
Análise de diagnóstico em modelos de regressão
ZAGA e ZAIG.

Juliana Scudilio Rodrigues

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito: /02/2016

Assinatura: _____

Juliana Scudilio Rodrigues

Análise de diagnóstico em modelos de regressão ZAGA e ZAIG.

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Mestra em Estatística – Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA.*

Área de Concentração: Estatística

Orientador: Prof. Dr. Gustavo Henrique de Araujo Pereira

UFSCAR – São Carlos
Fevereiro de 2016

Juliana Scudilio Rodrigues

Diagnostic analysis in ZAGA and ZAIG regression models.

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação - ICMC-USP and to the Departamento de Estatística – DEs-UFSCar, in partial fulfillment of the requirements for the degree of the Master Joint Graduate Program in Statistics DEs-UFSCar/ICMC-USP. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Gustavo Henrique de Araujo Pereira

UFSCAR – São Carlos
February 2016

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

R696a Rodrigues, Juliana Scudilio
Análise de diagnóstico em modelos de regressão
ZAGA e ZAIG / Juliana Scudilio Rodrigues. -- São
Carlos : UFSCar, 2016.
119 p.

Dissertação (Mestrado) -- Universidade Federal de
São Carlos, 2016.

1. Análise de diagnóstico. 2. Fundo de
investimento. 3. Modelos de regressão inflacionado
no zero. 4. Modelo ZAGA. 5. Modelo ZAIG. I. Título.



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 da candidata Juliana Scudilio Rodrigues, realizada em 10/03/2016:

Gustavo Henrique de A. Pereira

Prof. Dr. Gustavo Henrique de Araujo Pereira
UFSCar

Denise Aparecida Botter

Profa. Dra. Denise Aparecida Botter
USP

@Kelly

Profa. Dra. Maria Kelly Venezuela
Insper

Agradecimentos

A Deus por ter me dado saúde e força para superar as dificuldades.

À Instituição pelo ambiente criativo e amigável que proporciona. A Universidade Federal de São Carlos (UFSCAR) e a Universidade de São Paulo (USP-IME) em especial ao Departamento de Estatística de ambas as universidades, pela oportunidade de fazer o mestrado em Estatística.

Ao professor Dr. Gustavo Henrique de Araujo Pereira, pela orientação, pelo empenho dedicado à elaboração deste trabalho, pelo paciente trabalho de revisão da redação e pelo apoio e confiança nesses dois anos de caminhada.

Aos Professores Dr. Luís Aparecido, Dr. Afrânio M. Corrêa Vieira, Dra. Maria Kelly Venezuela e Dra. Denise Aparecida Botter pelas sugestões oportunas e correções feitas na qualificação e na defesa do mestrado.

Agradeço a todos os professores do departamento de estatística da USP e UFSCAR por me proporcionar o conhecimento não apenas racional, mas a manifestação do caráter e afetividade da educação no processo de formação profissional, por tanto que se dedicaram a mim, não somente por terem me ensinado, mas por terem me feito aprender. A palavra mestre, nunca fará justiça aos professores dedicados aos quais sem nominar terão os meus eternos agradecimentos.

Aos meus pais, pelo amor, incentivo e apoio incondicional. Agradeço a minha mãe Fatima Scudilio Rodrigues, heroína que me deu apoio, incentivo nas horas difíceis, de desânimo e cansaço. Ao meu pai Mario José Rodrigues que apesar de todas as dificuldades me fortaleceu e sempre esteve ao meu lado apoiando meu sonho.

Obrigada minha irmã, sobrinha e namorado que nos momentos de minha ausência dedicados ao estudo superior, sempre fizeram entender que o futuro é feito a partir da constante dedicação no presente!

Obrigada! Primos e tios pela contribuição valiosa. Meus agradecimentos aos amigos, companheiros de trabalhos e irmãos na amizade que fizeram parte da minha formação e que vão continuar presentes em minha vida.

A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

Resumo

Rodrigues, J. S. **Análise de diagnóstico em modelos de regressão ZAGA e ZAIG.**

Resíduos desempenham um papel importante na verificação do ajuste do modelo e na identificação de observações discrepantes e/ou influentes. Neste trabalho, estudamos duas classes de resíduos para os modelos de regressão gama inflacionados no zero (ZAGA) e gaussiana inversa inflacionados no zero (ZAIG). Essas classes de resíduos são uma função de um resíduo para o componente contínuo do modelo e da estimativa de máxima verossimilhança da probabilidade da observação assumir o valor zero. Estudos de simulação de Monte Carlo foram realizados para examinar as propriedades dessas classes de resíduos em ambos os modelos de regressão (ZAGA e ZAIG). Os resultados mostraram que um resíduo de uma dessas classes tem algumas propriedades semelhantes à da distribuição normal padrão nos modelos estudados. Além desse objetivo principal, descrevemos os modelos de regressão ZAGA e ZAIG, estudamos propriedades de alguns resíduos em modelos lineares generalizados com resposta gama e gaussiana inversa e discutimos outros aspectos de análise de diagnóstico nos modelos ZAGA e ZAIG. Para finalizar, foi feita uma aplicação com dados reais de fundos de investimentos, em que ajustamos o modelo ZAIG, para exemplificar os tópicos discutidos e mostrar a importância desses modelos e as vantagens de um dos resíduos estudados na análise de dados reais.

Palavras-chave: análise de diagnóstico, fundo de investimento, modelos de regressão inflacionado no zero, modelo ZAGA, modelo ZAIG, resíduo quantílico.

Abstract

Rodrigues, J. S. **Diagnostic analysis in ZAGA and ZAIG regression models.**

Residuals play an important role in checking model adequacy and in the identification of outliers and influential observations. In this paper, we studied two class of residuals for the zero adjusted gamma regression model (ZAGA) and the zero adjusted inverse Gaussian regression model (ZAIG). These classes of residuals are function of a residual for the continuous component of the model and the maximum likelihood estimate of the probability of the observation assuming the zero value. Monte Carlo simulation studies are performed to examine the properties of this class of residuals in both models (ZAGA and ZAIG). Results showed that a residual of one of these class has some similar properties to the standard normal distribution in the studied models. We also described ZAGA and ZAIG regression models, studied properties of some residuals in generalized linear models with response gamma and inverse Gaussian and discussed other aspects of diagnostic analysis in ZAGA and ZAIG models. To finsih, we presented a real dataset application from investment funds of Brazil. We fitted the ZAIG model to illustrate the topics discussed and showed the importance of these models and the advantages of one of the studied residuals in the analysis of real dataset.

Keywords: diagnostic analysis, inflated regression models, investment funds, quantile residual, ZAGA models, ZAIG models.

Sumário

Lista de Figuras	vi
Lista de Tabelas	vii
1 Introdução	1
2 Modelos de regressão ZAGA e ZAIG	3
2.1 Distribuições contínuas inflacionadas no zero	4
2.1.1 Distribuição ZAGA	5
2.1.2 Distribuição ZAIG	7
2.2 Modelos de regressão inflacionados no zero	8
2.2.1 Modelos de regressão gama inflacionados no zero	9
2.2.2 Modelos de regressão gaussiana inversa inflacionados no zero	11
2.3 Estimação nos modelos ZAGA e ZAIG	12
2.4 Intervalos de confiança e testes de hipóteses	13
2.5 Seleção de modelos	14
3 Análise de diagnóstico	15
3.1 Análise de diagnóstico em MLGs	15
3.1.1 Resíduo desvio	16
3.1.2 Resíduo de Pearson	17
3.1.3 Resíduo quantílico	17
3.1.4 Resíduo Anscombe	18
3.1.5 Resíduo Williams	19
3.2 Estudos de simulação	19
3.2.1 Modelo gama	20
3.2.2 Modelo gaussiana inversa	24
3.2.3 Considerações finais sobre os estudos de simulação nos MLGs	26
3.3 Análise de diagnóstico em modelos inflacionados ZAGA e ZAIG	28
3.3.1 Identificação de observações discrepantes - componente contínuo	28
3.4 Estudos de simulação para os modelos inflacionados no zero	33

3.4.1	Estudos de simulação para o modelo ZAGA	34
3.4.2	Estudos de simulação para o modelo ZAIG	37
3.4.3	Estudos de simulação para o caso do ajuste do modelo incorreto	40
3.5	Outros aspectos da análise de diagnóstico	41
3.5.1	Observações discrepantes - componente discreto	42
3.5.2	Pontos influentes	43
3.5.3	Gráficos de diagnóstico	44
3.5.4	Teste de qualidade de ajuste para o componente discreto	47
4	Aplicação	49
4.1	Descrição dos dados	49
4.2	Ajuste e análise de diagnóstico	53
4.3	Discussão sobre a identificação de observações positivas discrepantes.	58
5	Conclusões e trabalhos futuros	61
5.1	Próximos passos	62
A	Demonstrações dos resultados do Capítulo (2)	63
A.1	Modelo ZAGA	63
A.1.1	Obtenção da função escore para o modelo ZAGA	63
A.1.2	Obtenção da matriz de informação de Fisher - modelo ZAGA	64
A.2	Modelo ZAIG	66
A.2.1	Obtenção da função escore para o modelo ZAIG	66
A.2.2	Obtenção da matriz de informação de Fisher - modelo ZAIG	67
B	Tabelas dos estudos de simulação no modelo gama	69
C	Tabelas dos estudos de simulação no modelo gaussiana inversa	77
D	Tabelas dos estudos de simulação no modelo ZAGA	85
E	Tabelas dos estudos de simulação no modelo ZAIG	99
F	Tabelas do ajuste do modelo incorreto	113
	Referências Bibliográficas	116

Lista de Figuras

2.1	Gráfico da distribuição ZAGA para diferentes valores dos parâmetros	6
2.2	Gráfico da distribuição ZAIG para diferentes valores dos parâmetros	7
3.1	Gráfico dos resíduos para diferentes valores de α	31
4.1	Histograma da variável resposta - observações diferentes de zero.	51
4.2	Análise de diagnóstico componente contínuo do modelo ZAIG.	53
4.3	Análise de diagnóstico componente discreto do modelo ZAIG.	54
4.4	Análise de diagnóstico distância de Cook.	55
4.5	Gráfico envelope.	56
4.6	Resíduos para o componente contínuo do modelo ZAIG.	59

