	59,1	19,5	0,4	14,5	5,1	58,1	4.096	
	Stepwise	3,2	19,4	6,6	2,0	50,0	14,9	0,2	2,5	1,1	47,4	46	
	K passos	0,2	0,1	0,0	7,9	58,7	13,0	1,8	15,0	3,3	56,5	239	
	Dois passos	2,9	18,1	6,1	2,0	46,7	14,6	0,3	7,0	2,3	43,2	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	39,6	28,1	0,0	19,1	13,2	39,6	4.096
		Stepwise	0,0	0,0	0,1	0,0	51,1	35,2	0,0	8,0	5,6	51,0	49
		K passos	0,0	0,0	0,0	0,0	44,8	23,5	0,0	20,7	11,0	44,8	219
		Dois passos	0,0	0,0	0,1	0,0	42,5	29,1	0,0	16,5	11,8	42,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	45,5	26,4	0,0	17,9	10,2	45,5	4.096
		Stepwise	0,0	0,0	0,1	0,0	56,0	32,1	0,0	7,4	4,4	55,9	48
		K passos	0,0	0,0	0,0	0,0	51,0	21,2	0,0	19,2	8,6	51,0	221
		Dois passos	0,0	0,0	0,1	0,0	46,8	27,2	0,0	16,5	9,4	46,7	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	88,3	6,6	0,0	4,9	0,2	88,3	4.096
		Stepwise	0,0	1,4	0,2	0,0	90,7	6,7	0,0	0,9	0,0	90,5	46
		K passos	0,0	0,0	0,0	0,5	89,6	4,7	0,0	4,9	0,3	89,6	198
		Dois passos	0,0	1,1	0,2	0,0	88,7	6,5	0,0	3,2	0,3	88,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	67,5	17,3	0,0	12,4	2,8	67,5	4.096	
	Stepwise	0,0	0,3	0,1	0,0	75,9	19,3	0,0	3,6	0,7	75,7	47	
	K passos	0,0	0,0	0,0	0,0	72,6	12,5	0,0	12,5	2,4	72,6	205	
	Dois passos	0,0	0,3	0,1	0,0	69,5	18,0	0,0	10,1	2,1	69,2	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	42,5	27,4	0,0	17,4	12,7	42,5	4.096
		Stepwise	0,0	0,0	0,0	0,0	54,3	34,8	0,0	6,3	4,5	54,3	49
		K passos	0,0	0,0	0,0	0,0	47,7	22,3	0,0	20,1	9,9	47,7	219
		Dois passos	0,0	0,0	0,0	0,0	43,1	28,2	0,0	17,5	11,2	43,1	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	45,5	26,7	0,0	16,3	11,5	45,5	4.096
		Stepwise	0,0	0,0	0,0	0,0	57,2	32,9	0,0	6,1	3,9	57,2	49
		K passos	0,0	0,0	0,0	0,0	51,0	20,8	0,0	19,5	8,7	51,0	218
		Dois passos	0,0	0,0	0,0	0,0	45,6	26,9	0,0	17,7	9,9	45,6	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,6	4,0	0,0	3,3	0,1	92,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	95,4	4,1	0,0	0,5	0,0	95,4	46
		K passos	0,0	0,0	0,0	0,0	93,8	2,8	0,0	3,3	0,0	93,8	195
		Dois passos	0,0	0,0	0,0	0,0	93,2	4,1	0,0	2,6	0,1	93,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	73,3	14,8	0,0	9,6	2,3	73,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	80,9	16,4	0,0	2,3	0,3	80,9	47	
	K passos	0,0	0,0	0,0	0,0	77,6	10,3	0,0	10,5	1,6	77,6	206	
	Dois passos	0,0	0,0	0,0	0,0	74,2	15,3	0,0	9,0	1,5	74,2	128	

Tabela A.11: Resultados completos dos estudos de simulação para o cenário 11 da distribuição Gama.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	5,5	16,5	23,0	4,0	11,7	17,8	1,5	6,6	13,4	5,0	4.096
		Stepwise	15,8	36,2	31,1	2,3	5,8	6,2	0,7	1,0	0,9	2,6	37
		K passos	22,2	20,6	9,3	7,1	14,7	8,4	3,0	8,0	6,7	7,0	286
		Dois passos	13,4	32,5	29,2	2,8	7,7	8,7	1,0	1,9	2,9	2,7	128
	AICc	Todos modelos	13,6	27,2	22,2	4,8	14,3	8,0	1,6	4,6	3,7	6,2	4.096
		Stepwise	26,7	44,2	19,2	3,1	4,1	1,3	0,6	0,5	0,1	1,8	34
		K passos	39,2	25,4	4,7	8,5	12,4	2,3	2,3	4,0	1,2	6,4	273
		Dois passos	23,8	41,8	19,1	3,8	6,3	2,8	0,7	1,1	0,6	2,2	128
	BIC	Todos modelos	37,3	33,7	10,1	4,8	8,0	2,2	1,0	1,9	1,0	3,8	4.096
		Stepwise	55,8	35,6	5,9	1,6	0,9	0,1	0,1	0,0	0,0	0,4	27
		K passos	76,1	12,9	0,8	4,4	3,5	0,5	0,8	0,8	0,2	2,0	256
		Dois passos	52,8	35,8	6,9	2,0	1,7	0,5	0,2	0,1	0,0	0,6	128
HQC	Todos modelos	15,2	25,1	20,3	4,2	12,3	10,8	1,4	5,1	5,6	5,5	4.096	
	Stepwise	30,6	43,3	17,7	2,5	3,5	1,5	0,4	0,5	0,1	1,6	33	
	K passos	43,4	22,4	4,7	7,0	11,1	2,6	2,5	4,4	1,8	5,7	270	
	Dois passos	27,7	40,5	18,5	3,1	5,1	2,9	0,6	1,1	0,6	1,9	128	
n = 100	AIC	Todos modelos	0,2	3,0	3,7	0,3	30,5	31,7	0,0	12,8	17,8	27,9	4.096
		Stepwise	0,8	20,2	21,1	0,5	23,7	25,9	0,1	3,3	4,5	20,3	44
		K passos	0,5	4,4	2,9	1,3	36,8	23,1	0,5	16,9	13,7	33,0	256
		Dois passos	0,6	17,0	18,6	0,4	23,5	25,7	0,1	6,0	8,1	18,4	128
	AICc	Todos modelos	0,3	5,2	4,3	0,5	36,4	29,9	0,1	12,7	10,6	32,6	4.096
		Stepwise	1,5	27,2	19,6	0,7	26,4	19,6	0,2	2,9	2,0	22,3	43
		K passos	1,4	7,1	3,4	2,2	44,1	17,9	0,8	15,5	7,6	39,6	253
		Dois passos	1,2	23,9	17,2	0,7	26,7	20,0	0,2	5,3	4,6	20,7	128
	BIC	Todos modelos	5,0	26,9	3,7	4,5	45,8	8,4	0,4	4,2	1,1	42,5	4.096
		Stepwise	16,6	57,8	9,1	2,2	11,7	2,1	0,1	0,3	0,0	10,6	37
		K passos	19,4	21,0	1,3	10,2	40,1	2,8	0,9	3,9	0,4	37,0	265
		Dois passos	15,5	55,9	9,3	2,5	13,6	2,3	0,2	0,7	0,0	11,2	128
HQC	Todos modelos	1,2	10,7	5,4	1,0	45,5	20,8	0,4	9,6	5,4	41,0	4.096	
	Stepwise	4,2	41,2	17,7	1,4	22,8	10,3	0,1	1,6	0,7	20,0	41	
	K passos	4,4	12,5	3,3	4,7	48,3	11,0	1,1	11,2	3,6	43,8	256	
	Dois passos	3,7	38,0	16,7	1,4	24,4	10,9	0,2	2,9	1,7	19,8	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,8	36,7	0,0	13,9	16,6	32,8	4.096
		Stepwise	0,0	1,0	2,0	0,0	40,7	43,3	0,0	6,2	6,7	40,2	45
		K passos	0,0	0,0	0,0	0,0	43,7	25,7	0,0	17,6	12,9	43,6	226
		Dois passos	0,0	0,9	1,8	0,0	35,5	38,7	0,0	11,3	11,8	34,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,7	34,4	0,0	13,9	13,0	38,7	4.096
		Stepwise	0,0	1,3	2,2	0,0	46,3	39,4	0,0	5,9	4,9	45,8	45
		K passos	0,0	0,0	0,0	0,0	49,4	23,5	0,0	16,9	10,2	49,4	221
		Dois passos	0,0	1,2	2,0	0,0	40,8	35,6	0,0	11,4	9,0	40,2	128
	BIC	Todos modelos	0,0	1,0	0,2	0,0	83,5	9,8	0,0	5,2	0,3	83,5	4.096
		Stepwise	0,2	22,8	2,6	0,1	66,7	6,9	0,0	0,5	0,2	66,4	41
		K passos	0,3	0,9	0,1	0,8	88,4	4,6	0,0	4,8	0,3	88,3	213
		Dois passos	0,2	21,9	2,6	0,1	66,5	6,9	0,0	1,5	0,3	65,7	128
HQC	Todos modelos	0,0	0,0	0,1	0,0	62,0	22,3	0,0	11,4	4,2	62,0	4.096	
	Stepwise	0,0	7,1	2,9	0,0	63,5	22,6	0,0	2,8	1,0	63,1	43	
	K passos	0,0	0,1	0,1	0,1	71,2	13,7	0,0	12,0	2,7	71,1	214	
	Dois passos	0,0	6,4	2,7	0,0	61,1	21,4	0,0	6,0	2,4	60,1	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	37,3	33,4	0,0	15,4	13,9	37,3	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,9	43,5	0,0	6,6	5,9	43,9	45
		K passos	0,0	0,0	0,0	0,0	44,9	24,5	0,0	18,4	12,2	44,9	227
		Dois passos	0,0	0,0	0,0	0,0	37,7	37,1	0,0	12,8	12,4	37,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,2	32,7	0,0	15,5	12,6	39,2	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,2	42,0	0,0	6,4	5,4	46,2	45
		K passos	0,0	0,0	0,0	0,0	47,4	23,7	0,0	17,7	11,2	47,4	230
		Dois passos	0,0	0,0	0,0	0,0	39,9	36,0	0,0	12,6	11,4	39,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,0	5,1	0,0	2,6	0,3	92,0	4.096
		Stepwise	0,0	0,6	0,1	0,0	93,3	5,6	0,0	0,5	0,0	93,3	42
		K passos	0,0	0,0	0,0	0,0	93,7	2,9	0,0	3,3	0,2	93,7	195
		Dois passos	0,0	0,6	0,1	0,0	92,3	5,5	0,0	1,5	0,1	92,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,0	20,0	0,0	8,4	2,6	69,0	4.096	
	Stepwise	0,0	0,1	0,0	0,0	74,5	21,9	0,0	2,9	0,7	74,5	43	
	K passos	0,0	0,0	0,0	0,0	75,9	11,8	0,0	10,7	1,7	75,9	205	
	Dois passos	0,0	0,1	0,0	0,0	70,9	20,8	0,0	6,5	1,8	70,9	128	

Apêndice B

Resultados completos dos estudos de simulação - Distribuição Gaussiana Inversa

Tabela B.1: Resultados completos dos estudos de simulação para o cenário 1 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,5	0,3	1,0	6,1	19,9	24,1	3,7	14,6	29,8	14,2	4.096
		Stepwise	6,8	7,2	4,2	8,0	27,8	30,6	2,6	5,9	7,0	19,4	43
		K passos	3,4	0,9	0,3	9,7	26,3	18,1	5,5	18,7	17,1	19,3	247
		Dois passos	4,1	5,3	3,7	6,7	25,7	30,5	3,2	9,2	11,6	17,3	128
	AICc	Todos modelos	1,2	1,2	1,2	13,6	33,4	19,2	5,2	13,5	11,5	24,5	4.096
		Stepwise	13,4	9,9	3,1	15,2	33,8	16,1	3,1	3,7	1,7	24,3	40
		K passos	7,1	1,4	0,2	19,9	35,4	9,7	6,7	14,1	5,5	26,8	234
		Dois passos	9,3	7,2	3,0	13,8	34,2	18,4	4,1	6,8	3,2	24,0	128
	BIC	Todos modelos	6,1	2,2	0,9	23,4	35,9	11,6	5,2	8,6	6,1	27,0	4.096
		Stepwise	41,9	8,2	1,1	19,2	20,9	5,4	2,0	1,0	0,3	14,8	34
		K passos	25,7	1,1	0,1	28,8	27,1	3,9	5,0	7,0	1,4	20,9	222
		Dois passos	29,4	7,7	1,2	20,9	27,3	8,2	2,5	2,0	0,8	19,5	128
HQC	Todos modelos	1,2	1,1	1,3	12,2	29,3	20,7	4,9	12,8	16,5	21,7	4.096	
	Stepwise	17,8	9,7	3,3	13,2	29,8	17,9	2,7	3,5	2,1	21,1	40	
	K passos	8,7	1,3	0,2	18,5	31,5	11,3	6,7	13,9	7,9	23,9	241	
	Dois passos	11,6	7,5	3,2	12,2	31,5	20,2	3,5	6,0	4,4	21,9	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,6	30,1	33,7	0,1	15,3	20,2	29,0	4.096
		Stepwise	0,0	0,0	0,0	0,7	41,2	45,7	0,2	5,5	6,7	39,9	45
		K passos	0,0	0,0	0,0	0,9	38,0	25,0	0,4	19,4	16,3	37,1	235
		Dois passos	0,0	0,0	0,0	0,6	37,4	42,5	0,3	8,7	10,5	36,2	128
	AICc	Todos modelos	0,0	0,0	0,0	1,2	39,6	29,7	0,4	15,8	13,3	38,4	4.096
		Stepwise	0,0	0,0	0,0	1,4	50,7	38,5	0,3	4,8	4,4	48,9	44
		K passos	0,0	0,0	0,0	1,3	48,9	21,0	0,5	18,1	10,1	47,6	225
		Dois passos	0,0	0,0	0,0	1,1	47,1	36,4	0,4	7,8	7,1	45,4	128
	BIC	Todos modelos	0,0	0,0	0,0	7,3	70,5	13,2	0,6	6,6	1,8	68,2	4.096
		Stepwise	0,3	0,2	0,0	8,9	75,8	13,3	0,3	1,0	0,1	73,1	42
		K passos	0,1	0,0	0,0	9,3	74,3	7,3	0,7	7,7	0,7	73,0	202
		Dois passos	0,2	0,2	0,0	8,6	74,6	13,6	0,4	2,2	0,2	71,9	128
HQC	Todos modelos	0,0	0,0	0,0	2,4	52,1	24,4	0,5	13,2	7,4	50,2	4.096	
	Stepwise	0,1	0,0	0,0	3,1	62,1	29,8	0,2	3,0	1,8	59,8	44	
	K passos	0,0	0,0	0,0	3,2	59,8	15,8	0,7	15,2	5,4	58,2	211	
	Dois passos	0,1	0,0	0,0	2,8	59,2	29,2	0,3	5,4	3,0	57,1	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	31,8	36,6	0,0	15,4	16,2	31,8	4.096
		Stepwise	0,0	0,0	0,0	0,0	42,4	45,6	0,0	6,0	5,9	42,4	45
		K passos	0,0	0,0	0,0	0,0	42,7	26,0	0,0	18,3	13,0	42,7	227
		Dois passos	0,0	0,0	0,0	0,0	38,2	42,1	0,0	9,9	9,8	38,2	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	37,7	34,2	0,0	14,6	13,5	37,7	4.096
		Stepwise	0,0	0,0	0,0	0,0	47,4	42,6	0,0	5,8	4,2	47,4	45
		K passos	0,0	0,0	0,0	0,0	48,7	24,1	0,0	17,7	9,5	48,7	215
		Dois passos	0,0	0,0	0,0	0,0	43,2	39,5	0,0	9,8	7,5	43,2	128
	BIC	Todos modelos	0,0	0,0	0,0	0,2	84,8	9,2	0,0	5,4	0,4	84,8	4.096
		Stepwise	0,0	0,0	0,0	0,1	89,3	9,7	0,0	0,7	0,1	89,3	42
		K passos	0,0	0,0	0,0	0,1	89,7	4,6	0,0	5,2	0,4	89,7	197
		Dois passos	0,0	0,0	0,0	0,1	88,5	9,7	0,0	1,5	0,2	88,5	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,4	22,3	0,0	11,0	4,3	62,4	4.096	
	Stepwise	0,0	0,0	0,0	0,0	70,8	25,2	0,0	3,1	0,8	70,8	43	
	K passos	0,0	0,0	0,0	0,0	71,8	13,7	0,0	11,8	2,6	71,8	206	
	Dois passos	0,0	0,0	0,0	0,0	68,4	24,2	0,0	5,4	1,9	68,4	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	37,9	33,0	0,0	13,8	15,3	37,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	45,4	24,0	0,0	18,3	12,3	45,4	45
		K passos	0,0	0,0	0,0	0,0	45,4	24,0	0,0	18,3	12,3	45,4	222
		Dois passos	0,0	0,0	0,0	0,0	41,5	40,2	0,0	9,0	9,4	41,5	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	40,3	32,2	0,0	13,9	13,6	40,3	4.096
		Stepwise	0,0	0,0	0,0	0,0	48,0	42,1	0,0	5,5	4,4	48,0	45
		K passos	0,0	0,0	0,0	0,0	47,7	23,7	0,0	18,6	10,1	47,7	218
		Dois passos	0,0	0,0	0,0	0,0	44,0	38,6	0,0	9,2	8,1	44,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,0	5,3	0,0	2,6	0,1	92,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	94,0	5,6	0,0	0,4	0,0	94,0	42
		K passos	0,0	0,0	0,0	0,0	93,2	3,2	0,0	3,5	0,1	93,2	194
		Dois passos	0,0	0,0	0,0	0,0	93,7	5,5	0,0	0,7	0,0	93,7	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	68,8	19,5	0,0	8,9	2,8	68,8	4.096	
	Stepwise	0,0	0,0	0,0	0,0	75,9	21,4	0,0	2,2	0,5	75,9	43	
	K passos	0,0	0,0	0,0	0,0	76,4	11,1	0,0	10,8	1,7	76,4	204	
	Dois passos	0,0	0,0	0,0	0,0	73,9	21,0	0,0	4,1	1,0	73,9	128	

Tabela B.2: Resultados completos dos estudos de simulação para o cenário 2 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	6,1	15,7	19,8	3,8	12,1	19,0	1,4	6,9	15,2	4,5	4.096
		Stepwise	19,5	37,3	28,7	2,3	5,2	4,6	0,6	1,0	0,8	1,9	36
		K passos	24,7	20,1	8,1	7,7	14,1	7,7	3,3	8,1	6,3	5,9	281
		Dois passos	16,2	33,9	27,8	3,3	7,5	6,8	0,8	1,6	2,0	2,1	128
	AICc	Todos modelos	15,3	27,4	20,1	5,1	13,4	8,8	1,4	4,4	4,1	5,4	4.096
		Stepwise	31,3	42,9	17,4	2,8	3,6	1,2	0,4	0,4	0,0	1,5	33
		K passos	45,6	21,7	3,8	8,0	11,7	2,3	2,2	3,7	1,0	5,5	265
		Dois passos	27,6	40,8	18,2	3,7	5,8	2,2	0,6	0,8	0,2	1,9	128
	BIC	Todos modelos	39,2	30,5	10,0	4,7	8,4	2,7	0,8	2,1	1,6	4,1	4.096
		Stepwise	62,2	30,5	5,3	1,3	0,7	0,0	0,1	0,0	0,0	0,3	26
		K passos	79,3	10,4	0,7	4,5	3,3	0,3	0,9	0,6	0,1	1,7	252
		Dois passos	58,9	31,3	6,4	1,8	1,2	0,2	0,1	0,0	0,0	0,6	128
HQC	Todos modelos	17,0	24,2	18,3	4,0	13,4	10,6	1,3	4,3	6,9	5,4	4.096	
	Stepwise	36,3	40,8	15,8	2,4	2,8	1,2	0,4	0,2	0,1	1,1	32	
	K passos	50,2	18,5	3,7	7,2	10,2	2,3	2,5	3,9	1,6	4,7	259	
	Dois passos	32,6	38,8	17,6	3,0	4,5	2,1	0,4	0,7	0,3	1,5	128	
n = 100	AIC	Todos modelos	0,3	2,6	3,0	0,3	29,0	33,4	0,1	13,1	18,2	27,0	4.096
		Stepwise	1,3	20,0	20,9	0,6	23,8	25,6	0,1	3,2	4,4	19,8	44
		K passos	0,7	4,3	2,4	2,1	36,1	22,6	0,6	17,2	14,0	32,4	262
		Dois passos	0,9	17,1	18,2	0,6	23,9	26,0	0,2	5,7	7,4	18,4	128
	AICc	Todos modelos	0,3	4,7	4,0	0,9	37,0	28,7	0,2	12,5	11,7	32,9	4.096
		Stepwise	2,0	25,9	19,7	1,2	26,4	19,5	0,1	2,8	2,3	22,4	43
		K passos	1,6	6,5	2,9	3,2	43,2	18,3	0,8	15,9	7,8	38,9	256
		Dois passos	1,7	22,8	17,5	1,1	26,9	20,5	0,3	4,8	4,5	21,3	128
	BIC	Todos modelos	6,0	23,7	3,6	4,6	46,4	9,3	0,4	4,5	1,5	43,0	4.096
		Stepwise	21,5	53,3	8,2	2,7	11,8	2,1	0,1	0,4	0,0	10,6	36
		K passos	23,3	17,7	1,1	12,8	37,1	2,8	0,9	3,8	0,4	34,2	259
		Dois passos	19,7	52,1	8,4	2,9	13,6	2,5	0,1	0,7	0,0	11,3	128
HQC	Todos modelos	1,2	10,0	4,6	1,5	44,8	21,5	0,3	10,4	5,7	40,8	4.096	
	Stepwise	5,2	40,4	17,5	1,8	22,9	10,0	0,1	1,3	0,7	19,8	40	
	K passos	4,8	11,2	2,9	6,7	47,3	11,2	1,3	11,3	3,3	42,4	257	
	Dois passos	4,8	37,5	16,7	1,8	24,4	10,5	0,2	2,4	1,6	19,6	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	33,7	35,8	0,0	14,0	16,5	33,7	4.096
		Stepwise	0,0	1,2	2,1	0,0	40,7	43,5	0,0	6,3	6,1	40,3	45
		K passos	0,0	0,0	0,0	0,0	43,4	26,0	0,0	17,6	13,0	43,3	227
		Dois passos	0,0	1,1	1,7	0,0	35,5	39,2	0,0	11,3	11,1	35,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,3	35,0	0,0	14,1	12,6	38,3	4.096
		Stepwise	0,0	1,3	2,4	0,0	45,4	39,9	0,0	6,2	4,7	45,0	45
		K passos	0,0	0,0	0,0	0,0	48,4	24,6	0,0	17,3	9,7	48,4	223
		Dois passos	0,0	1,2	2,0	0,0	40,2	36,5	0,0	11,4	8,8	39,7	128
	BIC	Todos modelos	0,0	0,8	0,2	0,2	83,6	9,4	0,0	5,5	0,3	83,6	4.096
		Stepwise	0,6	21,1	2,4	0,1	67,9	7,2	0,0	0,6	0,1	67,4	41
		K passos	0,5	0,7	0,1	1,5	87,9	4,2	0,0	4,7	0,4	87,8	211
		Dois passos	0,5	20,0	2,3	0,1	68,2	7,3	0,0	1,4	0,2	67,4	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,6	22,2	0,0	11,5	3,7	62,6	4.096	
	Stepwise	0,0	6,9	2,7	0,0	63,7	22,6	0,0	3,2	0,9	63,1	43	
	K passos	0,0	0,1	0,0	0,1	71,6	13,7	0,0	11,8	2,7	71,5	212	
	Dois passos	0,0	6,0	2,4	0,0	61,2	21,7	0,0	6,5	2,2	60,2	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,9	33,8	0,0	15,1	14,2	36,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	44,5	43,2	0,0	6,6	5,6	44,5	45
		K passos	0,0	0,0	0,0	0,0	44,9	24,7	0,0	18,8	11,6	44,9	229
		Dois passos	0,0	0,0	0,0	0,0	38,2	37,4	0,0	12,9	11,5	38,2	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,6	32,9	0,0	14,4	13,1	39,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	46,7	41,8	0,0	6,4	5,0	46,7	45
		K passos	0,0	0,0	0,0	0,0	47,3	23,9	0,0	18,5	10,3	47,3	228
		Dois passos	0,0	0,0	0,0	0,0	40,5	36,4	0,0	12,5	10,5	40,5	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,9	5,3	0,0	2,6	0,2	91,9	4.096
		Stepwise	0,0	0,5	0,1	0,0	93,4	5,5	0,0	0,5	0,0	93,4	42
		K passos	0,0	0,0	0,0	0,0	93,4	3,2	0,0	3,4	0,1	93,4	195
		Dois passos	0,0	0,5	0,1	0,0	92,5	5,5	0,0	1,5	0,0	92,5	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,2	19,3	0,0	8,7	2,8	69,2	4.096	
	Stepwise	0,0	0,1	0,0	0,0	75,2	21,1	0,0	2,8	0,9	75,2	43	
	K passos	0,0	0,0	0,0	0,0	76,4	11,1	0,0	10,7	1,8	76,4	202	
	Dois passos	0,0	0,1	0,0	0,0	71,8	20,3	0,0	6,2	1,6	71,8	128	