Lista de Tabelas

3.1	Estruturas das tabelas apresentadas no Apêndice B e C	20
3.2	Alterações feitas nos cenários em relação ao cenário I.	21
3.3	Comparação da estatística de Anderson-Darling para $n=15$ e $n=50$ no modelo gama	23
3.4	Alterações feitas nos cenários em relação ao cenário I.	24
3.5	Comparação da estatística de Anderson-Darling para $n=15$ e $n=50$ no modelo gaussiana inversa.	27
3.6	Estrutura das tabelas apresentadas nos apêndices D e E	34
3.7	Alterações feitas nos cenários em relação ao cenário 1	35
3.8	Alterações feitas nos cenários em relação ao cenário 1	39
4.1	Medidas descritivas da variável resposta Y considerando apenas as observa- ções diferentes de zero.	51
4.2	Medidas descritivas das covariáveis.	52
4.3	Tabela de frequências das covariáveis categorizadas.	52
4.4	Modelo final para a variável taxa de crescimento do fundo referenciados DI existente em junho de 2013 a junho de 2015.	56
4.5	Tabela da variação percentual das estimativas dos parâmetros com a retirada das observações.	57
4.6	Valores das variáveis para a observação 167	60
4.7	Valores das variáveis para a observação # e *	60
B.1	Resultados da simulação para o modelo gama - cenário I.	70
B.2	Resultados da simulação para o modelo gama - cenário II.	71
B.3	Resultados da simulação para o modelo gama - cenário III.	72
B.4	Resultados da simulação para o modelo gama - cenário IV.	73
B.5	Resultados da simulação para o modelo gama - cenário V.	74
B.6	Resultados da simulação para o modelo gama - cenário VI.	75
B.7	Resultados da simulação para o modelo gama - cenário VII.	76
C.1	Resultados da simulação para o modelo gaussiana inversa - cenário I.	78
C.2	Resultados da simulação para o modelo gaussiana inversa - cenário II.	79

C.3	Resultados da simulação para o modelo gaussiana inversa - cenário III.	80
C.4	Resultados da simulação para o modelo gaussiana inversa - cenário IV.	81
C.5	Resultados da simulação para o modelo gaussiana inversa - cenário V.	82
C.6	Resultados da simulação para o modelo inversa gaussiana - cenário VI.	83
C.7	Resultados da simulação para o modelo gaussiana inversa - cenário VII.	84
D.1	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 1.	86
D.2	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 2a.	87
D.3	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - Cenário 2b.	88
D.4	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - Cenário 3.	89
D.5	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 4a.	90
D.6	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 4b.	91
D.7	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 5.	92
D.8	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 6a.	93
D.9	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 6b.	94
D.10	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 7a.	95
D.11	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 7b.	96
D.12	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 8a.	97
D.13	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 8b.	98
E.1	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 1.	100
E.2	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 2a.	101
E.3	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 2b.	102
E.4	Estatística Descritiva da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 3.	103

E.5	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 4a.	104
E.6	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 4b.	105
E.7	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 5.	106
E.8	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 6a.	107
E.9	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 6b.	108
E.10	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 7a.	109
E.11	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 7b.	110
E.12	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 8a.	111
E.13	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 8b.	112
F.1	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - cenário I : Gerando ZAIG e ajustando ZAGA.	114
F.2	Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - cenário II : Gerando ZAGA e ajustando ZAIG.	115

Capítulo 1

Introdução

Modelos usuais para dados assimétricos não admitem a ocorrência do valor zero com probabilidade positiva. Porém, em muitos casos ele pode estar presente na variável resposta. Por exemplo, no estudo de precipitações pluviométricas (Feuerverger, 1979) temos a ocorrência do valor zero na variável resposta quando não temos a ocorrência de chuva. Para esses casos, temos os modelos de regressão contínuos inflacionados no zero. Os modelos de regressão contínuos inflacionados no zero são usados quando a variável resposta pode assumir o valor zero com probabilidade positiva. A variável resposta desses modelos é discreta em zero e contínua em algum intervalo de números reais positivos. O foco desse trabalho é nesses modelos. Os modelos contínuos inflacionados no zero têm sido aplicados em diversas áreas como seguros (Resti et al., 2013), na área financeira na análise de cartões de créditos (Troiani, 2009), na inadimplência em empréstimos hipotecários (Tong et al., 2013), na saúde pública no estudo da mortalidade no trânsito (Venezuela et al., 2014) e na medicina nos estudos de doenças (Dias, 2014).

Os resíduos desempenham um papel importante na verificação da adequação do modelo e na identificação de "outliers" e observações influentes. Uma maneira possível para realizar a análise de diagnóstico em modelos de regressão contínuos inflacionados no zero é a realização de análise de resíduos separadamente para a parte discreta e para o componente contínuo do modelo. Neste caso, um resíduo é utilizado para cada componente do modelo. No entanto, este processo tem um inconveniente principal. Para cada observação positiva temos dois resíduos, sendo difícil assim identificar valores atípicos com base em um vetor de resíduos.

Uma alternativa possível para isso é utilizar um único resíduo para analisar conjuntamente os dois componentes do modelo. O resíduo quantílico proposto por Dunn e Smyth (1996) pode ser utilizado para este fim. No entanto, se a variável resposta assume um valor positivo próximo de zero, este resíduo não vai assumir um valor elevado, em módulo, a menos que a probabilidade de uma observação assumir o valor zero seja muito baixa (Pereira, Botter e Sandoval (2014)). Isso não é razoável em muitas situações práticas. Por exemplo, suponha que uma instituição financeira tenha interesse no total de investimento de cada cliente. Se um cliente tem uma alta renda e possui um baixo valor de investimento

é razoável imaginar que essa observação seja discrepante. Porém, se a probabilidade desse cliente não possuir investimento na instituição não for baixa, ou seja, se não for inferior a 2,5%, o resíduo quantílico não consideraria essa observação como discrepante. E é razoável imaginar que essa probabilidade pode não ser tão baixa, dado que clientes desse perfil podem investir em outras instituições financeiras.

Uma classe de resíduos para os modelos de regressão contínuos inflacionados de zero foi proposta por [Pereira, Botter e Sandoval \(2014\)](#). Esta classe de resíduos é função de um resíduo para o componente contínuo do modelo e da estimativa de máxima verossimilhança da probabilidade da observação assumir o valor zero. Resíduos desta classe não apresentam os problemas do resíduo quantílico e os autores concluíram que um resíduo desta classe tem boas propriedades nos modelos de regressão beta inflacionados no zero. Também definimos uma outra classe de resíduos baseada no resíduo proposto por [Ospina e Ferrari \(2012\)](#), que também é função de um resíduo para o componente contínuo do modelo e da estimativa de máxima verossimilhança da probabilidade da observação assumir o valor zero.

O principal objetivo deste trabalho é estudar e descrever técnicas de diagnóstico para os modelos de regressão gama inflacionados no zero ([Feuerverger, 1979](#)) e gaussiana inversa inflacionados no zero ([Heller et al., 2006](#)). Em particular, o principal interesse é estudar as propriedades dos resíduos da classe proposta por [Pereira, Botter e Sandoval \(2014\)](#) nesses modelos. Estudos de simulação de Monte Carlo são usados para este propósito. Para alcançar tal objetivo, esse trabalho está estruturado da seguinte forma. No Capítulo 2, iremos descrever os modelos inflacionados no zero e suas propriedades. No Capítulo 3, inicialmente descrevemos os resíduos utilizados nos modelos lineares generalizados com resposta gama e gaussiana inversa (Seção 3.1) e através de estudos de simulação de Monte Carlo, avaliamos as propriedades desses resíduos (Seção 3.2). Estudar as propriedades desses resíduos nos modelos contínuo com resposta gama e gaussiana inversa é importante, pois esses mesmos resíduos são utilizados na análise de diagnóstico nos modelos inflacionados. Na seção seguinte (Seção 3.3.1), descrevemos os resíduos utilizados nos modelos inflacionados no zero e utilizamos estudos de simulação de Monte Carlo, para analisar as propriedades desses resíduos (Seção 3.4). Na Seção 3.5, apresentamos outros aspectos da análise de diagnóstico em modelos inflacionados no zero, tais como: identificação de observações discrepantes no componente discreto (Seção 3.5.1), observações influentes (Seção 3.5.2), gráficos de diagnóstico (Seção 3.5.3) e teste de qualidade de ajuste para o componente discreto (Seção 3.5.4). No Capítulo 4, apresentamos o ajuste do modelo ZAIG aplicado a dados de fundo de investimentos do Brasil, com o propósito de exemplificar os tópicos discutidos anteriormente, mostrar a importância dos modelos apresentados e as vantagens da classe de resíduos proposta por [Pereira, Botter e Sandoval \(2014\)](#) na análise de dados reais. Por fim, as conclusões e os trabalhos futuros apresentamos no Capítulo 5.

Capítulo 2

Modelos de regressão ZAGA e ZAIG

Em muitas situações práticas nos deparamos com variáveis que podem assumir certos valores específicos com uma determinada probabilidade ou podem ter valores que pertencem a um determinado intervalo com uma determinada probabilidade, ou seja, são uma combinação de variáveis discretas e variáveis contínuas. Essas variáveis são conhecidas como variáveis mistas. Por exemplo, na área de seguros, no estudo do custo do sinistro de automóveis, é normal surgir variáveis desse tipo. Elas se comportam como variáveis discretas nos casos da não ocorrência de sinistro (custo zero) e se comportam como variáveis contínuas quando acontece o sinistro. Neste último caso, o custo da ocorrência varia no intervalo $(0; \infty)$, (Resti et al., 2013). Na área médica, é comum aparecer variáveis mistas em situações em que o interesse é dosar a concentração de alguma substância que pode ou não estar presente no sangue (Abdo et al., 2013) ou o volume de sangue recebido em uma cirurgia específica (Dias, 2014). Em microbiologia, é comum o interesse no estudo das variáveis que afetam a quantidade de fungos presentes em plantações de milhos e outros alimentos, sendo que em cada unidade amostral essa quantidade pode ser zero com probabilidade positiva ou qualquer outro valor contínuo (Rocha et al., 2009). Em pluviometria, o interesse recai na quantidade de chuva que cai em certa região, assumindo valor zero nos dias que não tem chuva ou qualquer outro valor contínuo nos dias em que temos chuva na região (Feuerverger, 1979).

Para o estudo de variáveis mistas que assumem o valor zero com probabilidade positiva, é necessário utilizar métodos que levam em conta suas peculiaridades, para assim obter um ajuste adequado. Uma metodologia para a modelagem de tal variável aleatória é utilizar os modelos de regressão Tweedie. Os modelos de regressão com variável resposta Tweedie pertencem a classe dos modelos de dispersão exponencial (Jørgensen, 1997). A classe Tweedie de modelos de dispersão exponencial é definida como uma distribuição da família exponencial com média μ e variância (μ^p/σ^2) , em que $p \in (-\infty, 0) \cup (1, \infty)$ (Dunn e Smyth, 2005). Jørgensen e Paes De Souza (1994) provam que para cada p existe uma distribuição de probabilidade. Devido a possibilidade de flexibilidade na distribuição da variável resposta, o modelo Tweedie é uma boa opção para modelar os valores iguais a zero. Essa metodologia não será abordada neste trabalho. Para maiores informações sobre os modelos Tweedie ver

Bortoluzzo et al. (2011), Pousinho (2013) e Ferreira (2005). Outra metodologia eficaz é a mistura entre duas distribuições para a variável resposta, uma distribuição degenerada no zero e uma distribuição contínua. Essas misturas de distribuição dão origem aos modelos de regressão inflacionados nos zero. Neste trabalho, temos interesse nos casos em que a variável resposta apresenta uma distribuição contínua no conjunto dos números reais positivos e assume o valor zero com probabilidade positiva.

O primeiro modelo de regressão inflacionado no zero para variáveis mistas foi proposto por Feuerverger (1979), que utilizou a distribuição gama para o estudo de precipitações pluviométricas. Após o trabalho de Feuerverger, vários trabalhos nessa área foram aparecendo, entre eles, Heller et al. (2006) que propôs um modelo que considerou a distribuição gaussiana inversa inflacionada no zero no ajuste do custo do sinistro de automóveis. Modelos com variável resposta do tipo proporção foram propostos por Cook et al. (2008), Hoff (2007), Ospina e Ferrari (2012) e Pereira et al. (2012) que propuseram os modelos de regressão beta inflacionados no zero, beta inflacionados no um, beta inflacionados no um ou no zero e beta inflacionados no zero, no um e em um ponto $c \in (0, 1)$.

O principal interesse deste capítulo é apresentar os modelos de regressão com resposta gama inflacionados no zero e gaussiana inversa inflacionados no zero para, no próximo capítulo (Capítulo 3), tratarmos sobre a análise de diagnóstico nesses modelos.

2.1 Distribuições contínuas inflacionadas no zero

As distribuições contínuas inflacionadas no zero são aquelas distribuições em que a variável resposta apresenta uma distribuição contínua no conjunto dos números reais positivos e pode assumir o valor zero com probabilidade positiva. Podemos definir essa classe de distribuições inflacionadas no zero da seguinte forma:

Definição 1. *Seja y uma variável aleatória com suporte em $\{0\} \cup (0, \infty)$. Dizemos que y pertence a classe de distribuições contínuas positivas triparmétrica inflacionadas no zero com vetor de parâmetro (α, μ, σ) , se sua função densidade de probabilidade é dada por*

$$g(y; \alpha, \mu, \sigma) = \begin{cases} 0, & \text{se } y < 0; \\ \alpha, & \text{se } y = 0 \\ (1 - \alpha)f(y; \mu, \sigma), & \text{se } y > 0, \end{cases} \quad (2.1)$$

em que $0 < \alpha < 1$ e $f(y; \mu, \sigma)$ é a função densidade de probabilidade de uma distribuição contínua com parâmetros μ e σ e suporte em $(0, \infty)$. Observe que $\alpha = P(y = 0)$.

A função densidade de y pode ser reescrita como:

$$g(y; \alpha, \mu, \sigma) = \{\alpha^{I_{(0)}(y)}(1 - \alpha)^{(1 - I_{(0)}(y))}\} \{f(y; \mu, \sigma)^{(1 - I_{(0)}(y))}\}, \quad (2.2)$$

sendo $I_{(0)}(y)$ a variável indicadora que recebe valor 1 se $y = 0$. Note que a função definida

em (2.2) pode ser fatorada em dois termos: o primeiro depende apenas de α e o segundo que depende de $(\mu; \sigma)$ e envolve apenas a parte contínua da variável.

Neste trabalho, iremos utilizar como $f(y; \mu, \sigma)$, a função densidade de probabilidade da distribuição gama e a função densidade de probabilidade da distribuição gaussiana inversa, definidas a seguir.

2.1.1 Distribuição ZAGA

A distribuição gama inflacionada no zero (*zero adjusted gamma*, ZAGA¹) é uma distribuição com três parâmetros, sendo uma extensão da distribuição gama, que permite a ocorrência do valor zero com probabilidade positiva.

A distribuição gama é uma distribuição de probabilidade contínua que tem dois parâmetros e em sua parametrização mais usual possui função densidade de probabilidade dada por

$$f(y; a, b) = \frac{a}{\Gamma(b)} (ay)^{(b-1)} e^{-ay} \quad (2.3)$$

em que $\Gamma(\cdot)$ é a função gama definida como $\Gamma(k) = \int_0^\infty t^{k-1} e^{-t} dt$, $a > 0$ e $b > 0$ são os parâmetros de escala e parâmetro de forma, respectivamente. Entretanto essa parametrização não é muito adequada para o uso em modelos de regressão, pois os parâmetros não são facilmente interpretáveis. Então definindo $\mu = b/a$ e $\sigma = \sqrt{1/b}$, obtém-se uma parametrização alternativa que é usada no pacote `gamlss` do software R (Stasinopoulos et al., 2014) dada por

$$f(y; \mu, \sigma) = \frac{\exp\left(-\frac{y}{\sigma^2\mu}\right) y^{\left(\frac{1}{\sigma^2}-1\right)}}{\Gamma\left(\frac{1}{\sigma^2}\right) (\mu\sigma^2)^{\frac{1}{\sigma^2}}}, \quad (2.4)$$

em que $y > 0$, $\mu > 0$ e $\sigma > 0$. Aqui $E(y) = \mu$ e $Var(y) = \sigma^2\mu^2$ são a média e a variância respectivamente da distribuição gama.

A distribuição gama inflacionada no zero é uma mistura da distribuição degenerada no zero e da distribuição gama. Utilizando a parametrização da equação (2.4), podemos defini-la da forma a seguir.

Definição 2. *Seja y uma variável aleatória com suporte em $\{0\} \cup (0; \infty)$. Dizemos que y tem distribuição gama inflacionada no zero com vetor de parâmetros (α, μ, σ) se sua função densidade de probabilidade é dada pela equação (2.1) e no lugar da função densidade de probabilidade $f(y; \mu, \sigma)$ consideramos a função de densidade de probabilidade da distribuição gama definida na equação (2.4).*

¹ O termo em inglês *zero adjusted gamma* - ZAGA significa distribuição gama ajustado no zero, que é o nome mais apropriado nesse caso, pelo fato da distribuição ZAGA ser uma extensão da distribuição gama que permite a inclusão do valor zero no seu suporte. Porém em português é mais comum o uso do termo gama inflacionada no zero, podendo levar a uma conclusão errônea de que possuímos uma inflação de zero.

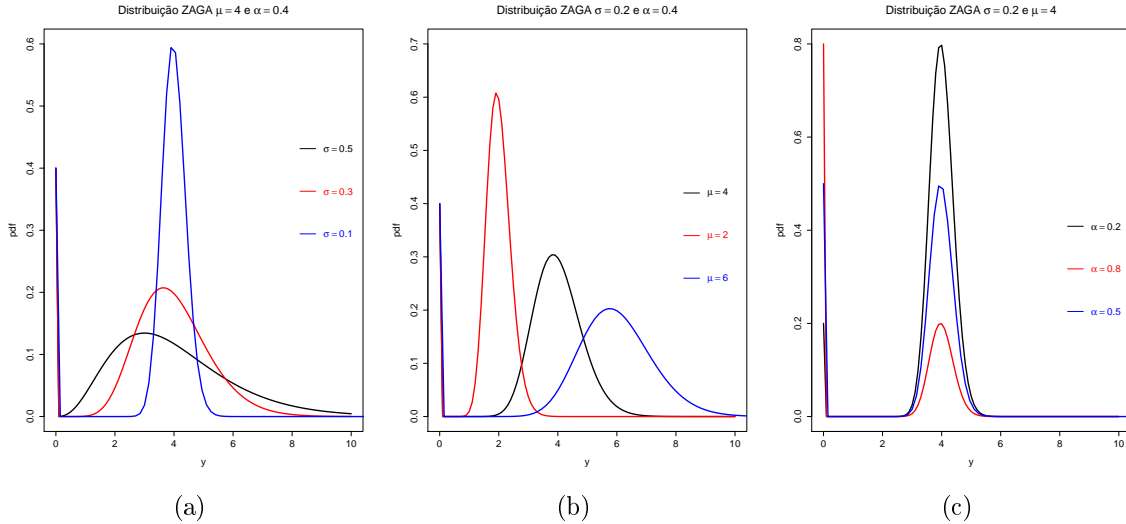


Figura 2.1: Gráfico da distribuição ZAGA para diferentes valores dos parâmetros

A esperança e a variância da distribuição ZAGA (Stasinopoulos et al., 2014) são dadas por

$$E(y) = (1 - \alpha)\mu \quad (2.5)$$

$$Var(y) = (1 - \alpha)\mu^2(\alpha + \sigma^2). \quad (2.6)$$

A distribuição ZAGA é uma distribuição assimétrica, pelo fato da distribuição gama ser assimétrica. Como dito, a distribuição gama inflacionada no zero possui três parâmetros, μ a média da distribuição gama, σ o parâmetro de escala ou parâmetro de dispersão e α é a probabilidade de ocorrer o valor zero. A Figura (2.1) apresenta gráficos da função densidade de probabilidade da distribuição ZAGA para diferentes valores dos parâmetros. Na Figura (2.1(a)) temos a função densidade da distribuição ZAGA para os parâmetros $\mu = 4$ e $\alpha = 0,4$ e variamos os valores de σ , considerando 0,1, 0,3 e 0,5. A variação do parâmetro σ , provoca uma alteração na escala do gráfico. Ao diminuirmos σ para 0,1 temos uma concentração dos valores próximo da média (curva azul), ao aumentarmos σ para 0,5 as observações estão mais espalhadas ao longo do suporte da distribuição (curva preta). Além disso, notamos que o aumento de valores para σ , aumenta a assimetria da distribuição, isso ocorre devido ao fato da assimetria da distribuição gama está associada ao parâmetro σ . Já na Figura (2.1(b)) consideramos $\sigma = 0,2$ e $\alpha = 0,4$ e variamos os valores de μ . A mudança na média altera os pontos de concentração dos dados. Observamos que quando a média é 4 os dados se concentram no ponto de $y = 4$ (curva preta), se $\mu = 2$ os dados se concentram em torno do ponto $y = 2$ (curva vermelha). A variação do parâmetro α , modifica a concentração de zeros na distribuição. Na Figura (2.1(c)) temos $\alpha = 0,5$ e podemos observar que a probabilidade de ocorrência do valor zero, eixo y da figura, é igual a $\alpha = P(y = 0) = 0,5$ (curva azul). Na Figura (2.1(c)) aumentamos α para 0,8 e a mudança faz aumentar a concentração de zeros na distribuição ZAGA (curva vermelha). Observamos também a flexibilidade da distribuição

ZAGA, dependendo do valor dos parâmetros temos densidades bem distintas.

2.1.2 Distribuição ZAIG

A distribuição gaussiana inversa inflacionada no zero (*zero adjusted inverse gaussian*, ZAIG²) é uma extensão da distribuição gaussiana inversa para as situações em que a variável de interesse assume o valor zero com probabilidade positiva. Existem diferentes parametrizações para a distribuição gaussiana inversa. Usaremos aqui a parametrização utilizada no pacote `gamlss` do software R (Stasinopoulos et al., 2014) cuja função densidade de probabilidade é dada por

$$f(y; \mu, \sigma) = \frac{1}{\sqrt{(2\pi y^3 \sigma^2)}} \exp \left\{ -\frac{(y - \mu)^2}{2\mu^2 \sigma^2 y} \right\} \quad (2.7)$$

em que $y > 0$, $\mu > 0$ e $\sigma > 0$. Aqui $E(y) = \mu$ e $Var(y) = \sigma^2 \mu^3$ são a média e a variância da função densidade da gaussiana inversa respectivamente.