Tabela B.3: Resultados completos dos estudos de simulação para o cenário 3 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,0	0,0	6,3	19,5	27,4	3,0	15,9	27,9	14,8	4.096
		Stepwise	0,0	0,0	0,0	8,4	34,0	41,8	2,1	6,3	7,3	25,3	45
		K passos	0,0	0,0	0,0	7,7	26,9	21,8	3,9	19,8	19,9	21,7	239
		Dois passos	0,0	0,0	0,0	6,8	30,1	40,3	2,5	8,9	11,4	22,9	128
	AICc	Todos modelos	0,0	0,0	0,0	12,9	36,2	19,3	4,3	14,6	12,7	26,8	4.096
		Stepwise	0,1	0,0	0,0	16,6	47,9	26,1	2,3	4,9	2,1	36,1	43
		K passos	0,0	0,0	0,0	15,7	42,4	13,2	5,0	16,9	6,8	34,7	220
		Dois passos	0,0	0,0	0,0	14,7	45,0	26,2	2,9	7,7	3,6	34,0	128
	BIC	Todos modelos	0,0	0,0	0,0	26,7	39,5	12,8	4,1	10,1	6,8	30,1	4.096
		Stepwise	2,5	0,1	0,0	30,6	47,6	14,8	1,7	2,1	0,5	36,9	40
		K passos	0,0	0,0	0,0	29,8	44,0	7,8	4,4	11,7	2,3	36,6	208
		Dois passos	0,0	0,0	0,0	28,4	48,8	16,6	2,2	3,1	0,9	38,2	128
HQC	Todos modelos	0,0	0,0	0,0	11,3	31,6	22,0	3,8	14,4	16,9	23,8	4.096	
	Stepwise	0,3	0,1	0,1	15,3	44,6	29,7	2,2	4,7	3,1	33,6	43	
	K passos	0,0	0,0	0,0	14,3	38,8	15,2	4,7	17,2	9,8	31,4	225	
	Dois passos	0,0	0,0	0,0	13,5	41,9	30,5	2,6	6,7	4,8	31,7	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,7	29,0	33,8	0,3	15,0	21,2	28,1	4.096
		Stepwise	0,0	0,0	0,0	0,6	41,5	49,1	0,1	4,0	4,6	40,6	45
		K passos	0,0	0,0	0,0	0,7	36,9	24,9	0,2	21,0	16,2	36,2	229
		Dois passos	0,0	0,0	0,0	0,6	39,8	48,3	0,1	5,2	5,9	38,9	128
	AICc	Todos modelos	0,0	0,0	0,0	1,0	38,9	31,0	0,2	15,8	13,1	37,5	4.096
		Stepwise	0,0	0,0	0,0	1,3	51,1	41,4	0,2	3,3	2,7	49,8	44
		K passos	0,0	0,0	0,0	1,3	46,6	22,1	0,3	19,8	9,9	45,8	230
		Dois passos	0,0	0,0	0,0	1,1	49,5	41,3	0,2	4,5	3,5	48,2	128
	BIC	Todos modelos	0,0	0,0	0,0	5,9	70,3	13,5	0,4	7,9	2,0	68,1	4.096
		Stepwise	0,0	0,0	0,0	7,9	77,4	14,2	0,1	0,5	0,1	75,0	42
		K passos	0,0	0,0	0,0	7,7	74,4	8,0	0,7	8,5	0,8	73,2	201
		Dois passos	0,0	0,0	0,0	7,6	77,1	14,4	0,1	0,7	0,1	74,7	128
HQC	Todos modelos	0,0	0,0	0,0	2,3	51,0	24,6	0,5	13,9	7,7	49,1	4.096	
	Stepwise	0,0	0,0	0,0	2,6	63,1	31,5	0,2	1,7	1,0	61,2	43	
	K passos	0,0	0,0	0,0	2,6	59,1	16,4	0,5	16,4	5,1	57,8	212	
	Dois passos	0,0	0,0	0,0	2,5	62,2	31,5	0,2	2,3	1,2	60,4	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,6	36,7	0,0	14,3	16,4	32,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,7	47,9	0,0	4,1	4,3	43,7	45
		K passos	0,0	0,0	0,0	0,0	42,8	26,2	0,0	18,4	12,6	42,8	227
		Dois passos	0,0	0,0	0,0	0,0	41,0	45,8	0,0	6,6	6,6	41,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	37,0	35,3	0,0	14,6	13,1	37,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	48,6	44,2	0,0	3,8	3,3	48,6	45
		K passos	0,0	0,0	0,0	0,0	48,0	24,0	0,0	17,8	10,2	48,0	233
		Dois passos	0,0	0,0	0,0	0,0	46,0	42,6	0,0	6,1	5,3	46,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,2	84,5	9,7	0,0	5,0	0,6	84,5	4.096
		Stepwise	0,0	0,0	0,0	0,2	89,5	10,0	0,0	0,2	0,0	89,5	42
		K passos	0,0	0,0	0,0	0,2	89,8	4,8	0,0	5,0	0,3	89,8	195
		Dois passos	0,0	0,0	0,0	0,2	89,3	10,1	0,0	0,4	0,0	89,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,4	22,1	0,0	10,5	5,0	62,4	4.096	
	Stepwise	0,0	0,0	0,0	0,0	71,6	26,0	0,0	1,8	0,5	71,6	43	
	K passos	0,0	0,0	0,0	0,0	72,0	14,1	0,0	11,5	2,4	72,0	204	
	Dois passos	0,0	0,0	0,0	0,0	70,2	25,6	0,0	3,1	1,1	70,2	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	37,1	32,6	0,0	14,3	16,0	37,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	47,7	45,3	0,0	3,5	3,5	47,7	45
		K passos	0,0	0,0	0,0	0,0	44,9	24,4	0,0	19,8	10,9	44,9	218
		Dois passos	0,0	0,0	0,0	0,0	46,4	44,1	0,0	4,7	4,9	46,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,4	32,4	0,0	14,2	14,0	39,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	50,3	43,2	0,0	3,5	3,0	50,3	44
		K passos	0,0	0,0	0,0	0,0	47,6	23,3	0,0	19,3	9,8	47,6	219
		Dois passos	0,0	0,0	0,0	0,0	49,0	42,1	0,0	4,7	4,2	49,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,2	5,4	0,0	3,3	0,1	91,2	4.096
		Stepwise	0,0	0,0	0,0	0,0	94,0	5,8	0,0	0,2	0,0	94,0	42
		K passos	0,0	0,0	0,0	0,0	93,0	3,2	0,0	3,8	0,0	93,0	194
		Dois passos	0,0	0,0	0,0	0,0	93,9	5,8	0,0	0,2	0,0	93,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,3	18,8	0,0	9,1	2,8	69,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	77,3	21,4	0,0	1,0	0,3	77,3	43	
	K passos	0,0	0,0	0,0	0,0	76,1	10,8	0,0	11,5	1,6	76,1	203	
	Dois passos	0,0	0,0	0,0	0,0	76,8	21,4	0,0	1,4	0,3	76,8	128	

Tabela B.4: Resultados completos dos estudos de simulação para o cenário 4 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	1,0	1,2	0,5	22,3	18,8	13,0	9,8	15,4	18,0	4,8	4.096
		Stepwise	3,2	1,9	0,9	28,6	23,6	17,4	10,2	8,3	5,8	6,9	41
		K passos	2,3	0,6	0,2	28,4	18,8	9,9	14,6	15,0	10,3	6,8	230
		Dois passos	2,0	0,9	0,5	25,5	21,7	18,1	10,7	11,0	9,4	6,1	128
	AICc	Todos modelos	2,3	1,4	0,8	39,1	21,9	7,1	11,8	10,7	4,9	6,1	4.096
		Stepwise	6,7	2,3	0,7	43,8	23,6	7,6	10,0	4,1	1,3	7,5	39
		K passos	4,5	0,9	0,2	46,0	18,7	4,2	14,7	8,4	2,4	7,4	221
		Dois passos	4,2	1,3	0,6	41,0	24,0	9,0	11,8	5,8	2,4	7,3	128
	BIC	Todos modelos	9,0	1,3	0,7	52,6	16,8	2,8	8,9	5,9	2,0	5,6	4.096
		Stepwise	22,5	2,1	0,2	52,8	12,8	2,4	5,8	1,2	0,2	4,8	34
		K passos	13,9	0,7	0,1	58,3	12,3	1,3	9,4	3,5	0,5	5,8	209
		Dois passos	13,9	1,5	0,1	55,8	16,3	3,5	6,5	2,1	0,3	6,0	128
HQC	Todos modelos	2,8	1,3	0,9	36,9	20,4	8,3	10,6	11,5	7,3	5,5	4.096	
	Stepwise	8,0	2,4	0,6	41,6	22,4	9,6	9,1	4,5	1,9	6,9	39	
	K passos	4,9	0,9	0,3	43,5	17,8	5,6	13,9	9,4	3,7	7,1	224	
	Dois passos	4,8	1,5	0,5	39,0	23,1	11,4	10,5	6,1	3,1	7,1	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	19,0	27,0	19,6	6,8	13,0	14,6	13,8	4.096
		Stepwise	0,0	0,0	0,0	20,6	34,1	24,6	5,7	8,7	6,3	17,9	43
		K passos	0,0	0,0	0,0	22,2	29,4	14,3	8,7	15,7	9,8	18,4	227
		Dois passos	0,0	0,0	0,0	18,8	31,6	23,9	6,5	10,8	8,5	16,9	128
	AICc	Todos modelos	0,0	0,0	0,0	26,7	30,2	15,8	8,2	11,0	8,1	15,9	4.096
		Stepwise	0,0	0,0	0,0	28,2	36,6	18,5	6,7	6,6	3,5	19,9	42
		K passos	0,0	0,0	0,0	30,3	32,2	9,8	9,7	12,7	5,4	20,5	216
		Dois passos	0,0	0,0	0,0	26,4	34,6	18,1	7,4	8,4	5,0	18,6	128
	BIC	Todos modelos	0,0	0,0	0,0	65,7	23,1	2,9	4,6	3,4	0,3	13,7	4.096
		Stepwise	0,0	0,0	0,0	69,8	23,2	3,0	3,2	0,8	0,1	14,3	37
		K passos	0,0	0,0	0,0	71,4	19,1	1,7	5,1	2,4	0,4	14,4	197
		Dois passos	0,0	0,0	0,0	68,4	23,3	3,5	3,7	1,0	0,2	14,1	128
HQC	Todos modelos	0,0	0,0	0,0	38,2	30,8	11,1	7,6	9,0	3,3	16,5	4.096	
	Stepwise	0,0	0,0	0,0	42,3	34,7	12,3	5,8	3,9	1,1	19,2	40	
	K passos	0,0	0,0	0,0	44,6	29,7	6,2	9,4	7,7	2,4	20,0	209	
	Dois passos	0,0	0,0	0,0	40,6	34,1	12,5	6,3	4,9	1,6	18,6	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	5,0	32,2	32,0	2,6	13,8	14,4	27,8	4.096
		Stepwise	0,0	0,0	0,0	5,9	38,0	37,3	1,6	8,8	8,4	33,0	45
		K passos	0,0	0,0	0,0	6,5	40,3	22,3	2,3	16,5	12,1	36,8	215
		Dois passos	0,0	0,0	0,0	5,5	35,6	35,7	1,9	10,8	10,6	31,0	128
	AICc	Todos modelos	0,0	0,0	0,0	6,1	36,7	30,2	2,6	13,7	10,7	31,6	4.096
		Stepwise	0,0	0,0	0,0	7,2	41,9	34,2	1,7	8,8	6,2	36,4	44
		K passos	0,0	0,0	0,0	8,1	44,2	20,2	2,3	16,3	9,0	40,2	215
		Dois passos	0,0	0,0	0,0	6,8	39,4	32,5	1,9	11,2	8,2	34,2	128
	BIC	Todos modelos	0,0	0,0	0,0	43,2	45,5	5,9	2,1	3,1	0,2	42,0	4.096
		Stepwise	0,0	0,0	0,0	46,0	46,5	5,2	1,1	1,2	0,1	43,2	39
		K passos	0,0	0,0	0,0	46,3	46,5	2,4	1,8	2,8	0,2	44,7	195
		Dois passos	0,0	0,0	0,0	45,1	46,6	5,4	1,3	1,5	0,1	43,3	128
HQC	Todos modelos	0,0	0,0	0,0	16,4	52,9	15,9	3,0	8,9	2,9	46,5	4.096	
	Stepwise	0,0	0,0	0,0	18,7	55,7	17,6	2,2	4,4	1,4	49,0	42	
	K passos	0,0	0,0	0,0	20,1	55,9	9,6	3,1	9,0	2,3	51,7	205	
	Dois passos	0,0	0,0	0,0	17,9	54,7	17,3	2,5	5,5	2,1	48,1	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,3	37,0	32,8	0,1	14,7	15,1	36,7	4.096
		Stepwise	0,0	0,0	0,0	0,4	41,6	39,7	0,1	9,8	8,5	41,2	45
		K passos	0,0	0,0	0,0	0,5	45,1	24,3	0,1	18,5	11,6	44,9	214
		Dois passos	0,0	0,0	0,0	0,3	39,0	37,4	0,1	12,2	11,0	38,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,3	40,2	31,6	0,1	14,5	13,3	39,8	4.096
		Stepwise	0,0	0,0	0,0	0,4	43,8	38,1	0,1	9,9	7,7	43,3	45
		K passos	0,0	0,0	0,0	0,5	47,6	23,2	0,1	18,2	10,4	47,3	211
		Dois passos	0,0	0,0	0,0	0,3	41,4	36,0	0,1	12,2	10,0	41,0	128
	BIC	Todos modelos	0,0	0,0	0,0	8,5	84,4	4,6	0,2	2,3	0,0	83,8	4.096
		Stepwise	0,0	0,0	0,0	10,1	83,7	5,0	0,3	0,9	0,0	83,1	42
		K passos	0,0	0,0	0,0	10,2	83,6	3,0	0,6	2,6	0,0	83,5	193
		Dois passos	0,0	0,0	0,0	10,0	83,5	5,0	0,3	1,2	0,0	82,8	128
HQC	Todos modelos	0,0	0,0	0,0	2,1	67,8	18,5	0,3	8,3	3,0	67,1	4.096	
	Stepwise	0,0	0,0	0,0	2,4	71,0	20,5	0,3	4,6	1,2	70,4	43	
	K passos	0,0	0,0	0,0	2,6	74,0	10,9	0,3	10,4	1,7	73,6	198	
	Dois passos	0,0	0,0	0,0	2,4	69,8	20,2	0,3	5,8	1,5	69,1	128	

Tabela B.5: Resultados completos dos estudos de simulação para o cenário 5 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,0	0,0	0,0	19,4	28,5	0,1	14,0	38,0	19,2	4.096
		Stepwise	6,2	35,6	34,6	0,3	8,8	12,3	0,0	0,9	1,2	7,6	38
		K passos	1,3	0,5	0,4	5,0	28,1	17,5	3,2	23,5	20,3	24,9	309
		Dois passos	4,6	29,6	30,9	0,5	11,9	16,8	0,1	2,1	3,5	9,0	128
	AICc	Todos modelos	0,0	0,1	0,2	0,0	40,0	25,1	0,2	15,9	18,5	39,5	4.096
		Stepwise	11,3	48,0	25,2	0,5	8,3	6,1	0,1	0,3	0,2	7,5	35
		K passos	4,5	1,6	0,4	13,7	40,0	9,7	3,9	19,1	7,0	34,8	300
		Dois passos	9,3	43,0	23,9	0,6	12,4	8,7	0,2	1,1	0,8	9,9	128
	BIC	Todos modelos	0,4	0,2	0,2	0,4	54,2	19,4	0,2	13,9	11,1	53,6	4.096
		Stepwise	36,1	49,4	9,1	0,5	3,4	1,4	0,0	0,0	0,1	3,0	29
		K passos	26,2	1,6	0,1	26,0	26,3	2,8	3,6	10,7	2,7	23,0	298
		Dois passos	32,8	48,0	10,5	0,7	5,6	2,1	0,1	0,0	0,1	4,7	128
HQC	Todos modelos	0,0	0,1	0,0	0,0	33,6	27,2	0,2	15,2	23,7	33,2	4.096	
	Stepwise	15,1	48,4	23,4	0,5	6,5	5,7	0,0	0,2	0,3	5,7	34	
	K passos	6,3	1,3	0,4	14,2	34,0	10,3	5,2	19,0	9,2	29,5	305	
	Dois passos	11,7	43,8	23,4	0,5	9,8	8,9	0,2	0,9	0,8	7,9	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	0,0	29,5	34,1	0,0	14,6	21,8	29,5	4.096
		Stepwise	0,0	3,8	8,5	0,0	38,9	45,6	0,0	1,4	1,9	38,3	45
		K passos	0,0	0,0	0,0	0,0	37,0	26,3	0,0	20,1	16,6	37,0	245
		Dois passos	0,0	3,1	6,7	0,0	37,7	45,0	0,0	3,2	4,3	36,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,0	32,1	0,0	15,5	13,4	39,0	4.096
		Stepwise	0,0	5,5	9,8	0,0	46,0	36,9	0,0	0,9	1,0	45,1	108
		K passos	0,0	0,0	0,0	0,1	46,6	23,1	0,0	19,4	10,7	46,6	243
		Dois passos	0,0	4,9	7,7	0,0	45,7	37,0	0,0	2,4	2,3	44,8	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	74,3	14,5	0,0	9,1	2,1	74,3	4.096
		Stepwise	3,6	26,9	15,2	0,0	45,6	8,5	0,0	0,2	0,0	45,2	40
		K passos	0,0	0,0	0,0	5,9	71,9	11,7	0,1	9,0	1,5	70,3	239
		Dois passos	1,8	25,6	13,2	0,0	49,4	9,6	0,0	0,3	0,0	48,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	50,4	27,0	0,0	14,7	7,9	50,4	4.096	
	Stepwise	0,2	11,0	13,1	0,0	49,4	25,4	0,0	0,4	0,4	48,8	43	
	K passos	0,0	0,0	0,0	0,7	57,9	19,3	0,0	15,5	6,5	57,5	238	
	Dois passos	0,1	10,1	11,0	0,0	50,5	26,2	0,0	1,1	0,9	49,8	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,4	35,1	0,0	15,1	17,4	32,4	4.096
		Stepwise	0,0	0,0	0,2	0,0	46,4	50,5	0,0	1,3	1,6	46,4	45
		K passos	0,0	0,0	0,0	0,0	42,8	26,1	0,0	19,7	11,5	42,8	232
		Dois passos	0,0	0,0	0,0	0,0	42,3	47,2	0,0	5,2	5,3	42,3	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	37,1	33,2	0,0	15,3	14,4	37,1	4.096
		Stepwise	0,0	0,1	0,2	0,0	51,6	45,7	0,0	1,4	1,1	51,6	45
		K passos	0,0	0,0	0,0	0,0	48,2	23,9	0,0	19,2	8,8	48,2	226
		Dois passos	0,0	0,0	0,0	0,0	47,7	43,3	0,0	5,1	3,9	47,7	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	84,7	9,2	0,0	5,5	0,6	84,7	4.096
		Stepwise	0,0	2,1	0,5	0,0	87,9	9,5	0,0	0,0	0,0	87,9	42
		K passos	0,0	0,0	0,0	0,1	89,7	4,6	0,0	5,2	0,4	89,7	198
		Dois passos	0,0	1,0	0,2	0,0	88,6	9,7	0,0	0,5	0,0	88,5	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,7	21,6	0,0	11,1	4,6	62,7	4.096	
	Stepwise	0,0	0,6	0,2	0,0	72,9	25,8	0,0	0,3	0,2	72,8	43	
	K passos	0,0	0,0	0,0	0,0	71,5	13,9	0,0	12,1	2,4	71,5	206	
	Dois passos	0,0	0,3	0,0	0,0	71,6	25,3	0,0	1,9	1,0	71,5	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	37,1	33,6	0,0	14,6	14,7	37,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	49,3	48,0	0,0	1,5	1,2	49,3	45
		K passos	0,0	0,0	0,0	0,0	44,5	24,3	0,0	20,3	11,0	44,5	223
		Dois passos	0,0	0,0	0,0	0,0	45,8	44,9	0,0	4,7	4,6	45,8	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	39,8	32,0	0,0	14,1	14,1	39,8	4.096
		Stepwise	0,0	0,0	0,0	0,0	51,9	45,8	0,0	1,3	1,0	51,9	45
		K passos	0,0	0,0	0,0	0,0	46,6	23,6	0,0	19,8	10,0	46,6	225
		Dois passos	0,0	0,0	0,0	0,0	48,3	43,0	0,0	4,7	3,9	48,3	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,0	5,3	0,0	3,5	0,2	91,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	94,3	5,7	0,0	0,0	0,0	94,3	42
		K passos	0,0	0,0	0,0	0,0	93,1	3,0	0,0	3,9	0,0	93,1	195
		Dois passos	0,0	0,0	0,0	0,0	94,2	5,7	0,0	0,1	0,0	94,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	68,3	19,0	0,0	9,4	3,3	68,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	78,1	21,5	0,0	0,3	0,1	78,1	43	
	K passos	0,0	0,0	0,0	0,0	76,2	10,6	0,0	11,4	1,8	76,2	202	
	Dois passos	0,0	0,0	0,0	0,0	76,9	21,3	0,0	1,4	0,4	76,9	128	

Tabela B.6: Resultados completos dos estudos de simulação para o cenário 6 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	10,8	18,6	18,6	4,9	9,3	16,8	2,4	5,6	13,0	1,1	4.096
		Stepwise	30,2	39,5	25,7	1,3	0,7	0,8	0,9	0,4	0,4	0,0	31
		K passos	24,9	16,6	6,2	14,6	9,6	5,5	10,0	7,7	4,9	0,9	305
		Dois passos	21,7	26,9	19,2	6,7	6,2	4,4	6,7	4,3	4,0	0,1	128
	AICc	Todos modelos	24,2	27,9	16,1	7,0	8,2	7,3	2,0	3,2	4,1	1,1	4.096
		Stepwise	41,0	40,4	16,0	1,4	0,2	0,1	0,6	0,3	0,0	0,0	29
		K passos	43,5	16,1	2,3	17,4	6,2	1,6	8,2	3,7	1,0	0,6	283
		Dois passos	32,9	29,0	12,6	9,3	4,2	1,4	6,8	2,5	1,2	0,0	128
	BIC	Todos modelos	47,7	28,5	7,6	4,5	4,3	2,8	1,2	1,3	2,1	0,7	4.096
		Stepwise	66,9	28,6	4,0	0,4	0,0	0,0	0,1	0,0	0,0	0,0	24
		K passos	72,2	7,1	0,5	12,7	1,6	0,3	4,7	0,7	0,2	0,1	258
		Dois passos	60,8	22,7	3,6	6,8	1,0	0,3	4,1	0,6	0,1	0,0	128
HQC	Todos modelos	25,2	25,7	14,9	6,0	7,9	8,7	1,8	3,9	5,9	0,9	4.096	
	Stepwise	46,2	38,5	13,2	1,2	0,1	0,1	0,4	0,2	0,1	0,0	27	
	K passos	46,5	14,3	2,1	16,2	5,4	1,9	8,3	3,8	1,5	0,5	280	
	Dois passos	37,2	28,3	11,6	8,2	3,2	1,6	6,1	2,4	1,4	0,0	128	
n = 100	AIC	Todos modelos	3,8	8,6	6,9	7,1	23,5	22,0	4,2	9,4	14,5	11,9	4.096
		Stepwise	20,6	41,4	33,5	0,7	2,1	1,5	0,1	0,1	0,1	0,7	36
		K passos	8,9	9,6	3,8	19,3	21,6	9,9	9,1	11,7	6,2	10,2	319
		Dois passos	19,1	38,7	32,7	1,1	2,9	2,5	0,5	1,1	1,3	0,6	128
	AICc	Todos modelos	6,1	11,2	8,0	10,6	26,1	17,7	3,7	8,9	7,7	13,7	4.096
		Stepwise	26,1	43,4	27,5	0,6	1,4	0,8	0,1	0,0	0,0	0,5	35
		K passos	12,6	11,7	3,2	24,7	21,8	6,7	7,8	8,6	2,9	10,6	304
		Dois passos	24,5	41,4	26,8	1,1	2,2	1,6	0,6	1,0	0,9	0,4	128
	BIC	Todos modelos	34,5	22,1	3,2	17,4	15,8	3,0	1,3	2,3	0,4	9,6	4.096
		Stepwise	67,1	28,9	4,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	27
		K passos	55,7	6,5	0,3	28,3	4,7	0,4	2,8	1,0	0,2	2,0	288
		Dois passos	65,3	28,7	4,3	0,6	0,2	0,1	0,4	0,3	0,0	0,0	128
HQC	Todos modelos	12,3	17,8	6,2	14,2	23,7	12,1	3,3	7,1	3,3	13,1	4.096	
	Stepwise	42,5	41,5	15,4	0,3	0,3	0,1	0,0	0,0	0,0	0,1	31	
	K passos	26,1	11,1	2,2	31,5	14,2	2,7	6,3	4,9	1,1	6,4	300	
	Dois passos	40,5	40,3	15,4	0,9	0,8	0,5	0,5	0,6	0,4	0,1	128	
n = 200	AIC	Todos modelos	0,2	2,2	1,7	2,5	31,7	33,1	1,0	13,3	14,3	28,9	4.096
		Stepwise	5,3	33,2	30,3	1,2	14,7	13,8	0,0	0,6	0,8	11,8	41
		K passos	1,2	2,2	1,3	10,3	37,1	19,6	3,6	16,1	8,5	31,3	282
		Dois passos	4,7	31,0	28,6	1,3	16,1	14,9	0,1	1,5	1,8	11,6	128
	AICc	Todos modelos	0,3	2,5	2,2	3,1	36,3	30,1	0,8	13,8	10,9	32,7	4.096
		Stepwise	6,5	36,6	29,0	1,6	14,0	11,4	0,1	0,3	0,5	11,5	41
		K passos	1,5	2,9	1,4	12,4	39,9	17,0	3,9	14,9	6,2	33,5	279
		Dois passos	5,9	34,3	27,7	1,6	15,6	12,4	0,1	1,1	1,2	11,6	128
	BIC	Todos modelos	9,7	14,2	1,5	18,6	48,6	3,9	0,8	2,5	0,2	45,3	4.096
		Stepwise	57,5	37,7	4,0	0,4	0,4	0,1	0,0	0,0	0,0	0,4	30
		K passos	24,1	4,2	0,1	54,3	13,0	0,8	1,9	1,5	0,1	11,2	295
		Dois passos	55,6	39,1	4,1	0,5	0,6	0,1	0,0	0,0	0,0	0,5	128
HQC	Todos modelos	1,8	6,8	2,0	7,5	52,2	15,4	1,1	10,1	3,1	47,0	4.096	
	Stepwise	24,1	50,8	15,6	1,2	6,4	1,9	0,0	0,0	0,0	5,4	36	
	K passos	6,6	5,2	0,8	33,5	36,0	5,5	4,4	7,0	1,1	30,4	280	
	Dois passos	22,0	49,4	15,9	1,7	8,5	2,4	0,0	0,1	0,1	6,9	128	
n = 400	AIC	Todos modelos	0,0	0,2	0,1	0,1	35,8	32,1	0,0	16,2	15,5	35,7	4.096
		Stepwise	0,1	15,6	14,9	0,2	33,2	33,0	0,0	1,3	1,7	31,8	44
		K passos	0,0	0,2	0,2	0,6	43,2	24,6	0,1	19,8	11,4	42,6	242
		Dois passos	0,1	14,7	14,0	0,2	32,6	32,8	0,0	2,7	2,8	30,6	128
	AICc	Todos modelos	0,0	0,2	0,1	0,1	39,2	30,1	0,0	15,7	14,6	39,0	4.096
		Stepwise	0,1	17,1	14,9	0,2	34,5	30,5	0,0	1,2	1,5	33,1	43
		K passos	0,0	0,2	0,2	0,7	46,0	23,1	0,1	19,7	10,0	45,4	243
		Dois passos	0,1	16,1	14,0	0,2	34,3	30,4	0,0	2,5	2,4	32,0	128
	BIC	Todos modelos	0,6	3,9	0,1	4,0	82,8	4,8	0,0	3,6	0,2	82,2	4.096
		Stepwise	9,8	72,3	5,1	0,8	11,5	0,5	0,0	0,0	0,0	11,4	38
		K passos	1,8	2,5	0,1	29,9	60,2	1,5	1,0	2,9	0,0	58,5	266
		Dois passos	9,3	71,6	5,0	0,9	12,6	0,5	0,0	0,0	0,0	12,4	128
HQC	Todos modelos	0,0	0,8	0,1	0,9	68,3	17,5	0,0	9,7	2,7	67,7	4.096	
	Stepwise	1,2	43,9	13,6	0,5	32,4	8,2	0,0	0,1	0,1	31,7	41	
	K passos	0,1	0,9	0,1	5,2	70,7	9,9	0,7	10,8	1,6	69,5	245	
	Dois passos	1,1	42,9	13,0	0,5	33,4	8,5	0,0	0,4	0,2	32,1	128	

Tabela B.7: Resultados completos dos estudos de simulação para o cenário 7 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 3$			# $\beta = 3$			# $\beta > 3$			Modelo Correto	Média de Modelos
			# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$	# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$	# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$		
n = 50	AIC	Todos modelos	0,3	3,0	4,5	0,7	10,4	17,0	0,6	13,1	50,4	7,7	4.096
		Stepwise	0,9	13,7	17,8	1,4	19,4	27,1	1,2	6,8	11,8	11,6	34
		K passos	3,3	5,7	2,3	3,2	20,8	10,4	2,4	26,6	25,3	14,4	275
		Dois passos	0,8	10,2	13,2	1,1	18,2	27,2	1,1	9,9	18,4	10,3	128
	AICc	Todos modelos	0,6	6,1	6,9	1,5	18,8	19,9	0,7	16,1	29,4	14,2	4.096
		Stepwise	2,1	21,1	17,6	3,1	23,6	19,9	1,5	6,1	5,1	14,6	31
		K passos	6,4	9,9	2,0	5,8	31,1	8,0	2,8	23,7	10,4	22,2	252
		Dois passos	1,8	16,5	14,1	2,8	24,5	21,1	1,6	9,1	8,4	14,0	128
	BIC	Todos modelos	8,2	19,0	6,5	4,2	27,8	12,2	0,6	10,6	10,9	21,7	4.096
		Stepwise	15,8	40,0	11,0	5,4	18,4	6,3	1,1	1,4	0,6	12,2	25
		K passos	31,3	14,6	0,8	9,1	27,9	2,3	1,9	10,3	1,7	22,2	223
		Dois passos	14,2	37,1	10,7	5,6	20,0	7,6	1,2	2,5	1,1	12,6	128
HQC	Todos modelos	1,2	7,7	6,3	2,0	18,6	18,5	0,6	14,9	30,2	14,2	4.096	
	Stepwise	4,2	25,3	17,6	3,4	22,7	16,9	1,3	4,5	4,2	14,2	30	
	K passos	10,0	11,2	2,0	6,5	30,4	6,5	2,4	21,9	9,2	22,1	245	
	Dois passos	3,6	21,3	14,7	3,3	23,6	18,5	1,5	6,9	6,6	13,9	128	
n = 100	AIC	Todos modelos	0,0	0,2	0,5	0,0	20,1	24,3	0,0	17,8	37,1	19,2	4.096
		Stepwise	0,1	1,9	2,5	0,2	26,4	39,2	0,1	11,9	17,7	24,6	36
		K passos	0,2	0,5	0,2	0,4	30,9	14,0	0,4	30,0	23,3	30,2	254
		Dois passos	0,1	1,3	1,6	0,2	22,9	34,4	0,1	15,7	23,8	20,9	128
	AICc	Todos modelos	0,0	0,3	0,4	0,2	26,3	26,0	0,1	18,9	27,8	25,4	4.096
		Stepwise	0,1	2,6	2,5	0,3	32,8	36,6	0,2	12,0	12,9	30,6	35
		K passos	0,3	0,7	0,2	0,6	38,4	13,2	0,5	29,4	16,7	37,5	239
		Dois passos	0,1	1,7	1,6	0,3	28,3	33,1	0,2	16,7	18,1	26,0	128
	BIC	Todos modelos	0,6	4,3	0,7	2,7	64,6	12,7	0,4	10,4	3,6	61,8	4.096
		Stepwise	1,8	16,6	3,0	3,0	60,1	11,6	0,4	2,8	0,7	56,1	30
		K passos	4,4	3,7	0,2	4,5	71,6	3,1	0,6	10,9	0,9	69,5	212
		Dois passos	1,5	14,7	2,8	2,9	60,5	11,8	0,5	4,4	1,0	55,7	128
HQC	Todos modelos	0,1	1,4	0,5	0,5	41,9	22,9	0,4	17,0	15,3	40,6	4.096	
	Stepwise	0,3	6,4	3,4	0,8	48,0	27,4	0,4	7,9	5,4	44,3	33	
	K passos	1,0	1,7	0,2	1,6	55,3	8,8	0,7	24,1	6,7	53,5	231	
	Dois passos	0,3	5,4	2,7	0,8	45,0	26,6	0,4	11,6	7,1	41,0	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	18,9	29,3	0,0	20,5	31,3	18,9	4.096
		Stepwise	0,0	0,0	0,1	0,0	28,6	39,6	0,0	12,4	19,2	28,5	36
		K passos	0,0	0,0	0,0	0,0	32,5	14,9	0,0	31,4	21,2	32,5	252
		Dois passos	0,0	0,0	0,0	0,0	23,1	33,5	0,0	17,7	25,6	23,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	21,4	29,2	0,0	21,9	27,5	21,4	4.096
		Stepwise	0,0	0,0	0,1	0,0	31,6	38,7	0,0	12,6	17,0	31,5	36
		K passos	0,0	0,0	0,0	0,0	36,2	14,1	0,0	31,5	18,1	36,2	246
		Dois passos	0,0	0,0	0,0	0,0	25,9	33,1	0,0	18,1	22,9	25,8	128
	BIC	Todos modelos	0,0	0,2	0,1	0,1	80,7	9,6	0,0	7,5	1,8	80,6	4.096
		Stepwise	0,0	0,9	0,2	0,1	85,2	9,8	0,0	3,1	0,7	84,9	31
		K passos	0,1	0,1	0,0	0,1	87,8	2,4	0,0	9,1	0,3	87,8	199
		Dois passos	0,0	0,8	0,2	0,1	83,2	9,7	0,0	5,2	0,8	83,0	128
HQC	Todos modelos	0,0	0,0	0,1	0,0	50,6	22,5	0,0	19,0	7,8	50,6	4.096	
	Stepwise	0,0	0,3	0,1	0,0	60,6	25,5	0,0	9,4	4,1	60,5	33	
	K passos	0,0	0,0	0,0	0,0	64,5	8,2	0,0	23,6	3,7	64,4	212	
	Dois passos	0,0	0,2	0,0	0,0	55,8	24,0	0,0	14,1	5,9	55,7	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	22,9	26,4	0,0	23,2	27,5	22,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	30,3	38,8	0,0	13,1	17,8	30,3	36
		K passos	0,0	0,0	0,0	0,0	33,0	15,0	0,0	32,3	19,7	33,0	240
		Dois passos	0,0	0,0	0,0	0,0	24,5	32,6	0,0	18,8	24,1	24,5	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	23,7	26,3	0,0	24,0	26,0	23,7	4.096
		Stepwise	0,0	0,0	0,0	0,0	31,7	38,4	0,0	13,3	16,6	31,7	36
		K passos	0,0	0,0	0,0	0,0	34,7	14,8	0,0	32,4	18,1	34,7	239
		Dois passos	0,0	0,0	0,0	0,0	25,9	32,3	0,0	18,8	23,0	25,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	88,0	6,1	0,0	5,6	0,3	88,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	91,5	6,8	0,0	1,6	0,1	91,5	31
		K passos	0,0	0,0	0,0	0,0	92,0	1,8	0,0	6,1	0,1	92,0	196
		Dois passos	0,0	0,0	0,0	0,0	90,0	6,7	0,0	3,1	0,2	90,0	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	58,7	20,0	0,0	15,6	5,7	58,7	4.096	
	Stepwise	0,0	0,0	0,0	0,0	66,2	23,6	0,0	7,6	2,7	66,2	32	
	K passos	0,0	0,0	0,0	0,0	70,2	7,4	0,0	19,8	2,6	70,2	206	
	Dois passos	0,0	0,0	0,0	0,0	61,7	21,9	0,0	12,0	4,4	61,7	128	

Tabela B.8: Resultados completos dos estudos de simulação para o cenário 8 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	6,1	5,0	6,8	4,3	14,9	20,3	3,4	10,0	29,2	7,4	4.096
		Stepwise	18,4	23,1	17,9	6,9	13,1	12,7	1,6	3,0	3,4	6,1	40
		K passos	19,3	8,6	4,1	9,8	18,0	11,4	5,2	12,1	11,4	9,2	255
		Dois passos	14,5	19,2	17,4	6,4	14,1	16,5	1,9	4,3	5,6	6,5	128
	AICc	Todos modelos	12,6	12,4	7,2	9,5	18,7	13,5	5,2	9,9	11,0	9,0	4.096
		Stepwise	33,1	26,4	11,0	9,7	11,3	4,6	1,7	1,7	0,6	5,9	37
		K passos	34,4	11,5	2,4	15,7	16,2	4,9	4,8	7,5	2,5	8,9	251
		Dois passos	26,8	24,5	11,1	10,8	13,9	6,6	2,2	3,2	1,0	7,0	128
	BIC	Todos modelos	32,2	14,2	4,9	12,2	15,1	6,7	3,8	5,5	5,4	7,0	4.096
		Stepwise	65,1	17,5	3,6	7,8	4,0	0,9	0,8	0,4	0,0	2,1	30
		K passos	65,8	6,1	0,8	13,9	7,1	0,9	2,8	2,3	0,4	4,2	229
		Dois passos	57,4	18,6	4,3	9,9	5,8	1,8	1,3	0,7	0,1	3,2	128
HQC	Todos modelos	13,2	10,3	6,7	8,4	16,1	14,4	4,6	9,7	16,6	8,3	4.096	
	Stepwise	36,7	25,5	10,7	8,5	9,8	5,0	1,4	1,6	0,8	5,2	36	
	K passos	37,0	10,4	2,5	14,3	14,4	5,5	4,2	7,7	3,9	7,8	246	
	Dois passos	30,0	23,8	11,3	9,3	12,0	7,5	1,8	2,8	1,5	6,3	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	1,4	28,1	33,9	0,8	15,5	20,3	26,7	4.096
		Stepwise	0,5	1,5	0,9	2,3	38,5	42,3	0,6	5,8	7,6	35,7	45
		K passos	0,0	0,0	0,0	2,7	35,1	24,7	1,4	19,9	16,2	32,9	237
		Dois passos	0,3	1,2	0,9	1,8	32,9	37,5	0,8	10,8	13,7	30,4	128
	AICc	Todos modelos	0,0	0,0	0,0	2,3	35,4	32,5	1,3	14,5	14,0	33,0	4.096
		Stepwise	1,1	2,1	0,8	3,4	46,6	35,6	0,6	5,2	4,7	43,4	44
		K passos	0,0	0,0	0,0	4,5	43,9	21,2	1,5	18,8	10,1	41,4	229
		Dois passos	0,8	1,6	0,7	2,9	41,0	32,1	1,2	10,7	9,0	38,0	128
	BIC	Todos modelos	0,2	0,0	0,0	11,9	64,7	13,0	1,3	6,7	2,2	60,7	4.096
		Stepwise	11,4	3,2	0,4	12,6	59,6	11,0	0,2	1,3	0,3	55,6	41
		K passos	1,7	0,1	0,0	21,4	60,5	6,2	1,7	7,3	1,0	57,3	216
		Dois passos	9,4	2,7	0,6	12,1	59,7	11,2	0,8	2,7	0,8	55,7	128
HQC	Todos modelos	0,1	0,0	0,0	5,0	46,9	25,8	1,4	13,0	7,8	43,5	4.096	
	Stepwise	2,5	3,0	0,9	5,8	56,0	26,1	0,4	3,3	2,1	51,7	43	
	K passos	0,2	0,0	0,0	8,1	54,5	14,9	1,7	15,6	4,9	51,0	221	
	Dois passos	2,2	2,6	0,8	5,2	52,4	24,9	1,0	7,0	3,9	48,2	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	29,4	31,7	0,0	17,4	21,5	29,4	4.096
		Stepwise	0,0	0,0	0,0	0,1	42,0	46,8	0,0	5,0	6,0	42,0	46
		K passos	0,0	0,0	0,0	0,0	37,1	25,1	0,1	20,7	17,0	37,1	241
		Dois passos	0,0	0,0	0,0	0,1	34,2	39,2	0,0	12,5	14,1	34,2	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	33,2	31,9	0,0	17,3	17,6	33,2	4.096
		Stepwise	0,0	0,0	0,0	0,1	46,4	44,0	0,0	4,5	5,0	46,3	45
		K passos	0,0	0,0	0,0	0,0	42,1	23,6	0,1	19,8	14,3	42,1	241
		Dois passos	0,0	0,0	0,0	0,1	38,4	37,4	0,0	12,3	11,9	38,3	128
	BIC	Todos modelos	0,0	0,0	0,0	1,7	82,5	9,8	0,1	5,1	0,8	82,4	4.096
		Stepwise	0,0	0,1	0,0	1,4	86,9	10,4	0,0	1,1	0,1	86,8	43
		K passos	0,0	0,0	0,0	1,3	85,9	5,3	0,1	6,7	0,7	85,8	198
		Dois passos	0,0	0,1	0,0	1,3	84,6	10,2	0,1	3,4	0,4	84,4	128
HQC	Todos modelos	0,0	0,0	0,0	0,1	59,2	22,1	0,0	12,1	6,5	59,2	4.096	
	Stepwise	0,0	0,0	0,0	0,3	68,2	27,6	0,0	2,9	1,0	68,1	44	
	K passos	0,0	0,0	0,0	0,3	65,6	14,9	0,1	15,1	4,0	65,6	216	
	Dois passos	0,0	0,0	0,0	0,3	62,7	25,7	0,0	8,2	3,1	62,7	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	30,9	34,4	0,0	15,8	18,9	30,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,4	47,1	0,0	4,4	5,2	43,4	45
		K passos	0,0	0,0	0,0	0,0	40,2	27,2	0,0	18,3	14,3	40,2	223
		Dois passos	0,0	0,0	0,0	0,0	38,6	42,8	0,0	8,9	9,7	38,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	32,6	34,7	0,0	16,0	16,7	32,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	45,6	45,8	0,0	4,2	4,5	45,6	45
		K passos	0,0	0,0	0,0	0,0	42,7	26,8	0,0	17,7	12,8	42,7	223
		Dois passos	0,0	0,0	0,0	0,0	41,0	41,6	0,0	8,6	8,9	41,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	89,4	7,0	0,0	3,4	0,2	89,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	92,8	7,0	0,0	0,1	0,0	92,8	42
		K passos	0,0	0,0	0,0	0,0	92,4	4,1	0,0	3,5	0,1	92,4	195
		Dois passos	0,0	0,0	0,0	0,0	92,3	7,0	0,0	0,6	0,1	92,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	63,8	21,4	0,0	11,6	3,2	63,8	4.096	
	Stepwise	0,0	0,0	0,0	0,0	74,9	22,7	0,0	1,8	0,6	74,9	43	
	K passos	0,0	0,0	0,0	0,0	72,2	14,2	0,0	11,5	2,0	72,2	203	
	Dois passos	0,0	0,0	0,0	0,0	72,8	22,1	0,0	3,8	1,3	72,8	128	