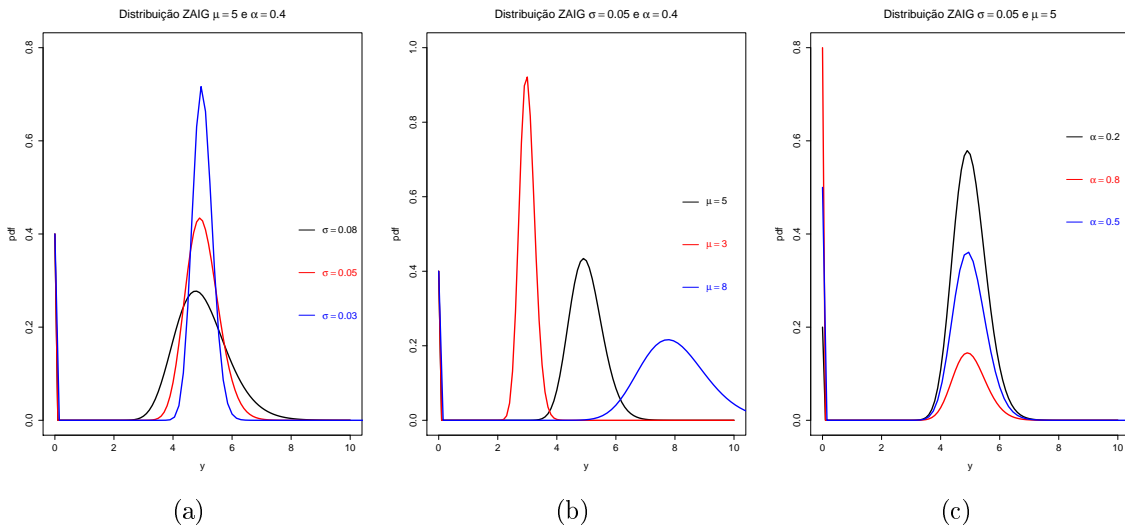


Figura 2.2: Gráfico da distribuição ZAIG para diferentes valores dos parâmetros

A distribuição gaussiana inversa inflacionada no zero é uma mistura da distribuição degenerada no zero e da distribuição gaussiana inversa. Utilizando a parametrização da equação (2.7), podemos defini-la da forma a seguir

Definição 3. *Seja y uma variável aleatória com suporte em $\{0\} \cup (0; \infty)$. Dizemos que y tem distribuição gaussiana inversa inflacionada no zero com vetor de parâmetros (α, μ, σ) se sua função densidade de probabilidade é dada pela equação (2.1) e no lugar da função*

²O termo em inglês *zero adjusted inverse gaussian*, ZAIG, significa gaussiana inversa ajustada no zero, ou seja, é a distribuição gaussiana inversa que tem seu suporte estendido para permitir a inclusão do valor zero. Porém o nome mais utilizado em português é gaussiana inversa inflacionada no zero, podendo levar a interpretação errônea de que a variável possui inflação de zero, mas no caso da gaussiana inversa inflacionada no zero temos apenas um ajuste no suporte para incluir o valor zero.

densidade $f(y; \mu; \sigma)$ utilizamos a função densidade da distribuição gaussiana inversa definida na equação (2.7)

A esperança e a variância da distribuição ZAIG (Stasinopoulos et al., 2014) são dadas por:

$$E(y) = (1 - \alpha)\mu \quad (2.8)$$

$$Var(y) = (1 - \alpha)\mu^2(\alpha + \mu\sigma^2). \quad (2.9)$$

A distribuição ZAIG é uma distribuição assimétrica como consequência direta da gaussiana inversa ser assimétrica. Os três parâmetros da distribuição representam a média da distribuição gaussiana inversa μ , σ o parâmetro de escala ou parâmetro de dispersão e α é a probabilidade de ocorrer o valor zero. Heller et al. (2006) mostrou que para dados com assimetria acentuada à direita, a distribuição gaussiana inversa demonstra superioridade na qualidade do ajuste com relação à distribuição gama. O fato ocorre devido a variância depender da forma cúbica da média enquanto na distribuição gama a dependência é quadrática. Na Figura(2.2) temos a função densidade de probabilidade da distribuição gaussiana inversa inflacionada no zero para diversos valores dos parâmetros μ, σ e α , observamos que dependendo dos valores dos parâmetros a função densidade de probabilidade da distribuição gaussiana inversa assume formas distintas.

2.2 Modelos de regressão inflacionados no zero

Muitas vezes nosso interesse recai na análise da relação entre as variáveis, ou seja, temos interesse na análise do efeito que uma ou mais variáveis explicativas (covariáveis) causam na variável de interesse (variável resposta). No nosso caso temos interesse em estudar a variável resposta em função de variáveis preditoras, em que a variável resposta têm distribuição gama ou gaussiana inversa inflacionada no zero. E para isso é necessário o uso de modelos de regressão. Os modelos de regressão ZAGA e ZAIG foram utilizados em diversas aplicações como na modelagem do custo da apólice de seguro (Resti et al., 2013) e na modelagem do volume de sangue utilizado em uma cirurgia (Dias, 2014). O modelo ZAIG foi utilizado por Heller et al. (2006), De Jong et al. (2008) e Bortoluzzo et al. (2011) em aplicações na modelagem de sinistros de seguradoras. Ferreira (2008) e Troiani (2009) também utilizaram o modelo ZAIG em uma aplicação na modelagem de perdas de uma carteira de crédito e Venezuela et al. (2014) utilizou no estudo da mortalidade no trânsito em um trabalho com dados longitudinais. O modelo ZAGA foi utilizado na modelagem da perda em caso de inadimplência de hipotecas residenciais (Tong et al., 2013) e nas precipitações pluviométricas (Feuerverger, 1979). A seguir, definimos de forma geral os modelos de regressão inflacionados no zero que iremos considerar neste trabalho.

Definição 4. Considere y_1, y_2, \dots, y_n variáveis aleatórias independentes, onde y_i tem dis-

tribuição inflacionada no zero como definido na equação (2.1). Os modelos de regressão contínuos positivos inflacionados no zero são definidos pela equação (2.1) e pelos seguintes componentes sistemáticos:

$$\begin{aligned} g_1(\alpha_i) &= \zeta_i \\ g_2(\mu_i) &= \eta_{i1} \\ g_3(\sigma_i) &= \eta_{i2} \end{aligned} \quad (2.10)$$

em que $\zeta_i = \mathbf{z}_i^\top \boldsymbol{\gamma}$, $\eta_{i1} = \mathbf{x}_i^\top \boldsymbol{\beta}$ e $\eta_{i2} = \mathbf{v}_i^\top \boldsymbol{\delta}$ são os preditores lineares, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \dots, \gamma_{p_\alpha})^\top$, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{p_\mu})^\top$ e $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_{p_\sigma})^\top$ são os vetores de parâmetros desconhecidos; $\mathbf{z}_i = (z_{i1}, z_{i2}, \dots, z_{ip_\alpha})^\top$, $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip_\mu})^\top$ e $\mathbf{v}_i = (v_{i1}, v_{i2}, \dots, v_{ip_\sigma})^\top$ são constantes que representam os valores das variáveis preditoras para cada um dos parâmetros da distribuição inflacionada no zero, $g_1(\cdot)$, $g_2(\cdot)$ e $g_3(\cdot)$ são funções de ligação estritamente monótonas e duas vezes diferenciáveis.

Neste trabalho, consideramos dois modelos inflacionados, os modelos gama inflacionados no zero e os modelos gaussianas inversa inflacionados no zero.

2.2.1 Modelos de regressão gama inflacionados no zero

Os modelos de regressão gama inflacionados no zero (Feuerverger, 1979) é definido da forma a seguir.

Definição 5. Considere y_1, y_2, \dots, y_n variáveis aleatórias independentes, onde y_i tem distribuição inflacionada no zero como definido na equação (2.1), tal que $f(y; \mu, \sigma)$ é a função densidade de probabilidade da distribuição gama (equação(2.4)). Os modelos de regressão gama inflacionados no zero são definidos pela equação (2.1) e pelos componentes sistemáticos definidos em (2.10).

Baseado numa amostra aleatória com n observações, o logaritmo da função de verossimilhança do modelo ZAGA é obtido a partir da equação (2.2) e é dado por:

$$l(\boldsymbol{\theta}) = l_1(\boldsymbol{\gamma}) + l_2(\boldsymbol{\beta}, \boldsymbol{\delta}) \quad (2.11)$$

em que

$$\begin{aligned} \boldsymbol{\theta} &= (\boldsymbol{\gamma}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top \\ l_1(\boldsymbol{\gamma}) &= \sum_{i=1}^n l_{i1}(\alpha_i), \\ l_2(\boldsymbol{\beta}, \boldsymbol{\delta}) &= \sum_{i=1; y_i \neq 0}^n l_{i2}(\mu_i, \sigma_i), \\ l_{i1}(\alpha_i) &= I_{(0)}(y_i) \ln(\alpha_i) + (1 - I_{(0)}(y_i)) \ln(1 - \alpha_i), \\ l_{i2}(\mu_i, \sigma_i) &= \ln \left(\Gamma \left(\frac{1}{\sigma_i^2} \right) \right) + \left(\frac{1}{\sigma_i^2} - 1 \right) \ln(y_i) - \frac{1}{\sigma_i^2} \ln(\sigma_i^2) - \frac{1}{\sigma_i^2} \ln(\mu_i) - \frac{y_i}{\sigma_i^2 \mu_i}. \end{aligned} \quad (2.12)$$

sendo $I_{(0)}(y_i)$ a função indicadora com valor 1 se $y = 0$ e 0 se $y \neq 0$. E os parâmetros μ_i , σ_i e α_i são definidos como funções de $\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ e $\boldsymbol{\delta}$ como definidas na equação (2.10), isto é,

$\alpha_i = g_1^{-1}(\zeta_i)$, $\mu_i = g_2^{-1}(\eta_{i1})$ e $\sigma_i = g_3^{-1}(\eta_{i2})$. Nota-se que $l_1(\boldsymbol{\gamma})$ é função apenas de $\boldsymbol{\gamma}$ e $l_2(\boldsymbol{\beta}, \boldsymbol{\delta})$ é função de $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$. Portanto $\boldsymbol{\gamma}$ e $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ são separáveis e podemos estimar $\boldsymbol{\gamma}$ separado da estimação de $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ e vice-versa (Pace e Salvan, 1997). Assim, a estimação por máxima verossimilhança sobre $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ é feita independente do vetor de parâmetros $\boldsymbol{\gamma}$. Observa-se ainda que o componente discreto $l_1(\boldsymbol{\gamma})$ envolve apenas o parâmetro utilizado para modelar a probabilidade de ocorrência do valor zero e $l_1(\boldsymbol{\gamma})$ é a função de log-verossimilhança de um modelo linear generalizado com resposta binária (Nelder e Wedderburn, 1972). E $l_2(\boldsymbol{\beta}, \boldsymbol{\delta})$ é a função de verossimilhança do modelo com resposta gama (Paula, 2013).