Tabela B.9: Resultados completos dos estudos de simulação para o cenário 9 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	1,4	2,8	3,3	4,8	25,6	21,9	1,7	15,3	23,2	10,7	4.096
		Stepwise	17,8	25,5	10,4	7,2	21,8	11,9	0,8	2,7	1,9	8,5	40
		K passos	6,7	4,0	1,7	11,4	30,5	13,0	4,2	18,0	10,4	15,3	267
		Dois passos	9,2	17,6	10,1	6,7	27,6	19,8	1,1	4,4	3,5	9,4	128
	AICc	Todos modelos	4,7	6,9	3,2	9,5	38,9	14,8	2,6	11,3	8,1	17,2	4.096
		Stepwise	32,1	29,3	6,1	8,7	17,4	4,4	0,5	1,1	0,4	6,8	36
		K passos	16,2	6,4	1,0	20,3	33,5	5,8	3,7	10,6	2,6	18,4	271
		Dois passos	18,9	23,3	7,8	11,2	26,4	8,4	1,1	2,1	0,8	9,6	128
	BIC	Todos modelos	15,2	10,5	1,9	11,4	39,7	9,5	1,8	6,6	3,4	17,5	4.096
		Stepwise	65,2	20,7	2,0	6,0	5,1	0,7	0,2	0,1	0,0	2,1	30
		K passos	42,3	5,3	0,4	23,8	19,0	2,0	2,0	4,8	0,4	10,8	264
		Dois passos	46,8	23,0	3,7	10,1	12,9	2,5	0,3	0,5	0,1	4,8	128
HQC	Todos modelos	4,8	6,4	3,2	8,5	34,7	15,9	2,3	12,4	11,8	15,5	4.096	
	Stepwise	37,1	28,0	6,0	7,4	14,6	4,8	0,5	1,1	0,4	5,8	36	
	K passos	17,7	5,9	1,2	19,2	30,2	6,5	3,8	11,6	3,8	16,6	269	
	Dois passos	21,9	23,2	8,0	9,7	24,1	9,4	0,9	2,0	0,8	8,6	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,1	0,4	37,8	24,8	0,2	19,2	17,5	26,6	4.096
		Stepwise	0,3	1,8	1,1	0,6	51,6	33,1	0,1	6,3	5,1	34,7	45
		K passos	0,1	0,1	0,0	0,9	45,2	19,7	0,1	21,4	12,5	35,4	236
		Dois passos	0,1	1,3	0,9	0,4	47,5	34,4	0,1	8,2	7,0	33,0	128
	AICc	Todos modelos	0,0	0,1	0,1	0,7	47,1	22,3	0,2	18,2	11,3	32,8	4.096
		Stepwise	0,5	3,0	1,2	0,9	59,6	26,4	0,1	5,5	2,7	40,4	44
		K passos	0,1	0,1	0,1	1,3	55,2	15,6	0,2	20,1	7,3	43,3	227
		Dois passos	0,2	2,0	0,9	0,9	57,0	27,9	0,1	7,3	3,6	39,3	128
	BIC	Todos modelos	0,3	0,8	0,0	5,0	78,0	7,5	0,5	6,3	1,6	54,4	4.096
		Stepwise	9,8	18,6	1,3	5,9	58,1	5,3	0,1	0,8	0,2	38,7	40
		K passos	2,5	1,5	0,0	10,1	74,3	3,6	0,6	7,1	0,4	57,8	226
		Dois passos	4,1	14,1	0,7	5,3	67,8	6,7	0,1	1,0	0,1	46,6	128
HQC	Todos modelos	0,0	0,4	0,0	1,3	60,6	17,4	0,6	13,9	5,8	41,9	4.096	
	Stepwise	1,4	7,4	1,6	2,1	66,2	17,1	0,1	2,9	1,1	44,3	43	
	K passos	0,3	0,6	0,0	2,8	65,8	11,1	0,7	15,6	3,1	51,9	229	
	Dois passos	0,8	5,2	1,1	2,0	66,4	19,0	0,1	4,3	1,1	45,4	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	35,9	33,0	0,0	14,4	16,7	34,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	47,9	41,2	0,0	5,8	5,1	45,9	45
		K passos	0,0	0,0	0,0	0,0	45,5	23,5	0,0	18,6	12,4	43,9	226
		Dois passos	0,0	0,0	0,0	0,0	43,1	39,7	0,0	8,9	8,3	41,3	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	41,7	30,0	0,0	14,7	13,6	39,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	52,6	38,0	0,0	5,5	3,9	50,0	45
		K passos	0,0	0,0	0,0	0,0	50,7	21,8	0,0	18,2	9,3	48,8	221
		Dois passos	0,0	0,0	0,0	0,0	48,2	36,7	0,0	8,5	6,5	46,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,3	87,6	6,3	0,0	5,6	0,2	81,4	4.096
		Stepwise	0,0	0,8	0,0	0,2	92,2	6,0	0,0	0,7	0,1	85,2	42
		K passos	0,0	0,0	0,0	0,1	91,0	3,4	0,0	5,2	0,2	87,4	197
		Dois passos	0,0	0,2	0,0	0,1	91,9	6,4	0,0	1,3	0,1	85,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	64,1	20,3	0,0	12,1	3,5	60,0	4.096	
	Stepwise	0,0	0,0	0,0	0,0	75,6	20,1	0,0	3,3	0,9	70,7	43	
	K passos	0,0	0,0	0,0	0,0	73,7	11,8	0,0	12,4	2,0	70,7	204	
	Dois passos	0,0	0,0	0,0	0,0	73,0	20,3	0,0	5,0	1,7	68,4	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	36,2	34,2	0,0	13,9	15,7	36,1	4.096
		Stepwise	0,0	0,0	0,0	0,0	44,8	42,0	0,0	7,5	5,8	44,7	45
		K passos	0,0	0,0	0,0	0,0	44,2	25,8	0,0	18,8	11,2	44,2	223
		Dois passos	0,0	0,0	0,0	0,0	39,8	39,5	0,0	10,9	9,8	39,8	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	38,7	33,8	0,0	14,0	13,5	38,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	47,3	40,7	0,0	7,1	4,9	47,2	45
		K passos	0,0	0,0	0,0	0,0	47,2	24,7	0,0	18,1	10,0	47,1	223
		Dois passos	0,0	0,0	0,0	0,0	42,0	38,6	0,0	10,7	8,7	42,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	90,5	5,6	0,0	3,4	0,5	89,4	4.096
		Stepwise	0,0	0,0	0,0	0,0	93,8	5,5	0,0	0,6	0,1	92,4	42
		K passos	0,0	0,0	0,0	0,0	93,4	3,2	0,0	3,2	0,2	93,3	194
		Dois passos	0,0	0,0	0,0	0,0	93,3	5,6	0,0	1,0	0,1	91,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	68,0	19,2	0,0	9,6	3,2	67,8	4.096	
	Stepwise	0,0	0,0	0,0	0,0	75,8	20,5	0,0	2,9	0,7	75,3	43	
	K passos	0,0	0,0	0,0	0,0	76,2	11,6	0,0	10,2	2,0	76,0	201	
	Dois passos	0,0	0,0	0,0	0,0	73,2	20,2	0,0	5,1	1,5	72,6	128	

Tabela B.10: Resultados completos dos estudos de simulação para o cenário 10 da distribuição Gaussiana Inversa.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$		
n = 50	AIC	Todos modelos	4,9	7,7	6,2	6,9	21,1	16,8	3,5	14,1	18,8	12,9	4.096
		Stepwise	39,4	30,4	15,4	4,3	4,7	3,3	1,1	0,8	0,8	2,1	38
		K passos	22,7	7,7	2,1	18,8	14,9	5,4	11,2	10,8	6,5	7,3	308
		Dois passos	31,6	26,8	14,2	6,4	9,0	5,6	2,2	2,3	2,0	3,0	128
	AICc	Todos modelos	14,9	14,0	6,4	14,7	22,7	9,0	4,6	8,5	5,2	14,8	4.096
		Stepwise	58,0	27,3	6,1	4,7	2,5	0,5	0,8	0,2	0,0	1,2	35
		K passos	45,8	6,5	0,5	23,8	8,5	1,4	8,2	4,3	0,9	4,4	285
		Dois passos	50,0	25,9	6,8	8,1	5,1	1,5	1,8	0,8	0,1	2,0	128
	BIC	Todos modelos	37,8	13,8	2,5	14,7	16,5	3,9	2,5	5,9	2,4	11,1	4.096
		Stepwise	84,1	12,2	1,6	1,5	0,3	0,0	0,2	0,0	0,0	0,1	28
		K passos	78,0	2,2	0,2	14,0	1,6	0,1	3,1	0,9	0,0	0,8	260
		Dois passos	79,1	13,8	1,9	3,4	1,0	0,1	0,6	0,2	0,0	0,5	128
HQC	Todos modelos	14,0	12,0	5,9	12,3	21,1	12,2	4,0	9,3	9,2	13,6	4.096	
	Stepwise	61,2	25,7	6,0	3,7	2,0	0,5	0,6	0,2	0,0	1,0	33	
	K passos	46,1	6,0	0,9	22,0	8,2	2,1	8,1	4,9	1,7	4,2	286	
	Dois passos	53,5	25,0	6,7	6,5	4,0	1,6	1,6	0,8	0,2	1,6	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,1	0,5	34,9	27,4	0,3	19,2	17,6	33,8	4.096
		Stepwise	3,6	15,5	11,7	1,0	35,3	26,3	0,2	3,4	3,0	31,8	47
		K passos	0,0	0,1	0,1	4,7	38,2	20,4	2,3	21,5	12,8	35,3	250
		Dois passos	2,7	13,3	10,0	1,0	34,3	26,6	0,3	6,3	5,4	29,7	128
	AICc	Todos modelos	0,0	0,3	0,3	1,3	46,3	23,2	0,5	17,2	10,9	44,9	4.096
		Stepwise	5,5	19,2	10,4	2,1	39,4	19,0	0,3	2,5	1,6	35,6	46
		K passos	0,2	0,2	0,1	9,2	46,3	15,5	3,1	19,1	6,3	42,7	255
		Dois passos	4,6	16,5	9,3	2,0	39,4	19,7	0,5	5,3	2,8	34,1	128
	BIC	Todos modelos	0,7	1,0	0,4	10,3	67,3	9,7	0,9	8,3	1,4	64,9	4.096
		Stepwise	44,3	27,5	3,8	4,7	17,2	2,2	0,1	0,2	0,1	16,1	39
		K passos	7,4	0,9	0,1	48,4	31,1	2,6	3,4	5,6	0,5	27,8	288
		Dois passos	38,1	29,8	4,2	5,6	19,1	2,5	0,2	0,4	0,1	17,4	128
HQC	Todos modelos	0,2	0,3	0,3	2,9	56,4	18,3	0,5	14,8	6,3	54,3	4.096	
	Stepwise	13,7	27,5	8,6	3,1	33,8	11,5	0,2	1,1	0,4	30,8	44	
	K passos	1,0	0,5	0,1	18,4	48,0	10,6	4,3	13,8	3,3	43,0	263	
	Dois passos	11,3	25,6	8,6	3,7	35,3	12,2	0,3	2,3	0,8	31,3	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	40,1	27,8	0,0	17,9	14,2	40,1	4.096
		Stepwise	0,0	0,3	0,4	0,0	51,7	37,2	0,0	6,2	4,3	51,2	48
		K passos	0,0	0,0	0,0	0,1	44,6	23,7	0,0	20,5	11,1	44,6	225
		Dois passos	0,0	0,2	0,2	0,0	44,8	32,7	0,0	12,9	9,2	44,3	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	46,8	25,1	0,0	16,9	11,2	46,8	4.096
		Stepwise	0,0	0,4	0,4	0,0	57,6	32,9	0,0	5,7	3,0	57,2	48
		K passos	0,0	0,0	0,0	0,1	50,9	21,2	0,0	19,7	8,1	50,9	217
		Dois passos	0,0	0,2	0,2	0,0	50,5	29,3	0,0	12,7	7,0	50,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,2	87,9	6,6	0,0	5,0	0,3	87,9	4.096
		Stepwise	1,3	7,1	0,9	0,3	83,6	6,2	0,0	0,5	0,0	82,9	46
		K passos	0,1	0,0	0,0	2,8	87,1	4,3	0,2	5,2	0,3	86,8	205
		Dois passos	0,8	5,4	0,7	0,3	84,9	6,3	0,0	1,6	0,1	83,9	128
HQC	Todos modelos	0,0	0,0	0,0	0,1	68,1	16,5	0,0	12,4	2,9	68,1	4.096	
	Stepwise	0,1	1,8	0,5	0,0	75,4	18,8	0,0	2,7	0,5	74,6	47	
	K passos	0,0	0,0	0,0	0,4	71,4	12,6	0,1	13,2	2,3	71,4	208	
	Dois passos	0,0	1,2	0,2	0,0	71,9	18,2	0,0	6,8	1,5	70,8	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	42,0	26,7	0,0	18,6	12,7	42,0	4.096
		Stepwise	0,0	0,0	0,0	0,0	54,9	36,4	0,0	5,7	3,0	54,9	48
		K passos	0,0	0,0	0,0	0,0	47,1	21,7	0,0	20,5	10,7	47,1	221
		Dois passos	0,0	0,0	0,0	0,0	48,7	32,7	0,0	11,8	6,9	48,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	43,9	25,9	0,0	18,5	11,7	43,9	4.096
		Stepwise	0,0	0,0	0,0	0,0	57,2	34,6	0,0	5,3	2,8	57,2	48
		K passos	0,0	0,0	0,0	0,0	50,0	20,6	0,0	19,9	9,5	50,0	220
		Dois passos	0,0	0,0	0,0	0,0	50,9	31,5	0,0	11,6	6,0	50,9	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,5	3,7	0,0	3,6	0,2	92,5	4.096
		Stepwise	0,0	0,1	0,0	0,0	95,2	4,4	0,0	0,3	0,0	95,2	46
		K passos	0,0	0,0	0,0	0,0	93,5	2,9	0,0	3,5	0,1	93,5	195
		Dois passos	0,0	0,0	0,0	0,0	94,5	4,4	0,0	1,1	0,0	94,5	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	72,3	15,0	0,0	10,3	2,4	72,3	4.096	
	Stepwise	0,0	0,0	0,0	0,0	81,6	15,8	0,0	2,3	0,3	81,6	47	
	K passos	0,0	0,0	0,0	0,0	76,6	10,3	0,0	11,8	1,4	76,6	203	
	Dois passos	0,0	0,0	0,0	0,0	78,8	15,2	0,0	5,1	0,9	78,8	128	

Apêndice C

Resultados completos dos estudos de simulação - Distribuição Normal

Tabela C.1: Resultados completos dos estudos de simulação para o cenário 1 da distribuição Normal.

n	Critério	Método	$\#\beta < 5$			$\#\beta = 5$			$\#\beta > 5$			Modelo Correto	Média de Modelos
			$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$		
n = 50	AIC	Todos modelos	3,5	9,5	7,5	3,8	17,8	22,5	2,6	11,6	21,2	10,8	4,096
		Stepwise	10,5	23,4	20,9	5,2	14,1	16,3	1,7	3,8	4,1	7,7	42
		K passos	10,8	10,2	5,5	9,3	20,8	13,2	4,4	13,0	12,9	12,7	276
		Dois passos	8,1	19,2	15,9	5,3	13,7	19,0	2,3	7,4	9,1	6,4	128
	AICc	Todos modelos	8,9	17,0	9,8	7,0	22,8	15,5	3,6	8,6	6,8	13,6	4,096
		Stepwise	20,9	31,2	14,1	8,3	13,7	6,5	2,1	2,2	0,9	7,9	38
		K passos	24,0	14,7	3,6	14,9	20,1	6,2	5,3	8,1	3,1	13,0	269
		Dois passos	17,3	26,3	12,0	8,5	15,9	9,3	3,1	5,3	2,4	8,3	128
	BIC	Todos modelos	27,0	24,5	6,3	9,0	17,1	5,9	3,1	3,9	3,2	11,6	4,096
		Stepwise	48,9	28,3	5,6	7,6	6,0	1,8	1,1	0,5	0,3	4,1	31
		K passos	56,4	10,8	1,0	13,7	10,4	1,4	2,9	2,6	0,8	7,5	251
		Dois passos	44,0	26,4	5,7	8,4	8,9	2,7	1,7	1,7	0,5	5,3	128
HQC	Todos modelos	9,0	15,7	8,7	6,4	20,4	17,0	3,3	9,7	9,8	12,3	4,096	
	Stepwise	23,6	29,7	14,5	7,1	12,4	7,4	1,7	2,3	1,3	6,9	37	
	K passos	27,1	12,8	3,8	13,0	18,7	6,9	4,9	8,4	4,4	12,0	268	
	Dois passos	20,0	25,5	12,5	7,2	14,0	10,0	2,5	5,1	3,1	7,5	128	
n = 100	AIC	Todos modelos	0,0	0,1	0,9	1,2	29,5	32,8	0,5	14,0	21,0	28,5	4,096
		Stepwise	0,3	4,0	5,4	0,8	35,7	39,1	0,3	6,3	8,2	33,5	46
		K passos	0,1	0,3	0,4	1,1	38,5	25,1	0,4	18,0	16,3	37,4	242
		Dois passos	0,2	3,0	4,5	0,6	29,4	34,1	0,5	12,7	15,2	27,0	128
	AICc	Todos modelos	0,0	0,3	1,4	1,4	38,9	29,9	0,8	14,1	13,2	37,4	4,096
		Stepwise	0,4	6,4	5,4	1,2	42,9	32,3	0,4	5,9	5,1	40,2	45
		K passos	0,1	0,6	0,6	2,0	48,5	20,6	0,6	17,1	9,9	46,8	234
		Dois passos	0,3	4,9	4,7	0,9	36,8	29,4	0,7	12,8	9,6	33,4	128
	BIC	Todos modelos	1,9	5,6	1,1	7,5	65,1	10,8	0,5	6,2	1,3	62,7	4,096
		Stepwise	7,9	26,8	4,2	6,6	44,9	7,8	0,2	1,4	0,2	42,5	41
		K passos	4,5	5,0	0,6	13,3	62,9	5,7	0,8	6,4	0,7	60,7	230
		Dois passos	7,3	24,6	3,8	6,7	45,2	7,7	0,6	3,2	0,9	41,8	128
HQC	Todos modelos	0,2	1,1	1,7	2,5	51,9	23,8	0,7	11,1	7,0	49,3	4,096	
	Stepwise	1,2	12,7	6,2	2,7	47,7	22,6	0,4	4,3	2,2	44,5	43	
	K passos	0,6	1,7	0,7	4,3	58,0	15,2	0,9	13,7	4,9	55,8	230	
	Dois passos	1,0	10,7	5,7	2,4	44,8	21,4	0,6	8,6	4,8	40,7	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	30,0	36,7	0,0	15,7	17,6	30,0	4,096
		Stepwise	0,0	0,1	0,0	0,0	40,0	44,2	0,0	7,5	8,1	40,0	46
		K passos	0,0	0,0	0,0	0,0	41,3	26,6	0,0	18,1	14,0	41,3	225
		Dois passos	0,0	0,1	0,0	0,0	33,5	38,2	0,0	13,8	14,4	33,5	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	35,5	34,8	0,0	15,3	14,4	35,5	4,096
		Stepwise	0,0	0,1	0,0	0,0	45,8	40,8	0,0	6,8	6,5	45,8	45
		K passos	0,0	0,0	0,0	0,0	47,6	24,0	0,0	17,6	10,7	47,6	219
		Dois passos	0,0	0,1	0,0	0,0	38,6	35,6	0,0	13,9	11,8	38,6	128
	BIC	Todos modelos	0,0	0,0	0,0	0,1	88,0	7,6	0,0	3,9	0,4	88,0	4,096
		Stepwise	0,0	1,2	0,2	0,2	87,9	9,1	0,0	1,3	0,1	87,8	43
		K passos	0,0	0,0	0,0	0,2	90,0	5,0	0,0	4,6	0,2	90,0	197
		Dois passos	0,0	1,2	0,2	0,2	85,3	8,9	0,0	3,8	0,3	85,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	63,4	21,8	0,0	11,0	3,8	63,4	4,096	
	Stepwise	0,0	0,2	0,1	0,1	68,3	25,3	0,0	4,6	1,4	68,3	44	
	K passos	0,0	0,0	0,0	0,1	70,9	13,9	0,0	12,1	2,9	70,9	206	
	Dois passos	0,0	0,2	0,1	0,1	63,1	23,6	0,0	9,8	3,2	63,0	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	34,4	33,5	0,0	13,0	19,1	34,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	43,1	42,0	0,0	6,3	8,6	43,1	46
		K passos	0,0	0,0	0,0	0,0	44,0	25,1	0,0	16,9	14,0	44,0	218
		Dois passos	0,0	0,0	0,0	0,0	34,9	35,3	0,0	14,4	15,4	34,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	36,6	33,3	0,0	12,5	17,6	36,6	4,096
		Stepwise	0,0	0,0	0,0	0,0	45,2	40,8	0,0	6,1	7,9	45,2	46
		K passos	0,0	0,0	0,0	0,0	46,5	24,1	0,0	16,6	12,7	46,5	218
		Dois passos	0,0	0,0	0,0	0,0	37,0	34,4	0,0	14,1	14,4	37,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,4	4,9	0,0	2,5	0,2	92,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	93,9	5,3	0,0	0,8	0,1	93,9	42
		K passos	0,0	0,0	0,0	0,0	94,7	2,7	0,0	2,4	0,1	94,7	195
		Dois passos	0,0	0,0	0,0	0,0	91,4	5,1	0,0	3,3	0,2	91,4	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	70,1	17,2	0,0	9,1	3,6	70,1	4,096	
	Stepwise	0,0	0,0	0,0	0,0	75,0	20,6	0,0	3,3	1,1	75,0	43	
	K passos	0,0	0,0	0,0	0,0	77,2	10,8	0,0	9,9	2,1	77,2	202	
	Dois passos	0,0	0,0	0,0	0,0	69,2	19,2	0,0	9,1	2,5	69,2	128	

Tabela C.2: Resultados completos dos estudos de simulação para o cenário 2 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	8,2	32,4	32,2	2,0	5,0	10,7	0,4	2,8	6,3	1,0	4.096
		Stepwise	15,3	40,8	37,0	1,3	2,2	2,3	0,2	0,5	0,4	0,4	35
		K passos	27,6	36,6	13,8	3,3	6,5	4,8	1,1	3,1	3,2	1,8	296
		Dois passos	13,4	37,8	35,1	1,4	3,8	4,8	0,4	1,5	1,9	0,8	128
	AICc	Todos modelos	16,2	42,3	28,5	1,7	4,4	3,1	0,4	1,6	1,8	1,0	4.096
		Stepwise	24,3	48,5	23,8	1,1	1,2	0,6	0,2	0,3	0,0	0,2	32
		K passos	46,9	36,4	5,9	2,9	4,2	1,4	0,7	1,2	0,4	1,1	281
		Dois passos	21,8	46,5	24,4	1,8	2,9	1,3	0,3	0,7	0,2	0,5	128
	BIC	Todos modelos	37,8	45,0	11,7	1,3	1,7	1,2	0,3	0,4	0,6	0,5	4.096
		Stepwise	48,3	42,3	8,8	0,3	0,2	0,0	0,0	0,1	0,0	0,1	25
		K passos	80,3	15,7	1,0	1,3	1,0	0,2	0,2	0,2	0,1	0,4	239
		Dois passos	45,3	43,3	9,7	0,5	0,6	0,3	0,1	0,1	0,0	0,1	128
HQC	Todos modelos	18,8	41,1	25,4	1,5	4,1	4,2	0,4	1,6	2,9	0,9	4.096	
	Stepwise	28,7	46,5	21,9	0,7	1,0	0,7	0,1	0,3	0,1	0,2	30	
	K passos	52,1	31,1	6,1	2,5	3,9	1,7	0,7	1,2	0,7	0,9	276	
	Dois passos	25,9	45,5	22,7	1,2	2,2	1,4	0,3	0,7	0,2	0,4	128	
n = 100	AIC	Todos modelos	0,9	20,1	29,3	0,5	16,2	17,7	0,3	5,9	9,1	9,6	4.096
		Stepwise	1,2	39,4	42,6	0,3	6,6	7,4	0,1	1,0	1,3	3,6	41
		K passos	2,7	30,4	15,7	0,9	23,6	13,2	0,3	7,2	5,9	13,7	275
		Dois passos	1,1	33,5	36,7	0,4	10,1	11,5	0,1	2,9	3,7	4,0	128
	AICc	Todos modelos	1,2	30,8	28,8	0,6	15,3	13,9	0,5	5,0	3,9	9,2	4.096
		Stepwise	2,0	46,7	38,5	0,5	6,6	4,6	0,1	0,6	0,5	3,4	40
		K passos	4,6	40,2	14,2	1,3	22,7	8,5	0,3	5,5	2,7	12,9	274
		Dois passos	1,7	41,0	33,8	0,6	10,5	8,4	0,2	2,0	1,8	4,1	128
	BIC	Todos modelos	11,7	68,8	11,8	0,6	5,3	0,9	0,1	0,6	0,2	2,9	4.096
		Stepwise	17,0	70,5	11,3	0,2	0,6	0,2	0,0	0,0	0,0	0,3	33
		K passos	39,2	49,2	3,1	1,4	5,7	0,5	0,1	0,7	0,0	3,3	283
		Dois passos	16,5	69,6	11,3	0,5	1,5	0,5	0,0	0,1	0,0	0,8	128
HQC	Todos modelos	3,5	45,9	24,4	0,9	12,9	7,9	0,3	3,0	1,2	7,6	4.096	
	Stepwise	5,3	60,4	27,8	0,6	3,8	1,8	0,0	0,2	0,1	1,9	37	
	K passos	12,9	50,3	10,1	1,4	17,2	4,0	0,2	2,7	1,2	10,1	273	
	Dois passos	5,0	56,5	26,3	0,6	6,6	3,8	0,1	0,8	0,3	2,5	128	
n = 200	AIC	Todos modelos	0,0	3,9	5,3	0,0	27,7	34,3	0,0	13,7	15,1	25,7	4.096
		Stepwise	0,0	18,6	20,1	0,0	24,8	27,0	0,0	4,3	5,1	22,4	44
		K passos	0,0	5,4	3,6	0,1	40,0	22,5	0,0	17,0	11,5	38,2	246
		Dois passos	0,0	15,3	16,8	0,0	24,1	26,8	0,0	8,0	8,9	19,5	128
	AICc	Todos modelos	0,0	4,7	6,3	0,0	32,9	31,0	0,0	12,8	12,3	30,4	4.096
		Stepwise	0,0	22,1	20,1	0,0	27,0	23,5	0,0	3,6	3,8	24,6	44
		K passos	0,0	6,8	3,9	0,1	44,8	19,8	0,0	15,4	9,2	42,4	244
		Dois passos	0,0	18,3	16,8	0,0	26,6	23,6	0,0	7,5	7,1	21,8	128
	BIC	Todos modelos	0,2	44,9	4,9	0,0	44,3	4,1	0,0	1,6	0,0	42,5	4.096
		Stepwise	0,7	78,5	7,7	0,0	11,8	1,1	0,0	0,2	0,0	11,5	37
		K passos	2,4	46,1	1,6	0,4	45,8	1,3	0,0	2,3	0,0	43,7	248
		Dois passos	0,6	76,4	7,4	0,0	13,6	1,5	0,0	0,5	0,1	12,6	128
HQC	Todos modelos	0,0	17,0	7,0	0,0	48,3	17,0	0,0	7,8	2,9	45,4	4.096	
	Stepwise	0,1	47,3	17,2	0,0	24,7	8,9	0,0	1,3	0,5	22,3	41	
	K passos	0,1	18,8	4,2	0,1	57,2	8,3	0,0	9,6	1,7	54,0	241	
	Dois passos	0,1	42,9	16,0	0,0	27,3	9,4	0,0	3,1	1,2	22,7	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	34,4	33,5	0,0	13,0	19,1	34,4	4.096
		Stepwise	0,0	2,3	2,6	0,0	41,1	39,9	0,0	5,9	8,2	40,6	46
		K passos	0,0	0,1	0,1	0,0	44,0	24,9	0,0	17,0	13,8	44,0	221
		Dois passos	0,0	1,8	2,1	0,0	34,0	34,2	0,0	13,4	14,5	33,1	128
	AICc	Todos modelos	0,0	0,0	0,1	0,0	36,6	33,2	0,0	12,5	17,6	36,6	4.096
		Stepwise	0,0	2,7	2,5	0,0	43,1	38,6	0,0	5,7	7,6	42,5	45
		K passos	0,0	0,1	0,1	0,0	46,6	23,9	0,0	16,7	12,6	46,5	221
		Dois passos	0,0	2,1	1,9	0,0	36,0	33,3	0,0	13,2	13,5	35,0	128
	BIC	Todos modelos	0,0	3,7	0,1	0,0	88,8	4,8	0,0	2,4	0,2	88,8	4.096
		Stepwise	0,0	35,1	2,4	0,0	58,7	3,2	0,0	0,5	0,0	58,5	41
		K passos	0,0	4,0	0,0	0,0	90,9	2,7	0,0	2,3	0,1	90,8	217
		Dois passos	0,0	33,6	2,4	0,0	58,7	3,2	0,0	2,0	0,2	57,8	128
HQC	Todos modelos	0,0	0,7	0,1	0,0	69,4	17,2	0,0	9,1	3,5	69,4	4.096	
	Stepwise	0,0	12,6	3,3	0,0	62,8	17,4	0,0	2,8	1,0	62,5	43	
	K passos	0,0	0,6	0,1	0,0	76,7	10,7	0,0	9,8	2,1	76,6	211	
	Dois passos	0,0	11,4	3,1	0,0	59,3	16,4	0,0	7,5	2,2	58,0	128	

Tabela C.3: Resultados completos dos estudos de simulação para o cenário 3 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	1,0	1,2	5,4	22,7	21,9	3,4	16,1	28,3	15,6	4,096
		Stepwise	1,9	5,0	4,1	8,1	28,2	32,9	3,4	7,4	9,1	18,6	45
		K passos	1,2	1,3	1,0	8,7	26,4	18,7	5,5	17,8	19,3	19,7	250
		Dois passos	1,3	3,8	3,0	6,6	22,0	28,6	4,0	13,4	17,3	14,0	128
	AICc	Todos modelos	0,6	2,9	1,0	13,3	32,4	18,7	5,8	15,7	9,6	22,8	4,096
		Stepwise	5,5	8,3	3,4	16,0	35,6	18,5	4,6	5,5	2,6	24,3	42
		K passos	4,0	2,4	0,9	19,6	33,7	11,5	7,8	14,0	6,1	26,4	234
		Dois passos	4,5	6,2	2,7	13,9	31,4	18,6	5,7	11,1	5,9	20,5	128
	BIC	Todos modelos	5,4	4,6	1,1	22,5	35,5	10,4	6,7	8,5	5,3	26,5	4,096
		Stepwise	22,7	10,1	2,1	23,0	27,9	7,9	3,5	1,9	0,9	20,3	38
		K passos	15,9	2,5	0,6	30,1	30,0	4,8	6,7	7,4	2,0	24,3	225
		Dois passos	18,4	8,8	2,3	22,0	27,9	8,9	5,1	4,9	1,7	19,7	128
HQC	Todos modelos	0,6	2,6	1,0	11,8	29,3	18,7	4,9	16,5	14,6	20,4	4,096	
	Stepwise	6,5	8,3	3,7	14,5	33,0	20,9	3,7	5,8	3,6	22,1	42	
	K passos	4,5	2,2	0,8	17,7	31,2	13,3	7,2	14,3	8,8	24,3	235	
	Dois passos	5,2	6,6	2,9	12,4	29,2	21,0	4,6	10,6	7,3	18,8	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	1,2	29,2	33,3	0,5	14,3	21,5	28,5	4,096
		Stepwise	0,0	0,1	0,0	0,9	38,5	43,9	0,3	7,3	8,9	37,4	46
		K passos	0,0	0,0	0,0	1,0	37,8	25,9	0,4	18,4	16,5	37,1	234
		Dois passos	0,0	0,1	0,0	0,6	30,7	36,8	0,5	14,4	17,0	29,8	128
	AICc	Todos modelos	0,0	0,0	0,0	1,4	39,3	30,6	0,8	14,5	13,4	38,0	4,096
		Stepwise	0,0	0,1	0,1	1,4	47,7	37,2	0,4	7,3	5,7	46,1	45
		K passos	0,0	0,0	0,0	1,7	48,1	21,9	0,5	17,7	10,0	47,0	227
		Dois passos	0,0	0,1	0,1	0,9	39,3	32,7	0,8	15,0	11,1	37,9	128
	BIC	Todos modelos	0,0	0,0	0,0	8,6	70,7	11,8	0,6	6,8	1,5	68,2	4,096
		Stepwise	0,2	0,8	0,3	9,4	73,6	12,7	0,3	2,2	0,5	71,1	42
		K passos	0,0	0,0	0,0	9,9	73,8	7,4	1,0	7,0	0,9	72,2	205
		Dois passos	0,2	0,7	0,2	8,8	70,5	12,1	0,8	5,2	1,4	68,1	128
HQC	Todos modelos	0,0	0,0	0,0	2,6	52,5	25,0	0,8	11,9	7,2	50,9	4,096	
	Stepwise	0,0	0,2	0,1	3,0	58,7	29,2	0,5	5,6	2,8	56,5	44	
	K passos	0,0	0,0	0,0	3,2	59,6	16,7	0,8	14,5	5,3	58,3	216	
	Dois passos	0,0	0,2	0,1	2,6	52,5	26,7	0,8	11,2	5,9	50,6	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	30,0	36,7	0,0	15,7	17,6	30,0	4,096
		Stepwise	0,0	0,0	0,0	0,0	40,1	44,2	0,0	7,5	8,1	40,1	46
		K passos	0,0	0,0	0,0	0,0	41,3	26,6	0,0	18,1	14,0	41,3	225
		Dois passos	0,0	0,0	0,0	0,0	33,6	38,2	0,0	13,8	14,4	33,6	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	35,5	34,8	0,0	15,3	14,4	35,5	4,096
		Stepwise	0,0	0,0	0,0	0,0	45,9	40,8	0,0	6,8	6,5	45,8	45
		K passos	0,0	0,0	0,0	0,0	47,6	24,0	0,0	17,6	10,7	47,6	219
		Dois passos	0,0	0,0	0,0	0,0	38,7	35,6	0,0	13,9	11,8	38,6	128
	BIC	Todos modelos	0,0	0,0	0,0	0,1	88,0	7,6	0,0	3,9	0,4	88,0	4,096
		Stepwise	0,0	0,0	0,0	0,2	89,1	9,3	0,0	1,4	0,1	89,1	43
		K passos	0,0	0,0	0,0	0,2	90,0	5,0	0,0	4,6	0,2	90,0	196
		Dois passos	0,0	0,0	0,0	0,2	86,5	9,2	0,0	3,9	0,3	86,5	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	63,4	21,8	0,0	11,0	3,8	63,4	4,096	
	Stepwise	0,0	0,0	0,0	0,1	68,6	25,4	0,0	4,6	1,4	68,5	44	
	K passos	0,0	0,0	0,0	0,1	71,0	13,9	0,0	12,1	2,9	70,9	206	
	Dois passos	0,0	0,0	0,0	0,1	63,2	23,7	0,0	9,8	3,2	63,2	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	34,4	33,5	0,0	13,0	19,1	34,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	43,1	42,0	0,0	6,3	8,6	43,1	46
		K passos	0,0	0,0	0,0	0,0	44,0	25,1	0,0	16,9	14,0	44,0	218
		Dois passos	0,0	0,0	0,0	0,0	34,9	35,3	0,0	14,4	15,4	34,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	36,6	33,3	0,0	12,5	17,6	36,6	4,096
		Stepwise	0,0	0,0	0,0	0,0	45,2	40,8	0,0	6,1	7,9	45,2	46
		K passos	0,0	0,0	0,0	0,0	46,5	24,1	0,0	16,6	12,7	46,5	218
		Dois passos	0,0	0,0	0,0	0,0	37,0	34,4	0,0	14,1	14,4	37,0	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,4	4,9	0,0	2,5	0,2	92,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	93,9	5,3	0,0	0,8	0,1	93,9	42
		K passos	0,0	0,0	0,0	0,0	94,7	2,7	0,0	2,4	0,1	94,7	195
		Dois passos	0,0	0,0	0,0	0,0	91,4	5,1	0,0	3,3	0,2	91,4	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	70,1	17,2	0,0	9,1	3,6	70,1	4,096	
	Stepwise	0,0	0,0	0,0	0,0	75,0	20,6	0,0	3,3	1,1	75,0	43	
	K passos	0,0	0,0	0,0	0,0	77,2	10,8	0,0	9,9	2,1	77,2	202	
	Dois passos	0,0	0,0	0,0	0,0	69,2	19,2	0,0	9,1	2,5	69,2	128	

Tabela C.4: Resultados completos dos estudos de simulação para o cenário 4 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	8,8	9,3	6,0	17,7	14,9	12,3	10,0	9,5	11,5	3,2	4.096
		Stepwise	16,8	12,7	8,2	21,5	14,0	10,6	8,0	4,7	3,5	2,9	40
		K passos	16,8	6,8	2,9	25,4	14,3	7,3	11,1	8,4	7,0	4,1	235
		Dois passos	13,5	9,0	6,1	20,1	14,5	13,6	9,2	7,7	6,4	2,7	128
	AICc	Todos modelos	19,6	12,8	4,6	26,0	14,3	5,9	8,5	5,5	2,8	3,1	4.096
		Stepwise	31,1	14,3	4,4	26,9	10,8	3,2	7,0	1,8	0,6	2,4	36
		K passos	32,3	7,0	1,4	32,6	10,2	2,0	9,9	3,5	1,1	3,3	238
		Dois passos	25,6	11,3	3,6	27,8	12,5	5,0	9,2	3,6	1,4	2,7	128
	BIC	Todos modelos	49,1	8,4	3,1	22,0	7,1	1,7	5,5	2,3	0,8	2,1	4.096
		Stepwise	61,0	8,6	1,5	20,2	3,8	0,8	3,5	0,5	0,1	0,9	30
		K passos	60,1	3,3	0,3	25,5	3,7	0,5	5,0	1,3	0,2	1,1	239
		Dois passos	54,9	7,4	1,7	23,6	5,6	1,4	4,3	0,7	0,2	1,3	128
HQC	Todos modelos	21,3	10,8	4,8	23,4	14,2	6,4	8,2	6,2	4,7	2,9	4.096	
	Stepwise	34,1	13,7	4,6	24,5	10,3	3,6	6,5	1,9	0,9	2,2	36	
	K passos	34,0	6,4	1,6	30,1	10,1	2,7	9,4	4,0	1,6	3,2	239	
	Dois passos	28,6	10,9	3,7	24,8	12,3	5,8	8,4	3,5	2,0	2,6	128	
n = 100	AIC	Todos modelos	0,2	0,6	0,5	19,7	26,6	18,8	6,9	14,9	11,8	13,3	4.096
		Stepwise	0,7	1,2	1,1	20,8	32,5	21,7	6,9	9,1	6,0	16,1	43
		K passos	0,5	0,6	0,2	22,8	29,8	13,3	10,0	14,1	8,7	17,7	225
		Dois passos	0,5	0,8	0,7	18,2	28,7	21,1	8,5	12,4	9,1	14,1	128
	AICc	Todos modelos	0,5	1,0	0,8	25,9	30,8	14,3	8,7	12,2	5,8	16,0	4.096
		Stepwise	1,2	1,7	1,1	27,7	34,3	15,4	7,8	7,4	3,3	17,0	42
		K passos	0,9	0,7	0,3	30,9	31,5	9,2	10,8	11,2	4,4	19,3	218
		Dois passos	0,9	1,2	0,7	25,1	31,3	15,7	9,8	10,4	4,9	15,4	128
	BIC	Todos modelos	9,5	2,8	0,4	60,1	18,0	2,2	4,4	2,1	0,5	11,6	4.096
		Stepwise	14,6	3,3	0,2	58,8	16,2	2,3	3,4	1,0	0,2	9,8	37
		K passos	12,7	1,4	0,0	62,6	14,7	1,3	5,2	1,9	0,2	10,5	216
		Dois passos	12,4	2,6	0,2	59,2	16,9	2,7	4,1	1,6	0,3	10,1	128
HQC	Todos modelos	2,4	1,3	0,8	38,5	28,5	10,1	8,0	7,9	2,5	15,5	4.096	
	Stepwise	3,7	2,8	0,8	41,1	29,8	9,3	6,7	4,3	1,5	15,8	40	
	K passos	3,1	1,0	0,2	44,4	27,1	5,6	9,7	6,7	2,0	17,7	215	
	Dois passos	3,1	2,1	0,5	38,9	29,1	9,8	8,4	5,9	2,3	15,0	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	5,0	32,0	30,8	3,0	14,7	14,5	26,2	4.096
		Stepwise	0,0	0,0	0,0	6,0	37,0	35,0	2,3	10,7	9,1	31,1	45
		K passos	0,0	0,0	0,0	7,4	39,6	22,0	3,0	16,6	11,4	35,5	217
		Dois passos	0,0	0,0	0,0	5,3	33,3	32,2	2,8	14,0	12,4	27,9	128
	AICc	Todos modelos	0,0	0,0	0,0	6,5	36,0	28,7	3,1	14,2	11,5	29,9	4.096
		Stepwise	0,0	0,0	0,0	7,5	40,7	31,7	2,4	10,9	6,8	34,4	44
		K passos	0,0	0,0	0,0	8,9	43,6	19,5	3,1	16,3	8,5	39,0	214
		Dois passos	0,0	0,0	0,0	6,6	37,3	29,4	2,9	14,0	9,7	31,4	128
	BIC	Todos modelos	0,0	0,1	0,0	45,0	47,5	3,6	1,5	2,3	0,0	42,4	4.096
		Stepwise	0,0	0,1	0,0	47,0	45,9	3,9	1,8	1,2	0,1	40,8	39
		K passos	0,1	0,0	0,0	48,1	44,8	2,4	2,1	2,6	0,0	42,3	194
		Dois passos	0,0	0,1	0,0	45,5	45,6	4,1	2,4	2,2	0,1	40,7	128
HQC	Todos modelos	0,0	0,0	0,0	21,6	47,8	14,5	3,5	9,7	2,9	41,6	4.096	
	Stepwise	0,0	0,0	0,0	21,1	53,1	16,3	2,8	5,2	1,5	45,8	42	
	K passos	0,0	0,0	0,0	22,5	53,0	9,1	3,6	9,8	2,0	48,4	203	
	Dois passos	0,0	0,0	0,0	20,1	50,7	16,0	3,1	7,9	2,2	43,7	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,3	34,1	33,7	0,1	14,2	17,6	33,7	4.096
		Stepwise	0,0	0,0	0,0	0,5	39,3	39,0	0,1	9,9	11,3	39,0	45
		K passos	0,0	0,0	0,0	0,5	43,6	24,6	0,2	17,1	14,1	43,4	220
		Dois passos	0,0	0,0	0,0	0,4	35,1	35,8	0,2	14,0	14,5	34,9	128
	AICc	Todos modelos	0,0	0,0	0,0	0,4	36,9	33,7	0,2	13,2	15,6	36,5	4.096
		Stepwise	0,0	0,0	0,0	0,5	42,2	37,3	0,2	9,5	10,3	41,9	45
		K passos	0,0	0,0	0,0	0,6	46,6	23,7	0,2	16,4	12,6	46,4	223
		Dois passos	0,0	0,0	0,0	0,5	37,7	34,3	0,2	13,9	13,5	37,4	128
	BIC	Todos modelos	0,0	0,0	0,0	10,4	82,9	4,1	0,0	2,3	0,3	82,2	4.096
		Stepwise	0,0	0,0	0,0	10,0	83,9	4,4	0,2	1,4	0,1	83,0	42
		K passos	0,0	0,0	0,0	10,0	84,9	2,3	0,3	2,3	0,2	84,3	195
		Dois passos	0,0	0,0	0,0	9,8	82,8	4,4	0,4	2,5	0,2	81,9	128
HQC	Todos modelos	0,0	0,0	0,0	2,7	68,6	17,0	0,5	8,3	2,9	67,8	4.096	
	Stepwise	0,0	0,0	0,0	2,6	70,8	18,9	0,3	5,7	1,8	69,9	43	
	K passos	0,0	0,0	0,0	2,9	75,0	10,5	0,4	9,3	1,9	74,4	200	
	Dois passos	0,0	0,0	0,0	2,4	67,8	18,0	0,5	8,6	2,7	67,0	128	