A função escore para o modelo ZAGA é obtida derivando o logaritmo da função de verossimilhança (equação (2.11)) em relação a cada um dos componentes de $\boldsymbol{\theta}$. Pela separabilidade dos parâmetros $\boldsymbol{\gamma}$ e $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$, podemos obter de forma independente a função escore para $\boldsymbol{\gamma}$ e a função escore para $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$. Assim a função escore para o modelo ZAGA é definida como:

$$U(\boldsymbol{\theta}) = (U_\gamma(\boldsymbol{\gamma})^\top, U_\beta(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top, U_\delta(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top)^\top \quad (2.13)$$

em que

$$\begin{aligned} U_\gamma &= Z^\top PG(\mathbf{y}^c - \boldsymbol{\alpha}), \\ U_\beta &= X^\top TD(\mathbf{y} - \boldsymbol{\mu}), \\ U_\delta &= V^\top Q(\mathbf{m} - \boldsymbol{\sigma}^*), \end{aligned} \quad (2.14)$$

$$\begin{aligned} Z &= (\mathbf{z}_1, \dots, \mathbf{z}_n)^\top, P = \text{diag}\left(\frac{1}{\alpha_1(1-\alpha_1)}, \dots, \frac{1}{\alpha_n(1-\alpha_n)}\right), G = \text{diag}\left(\frac{d\alpha_1}{d\zeta_1}, \dots, \frac{d\alpha_n}{d\zeta_n}\right), \mathbf{y}^c = \\ &= (I_{(0)}(y_1), \dots, I_{(0)}(y_n)), \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^\top, \mathbf{y} = (y_1, \dots, y_n)^\top, \boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top, \\ D &= \text{diag}\left(\frac{1}{\mu_1^2 \sigma_1^2}, \dots, \frac{1}{\mu_n^2 \sigma_n^2}\right), T = \text{diag}\left(\frac{d\mu_1}{d\eta_{11}}, \dots, \frac{d\mu_n}{d\eta_{n1}}\right), X = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top, V = (\mathbf{v}_1, \dots, \mathbf{v}_n)^\top, \\ Q &= \text{diag}\left(\frac{d\sigma_1}{d\eta_{12}}, \dots, \frac{d\sigma_n}{d\eta_{n2}}\right), \boldsymbol{\sigma}^* = (-\log(\sigma_1^2), \dots, -\log(\sigma_n^2)) \text{ e } \mathbf{m} = \left(\psi\left(\frac{1}{\sigma_1^2}\right) \log\left(\frac{\mu_1}{y_1 \sigma_1^2}\right) + \frac{(y_1 - \mu_1)}{\mu_1}, \right. \\ &\dots, \left. \psi\left(\frac{1}{\sigma_n^2}\right) \log\left(\frac{\mu_n}{y_n \sigma_n^2}\right) + \frac{(y_n - \mu_n)}{\mu_n}\right)^\top \text{ e tal que } \psi\left(\frac{1}{\sigma_i^2}\right) = \frac{\Gamma'\left(\frac{1}{\sigma_i^2}\right)}{\Gamma\left(\frac{1}{\sigma_i^2}\right)}. \end{aligned}$$

Devido a separabilidade dos parâmetros $\boldsymbol{\gamma}$ e $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ temos que os componentes de $U(\boldsymbol{\theta})$ correspondente ao parâmetro $\boldsymbol{\gamma}$ é função apenas do parâmetro $\boldsymbol{\gamma}$, enquanto os componentes correspondentes ao parâmetros $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ são função de $\boldsymbol{\beta}$ e $\boldsymbol{\delta}$. Apresentamos os cálculos algébricos no Apêndice A.1.1.

Obtém-se cada elemento da matriz de informação de Fisher calculando o valor esperado da segunda derivada do logaritmo da função de verossimilhança em relação a cada um dos parâmetros e multiplicamos esse valor por -1 , ou seja, definimos como:

$$k(\boldsymbol{\theta}) = -E\left(\frac{\partial^2}{\partial \boldsymbol{\theta}^2} l(\boldsymbol{\theta})\right). \quad (2.15)$$

A matriz de informação de Fisher para o modelo ZAGA é obtida a partir da função escore (2.13) (Apêndice A.1.1) e é dada por:

$$k(\boldsymbol{\theta}) = \begin{pmatrix} K_{\gamma\gamma} & 0 \\ 0 & K_{\nu} \end{pmatrix} \quad (2.16)$$

sendo $K_{\gamma\gamma}$ é a matriz de informação de Fisher para γ e

$$K_{\nu} = \begin{pmatrix} K_{\beta\beta} & 0 \\ 0 & K_{\delta\delta} \end{pmatrix} \quad (2.17)$$

Devido a ortogonalidade de $\boldsymbol{\mu}$ e $\boldsymbol{\sigma}$, temos diretamente a ortogonalidade entre $\boldsymbol{\beta}$ e $\boldsymbol{\delta}$. Assim a matriz de informação de Fisher para K_{ν} é uma matriz diagonal.

Após alguns cálculos que estão apresentados no Apêndice A.1.1, obtemos:

$$K(\boldsymbol{\theta}) = \begin{pmatrix} k_{\gamma\gamma} & 0 & 0 \\ 0 & k_{\beta\beta} & 0 \\ 0 & 0 & k_{\delta\delta} \end{pmatrix} \quad (2.18)$$

em que

$$\begin{aligned} k_{\gamma\gamma} &= Z^{\top}(PG^2)Z, \\ k_{\beta\beta} &= X^{\top}(\Delta DT^2)X, \\ k_{\delta\delta} &= V^{\top}(\Delta AQ^2)V, \end{aligned} \quad (2.19)$$

em que $\Delta = \text{diag}((1 - \alpha_1), (1 - \alpha_2), \dots, (1 - \alpha_n))$, $A = \text{diag}(a_1, a_2, \dots, a_n)$, $a_i = \frac{4}{\sigma_i^6}$ $\left(\psi'(\frac{1}{\sigma_i^2}) - \sigma_i^2\right)$ e demais vetores e matrizes conforme definido em (2.14).

Nota-se pelas equações (2.2.1) e (2.19) que o parâmetro γ são ortogonais a $(\boldsymbol{\beta}^{\top}, \boldsymbol{\delta}^{\top})^{\top}$ e portanto seus estimadores de máxima verossimilhança são assintoticamente independentes.

2.2.2 Modelos de regressão gaussiana inversa inflacionados no zero

Os modelos de regressão gaussiana inversa inflacionados no zero (Heller et al., 2006) é definido da forma a seguir.

Definição 6. Considere y_1, y_2, \dots, y_n variáveis aleatórias independentes, onde y_i tem distribuição inflacionada no zero como definido na equação (2.1), tal que $f(y; \mu, \sigma)$ é a função densidade de probabilidade da distribuição gaussiana inversa (equação(2.7)). Os modelos de regressão gaussiana inversa inflacionados no zero são definidos pela equação (2.1) e pelos componentes sistemáticos definidos em (2.10).

Baseado numa amostra aleatória com n observações, o logaritmo da função da verossimilhança dos modelos ZAIG é obtido a partir da equação (2.2) e é dado por:

$$l(\boldsymbol{\theta}) = l_1(\gamma) + l_2(\boldsymbol{\beta}, \boldsymbol{\delta}) \quad (2.20)$$

em que

$$\begin{aligned}
\theta &= (\boldsymbol{\gamma}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top, \\
l_1(\boldsymbol{\gamma}) &= \sum_{i=1}^n l_{i1}(\alpha_i), \\
l_2(\boldsymbol{\beta}, \boldsymbol{\delta}) &= \sum_{i=1; y_i \neq 0}^n l_{i2}(\mu_i, \sigma_i), \\
l_{i1}(\alpha_i) &= I_{(0)}(y_i) \ln(\alpha_i) + (1 - I_{\{y_i=0\}}) \ln(1 - \alpha_i), \\
l_{i2}(\mu_i, \sigma_i) &= \left\{ -\ln \left(\sqrt{2\pi y_i^3} \right) - \ln(\sigma_i) - \frac{(y_i - \mu_i)^2}{2y_i \mu_i^2 \sigma_i^2} \right\}.
\end{aligned} \tag{2.21}$$

Como no logaritmo da função da verossimilhança do modelo ZAGA, no modelo ZAIG temos que $\boldsymbol{\gamma}$ é separável $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ e assim, podemos obter de forma independente a função escore para $\boldsymbol{\gamma}$ e a função escore para $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$. Definimos a função escore para o modelo ZAIG como:

$$U(\boldsymbol{\theta}) = (U_\gamma(\boldsymbol{\gamma})^\top, U_\beta(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top, U_\delta(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top)^\top \tag{2.22}$$

em que

$$\begin{aligned}
U_\gamma &= Z^\top P G(\mathbf{y}^c - \boldsymbol{\alpha}), \\
U_\beta &= X^\top T C(\mathbf{y} - \boldsymbol{\mu}), \\
U_\delta &= V^\top Q M(\mathbf{s} - \boldsymbol{\sigma}^*),
\end{aligned} \tag{2.23}$$

em que $C = \text{diag} \left(\frac{1}{\mu_1^3 \sigma_1^2}, \dots, \frac{1}{\mu_n^3 \sigma_n^2} \right)$, $M = \text{diag} \left(\frac{1}{\sigma_1^3}, \dots, \frac{1}{\sigma_n^3} \right)$, $s = \left(\frac{(y_1 - \mu_1)^2}{y_1 \mu_1^2}, \dots, \frac{(y_n - \mu_n)^2}{y_n \mu_n^2} \right)$, $\boldsymbol{\sigma}^* = (\sigma_1^2, \dots, \sigma_n^2)^\top$ e demais vetores e matrizes conforme definido anteriormente. Os cálculos algébricos são apresentadas no Apêndice A.1.2.

Pela separabilidade dos parâmetros $\boldsymbol{\gamma}$ e $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$, o estimador de máxima verossimilhança de $\boldsymbol{\gamma}$ é obtido independentemente de $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top)^\top$ e vice-versa. Assim a matriz de Informação de Fisher é definida como

$$K(\boldsymbol{\theta}) = \begin{pmatrix} K_{\gamma\gamma} & 0 & 0 \\ 0 & k_{\beta\beta} & 0 \\ 0 & 0 & k_{\delta\delta} \end{pmatrix} \tag{2.24}$$

em que

$$\begin{aligned}
k_{\gamma\gamma} &= Z^\top (P G^2) Z, \\
k_{\beta\beta} &= X^\top (\Delta C T^2) X, \\
k_{\delta\delta} &= V^\top (\Delta B Q^2) V,
\end{aligned} \tag{2.25}$$

em que $B = \text{diag}(b_1, b_2, \dots, b_n)$, $b_i = \frac{-2}{\sigma_i^2}$ e demais vetores e matrizes conforme definidos anteriormente.