Tabela C.5: Resultados completos dos estudos de simulação para o cenário 5 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,1	3,2	7,0	0,2	17,5	28,2	0,1	12,2	31,5	15,9	4.096
		Stepwise	0,7	35,0	42,7	0,1	7,1	11,0	0,0	0,9	2,5	5,6	39
		K passos	0,7	5,4	4,7	0,8	28,3	19,3	0,8	21,5	18,5	25,6	296
		Dois passos	0,6	31,3	38,4	0,1	7,3	12,7	0,0	3,2	6,3	4,3	128
	AICc	Todos modelos	0,2	8,3	9,0	0,3	32,9	23,2	0,1	11,9	14,1	28,9	4.096
		Stepwise	1,9	50,5	34,6	0,1	7,2	4,7	0,1	0,3	0,5	6,4	36
		K passos	2,6	12,9	4,3	1,6	43,9	10,4	0,8	18,0	5,5	38,4	292
		Dois passos	1,5	46,1	32,5	0,2	8,9	6,2	0,1	2,5	2,0	5,7	128
	BIC	Todos modelos	2,1	19,6	8,3	0,2	37,7	15,4	0,3	9,7	6,7	34,2	4.096
		Stepwise	6,9	69,1	19,7	0,1	2,7	1,2	0,0	0,1	0,1	2,5	30
		K passos	14,9	20,4	2,5	3,5	41,1	3,7	1,0	11,3	1,6	36,9	279
		Dois passos	6,4	66,6	19,5	0,2	4,2	1,9	0,1	0,8	0,4	2,7	128
HQC	Todos modelos	0,2	7,1	8,3	0,2	28,7	24,3	0,1	12,1	19,0	25,7	4.096	
	Stepwise	2,2	52,5	33,4	0,1	5,7	4,8	0,0	0,4	0,8	5,0	35	
	K passos	3,0	11,6	4,0	1,7	39,9	11,3	1,0	19,5	7,9	35,2	295	
	Dois passos	1,9	48,7	31,8	0,2	6,8	6,2	0,0	2,0	2,4	4,4	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,2	0,0	29,2	33,5	0,0	15,9	21,2	29,1	4.096
		Stepwise	0,0	25,9	33,6	0,0	16,2	20,8	0,0	1,5	1,9	14,8	43
		K passos	0,0	0,1	0,1	0,0	36,4	27,1	0,0	19,5	16,9	36,3	277
		Dois passos	0,0	21,4	28,7	0,0	16,2	20,8	0,0	5,6	7,4	11,1	128
	AICc	Todos modelos	0,0	0,1	0,2	0,0	38,8	30,4	0,0	17,2	13,3	38,7	4.096
		Stepwise	0,0	32,3	32,1	0,0	17,9	15,5	0,0	1,2	1,0	16,6	42
		K passos	0,0	0,1	0,1	0,0	47,5	22,8	0,0	19,0	10,4	47,4	272
		Dois passos	0,0	27,3	28,5	0,0	18,3	16,2	0,0	5,3	4,4	12,8	128
	BIC	Todos modelos	0,0	1,2	0,3	0,0	75,0	13,3	0,0	8,5	1,7	74,6	4.096
		Stepwise	0,3	71,4	17,1	0,0	9,0	1,9	0,0	0,1	0,0	8,7	37
		K passos	0,0	1,1	0,2	1,0	79,4	8,8	0,1	8,2	1,2	78,9	266
		Dois passos	0,3	69,3	16,6	0,0	10,4	2,5	0,0	0,8	0,1	7,9	128
HQC	Todos modelos	0,0	0,4	0,2	0,0	52,2	25,2	0,0	13,2	8,8	52,2	4.096	
	Stepwise	0,0	48,4	27,0	0,0	15,6	8,3	0,0	0,4	0,4	14,7	40	
	K passos	0,0	0,2	0,1	0,1	61,4	17,0	0,0	15,3	5,7	61,3	270	
	Dois passos	0,0	43,8	25,2	0,0	17,4	9,1	0,0	2,8	1,6	12,3	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	29,4	36,2	0,0	16,2	18,2	29,4	4.096
		Stepwise	0,0	4,7	5,3	0,0	40,5	42,7	0,0	3,1	3,6	40,2	46
		K passos	0,0	0,0	0,0	0,0	41,7	25,4	0,0	19,6	13,3	41,7	240
		Dois passos	0,0	3,9	4,6	0,0	32,4	34,5	0,0	11,9	12,7	31,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	35,1	34,7	0,0	16,0	14,2	35,1	4.096
		Stepwise	0,0	5,5	5,1	0,0	44,2	39,6	0,0	2,9	2,6	43,9	45
		K passos	0,0	0,0	0,0	0,0	46,6	23,6	0,0	19,2	10,5	46,6	235
		Dois passos	0,0	4,7	4,7	0,0	35,9	32,6	0,0	12,0	10,1	34,8	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	87,4	7,7	0,0	4,5	0,4	87,4	4.096
		Stepwise	0,0	37,0	4,1	0,0	53,0	5,8	0,0	0,1	0,0	52,8	41
		K passos	0,0	0,0	0,0	0,0	90,2	4,2	0,0	5,4	0,2	90,2	223
		Dois passos	0,0	36,2	3,8	0,0	52,0	5,9	0,0	2,0	0,1	51,0	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	62,4	22,3	0,0	11,7	3,6	62,4	4.096	
	Stepwise	0,0	15,9	5,7	0,0	56,0	20,8	0,0	1,1	0,4	55,8	43	
	K passos	0,0	0,0	0,0	0,0	70,4	13,6	0,0	13,6	2,3	70,4	224	
	Dois passos	0,0	14,5	5,4	0,0	51,4	19,4	0,0	7,0	2,3	50,0	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	33,7	34,1	0,0	13,8	18,4	33,7	4.096
		Stepwise	0,0	0,2	0,2	0,0	46,0	46,5	0,0	3,1	4,0	46,0	46
		K passos	0,0	0,0	0,0	0,0	43,2	26,0	0,0	17,5	13,3	43,2	230
		Dois passos	0,0	0,1	0,2	0,0	33,3	35,4	0,0	15,7	15,3	33,2	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	36,0	33,6	0,0	13,9	16,5	36,0	4.096
		Stepwise	0,0	0,3	0,3	0,0	48,2	45,0	0,0	2,8	3,5	48,1	46
		K passos	0,0	0,0	0,0	0,0	45,6	25,3	0,0	17,1	12,1	45,6	227
		Dois passos	0,0	0,2	0,2	0,0	35,4	34,5	0,0	15,6	14,2	35,2	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	92,2	4,8	0,0	2,8	0,2	92,2	4.096
		Stepwise	0,0	5,5	0,5	0,0	89,1	4,8	0,0	0,1	0,0	89,1	42
		K passos	0,0	0,0	0,0	0,0	94,9	2,5	0,0	2,5	0,2	94,9	200
		Dois passos	0,0	5,4	0,5	0,0	86,4	4,5	0,0	3,0	0,2	86,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	70,0	17,2	0,0	9,1	3,7	70,0	4.096	
	Stepwise	0,0	1,1	0,4	0,0	76,1	21,0	0,0	1,0	0,3	76,1	44	
	K passos	0,0	0,0	0,0	0,0	76,8	11,3	0,0	9,9	1,9	76,8	210	
	Dois passos	0,0	0,9	0,3	0,0	68,0	18,7	0,0	9,4	2,7	67,7	128	

Tabela C.6: Resultados completos dos estudos de simulação para o cenário 6 da distribuição Normal.

n	Critério	Método	$\#\beta < 5$			$\#\beta = 5$			$\#\beta > 5$			Modelo Correto	Média de Modelos
			$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$	$\#\gamma < 3$	$\#\gamma = 3$	$\#\gamma > 3$		
n = 50	AIC	Todos modelos	13,9	27,9	26,7	2,4	8,1	12,0	0,7	2,8	5,5	1,4	4.096
		Stepwise	26,4	41,6	29,1	0,5	1,1	1,1	0,0	0,2	0,1	0,1	33
		K passos	36,3	27,2	8,8	7,1	8,1	3,9	2,9	3,2	2,6	0,9	305
		Dois passos	24,3	38,1	29,1	1,3	3,0	2,5	0,3	0,7	0,8	0,4	128
	AICc	Todos modelos	26,1	36,4	22,2	2,7	4,9	3,6	0,8	1,9	1,4	0,6	4.096
		Stepwise	37,8	43,5	17,7	0,4	0,4	0,1	0,0	0,1	0,0	0,1	30
		K passos	55,6	25,5	3,6	7,4	3,8	0,8	1,9	1,2	0,3	0,6	305
		Dois passos	35,5	41,5	18,5	1,5	1,7	0,7	0,2	0,3	0,0	0,3	128
	BIC	Todos modelos	50,3	35,5	8,1	2,3	1,9	0,9	0,2	0,5	0,3	0,3	4.096
		Stepwise	61,9	33,1	4,9	0,0	0,1	0,0	0,0	0,0	0,0	0,0	24
		K passos	83,0	11,6	0,4	3,3	0,7	0,1	0,7	0,2	0,0	0,2	260
		Dois passos	60,7	33,0	5,5	0,4	0,2	0,1	0,0	0,1	0,0	0,1	128
HQC	Todos modelos	28,2	33,8	20,6	2,4	5,1	5,3	0,8	1,7	2,1	0,8	4.096	
	Stepwise	42,3	41,3	15,4	0,3	0,5	0,1	0,0	0,1	0,0	0,2	29	
	K passos	58,5	22,6	3,5	6,7	3,6	1,1	2,1	1,5	0,4	0,7	291	
	Dois passos	40,5	39,4	16,4	0,9	1,5	0,8	0,2	0,2	0,1	0,3	128	
n = 100	AIC	Todos modelos	7,6	19,3	22,8	5,5	16,2	15,1	1,2	6,1	6,2	5,9	4.096
		Stepwise	17,1	42,0	35,0	1,0	2,2	2,5	0,0	0,2	0,1	0,8	38
		K passos	18,2	23,3	9,5	10,5	17,6	7,9	3,7	5,8	3,5	7,3	312
		Dois passos	15,1	37,1	31,2	2,1	5,8	6,3	0,3	0,8	1,2	0,9	128
	AICc	Todos modelos	12,4	24,2	22,1	6,6	15,2	10,2	1,7	4,3	3,3	6,0	4.096
		Stepwise	21,0	46,3	28,6	1,0	1,8	1,1	0,0	0,1	0,1	0,7	37
		K passos	25,6	26,0	7,5	11,5	15,9	4,3	3,2	4,5	1,5	6,8	313
		Dois passos	19,1	42,4	26,5	2,1	4,6	3,8	0,3	0,6	0,5	0,8	128
	BIC	Todos modelos	47,6	37,3	5,9	3,6	4,3	0,5	0,5	0,3	0,0	1,9	4.096
		Stepwise	59,1	35,4	5,4	0,0	0,1	0,0	0,0	0,0	0,0	0,0	28
		K passos	76,4	13,8	0,6	6,4	2,0	0,1	0,5	0,2	0,0	1,0	280
		Dois passos	58,6	35,0	5,6	0,3	0,3	0,1	0,0	0,0	0,0	0,1	128
HQC	Todos modelos	22,0	34,5	15,1	6,8	11,1	5,4	1,2	2,7	1,2	4,5	4.096	
	Stepwise	34,0	48,1	16,6	0,4	0,7	0,2	0,0	0,0	0,0	0,2	33	
	K passos	45,7	23,3	4,2	11,3	9,1	1,6	2,4	2,0	0,5	4,0	297	
	Dois passos	32,4	45,8	16,9	1,5	2,3	0,8	0,1	0,3	0,0	0,3	128	
n = 200	AIC	Todos modelos	0,5	9,7	10,7	2,9	24,2	26,4	1,1	11,1	13,4	18,2	4.096
		Stepwise	3,6	35,0	35,5	1,4	11,5	10,9	0,1	0,8	1,2	8,4	42
		K passos	2,3	12,7	5,9	6,8	32,2	16,3	2,7	13,2	8,0	24,4	276
		Dois passos	3,1	29,7	30,9	1,6	13,9	13,0	0,3	3,4	4,0	7,4	128
	AICc	Todos modelos	0,7	11,6	10,9	3,1	27,0	24,9	1,3	10,5	10,0	20,8	4.096
		Stepwise	4,7	38,2	33,6	1,4	11,5	8,8	0,1	0,7	1,0	8,3	41
		K passos	3,2	14,7	5,8	7,9	34,1	13,9	2,6	12,0	5,8	25,9	275
		Dois passos	3,9	33,1	29,8	1,8	14,0	11,0	0,4	3,2	2,9	7,7	128
	BIC	Todos modelos	18,1	45,4	4,5	7,7	21,2	1,1	0,5	1,3	0,2	18,7	4.096
		Stepwise	36,7	55,9	5,5	0,6	1,2	0,1	0,0	0,0	0,0	0,9	32
		K passos	48,0	19,3	0,2	18,7	11,9	0,3	0,9	0,8	0,0	10,2	289
		Dois passos	36,3	55,0	5,5	1,0	2,0	0,1	0,0	0,1	0,0	1,4	128
HQC	Todos modelos	5,2	27,4	8,8	7,0	32,3	11,0	0,8	5,6	1,9	26,5	4.096	
	Stepwise	14,5	57,1	18,9	1,2	6,0	2,0	0,0	0,2	0,1	4,7	37	
	K passos	15,4	23,8	2,7	16,8	29,3	3,5	2,0	5,4	1,1	23,4	288	
	Dois passos	13,6	53,1	18,2	1,8	9,3	2,7	0,2	0,7	0,4	5,6	128	
n = 400	AIC	Todos modelos	0,0	2,4	2,9	0,1	33,1	33,0	0,0	11,4	17,1	32,0	4.096
		Stepwise	0,2	22,4	23,1	0,1	24,2	24,6	0,0	2,2	3,2	22,1	44
		K passos	0,1	3,5	1,8	0,5	42,0	23,9	0,1	15,9	12,2	40,7	252
		Dois passos	0,2	17,2	18,3	0,1	23,5	23,6	0,1	8,0	9,1	17,0	128
	AICc	Todos modelos	0,0	3,2	3,0	0,1	34,1	32,2	0,0	11,4	16,0	32,9	4.096
		Stepwise	0,2	24,3	22,9	0,2	24,7	22,8	0,0	2,0	3,0	22,4	44
		K passos	0,1	4,0	1,9	0,6	43,6	23,0	0,1	15,9	10,7	42,3	252
		Dois passos	0,2	19,0	18,0	0,1	24,1	22,7	0,1	7,8	8,0	17,5	128
	BIC	Todos modelos	2,4	29,0	1,0	2,3	60,4	2,9	0,0	1,9	0,1	59,2	4.096
		Stepwise	6,8	78,8	4,8	0,5	8,7	0,5	0,0	0,0	0,0	8,4	37
		K passos	10,7	22,2	0,4	15,5	48,4	1,2	0,4	1,3	0,1	47,3	271
		Dois passos	6,6	77,1	4,6	0,5	10,3	0,5	0,1	0,2	0,1	9,0	128
HQC	Todos modelos	0,1	10,7	2,8	0,7	60,2	15,5	0,1	7,3	2,6	58,4	4.096	
	Stepwise	1,1	55,6	14,7	0,4	21,1	6,4	0,0	0,6	0,2	19,9	40	
	K passos	1,6	11,9	1,6	3,8	62,5	8,5	0,4	8,3	1,5	60,4	243	
	Dois passos	1,0	50,8	13,6	0,4	24,1	6,7	0,0	2,4	1,0	19,2	128	

Tabela C.7: Resultados completos dos estudos de simulação para o cenário 7 da distribuição Normal.

n	Critério	Método	# $\beta < 3$			# $\beta = 3$			# $\beta > 3$			Modelo Correto	Média de Modelos
			# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$	# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$	# $\gamma < 2$	# $\gamma = 2$	# $\gamma > 2$		
n = 50	AIC	Todos modelos	0,6	9,5	15,2	0,4	10,6	22,1	0,9	9,9	30,8	4,0	4.096
		Stepwise	1,4	20,9	31,2	1,0	12,7	17,7	0,9	5,4	8,7	5,0	33
		K passos	4,8	16,9	7,7	2,5	21,2	10,8	1,6	17,6	17,0	9,7	267
		Dois passos	1,3	15,9	23,0	0,9	13,6	21,3	0,9	8,0	15,0	4,5	128
	AICc	Todos modelos	2,1	17,2	18,0	1,5	17,1	17,8	1,2	9,0	16,1	7,6	4.096
		Stepwise	2,5	31,4	29,0	2,2	14,6	11,8	1,2	3,9	3,3	6,1	30
		K passos	8,5	25,8	5,8	4,8	25,5	7,5	2,0	13,7	6,4	11,5	270
		Dois passos	2,4	24,8	23,2	2,2	16,8	15,3	1,2	7,4	6,9	5,6	128
	BIC	Todos modelos	11,8	38,8	15,9	2,6	13,8	7,6	0,5	4,8	4,2	6,9	4.096
		Stepwise	14,3	54,0	17,2	3,1	6,9	2,5	0,6	1,0	0,4	3,3	24
		K passos	35,5	34,7	2,0	4,9	15,7	1,1	0,9	4,6	0,7	7,8	231
		Dois passos	13,6	49,8	17,3	3,0	9,3	3,4	0,6	2,2	0,7	3,4	128
HQC	Todos modelos	4,0	20,6	17,5	1,8	16,2	15,1	1,0	7,7	16,1	7,0	4.096	
	Stepwise	4,4	37,3	28,1	2,5	11,8	9,2	1,0	3,1	2,6	5,1	28	
	K passos	13,2	29,0	5,3	4,8	22,4	5,9	1,7	11,5	6,1	10,3	252	
	Dois passos	4,2	31,7	24,1	2,3	14,1	11,7	1,0	5,6	5,4	4,8	128	
n = 100	AIC	Todos modelos	0,0	3,0	5,9	0,0	18,9	23,2	0,2	16,6	32,2	14,4	4.096
		Stepwise	0,0	11,4	14,5	0,1	20,7	28,5	0,3	9,5	15,1	16,5	35
		K passos	0,3	7,1	3,0	0,4	29,4	13,8	0,5	25,5	20,0	23,1	262
		Dois passos	0,0	7,4	9,7	0,1	18,0	25,4	0,2	15,4	23,7	12,8	128
	AICc	Todos modelos	0,0	5,4	7,1	0,1	22,5	24,2	0,3	16,9	23,5	17,7	4.096
		Stepwise	0,1	14,3	15,4	0,1	24,6	25,3	0,3	9,2	10,7	19,8	34
		K passos	0,3	10,1	3,2	0,5	34,6	12,5	0,5	24,3	13,9	27,3	251
		Dois passos	0,1	9,3	10,3	0,1	22,5	24,2	0,3	15,8	17,4	15,8	128
	BIC	Todos modelos	2,2	36,3	8,0	0,9	36,9	8,1	0,1	5,7	1,8	31,1	4.096
		Stepwise	3,2	54,9	9,8	1,6	23,1	5,0	0,2	1,9	0,4	19,6	27
		K passos	10,6	36,9	1,1	2,7	40,4	1,5	0,2	6,2	0,3	34,3	232
		Dois passos	3,2	49,2	8,9	1,5	26,4	5,8	0,2	4,0	0,8	21,0	128
HQC	Todos modelos	0,5	13,7	8,1	0,6	33,3	20,5	0,1	12,7	10,5	27,4	4.096	
	Stepwise	0,5	27,9	14,5	0,7	29,2	16,6	0,2	6,5	3,8	23,6	31	
	K passos	2,2	19,4	2,4	1,3	44,2	6,8	0,3	18,6	4,8	36,0	244	
	Dois passos	0,5	22,4	11,7	0,7	29,3	17,2	0,2	11,2	6,9	22,0	128	
n = 200	AIC	Todos modelos	0,0	0,6	0,7	0,0	16,0	27,9	0,0	23,1	31,7	15,4	4.096
		Stepwise	0,0	1,5	2,6	0,0	25,1	37,7	0,0	13,9	19,3	24,6	37
		K passos	0,0	0,8	0,5	0,0	31,6	14,7	0,0	30,9	21,5	31,3	243
		Dois passos	0,0	1,1	1,7	0,0	19,0	29,7	0,0	20,1	28,5	18,4	128
	AICc	Todos modelos	0,0	0,7	1,1	0,0	19,1	28,4	0,0	21,9	28,8	18,4	4.096
		Stepwise	0,0	1,8	2,7	0,0	27,9	37,0	0,0	14,0	16,5	27,4	36
		K passos	0,0	1,0	0,5	0,0	35,3	14,0	0,0	31,0	18,1	34,9	241
		Dois passos	0,0	1,3	1,8	0,0	21,5	29,8	0,0	20,7	24,8	20,8	128
	BIC	Todos modelos	0,0	10,4	2,0	0,0	73,5	7,5	0,0	6,2	0,4	71,9	4.096
		Stepwise	0,1	24,3	3,3	0,0	62,2	7,3	0,0	2,4	0,3	60,8	29
		K passos	0,5	11,6	0,3	0,0	77,7	2,4	0,0	7,3	0,2	75,9	209
		Dois passos	0,1	21,7	2,8	0,0	61,8	7,4	0,0	5,4	0,7	59,6	128
HQC	Todos modelos	0,0	3,5	0,9	0,0	48,5	20,8	0,0	18,0	8,3	47,6	4.096	
	Stepwise	0,0	7,5	3,4	0,0	53,0	23,6	0,0	8,3	4,2	51,6	32	
	K passos	0,0	3,6	0,7	0,0	62,4	7,9	0,0	20,8	4,4	61,4	225	
	Dois passos	0,0	6,1	2,6	0,0	48,0	22,0	0,0	14,4	6,9	46,2	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	21,7	28,7	0,0	18,7	30,9	21,7	4.096
		Stepwise	0,0	0,0	0,1	0,0	27,9	38,2	0,0	13,4	20,4	27,9	37
		K passos	0,0	0,0	0,0	0,0	34,1	15,1	0,0	30,8	20,0	34,1	245
		Dois passos	0,0	0,0	0,1	0,0	20,3	28,6	0,0	20,9	30,1	20,3	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	23,6	28,1	0,0	19,3	29,0	23,6	4.096
		Stepwise	0,0	0,0	0,1	0,0	29,6	37,7	0,0	13,6	19,1	29,6	37
		K passos	0,0	0,0	0,0	0,0	36,0	14,5	0,0	31,0	18,4	36,0	246
		Dois passos	0,0	0,0	0,1	0,0	21,8	28,4	0,0	21,2	28,4	21,8	128
	BIC	Todos modelos	0,0	0,7	0,0	0,0	86,0	6,4	0,0	6,3	0,6	86,0	4.096
		Stepwise	0,0	1,6	0,1	0,0	89,0	6,2	0,0	2,7	0,3	88,9	31
		K passos	0,0	0,6	0,0	0,0	91,6	1,6	0,0	5,9	0,3	91,6	197
		Dois passos	0,0	1,5	0,1	0,0	85,8	6,0	0,0	6,0	0,5	85,6	128
HQC	Todos modelos	0,0	0,0	0,1	0,0	58,4	18,3	0,0	15,9	7,3	58,4	4.096	
	Stepwise	0,0	0,4	0,3	0,0	64,1	22,4	0,0	9,5	3,4	64,0	33	
	K passos	0,0	0,1	0,0	0,0	71,1	6,8	0,0	19,2	2,8	71,0	209	
	Dois passos	0,0	0,3	0,2	0,0	56,4	19,9	0,0	17,2	5,9	56,3	128	

Tabela C.8: Resultados completos dos estudos de simulação para o cenário 8 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	6,7	15,3	19,0	3,8	12,5	18,7	1,8	6,9	15,3	4,9	4.096
		Stepwise	16,8	28,9	24,6	4,9	10,8	7,9	1,7	2,3	2,2	4,2	40
		K passos	20,9	17,6	8,9	8,4	15,8	9,9	3,2	8,3	7,0	6,9	262
		Dois passos	13,0	24,0	22,6	5,2	12,2	11,6	1,9	4,1	5,4	4,5	128
	AICc	Todos modelos	16,2	26,5	16,5	7,0	13,8	9,5	2,4	4,0	4,1	6,0	4.096
		Stepwise	29,9	34,0	15,6	7,4	7,1	2,9	1,5	1,1	0,4	3,0	36
		K passos	39,7	19,5	5,3	12,0	12,0	3,6	2,9	3,7	1,4	5,6	250
		Dois passos	25,2	29,8	16,2	7,9	10,6	4,9	1,9	2,1	1,4	4,2	128
	BIC	Todos modelos	40,7	27,1	9,3	5,9	7,6	3,7	1,7	2,0	2,0	3,3	4.096
		Stepwise	60,3	26,2	5,7	4,7	1,9	0,6	0,4	0,1	0,0	0,9	29
		K passos	73,8	10,0	1,1	7,2	4,4	0,6	1,4	1,1	0,3	2,3	244
		Dois passos	54,3	27,0	6,9	5,5	3,8	1,1	0,7	0,4	0,2	1,7	128
HQC	Todos modelos	16,9	24,3	15,2	6,1	13,7	10,9	1,8	4,9	6,2	6,0	4.096	
	Stepwise	33,4	32,1	15,5	6,6	6,3	3,4	1,4	0,8	0,6	2,6	35	
	K passos	42,8	17,2	4,9	10,3	11,4	3,9	2,6	4,4	2,3	5,4	249	
	Dois passos	28,1	29,0	16,5	6,9	9,4	5,1	1,6	1,5	1,9	3,8	128	
n = 100	AIC	Todos modelos	0,1	0,1	0,2	0,7	28,7	31,7	0,5	15,7	22,3	27,2	4.096
		Stepwise	1,4	9,7	10,2	1,1	28,4	34,6	0,6	6,3	7,7	24,7	45
		K passos	0,3	0,4	0,1	2,7	36,0	23,4	1,8	18,3	17,0	33,5	248
		Dois passos	1,1	8,1	8,7	1,1	23,7	31,0	0,7	12,0	13,6	19,5	128
	AICc	Todos modelos	0,1	0,4	0,4	1,2	40,3	27,8	1,2	13,6	15,0	37,7	4.096
		Stepwise	2,2	12,5	10,0	1,6	35,4	27,6	0,7	5,1	5,0	30,7	44
		K passos	0,7	0,7	0,3	4,1	45,2	19,4	1,9	17,2	10,5	41,8	248
		Dois passos	1,8	10,7	8,7	1,5	30,7	25,2	1,0	11,3	9,3	25,6	128
	BIC	Todos modelos	3,1	3,2	1,0	10,0	61,5	10,0	1,3	6,8	3,1	58,6	4.096
		Stepwise	16,9	34,2	6,1	6,2	29,5	5,6	0,4	0,9	0,2	26,9	39
		K passos	8,1	2,7	0,4	23,4	51,9	3,9	2,6	6,3	0,9	49,0	245
		Dois passos	15,2	31,5	5,9	6,6	31,0	5,9	1,0	2,4	0,6	27,6	128
HQC	Todos modelos	0,3	0,6	0,8	3,5	50,3	21,9	1,1	12,2	9,3	47,6	4.096	
	Stepwise	5,2	19,8	10,6	3,8	36,5	17,7	0,7	3,3	2,3	32,2	42	
	K passos	1,8	0,9	0,5	9,2	52,5	13,1	3,2	13,2	5,5	49,0	246	
	Dois passos	4,6	17,8	9,9	3,5	34,4	17,0	1,1	7,2	4,5	29,4	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,1	31,2	37,3	0,2	11,5	19,7	31,2	4.096
		Stepwise	0,0	0,8	0,8	0,0	39,6	42,2	0,1	6,9	9,6	39,5	46
		K passos	0,0	0,0	0,0	0,0	40,1	26,0	0,1	17,1	16,7	40,1	239
		Dois passos	0,0	0,6	0,7	0,0	30,7	33,0	0,1	15,8	19,2	30,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,2	36,1	35,8	0,1	11,6	16,2	36,1	4.096
		Stepwise	0,0	1,0	0,9	0,1	44,3	39,2	0,0	7,0	7,6	44,2	46
		K passos	0,0	0,0	0,0	0,1	45,2	24,7	0,1	16,9	13,0	45,2	230
		Dois passos	0,0	0,7	0,8	0,0	35,3	30,9	0,1	15,9	16,4	34,9	128
	BIC	Todos modelos	0,1	0,1	0,1	1,4	83,3	9,4	0,0	4,9	0,7	82,9	4.096
		Stepwise	0,4	11,1	1,3	1,4	75,9	7,9	0,0	1,6	0,3	75,3	42
		K passos	0,1	0,2	0,0	2,4	86,8	4,8	0,1	5,1	0,4	86,5	208
		Dois passos	0,4	10,8	1,3	1,2	72,1	7,3	0,2	5,9	0,9	71,4	128
HQC	Todos modelos	0,1	0,1	0,1	0,5	60,5	23,6	0,0	10,4	4,7	60,4	4.096	
	Stepwise	0,0	3,0	1,7	0,4	65,5	22,9	0,0	4,4	2,0	65,0	44	
	K passos	0,0	0,1	0,1	0,5	68,6	14,4	0,1	12,4	3,7	68,6	217	
	Dois passos	0,0	2,7	1,5	0,3	56,7	20,3	0,1	13,4	5,0	56,0	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	31,6	36,7	0,0	14,2	17,5	31,6	4.096
		Stepwise	0,0	0,0	0,0	0,0	41,5	44,8	0,0	6,3	7,3	41,5	46
		K passos	0,0	0,0	0,0	0,0	42,4	26,1	0,0	18,0	13,4	42,4	227
		Dois passos	0,0	0,0	0,0	0,0	34,1	37,4	0,0	13,7	14,8	34,1	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	34,2	36,1	0,0	14,4	15,3	34,2	4.096
		Stepwise	0,0	0,0	0,0	0,0	43,7	43,5	0,0	6,4	6,4	43,7	46
		K passos	0,0	0,0	0,0	0,0	44,9	25,5	0,0	18,0	11,5	44,9	224
		Dois passos	0,0	0,0	0,0	0,0	36,3	36,5	0,0	13,7	13,5	36,3	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,1	5,3	0,0	3,3	0,3	91,1	4.096
		Stepwise	0,0	0,3	0,0	0,0	93,1	5,8	0,0	0,7	0,0	93,1	42
		K passos	0,0	0,0	0,0	0,0	93,9	2,7	0,0	3,2	0,1	93,9	197
		Dois passos	0,0	0,3	0,0	0,0	91,2	5,7	0,0	2,6	0,2	91,2	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	68,6	20,2	0,0	8,0	3,2	68,6	4.096	
	Stepwise	0,0	0,1	0,0	0,0	74,1	22,2	0,0	2,6	1,0	74,1	44	
	K passos	0,0	0,0	0,0	0,0	76,3	12,0	0,0	9,6	2,2	76,3	203	
	Dois passos	0,0	0,1	0,0	0,0	68,2	21,1	0,0	8,4	2,2	68,2	128	

Tabela C.9: Resultados completos dos estudos de simulação para o cenário 9 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	4,9	17,7	18,3	3,2	21,2	17,3	0,7	7,7	9,0	5,5	4.096
		Stepwise	19,5	45,6	20,4	3,2	6,4	3,2	0,7	0,6	0,3	1,2	37
		K passos	16,0	23,7	7,5	8,9	22,6	8,5	2,4	6,4	4,0	7,4	292
		Dois passos	13,1	36,4	21,7	3,7	11,7	8,2	1,0	2,0	2,2	1,9	128
	AICc	Todos modelos	11,2	31,1	15,9	6,0	20,1	9,2	0,6	3,9	2,0	5,7	4.096
		Stepwise	30,9	48,8	12,2	3,2	3,5	0,9	0,3	0,2	0,0	0,8	34
		K passos	30,3	30,6	3,5	10,9	17,7	2,2	1,4	2,9	0,6	7,2	291
		Dois passos	23,4	44,7	13,9	4,8	8,3	3,0	0,6	1,1	0,2	1,7	128
	BIC	Todos modelos	31,3	37,5	7,8	4,5	12,9	3,1	0,2	1,7	1,0	3,8	4.096
		Stepwise	58,4	37,0	3,1	1,1	0,2	0,1	0,1	0,0	0,0	0,1	29
		K passos	63,5	19,9	0,9	7,2	6,6	0,4	0,7	0,8	0,1	3,2	273
		Dois passos	50,7	39,2	4,6	2,9	1,9	0,4	0,2	0,1	0,1	0,5	128
HQC	Todos modelos	11,9	29,2	15,5	5,6	18,8	10,9	0,4	4,8	2,9	5,3	4.096	
	Stepwise	35,1	46,9	11,2	2,6	2,9	0,9	0,1	0,2	0,0	0,7	33	
	K passos	33,0	28,0	3,6	10,0	16,8	2,7	1,4	3,3	1,2	6,9	294	
	Dois passos	26,7	43,6	13,9	3,9	6,9	3,2	0,5	0,8	0,4	1,4	128	
n = 100	AIC	Todos modelos	0,1	5,0	4,4	0,6	38,6	26,7	0,2	11,3	13,1	23,9	4.096
		Stepwise	1,1	26,7	16,5	0,7	29,9	19,1	0,2	4,2	1,6	15,7	42
		K passos	0,6	7,5	2,7	1,1	45,5	17,9	0,4	16,4	7,9	29,9	262
		Dois passos	0,5	21,3	14,9	0,5	29,6	21,9	0,3	7,1	3,9	15,8	128
	AICc	Todos modelos	0,3	7,4	4,8	1,4	47,8	21,5	0,3	9,1	7,4	29,6	4.096
		Stepwise	1,7	34,3	14,9	1,1	31,5	12,9	0,2	2,9	0,6	16,6	42
		K passos	0,9	10,9	2,6	2,0	53,8	12,4	0,5	13,1	3,9	35,0	264
		Dois passos	1,1	28,6	13,3	0,8	33,7	14,8	0,3	5,5	1,9	18,0	128
	BIC	Todos modelos	5,3	33,3	4,2	3,0	46,3	5,8	0,0	1,4	0,7	30,5	4.096
		Stepwise	16,4	64,8	5,3	2,0	10,5	0,8	0,0	0,1	0,0	5,5	35
		K passos	16,3	30,2	1,1	7,7	40,4	1,4	0,5	2,3	0,2	27,2	297
		Dois passos	12,7	63,6	5,9	2,3	13,7	1,3	0,0	0,4	0,0	7,5	128
HQC	Todos modelos	1,0	14,9	4,7	2,4	53,2	15,1	0,1	5,2	3,4	32,5	4.096	
	Stepwise	4,6	49,7	12,5	1,5	24,5	5,6	0,2	1,3	0,1	12,6	39	
	K passos	2,6	18,4	2,1	4,0	54,8	7,5	0,5	8,6	1,6	36,1	271	
	Dois passos	3,2	43,5	11,7	1,4	29,1	7,8	0,2	2,6	0,5	16,0	128	
n = 200	AIC	Todos modelos	0,0	0,1	0,1	0,0	35,2	34,6	0,0	15,3	14,7	33,7	4.096
		Stepwise	0,0	3,4	2,8	0,0	44,1	37,4	0,0	6,8	5,5	41,1	45
		K passos	0,0	0,2	0,2	0,0	45,7	24,8	0,0	18,2	10,9	43,8	229
		Dois passos	0,0	2,4	2,1	0,0	38,5	33,9	0,0	11,9	11,3	35,7	128
	AICc	Todos modelos	0,0	0,2	0,4	0,0	40,3	32,9	0,0	15,4	10,8	38,3	4.096
		Stepwise	0,0	4,2	2,8	0,0	48,6	33,9	0,0	6,3	4,2	44,9	45
		K passos	0,0	0,4	0,2	0,0	51,3	22,5	0,0	17,0	8,7	48,8	223
		Dois passos	0,0	2,7	2,1	0,0	43,4	31,3	0,0	11,8	8,7	39,6	128
	BIC	Todos modelos	0,0	6,5	0,6	0,0	82,7	6,5	0,0	3,5	0,2	75,8	4.096
		Stepwise	0,2	42,0	2,3	0,1	51,3	3,5	0,0	0,6	0,0	45,0	40
		K passos	0,2	7,0	0,1	0,4	86,0	2,7	0,0	3,5	0,1	80,0	220
		Dois passos	0,1	32,8	2,0	0,1	59,4	4,3	0,0	1,3	0,0	52,1	128
HQC	Todos modelos	0,0	1,3	0,4	0,0	65,0	20,1	0,0	10,5	2,7	61,4	4.096	
	Stepwise	0,0	15,1	3,5	0,0	61,5	16,5	0,0	2,7	0,7	55,4	42	
	K passos	0,0	1,8	0,3	0,0	73,8	11,1	0,0	10,9	2,1	69,8	216	
	Dois passos	0,0	11,2	2,8	0,0	61,4	17,0	0,0	5,9	1,7	55,0	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	35,9	34,0	0,0	14,6	15,5	35,8	4.096
		Stepwise	0,0	0,1	0,0	0,0	44,7	39,7	0,0	7,8	7,6	44,2	45
		K passos	0,0	0,0	0,0	0,0	44,9	24,9	0,0	17,8	12,4	44,8	227
		Dois passos	0,0	0,0	0,0	0,0	37,4	34,6	0,0	13,8	14,3	37,0	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	37,8	33,6	0,0	14,5	14,1	37,7	4.096
		Stepwise	0,0	0,1	0,0	0,0	47,0	38,4	0,0	7,6	6,8	46,5	45
		K passos	0,0	0,0	0,0	0,0	47,3	24,3	0,0	17,3	11,1	47,2	223
		Dois passos	0,0	0,0	0,0	0,0	39,6	33,6	0,0	13,6	13,1	39,2	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	93,1	4,5	0,0	2,3	0,1	91,8	4.096
		Stepwise	0,0	8,6	0,4	0,0	85,6	4,9	0,0	0,4	0,0	81,4	42
		K passos	0,0	0,0	0,0	0,0	94,3	2,9	0,0	2,6	0,1	93,6	199
		Dois passos	0,0	2,7	0,1	0,0	90,7	5,1	0,0	1,4	0,1	87,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	69,3	18,0	0,0	10,4	2,3	68,9	4.096	
	Stepwise	0,0	1,0	0,2	0,0	74,6	19,7	0,0	3,5	1,0	72,3	43	
	K passos	0,0	0,0	0,0	0,0	76,7	11,6	0,0	10,1	1,6	76,3	207	
	Dois passos	0,0	0,5	0,1	0,0	70,2	19,0	0,0	7,8	2,4	68,2	128	