Na próxima seção discutiremos o processo de estimação dos modelos ZAGA e ZAIG.

2.3 Estimação nos modelos ZAGA e ZAIG

Os modelos com resposta gama inflacionados no zero e gaussiana inversa inflacionados no zero pertencem a uma família mais ampla de modelos de regressão, conhecida como modelos aditivos generalizados para locação, escala e forma (GAMLSS, *generalized additive models*

location, scale and shape), (Stasinopoulos e Rigby, 2007). Os parâmetros das distribuições, nessa família, podem ser modelados através de preditores lineares ou através de funções não paramétricas (tipicamente suavizadores lineares) (Ospina, 2008).

Da seção anterior temos que o sistema de equação $U(\boldsymbol{\theta}) = 0$ de ambos os modelos ZAGA e ZAIG não tem solução algébrica. Portanto, os estimadores de máxima verossimilhança dos parâmetros dos modelos ZAGA e ZAIG devem ser obtidos por métodos numéricos.

No pacote "gamlss" do software R, dois algoritmos podem ser utilizados para a maximização da função de verossimilhança. O primeiro algoritmo (**CG**) é uma generalização do algoritmo Cole e Green (1992) e usa as derivadas de primeira e segunda ordem (o valor esperado ou aproximado) e também as derivadas cruzadas do logaritmo da função da verossimilhança em relação aos parâmetros da distribuição. O segundo algoritmo conhecido como **RS** é mais simples do que o **CG** e não utiliza o valor esperado das derivadas cruzadas. **RS** é uma generalização do algoritmo proposto por Rigby e Stasinopoulos (1996) para ajuste da média e dispersão em modelos aditivos. Ainda, existe um terceiro algoritmo implementado no pacote denominado, *mixed()*, como o nome já diz, esse algoritmo é uma mistura dos dois outros algoritmo, *CG* e *RS*, sendo que ele começa utilizando o algoritmo *RS* e termina usando o algoritmo *CG* (para maiores informações sobre os algoritmo veja, Rigby e Stasinopoulos (2005)). O objetivo desses algoritmos é maximizar o logaritmo da função da verossimilhança dos modelos ZAGA e ZAIG definidas nas equações (2.11) e (2.20), respectivamente. Em todos os estudos de simulação de Monte Carlo, realizados nesse trabalho, foi utilizado o método RS (Stasinopoulos e Rigby, 2007) do pacote "gamlss" do software R.

Na próxima seção discutiremos os intervalos de confiança e os teste de hipótese nos modelos ZAGA e ZAIG.

2.4 Intervalos de confiança e testes de hipóteses

Assim como nos modelos lineares generalizados, nos modelos ZAGA e ZAIG testes de hipóteses, podem ser feitos utilizando-se as propriedades do estimador de máxima verossimilhança. Temos que, sob condições de regularidade,

$$\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{D} N_p(0, J(\boldsymbol{\theta})^{-1}), \quad (2.26)$$

em que $J(\boldsymbol{\theta}) = \lim_{n \rightarrow \infty} K(\boldsymbol{\theta})/n$, $K(\boldsymbol{\theta})$ igual ao definido nas equações (2.2.1) e (2.24) e $p = p_\alpha + p_\mu + p_\sigma$. Usando a equação (2.26) e as expressões para o logaritmo da função de verossimilhança, da função score e da informação de Fisher apresentadas na Seção 2.2, obtém-se expressões para intervalos de confiança e testes de hipóteses usando as estatísticas de Wald, score, razão de verossimilhanças e gradiente para os parâmetros dos modelos ZAGA e ZAIG.

O pacote *gamlss* apresenta o valor da estatística do teste de Wald e o correspondente valor-p para testar se cada um dos parâmetros do modelo é ou não igual a zero. Pode-se

realizar ainda facilmente usando o pacote, o teste da razão de verossimilhanças para testar se um conjunto de m parâmetros do modelo são todos iguais a zero. Para isso, ajusta-se o modelo completo (que contém todos os parâmetros em estudo) e o modelo reduzido que não contém os m parâmetros em teste. O pacote calcula o desvio global de cada modelo que é simplesmente menos duas vezes o logaritmo da função de verossimilhança do modelo. A diferença entre o desvio global do modelo reduzido e do modelo completo é o valor da estatística do teste da razão de verossimilhanças. Sob a hipótese nula de que todos os m parâmetros em teste são iguais a zero, a estatística do teste da razão de verossimilhanças tem distribuição assintótica qui-quadrado com m graus de liberdade. As estatísticas Wald, score e gradiente também têm distribuição assintótica qui-quadrado, mas o cálculo dessas estatística não são feitas pelo pacote `gamlss`.

2.5 Seleção de modelos

Conforme mencionado anteriormente os modelos ZAGA e ZAIG pertencem a uma família mais ampla de modelos de regressão, os modelos GAMLSS. O pacote do R que permite o ajuste desses modelos apresenta diversas funções para a seleção de variáveis explicativas. Uma dessas funções é construída com base no critério de informação de Akaike generalizado (*Generalized Akaike information criterion*- GAIC) (Rigby e Stasinopoulos, 1996). Nesta dissertação consideramos essa função GAIC para a seleção de modelos, cuja definição apresentamos a seguir:

$$GAIC(\#) = GD + \#m \quad (2.27)$$

em que GD é o desvio global dado por $-2l(\boldsymbol{\theta})$ e $\#$ uma penalidade que é fixa e é aplicada a cada um dos m parâmetros utilizados no modelo. Na literatura existem algumas sugestões sobre a penalidade $\#$, entre elas $\#= 2$ (AIC, Akaike (1974)) e $\#= \log(n)$ (BIC, Schwarz et al. (1978)). Aqui utilizaremos $\#= 2$.

Considerando o método "backward" em que subtraímos gradativamente as variáveis do modelo com todas as covariáveis, pelo critério GAIC o modelo sugerido será aquele com menor valor para o GAIC. Em relação aos testes de hipóteses mencionados na Seção 2.4, o GAIC tem a vantagem de permitir a comparação de modelos não encaixados. Na aplicação apresentada no Capítulo 4, consideramos o AIC na seleção de variáveis explicativas.

Os resultados apresentados nesse capítulo permitem o ajuste dos modelos ZAGA e ZAIG a conjuntos de dados reais e também a realização de análise inferencial usando o modelo obtido. Porém, as conclusões são válidas apenas se o ajuste do modelo for satisfatório. Assim é necessário fazer análise de diagnóstico do modelo ajustado. No Capítulo 3, discutimos diversos aspectos da análise de diagnóstico em modelos ZAGA e ZAIG e estudamos as propriedades de alguns resíduos para essa classe de modelos utilizando simulação de Monte Carlo.

Capítulo 3

Análise de diagnóstico

Na análise de diagnóstico, verificamos possíveis problemas nas suposições feitas para o modelo, como problemas na escolha da função de ligação ou na escolha das variáveis preditoras, ou até problemas com observações discrepantes ("outliers") que interferem nas estimativas dos parâmetros do modelo. Na prática podem existir diferentes falhas no ajuste do modelo, por isso a análise de diagnóstico é uma etapa muito importante.

Na Seção 3.1, iremos definir alguns resíduos utilizados nos modelos lineares generalizados com resposta gama e gaussiana inversa. Iremos ainda analisar as propriedades desses resíduos por meio de estudos de simulação de Monte Carlo (Seção 3.2). A análise do comportamento desses resíduos é importante, pois os mesmos serão utilizados nos modelos inflacionados no zero. Na Seção 3.3.1, iremos descrever e definir duas classes de resíduos utilizados nos modelos inflacionados no zero e iremos por meio de estudos de simulação de Monte Carlo analisar suas propriedades (Seção 3.4). Nas próximas seções, discutimos a análise de observações discrepantes no componente discreto (Seção 3.5.1), pontos influentes nos modelos inflacionados no zero (Seção 3.3.1), gráficos de diagnóstico nos modelos ZAGA e ZAIG (Seção 3.5.3) e teste de ajuste de qualidade para o componente discreto (Seção 3.5.4).

3.1 Análise de diagnóstico em MLGs

Nos modelos lineares generalizados (MLGs) proposto por [Nelder e Wedderburn \(1972\)](#), a média da distribuição da variável resposta é relacionada com a parte sistemática (não aleatória, ou preditor linear) por meio de uma função denominada função de ligação, ou seja,

$$g(\mu_i) = \eta_i \tag{3.1}$$

em que $g(\cdot)$ é uma função monótona e duas vezes diferenciável, $\eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ é o preditor linear, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$, $p < n$, é um vetor de parâmetros desconhecidos que temos interesse em estimar, $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})^\top$ são as covariáveis, $i = 1, 2, \dots, n$.

Em relação ao modelo linear geral, temos uma maior flexibilidade na relação entre a média da variável resposta e o preditor linear, além de ter mais opções para a distribuição

da variável resposta, permitindo que a distribuição da mesma pertença à família exponencial de distribuições. Entre as distribuições mais conhecidas que pertencem a família exponencial estão as distribuições normal, poisson, binomial, gama e gaussiana inversa.

A análise de diagnóstico nos modelos lineares generalizados foi amplamente estudada e vários resíduos foram propostos para esses modelos. Entre os trabalhos que estudaram esses resíduos estão [Atkinson \(1985\)](#), [Cordeiro e Demétrio \(2008\)](#), [McCullagh e Nelder \(1989\)](#), [Williams \(1987\)](#), [Dunn e Smyth \(1996\)](#), [Anscombe \(1953\)](#) e [Paula \(2013\)](#). Nesta seção, abordaremos apenas os resíduos para os modelos com resposta gama e gaussiana inversa. A seguir definiremos os resíduos desvio, Pearson, William, Anscombe e quantílico, que são aqueles que utilizamos neste trabalho.

Vale ressaltar que nesse capítulo utilizamos o parâmetro de escala σ ao invés do parâmetro de dispersão ϕ , mais utilizado ao denotar a classe das distribuições pertencentes a família exponencial. Temos que $\phi = 1/\sigma^2$ e escolhemos utilizar σ nessa seção, pois é a notação mais utilizada nos modelos inflacionados no zero e no pacote `gamlss` do software R.

3.1.1 Resíduo desvio

Nos modelos lineares generalizados, um dos resíduos mais utilizados é definido a partir dos componentes da função desvio. A função desvio mede a diferença entre as log-verossimilhanças do modelo saturado e do modelo em estudo, dado $\hat{\beta}$. Definimos o componente do desvio como

$$d(y_i; \hat{\mu}_i) = \text{sign}(y_i - \hat{\mu}_i) \sqrt{2(\tilde{l}_i - \hat{l}_i)}, \quad (3.2)$$

em que \tilde{l}_i e \hat{l}_i são as log-verossimilhanças do modelo saturado (com o número de parâmetros máximo) e do modelo estudado respectivamente. A função $\text{sign}(y_i - \hat{\mu}_i)$ é definida como:

$$\text{sign}(y_i - \hat{\mu}_i) = \begin{cases} -1, & (y_i - \hat{\mu}_i) < 0 \\ 1, & (y_i - \hat{\mu}_i) > 0. \end{cases}$$

Definição 7. *Utilizando o componente do desvio definido em (3.2), definimos o resíduo desvio padronizado como*

$$r_i^{\text{dev}} = \frac{d(y_i; \hat{\mu}_i)}{\hat{\sigma} \sqrt{1 - \hat{h}_{ii}}}, \quad (3.3)$$

em que $\hat{\mu}_i = g^{-1}(\hat{\eta}_i)$; $\hat{\eta}_i = \mathbf{x}_i^T \hat{\beta}$ e \hat{h}_{ii} é o i -ésimo elemento da diagonal principal da matriz $\hat{H} = \hat{W}^{1/2} X (X^T \hat{W} X)^{-1} X^T \hat{W}^{1/2}$, em que $\hat{W} = \text{diag}\{\hat{w}_1, \dots, \hat{w}_n\}$ é a matriz de pesos estimados, sendo o peso \hat{w}_i definido como $\hat{w}_i = \frac{1}{V_i} \left(\frac{d\mu_i}{d\eta_i}\right)^2$ e $V_i = \frac{d\mu_i}{d\varphi_i}$ é a função de variância, tal que $\varphi_i = 1/\mu_i$ no modelo gama e $\varphi_i = 1/(2\mu_i^2)$ no modelo gaussiana inversa.

Podemos observar que o resíduo r_i^{dev} representa a distância na escala da log-verossimilhança entre o valor observado e o valor ajustado ($y_i - \hat{\mu}_i$). Nos modelos de regressão com resposta

gama o desvio é definido como:

$$d(y_i; \hat{\mu}_i) = \left\{ -\log\left(\frac{y_i}{\hat{\mu}_i}\right) + \frac{(y_i - \hat{\mu}_i)^2}{y_i \hat{\mu}_i^2} \right\}, \quad (3.4)$$

em que $w_i = \mu_i^2 \left(\frac{d\mu_i}{d\eta_i}\right)^2$.

O desvio para os modelos de regressão com resposta gaussiana inversa é definida como

$$d(y_i; \hat{\mu}_i) = \left\{ \frac{(y_i - \hat{\mu}_i)}{y_i \hat{\mu}_i^2} \right\}, \quad (3.5)$$

em que $w_i = \mu_i^3 \left(\frac{d\mu_i}{d\eta_i}\right)^2$.

3.1.2 Resíduo de Pearson

O resíduo de Pearson é outro resíduo muito utilizado nos modelos lineares generalizados baseado na diferença entre valor observado e o valor ajustado.

Definição 8. *A versão padronizada do resíduo de Pearson é definida por:*

$$r_i^{pea} = \frac{(y_i - \hat{\mu}_i)}{\hat{\sigma} \sqrt{\hat{V}_i(1 - \hat{h}_{ii})}}, \quad (3.6)$$

Nos modelos com resposta gama V_i é definido como $V_i = \mu_i^2$ e nos modelos com resposta gaussiana inversa $V_i = \mu_i^3$,

O resíduo de Pearson é muito simples de ser calculado, porém tem a desvantagem de ter a distribuição assimétrica em modelos com resposta diferente da normal (Cordeiro e Demétrio, 2008).

3.1.3 Resíduo quantílico

Considere a função de distribuição $F(y; \mu, \sigma)$. Se $F(\cdot)$ é contínua, então $F(y_i; \mu_i, \sigma)$ é uniformemente distribuída no intervalo unitário. O resíduo quantílico é uma classe de resíduo definida para os modelos com resposta contínua (Dunn e Smyth, 1996).

Definição 9. *Definimos o resíduo quantílico como:*

$$r_i^{qu} = \Phi^{-1}\{F(y_i; \hat{\mu}_i, \hat{\sigma})\}, \quad (3.7)$$

em que $\Phi(\cdot)$ é a função de distribuição acumulada da normal padrão.

No nosso caso consideramos como $F(\cdot)$ a função de distribuição acumulada da distribuição gama e gaussiana inversa.

Proposição 1. *Se μ_i e σ são conhecidos, então r_i^{qu} converge em distribuição para a normal padrão.*

Demonstração.

$$P(F(y_i; \mu_i; \sigma) < k) = P(y_i < F^{-1}(k)) = F(F^{-1}(k)) = k. \quad (3.8)$$

$$(3.9)$$

Então

$$F(y_i; \mu_i, \sigma) \sim U(0, 1). \quad (3.10)$$

Portanto,

$$\Phi^{-1}\{F(y_i; \mu_i, \sigma)\} \sim N(0, 1). \quad (3.11)$$

□

Substituindo μ_i e σ por suas estimativas $\hat{\mu}_i$ e $\hat{\sigma}$, espera-se que sua distribuição seja próxima da normal padrão. Apesar do resíduo r_i^{qu} ser facilmente implementado, não existe na literatura estudos sobre esses resíduos nos MLGs. Na próxima seção (Seção 3.2) analisamos as propriedades desses resíduos nos MLGs.

3.1.4 Resíduo Anscombe

Anscombe (1953) sugeriu um resíduo baseado na ideia de transformar a variável observada y_i , com o objetivo de que o resíduo baseado nessa função transformada se aproxime o máximo possível da distribuição normal padrão.

Definição 10. Considere $\psi(\cdot)$ uma função utilizada para normalizar a distribuição da variável resposta y . Definimos o resíduo Anscombe como

$$r_i^{ans} = \frac{\{\psi(y_i) - \psi(\hat{\mu}_i)\}}{\hat{\sigma}\hat{V}^{1/2}(\hat{\mu})\psi'(\hat{\mu}_i)}. \quad (3.12)$$

Nos modelos lineares generalizados, definimos a função ψ como:

$$\psi(\mu) = \int_0^\mu V^{-1/3}(t)dt. \quad (3.13)$$

Particularmente temos que o resíduo de Anscombe para o modelo gama é definido como:

$$r_i^{ans} = \frac{\{3y_i^{1/3} - 3\hat{\mu}_i^{1/3}\}}{\hat{\sigma}\hat{V}^{1/2}(\hat{\mu})\hat{\mu}_i^{(-2/3)}}. \quad (3.14)$$

O resíduo de Anscombe para o modelo gaussiana inversa é definido como:

$$r_i^{ans} = \frac{\{\log(y_i) - \log(\hat{\mu}_i)\}}{\hat{\sigma}\hat{V}^{1/2}(\hat{\mu}_i)(1/\hat{\mu}_i)}. \quad (3.15)$$

Nota-se que os dados observados e os dados ajustados têm a mesma função normalizadora.

3.1.5 Resíduo Williams

Baseado no resíduo de Pearson e desvio, Williams (1987) propôs um resíduo para os MLGs.

Definição 11. *O resíduo Williams é definido como:*

$$r_i^{wil} = \text{sign}(y_i - \hat{\mu}_i) \{ (1 - \hat{h}_{ii})(r_i^{dev})^2 + \hat{h}_{ii}(r_i^{pea})^2 \}^{1/2}. \quad (3.16)$$

Observa-se que o resíduo proposto por Williams (1987) é uma espécie de média ponderada dos resíduos de Pearson e de desvio. O autor verificou por meio de estudos de simulação que r_i^{wil} tem variância excedendo um, média próxima de zero, assimetria desprezível e um excesso de curtose (Paula, 2013).

3.2 Estudos de simulação

Nesta seção analisamos e comparamos os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} no modelo de regressão com resposta gama e gaussiana inversa em que $g(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$, $i = 1, \dots, n$, por meio de estudos de simulação de Monte Carlo (MC). As covariáveis foram obtidas a partir das distribuições uniforme padrão, normal (N), gaussiana inversa (IG) e gama (GA), o que será explicado a seguir, e permaneceram constantes ao longo das simulações.

Consideramos amostras de tamanho $n = 15$ e $n = 50$ em todos os cenários e foram geradas $M = 5000$ réplicas de Monte Carlo. Após o ajuste dos modelos foram calculados os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} e algumas medidas de interesse foram calculadas para cada resíduo, como média, variância, assimetria, curtose e a estatística de Anderson Darling (AD). O teste de Anderson-Darling é um teste estatístico para verificar se uma dada amostra de dados é tirada a partir de uma dada distribuição de probabilidade. Utilizaremos essa medida para analisar quais dos resíduos possui uma distribuição mais próxima da normal padrão, ou seja, aquele resíduo com menor estatística AD é o resíduo que mais se aproxima da distribuição normal padrão (Anderson e Darling, 1952).

É bastante interessante que a distribuição dos resíduos utilizados na análise de diagnóstico se aproxime da normal padrão, pois assim teremos resíduos com propriedades conhecidas e os gráficos de diagnóstico serão mais facilmente interpretáveis. Assim, utilizaremos essas medidas para decidir quais resíduos têm distribuição mais próxima da normal padrão, aquele resíduo que tiver distribuição mais próxima da normal padrão será considerada um resíduo com boas propriedades¹. Para a obtenção dos resultados calculamos a média, variância, assimetria, curtose e estatística de Anderson Darling (AD) do r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} para as $M = 5000$ réplicas de Monte Carlo. A construção das tabelas para analisar os resultados

¹Na análise de diagnóstico nos MLGs, os resíduos foram desenvolvidos para ter distribuição mais próxima da distribuição normal, mas temos na literatura resíduos com boas propriedades, cuja distribuição é diferente da distribuição normal padrão.

de cada cenário, foram construídas da seguinte maneira: para cada réplica de MC calculamos os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} e depois para cada observação i das 5000 réplicas de MC calculamos a média, variância, assimetria, curtose e estatística de Anderson Darling (AD). Por exemplo, para primeira observação da amostra $i = 1$, calculamos o resíduo desvio para as 5000 réplicas de MC. Após o cálculo dos resíduos, calculamos a média do resíduo desvio para a observação 1, a variância do resíduo desvio para a observação e assim por diante, como apresentamos na Tabela 3.1. Nos estudos de simulação foi utilizado o software estatístico R

Tabela 3.1: Estruturas das tabelas apresentadas no Apêndice B e C

replica \ resíduo	r_1	r_2	r_3	...	r_{15}
1	r_{11}	r_{12}	r_{13}	...	r_{115}
2	r_{21}	r_{22}	r_{23}	...	r_{215}
3	r_{31}	r_{32}	r_{33}	...	r_{315}
4	r_{41}	r_{42}	r_{43}	...	r_{415}
⋮	⋮	⋮	⋮	...	⋮
4999	r_{49991}	r_{49992}	r_{49993}	...	r_{499915}
5000	r_{50001}	r_{50002}	r_{50003}	...	r_{500015}
média	\bar{r}_1	\bar{r}_2	\bar{r}_3	...	\bar{r}_{15}
variância	$var(r_1)$	$var(r_2)$	$var(r_3)$...	$var(r_{15})$
assimetria	$assim(r_1)$	$assim(r_2)$	$assim(r_3)$...	$assim(r_{15})$
curtose	$cur(r_1)$	$cur(r_2)$	$cur(r_3)$...	$cur(r_{15})$
AD	$AD(r_1)$	$AD(r_2)$	$AD(r_3)$...	$AD(r_{15})$

e utilizou-se o pacote "GAMLSS" para a estimação dos parâmetros e geração das variáveis (mais informações sobre o gamlss em [Stasinopoulos et al. \(2014\)](#)).