Tabela C.10: Resultados completos dos estudos de simulação para o cenário 10 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$	# $\gamma < 4$	# $\gamma = 4$	# $\gamma > 4$		
n = 50	AIC	Todos modelos	10,9	26,5	20,2	3,7	8,1	12,0	1,5	7,5	9,6	3,0	4.096
		Stepwise	24,9	41,0	25,4	2,0	2,9	2,5	0,3	0,6	0,3	1,0	39
		K passos	31,3	23,4	7,9	8,4	10,2	4,9	3,7	6,1	4,1	3,7	299
		Dois passos	21,6	37,2	23,5	3,0	5,2	5,0	0,8	2,2	1,4	1,4	128
	AICc	Todos modelos	23,3	35,3	15,3	5,6	8,2	4,9	1,9	3,6	1,9	3,2	4.096
		Stepwise	39,9	42,9	13,0	2,0	1,3	0,5	0,2	0,1	0,0	0,4	36
		K passos	58,7	19,6	2,2	8,3	5,5	1,0	2,3	1,9	0,4	1,8	292
		Dois passos	34,9	41,7	13,3	3,5	3,3	1,5	1,1	0,6	0,2	1,0	128
	BIC	Todos modelos	46,9	31,9	7,6	3,7	4,6	1,8	1,3	1,3	0,9	1,8	4.096
		Stepwise	62,9	31,7	4,3	0,7	0,2	0,1	0,1	0,0	0,0	0,2	29
		K passos	84,1	8,4	0,5	4,1	1,3	0,2	0,7	0,6	0,1	0,7	278
		Dois passos	59,7	32,7	5,1	1,5	0,5	0,2	0,1	0,2	0,1	0,3	128
HQC	Todos modelos	23,2	32,7	15,1	4,9	7,6	6,7	1,7	4,5	3,6	2,7	4.096	
	Stepwise	42,1	40,7	13,5	1,7	1,0	0,6	0,3	0,1	0,1	0,4	34	
	K passos	57,8	18,8	2,7	7,6	5,9	1,5	2,5	2,4	0,8	1,9	286	
	Dois passos	37,3	39,9	14,2	2,8	2,5	1,7	0,9	0,6	0,3	0,7	128	
n = 100	AIC	Todos modelos	0,8	9,3	9,1	0,8	30,5	22,4	0,9	14,2	12,0	24,7	4.096
		Stepwise	3,2	42,3	31,7	0,7	10,7	8,9	0,1	1,4	1,0	7,1	45
		K passos	1,9	11,7	5,0	4,0	35,7	14,6	1,7	16,4	9,1	28,2	287
		Dois passos	2,7	35,8	27,4	0,9	14,3	11,7	0,2	3,9	3,0	6,6	128
	AICc	Todos modelos	1,2	14,8	9,2	1,2	36,2	17,6	0,9	12,9	6,0	29,8	4.096
		Stepwise	4,8	51,4	25,8	0,9	10,3	5,4	0,0	0,8	0,5	6,6	44
		K passos	4,6	16,7	4,7	6,7	38,3	10,3	1,7	13,0	4,1	30,3	284
		Dois passos	4,4	44,9	23,0	1,1	14,3	7,5	0,3	2,8	1,7	6,5	128
	BIC	Todos modelos	13,2	44,5	5,0	3,6	26,1	3,9	0,4	2,7	0,6	21,9	4.096
		Stepwise	33,9	57,2	6,9	0,8	1,0	0,2	0,0	0,0	0,0	0,7	36
		K passos	45,9	20,5	1,1	12,8	15,3	1,0	1,3	1,9	0,3	12,3	301
		Dois passos	33,1	56,5	7,0	1,1	1,6	0,4	0,0	0,2	0,0	0,8	128
HQC	Todos modelos	2,8	25,2	9,3	1,9	36,4	11,6	0,9	9,1	2,8	30,5	4.096	
	Stepwise	11,6	59,9	19,2	1,1	5,8	2,0	0,0	0,3	0,1	3,8	41	
	K passos	12,7	22,7	3,4	11,0	32,1	6,1	1,8	8,3	1,9	25,2	298	
	Dois passos	11,0	55,2	18,2	1,3	9,5	3,2	0,1	1,2	0,2	4,5	128	
n = 200	AIC	Todos modelos	0,0	0,7	0,4	0,0	38,0	29,1	0,0	18,8	13,0	37,7	4.096
		Stepwise	0,0	13,0	9,8	0,0	38,3	28,9	0,0	6,0	4,0	36,0	48
		K passos	0,0	0,6	0,3	0,1	45,2	22,4	0,0	20,3	11,2	44,9	234
		Dois passos	0,0	10,6	8,5	0,0	33,7	25,7	0,0	12,9	8,6	29,6	128
	AICc	Todos modelos	0,0	1,1	0,5	0,0	42,8	27,0	0,0	19,1	9,5	42,6	4.096
		Stepwise	0,0	15,4	9,8	0,0	41,1	25,4	0,0	5,5	2,7	38,7	47
		K passos	0,0	0,8	0,4	0,1	50,8	19,7	0,0	19,4	8,6	50,5	235
		Dois passos	0,0	12,9	8,6	0,0	36,6	23,1	0,0	12,4	6,4	32,1	128
	BIC	Todos modelos	0,2	12,0	1,0	0,1	77,8	4,6	0,0	3,9	0,4	76,8	4.096
		Stepwise	0,9	68,7	5,1	0,1	23,0	1,8	0,0	0,4	0,0	22,4	42
		K passos	2,4	10,9	0,5	2,8	76,4	2,7	0,0	4,2	0,2	75,3	244
		Dois passos	0,9	66,9	4,8	0,1	23,9	2,0	0,0	1,2	0,1	22,2	128
HQC	Todos modelos	0,0	3,0	0,7	0,0	64,2	16,4	0,0	12,5	3,2	63,6	4.096	
	Stepwise	0,1	36,7	8,9	0,0	41,0	10,9	0,0	1,8	0,5	38,9	45	
	K passos	0,1	2,7	0,6	0,4	70,6	10,3	0,0	12,9	2,3	69,9	233	
	Dois passos	0,1	33,8	8,5	0,0	40,1	10,2	0,0	5,6	1,6	35,7	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	40,2	27,2	0,0	17,2	15,4	40,2	4.096
		Stepwise	0,0	0,8	0,9	0,0	51,7	33,6	0,0	6,7	6,3	51,5	49
		K passos	0,0	0,0	0,0	0,0	46,6	22,1	0,0	19,2	12,1	46,6	224
		Dois passos	0,0	0,7	0,6	0,0	40,4	27,4	0,0	18,0	12,9	40,1	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	42,2	27,3	0,0	17,5	13,0	42,2	4.096
		Stepwise	0,0	1,0	0,8	0,0	54,1	32,0	0,0	6,6	5,4	53,9	49
		K passos	0,0	0,0	0,0	0,0	49,4	21,3	0,0	18,8	10,4	49,4	222
		Dois passos	0,0	0,8	0,5	0,0	42,7	26,0	0,0	18,1	11,8	42,4	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	93,3	4,0	0,0	2,5	0,2	93,3	4.096
		Stepwise	0,0	19,6	1,0	0,0	75,5	3,2	0,0	0,6	0,1	75,5	46
		K passos	0,0	0,2	0,0	0,0	94,6	2,5	0,0	2,6	0,1	94,6	208
		Dois passos	0,0	19,0	1,0	0,0	74,0	3,1	0,0	2,8	0,2	73,3	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	74,2	14,1	0,0	9,8	1,9	74,2	4.096	
	Stepwise	0,0	5,5	1,4	0,0	75,5	14,7	0,0	2,4	0,5	75,4	47	
	K passos	0,0	0,0	0,0	0,0	77,3	10,6	0,0	10,2	1,8	77,3	208	
	Dois passos	0,0	4,9	1,2	0,0	68,4	13,7	0,0	10,1	1,7	67,7	128	

Tabela C.11: Resultados completos dos estudos de simulação para o cenário 11 da distribuição Normal.

n	Critério	Método	# $\beta < 5$			# $\beta = 5$			# $\beta > 5$			Modelo Correto	Média de Modelos
			# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$	# $\gamma < 3$	# $\gamma = 3$	# $\gamma > 3$		
n = 50	AIC	Todos modelos	0,0	0,0	0,0	6,0	21,8	21,9	6,4	16,5	27,4	14,6	4,096
		Stepwise	0,0	0,0	0,0	9,2	27,4	30,9	7,3	11,8	13,5	16,8	45
		K passos	0,0	0,0	0,0	9,3	24,5	18,1	8,4	18,7	20,9	17,3	240
		Dois passos	0,0	0,0	0,0	6,4	19,2	25,1	8,1	18,6	22,6	11,4	128
	AICc	Todos modelos	0,0	0,0	0,0	15,3	33,2	16,7	9,1	16,0	9,7	21,7	4,096
		Stepwise	0,0	0,0	0,0	20,0	38,1	18,0	10,7	9,1	4,0	23,8	43
		K passos	0,0	0,0	0,0	20,4	34,2	11,6	12,5	15,2	6,1	25,4	217
		Dois passos	0,0	0,0	0,0	15,2	29,9	15,7	12,9	17,3	8,9	17,9	128
	BIC	Todos modelos	0,0	0,0	0,0	29,9	34,5	9,9	10,7	10,0	5,0	23,5	4,096
		Stepwise	0,0	0,0	0,0	36,5	37,0	10,4	10,0	4,8	1,4	24,4	40
		K passos	0,0	0,0	0,0	36,4	33,4	6,7	12,2	8,7	2,6	25,9	209
		Dois passos	0,0	0,0	0,0	31,3	33,5	10,3	12,0	9,5	3,4	22,1	128
HQC	Todos modelos	0,0	0,0	0,0	14,0	30,1	17,8	8,7	15,3	14,1	20,2	4,096	
	Stepwise	0,0	0,0	0,0	18,6	35,4	21,7	9,1	9,4	5,7	22,1	43	
	K passos	0,0	0,0	0,0	18,7	31,6	13,5	11,2	15,8	9,3	23,2	222	
	Dois passos	0,0	0,0	0,0	14,3	27,5	19,2	11,0	16,7	11,4	16,7	128	
n = 100	AIC	Todos modelos	0,0	0,0	0,0	1,4	30,0	32,7	0,8	14,7	20,4	29,2	4,096
		Stepwise	0,0	0,0	0,0	1,2	34,9	40,3	0,7	10,1	12,8	33,8	47
		K passos	0,0	0,0	0,0	1,1	36,8	27,1	1,0	16,6	17,3	36,0	245
		Dois passos	0,0	0,0	0,0	0,9	23,1	27,6	1,0	21,2	26,3	22,1	128
	AICc	Todos modelos	0,0	0,0	0,0	2,1	39,9	30,1	1,1	13,5	13,3	38,6	4,096
		Stepwise	0,0	0,0	0,0	1,9	44,0	34,7	1,0	10,0	8,5	42,2	46
		K passos	0,0	0,0	0,0	2,0	47,1	22,9	1,2	15,8	11,0	46,0	230
		Dois passos	0,0	0,0	0,0	1,4	30,7	24,4	1,3	22,9	19,3	29,3	128
	BIC	Todos modelos	0,0	0,0	0,0	10,9	69,1	10,7	1,5	5,8	2,0	66,5	4,096
		Stepwise	0,0	0,0	0,0	11,6	69,7	12,0	1,6	4,0	1,1	66,8	43
		K passos	0,0	0,0	0,0	11,7	71,7	6,9	2,0	6,6	1,1	69,9	206
		Dois passos	0,0	0,0	0,0	9,8	58,8	10,2	3,2	14,9	3,2	56,3	128
HQC	Todos modelos	0,0	0,0	0,0	3,3	52,5	24,5	1,5	10,4	7,8	50,9	4,096	
	Stepwise	0,0	0,0	0,0	3,8	55,8	27,0	1,2	7,7	4,6	53,3	45	
	K passos	0,0	0,0	0,0	4,1	58,5	17,3	1,5	12,4	6,2	56,9	221	
	Dois passos	0,0	0,0	0,0	3,1	41,4	20,9	1,8	21,3	11,5	39,7	128	
n = 200	AIC	Todos modelos	0,0	0,0	0,0	0,0	30,4	37,2	0,0	15,4	17,0	30,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	36,5	40,7	0,0	10,9	12,0	36,4	47
		K passos	0,0	0,0	0,0	0,0	40,1	28,0	0,0	17,2	14,7	40,1	229
		Dois passos	0,0	0,0	0,0	0,0	24,7	28,7	0,0	22,4	24,1	24,7	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	34,7	36,7	0,0	15,0	13,6	34,7	4,096
		Stepwise	0,0	0,0	0,0	0,0	40,9	38,6	0,0	10,9	9,6	40,9	46
		K passos	0,0	0,0	0,0	0,0	45,3	26,0	0,0	16,9	11,8	45,3	231
		Dois passos	0,0	0,0	0,0	0,0	28,3	27,2	0,0	23,3	21,1	28,3	128
	BIC	Todos modelos	0,0	0,0	0,0	0,1	86,7	8,1	0,0	4,5	0,6	86,7	4,096
		Stepwise	0,0	0,0	0,0	0,2	86,8	9,1	0,0	3,4	0,5	86,8	43
		K passos	0,0	0,0	0,0	0,2	89,5	4,9	0,0	5,0	0,4	89,5	201
		Dois passos	0,0	0,0	0,0	0,2	77,1	8,2	0,0	13,0	1,4	77,1	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	61,3	23,1	0,0	11,8	3,8	61,3	4,096	
	Stepwise	0,0	0,0	0,0	0,1	66,1	24,1	0,0	6,7	2,9	66,1	45	
	K passos	0,0	0,0	0,0	0,1	70,1	14,9	0,0	11,5	3,4	70,1	211	
	Dois passos	0,0	0,0	0,0	0,1	52,5	19,3	0,0	20,2	7,9	52,5	128	
n = 400	AIC	Todos modelos	0,0	0,0	0,0	0,0	32,9	35,6	0,0	14,9	16,6	32,9	4,096
		Stepwise	0,0	0,0	0,0	0,0	39,4	39,6	0,0	10,0	11,0	39,4	47
		K passos	0,0	0,0	0,0	0,0	43,1	26,2	0,0	16,6	14,2	43,1	234
		Dois passos	0,0	0,0	0,0	0,0	25,4	26,0	0,0	23,8	24,8	25,4	128
	AICc	Todos modelos	0,0	0,0	0,0	0,0	35,4	35,3	0,0	14,6	14,7	35,4	4,096
		Stepwise	0,0	0,0	0,0	0,0	41,8	38,8	0,0	9,8	9,6	41,8	47
		K passos	0,0	0,0	0,0	0,0	45,3	25,9	0,0	16,7	12,1	45,3	223
		Dois passos	0,0	0,0	0,0	0,0	27,3	25,7	0,0	24,3	22,7	27,3	128
	BIC	Todos modelos	0,0	0,0	0,0	0,0	91,5	5,7	0,0	2,8	0,0	91,5	4,096
		Stepwise	0,0	0,0	0,0	0,0	92,8	5,3	0,0	1,7	0,2	92,8	43
		K passos	0,0	0,0	0,0	0,0	94,1	2,8	0,0	2,8	0,3	94,1	200
		Dois passos	0,0	0,0	0,0	0,0	83,7	4,8	0,0	10,8	0,7	83,7	128
HQC	Todos modelos	0,0	0,0	0,0	0,0	70,6	18,7	0,0	8,2	2,5	70,6	4,096	
	Stepwise	0,0	0,0	0,0	0,0	72,0	20,0	0,0	6,1	2,0	72,0	44	
	K passos	0,0	0,0	0,0	0,0	76,5	12,0	0,0	9,5	2,1	76,5	210	
	Dois passos	0,0	0,0	0,0	0,0	57,0	16,3	0,0	21,0	5,6	57,0	128	

Referências Bibliográficas

- Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in B. N. Petrov e F. Csaki, eds, 'Second International Symposium on Information Theory', Akademiai Kiado, Budapest, pp. 267–281. [11](#)
- Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control* **19**(6), 716–723. [1](#), [11](#)
- Akaike, H. (1978), 'A bayesian analysis of the minimum aic procedure', *Annals of the Institute of Statistical Mathematics A* **30**, 9–14. [1](#), [12](#)
- Atkinson, A. C. (1985), Plots, transformations and regression: an introduction to graphical methods of diagnostic regression analysis, Technical report. [8](#)
- Bayer, F. e Cribari-Neto, F. (2015a), 'Bootstrap-based model selection criteria for beta regressions', *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research* **24**(4), 776–795. [10](#)
- Bayer, F. M. e Cribari-Neto, F. (2015b), 'Model selection criteria in beta regression with varying dispersion', *Communications in Statistics-Simulation and Computation* **46**(1), 729–746. [2](#), [14](#), [44](#)
- Bolfarine, H. e Sandoval, M. C. (2001), *Introdução à inferência estatística*, Vol. 2, SBM, São Paulo, Brasil. [1](#)
- Box, G. E. P. e Cox, D. R. (1964), 'An analysis of transformations', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 211–252. [1](#)
- Breiman, L. (2001), 'Random forests', *Machine learning* **45**(1), 5–32. [31](#)
- Burnham, K. P. e Anderson, D. R. (2003), *Model selection and multimodel inference: a practical information-theoretic approach*, Springer Science & Business Media, Nova York, Estados Unidos. [11](#), [12](#)
- Cavalaro, L. L. (2016), Seleção de variáveis em modelos de regressão normais com dispersão variável, Technical report, Universidade Federal de São Carlos, São Carlos, Brasil. [10](#)

- Claeskens, G. e Hjort, N. L. (2008), *Model selection and model averaging*, Vol. 330, Cambridge University Press, Cambridge, Inglaterra. 2
- Cole, T. J. e Green, P. J. (1992), 'Smoothing reference centile curves: the lms method and penalized likelihood', *Statistics in medicine* **11**(10), 1305–1319. 8
- Cottet, R., Kohn, R. J. e Nott, D. J. (2008), 'Variable selection and model averaging in semi-parametric overdispersed generalized linear models', *Journal of the American Statistical Association* **103**(482). 10
- Davison, A., Gigli e A (1989), 'Deviance residuals and normal scores plots', *Biometrika* **76**(2), 211–221. 8
- Demirtas, H. e Allozi, R. (2018), 'Multirng: Multivariate pseudo-random number generation', *Journal of Statistical Software* . 19
- Dunn, P. K. e Smyth, G. K. (1996), 'Randomized quantile residuals', *Journal of Computational and Graphical Statistics* **5**(3), 236–244. 8
- Efron, B. e Tibshirani, R. (1997), 'Improvements on cross-validation: the 632+ bootstrap method', *Journal of the American Statistical Association* **92**(438), 548–560. 2
- Freund, Y., Schapire, R. E. et al. (1996), Experiments with a new boosting algorithm, in 'icml', Vol. 96, Citeseer, pp. 148–156. 40
- Friedman, J. H. (2001), 'Greedy function approximation: a gradient boosting machine', *Annals of statistics* . 40
- Friedman, J., Hastie, T. e Tibshirani, R. (2001), *The elements of statistical learning*, Vol. 1, Springer series in statistics, Berlim, Alemanha. 39
- Friedman, J., Hastie, T. e Tibshirani, R. (2010), 'Regularization paths for generalized linear models via coordinate descent', *Journal of statistical software* **33**(1), 1–22. 40
- Golfe.tur (2019), Disponível em: < <http://www.golfe.tur.br/index.php> >. Acessado em: 28 jan. 2019. 32, 33
- Hannan, E. J. e Quinn, B. G. (1979), 'The determination of the order of an autoregression', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 190–195. 1, 13
- Hoerl, A. E. e Kennard, R. W. (1970), 'Ridge regression: Biased estimation for nonorthogonal problems', *Technometrics* **12**(1), 55–67. 39
- Hofner, B., Mayr, A. e Schmid, M. (2014), 'gamboostlss: An r package for model building and variable selection in the gamlss framework', *arXiv preprint:1407.1774* . 41

- Hurvich, C. M. e Tsai, C.-L. (1989), 'Regression and time series model selection in small samples', *Biometrika* **76**(2), 297–307. [1](#), [13](#)
- James, G., Witten, D., Hastie, T. e Tibshirani, R. (2013), *An introduction to statistical learning*, Vol. 112, Springer Science & Business Media, Nova York, Estados Unidos. [40](#)
- Kutner, M. H., Nachtsheim, C. e Neter, J. (2004), *Applied linear regression models*, 5 edn, McGraw-Hill Irwin, Nova York, Estados Unidos. [9](#), [17](#)
- Mayr, A., Fenske, N., Hofner, B., Kneib, T. e Schmid, M. (2012), 'Generalized additive models for location, scale and shape for high dimensional data: a flexible approach based on boosting', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **61**(3). [40](#), [41](#)
- McCullagh, P. e Nelder, J. A. (1989), *Generalized linear models*, Vol. 37, Chapman and Hall, Londres, Inglaterra. [8](#)
- Nelder, J. A. e Wedderburn, R. W. M. (1972), 'Generalized linear models', *Journal of the Royal Statistical Society* **135**(3), "370–384". [1](#)
- Paula, G. A. (2004), *Modelos de regressão: com apoio computacional*, IME-USP, São Paulo, Brasil. [3](#), [7](#), [8](#)
- Paula, G. A. (2013), 'On diagnostics in double generalized linear models', *Computational Statistics & Data Analysis* **68**, 44–51. [8](#)
- R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. Acessado em: 21 abr. 2019. [2](#)
- Rigby, R. A. e Stasinopoulos, D. M. (2005), 'Generalized additive models for location, scale and shape', *Journal of the Royal Statistical Society: Series C (Applied Statistics)* **54**(3), 507–554. [8](#), [18](#), [45](#)
- Rigby, R. A. e Stasinopoulos, M. D. (1996), Mean and dispersion additive models, in 'Statistical theory and computational aspects of smoothing', Springer, pp. 215–230. [8](#)
- Schwarz, G. (1978), 'Estimating the dimension of a model', *The annals of statistics* **6**(2), 461–464. [1](#), [12](#), [13](#)
- Sen, P. K., Singer, J. M. e de Lima, A. C. P. (2010), *From finite sample to asymptotic methods in statistics*, Cambridge University Press, Cambridge, Inglaterra. [8](#)
- Shibata, R. (1976), 'Selection of the order of an autoregressive model by Akaike's information criterion', *Biometrika* **63**(1), 117–126. [13](#)
- Smyth, G. K. (1989), 'Generalized linear models with varying dispersion', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 47–60. [1](#), [3](#), [5](#)

- SportRegras* (2019), Disponível em: < <https://sportsregras.com/golfe-regras-historia/> >. Acessado em: 28 jan. 2019. 32, 33
- sportsillustrated.cnn (2004), 'Pga', Disponível em: < <http://users.stat.ufl.edu/~winner/datasets.html> >. Acessado em: 6 dez. 2018. 33
- Stasinopoulos, D. M. e Rigby, R. A. (2007), 'Generalized additive models for location scale and shape (gamlss) in r', *Journal of Statistical Software* 23(7), 1–46. 3, 18
- Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V. e De Bastiani, F. (2017), *Flexible regression and smoothing: using gamlss in R*, Chapman and Hall/CRC Press, Boca Raton, Estados Unidos. 2, 17, 44
- Terrell, G. R. (2002), 'The gradient statistic', *Computing Science and Statistics* 34(34), 206–215. 8
- Tibshirani, R. (1996), 'Regression shrinkage and selection via the lasso', *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 267–288. 39
- Wang, Z. (2019), 'bst: Gradient boosting', Disponível em: < <https://CRAN.R-project.org/package=bst> >. R package. Acessado em: 28 jan. 2019. 40
- Wu, L.-C., Zhang, Z.-Z. e Xu, D.-K. (2012), 'Variable selection for joint mean and dispersion models of the Lognormal Distribution', *Hacettepe Journal of Mathematics and Statistics* 41(2). 10
- Wu, L. e Li, H. (2012), 'Variable selection for joint mean and dispersion models of the Inverse Gaussian distribution', *Metrika* 75(6), 795–808. 10
- Xu, D., Zhang, Z. e Wu, L. (2014), 'Variable selection in high-dimensional double generalized linear models', *Statistical Papers* 55(2), 327–347. 10