3.2.1 Modelo gama

Nos 6 primeiros cenários consideramos a seguinte estrutura para o modelo com resposta gama:

$$\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}. \quad (3.17)$$

Diversos cenários foram considerados nos estudos de simulação de Monte Carlo, para avaliar a performance desses resíduos. Nos 4 primeiros cenários consideramos a distribuição uniforme padrão para gerar as variáveis preditoras. No primeiro cenário consideramos $\beta_0 = 3, \beta_1 = 2, \beta_2 = 1$ e $\sigma = 0,1$ resultando em $\mu \in (20,085; 403,429)$. No segundo e terceiro cenários consideramos os mesmos valores dos $\beta_j, j = 0, 1, 2$ do cenário I. Porém variamos o valor de σ , considerando $\sigma = 0,05$ e $\sigma = 0,5$ nos cenários II e III respectivamente. No quarto cenário, variamos os valores de μ_i para próximo de zero e coeficiente de variação igual ao cenário I, ou seja, consideramos $\beta_0 = -3, \beta_1 = 1,5, \beta_2 = 1$, resultando em $\mu \in (0,049; 0,606)$ e $\sigma = 0,1$. No quinto e sexto cenários mudamos a distribuição das variáveis preditoras, considerando $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$ no cenário 5 e $x_1 \sim N(0,5; 0,25^2)$ e

$x_2 \sim IG(0,4;2)$ no cenário VI. Em ambos os cenários (cenários V e VI) consideramos $\beta_0 = 3, \beta_1 = 2, \beta_2 = 1$ e $\sigma = 0,1$.

No sétimo cenário alteramos a estrutura do modelo (3.19) em que mudamos a função de ligação log para a função de ligação canônica do modelo gama (função inversa):

$$\frac{1}{\mu_i} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}. \quad (3.18)$$

em que $\beta_0 = 0,0025, \beta_1 = 0,04, \beta_2 = 0,01$ e $\sigma = 0,1$ resultando em $\mu \in (20,22; 190,10)$. Na Tabela 3.2 apresentamos um resumo das alterações feitas em cada cenário em relação ao cenário I. A escolha da estrutura do modelo e seus coeficientes foram definidos para serem os mais próximos de dados reais. A função de ligação log foi escolhida, pois essa é a função mais utilizada nesses tipo de modelo. E optamos por mudar a função de ligação no cenário VII, para ver o comportamento dos resíduos quando alteramos a função de ligação.

Tabela 3.2: Alterações feitas nos cenários em relação ao cenário I.

Cenário	Alterações
II	Diminuímos o valor de σ
III	Aumentamos o valor de σ
IV	Valores de μ próximo de zero
V	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$
VI	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim IG(0,4; 2)$
VII	Mudança na função de ligação: função de ligação invesa

Os resultados dos estudos de simulação para o tamanho de amostra $n = 15$ são apresentados nas tabelas de B1 a B7 (apresentadas no Apêndice B).

No cenário I, a variável Y tem média, μ entre 20,08 e 403,43; desvio padrão entre 2,00 e 40,34, assimetria de 0,2 e coeficiente de variação de 0,1. Os resultados dos estudos de simulação de MC para esse cenário sugerem que em geral, todos os resíduos têm média, assimetria e curtose próximo da normal padrão. Mas aqueles resíduos que tiveram suas médias mais próximas de zero foram os resíduos r_i^{pea} e r_i^{qu} , sendo r_i^{pea} aquele com a média mais próxima de zero. A variância próxima de 1 notamos mais nos resíduos de r_i^{pea} , r_i^{dev} e r_i^{wil} . A assimetria mais próxima de zero notamos no resíduo r_i^{wil} , r_i^{dev} e r_i^{qu} . O excesso de curtose não é muito próximo de zero para nenhum dos resíduos apresentados e todos os resíduos forneceram resultados de curtose muito próximos. Analisando a estatística AD notamos que aqueles que possuem menor estatística AD são os resíduos r_i^{pea} , r_i^{dev} e r_i^{wil} . Os resíduos r_i^{qu} e r_i^{ans} têm AD consideravelmente superiores principalmente por apresentar variância consideravelmente inferior a 1.

Ao diminuirmos a variância (cenário II, Tabela B.2) não notamos grandes diferenças entre os resultados do estudo de simulação do cenário I e cenário II. Ou seja, como no cenário I temos que todos os resíduos possuem médias próximas de zero tendo uma pequena melhora nas médias quando comparada ao cenário I, a variância próxima de 1 notamos nos resíduos r_i^{pea} , r_i^{dev} e r_i^{wil} . O excesso de curtose é notado para todos os resíduos e as menores estatísticas

AD são dos resíduos r_i^{pea} , r_i^{dev} e r_i^{wil} . Já o aumento da variância, cenário III (Tabela B.3) provoca uma piora nos resultados dos resíduos em geral. Ao aumentarmos a variância temos que os resíduos r_i^{dev} , r_i^{wil} , r_i^{ans} têm um leve aumento na média, não sendo tão próximas de zero como nos cenários anteriores. E para todos os resíduos, o aumento da variância provoca uma piora na assimetria dos resíduos, ou seja, nesse cenário todos os resíduos apresentam uma assimetria não tão próxima de zero como no cenário I. O resíduo que possui a menor estatística AD no cenário III é o r_i^{qu} . Vale ressaltar ainda que o resíduo r_i^{qu} foi o resíduo que sofreu menor influência do aumento da variância. Uma possível explicação para a piora dos resultados dos resíduos com o aumento de σ , se deve ao fato que o aumento do valor de σ aumenta a assimetria da distribuição da variável resposta (Figura 2.1).

Na Tabela B.4, apresentamos o cenário IV, onde mantemos o mesmo coeficiente de variação do cenário I e diminuimos o parâmetro μ para próximo de zero e nas Tabelas B.5 e B.6 (cenários V e VI) mudamos a função geradora das variáveis preditoras. Essas mudanças não parecem afetar substancialmente os resultados da simulação, sendo todos os resultados para esses cenários bem próximos aos resultados apresentados para o cenário I.

No último cenário, para o modelo com resposta gama (Tabela B.7), mudamos a função de ligação do modelo para a função inversa e mais uma vez os resultados apresentados para esse cenário são similares aos resultados do cenário I.

Na Tabela 3.3, temos a comparação das medidas descritivas da estatística de Anderson-Darling para $n = 15$ e $n = 50$ para todos os cenários. Observamos que para amostras de tamanho 15 temos que a distribuição do resíduo de Pearson padronizado é mais próxima da normal padrão do que dos demais resíduos exceto no cenário que aumentamos a variância, (cenário 3) onde encontramos o r_i^{qu} como o resíduo cuja distribuição mais se aproxima da normal padrão. Para $n = 50$, os resultados sugerem que em geral o r_i^{qu} nesses modelos tem distribuição consideravelmente mais próxima da normal padrão do que dos demais resíduos. Note porém que nos cenários 6 e 7 para $n = 50$ (Tabela 3.3), o valor máximo de r_i^{qu} é consideravelmente maior que o valor máximo dos outros resíduos. Isso, ocorre porque para uma observação temos um valor ajustado muito maior/menor dos que os demais valores ajustados, assim a variabilidade do resíduo r_i^{qu} para essa observação fica próxima de 0,15, bem afastado da variância da normal padrão. Assim, o resíduo quantílico deve ser usado com cautela em situações em que há observações com valores ajustados muito superiores ou muito inferiores as demais.

Concluimos que para o modelo com resposta gama o resíduo cuja distribuição mais se aproxima da normal padrão é o r_i^{pea} quando o tamanho da amostra é pequena ($n = 15$). E para amostras de tamanhos moderados ($n = 50$) ou grandes o melhor resíduo parece ser o r_i^{qu} . E o resíduo r_i^{qu} é aquele cuja estatística AD apresenta maior redução quando o tamanho da amostra aumenta.

Tabela 3.3: Comparação da estatística de Anderson-Darling para $n=15$ e $n=50$ no modelo gama

Cenário	σ	n	Resíduos	Desvio						
				media	Padrão	Mínimo	quantiles			
							Q1	Q2	Q3	Máximo
1	0,1	15	r_i^{dev}	4,89	2,33	1,38	2,91	4,59	6,40	8,82
			r_i^{pea}	4,10	1,82	1,86	2,42	3,51	5,58	6,92
			r_i^{qu}	15,26	15,52	1,57	5,53	7,70	19,54	49,03
			r_i^{ans}	16,39	14,87	3,69	7,17	9,88	19,22	49,88
			r_i^{wil}	4,45	2,23	1,18	2,55	4,36	5,89	8,07
1	0,1	50	r_i^{dev}	3,38	2,19	0,43	1,68	3,14	4,52	11,37
			r_i^{pea}	2,75	1,32	0,81	1,75	2,49	3,42	7,49
			r_i^{qu}	2,03	1,46	0,25	0,90	1,63	2,84	6,78
			r_i^{ans}	4,16	2,46	0,78	2,12	3,67	5,61	13,52
			r_i^{wil}	3,20	2,10	0,39	1,61	2,96	4,27	10,90
2	0,05	15	r_i^{dev}	4,05	2,36	1,19	2,47	3,66	4,71	10,98
			r_i^{pea}	3,94	1,98	1,18	2,59	3,76	4,31	9,79
			r_i^{qu}	15,39	17,07	2,39	4,67	8,94	16,56	65,91
			r_i^{ans}	15,72	17,03	1,86	4,65	9,17	17,95	65,34
			r_i^{wil}	3,94	2,25	1,14	2,45	3,70	4,51	10,57
2	0,05	50	r_i^{dev}	1,66	1,14	0,31	0,80	1,31	2,07	4,70
			r_i^{pea}	1,49	0,74	0,49	0,91	1,32	1,93	3,44
			r_i^{qu}	1,96	1,36	0,40	1,19	1,48	2,45	8,01
			r_i^{ans}	2,51	1,50	0,33	1,33	2,35	3,57	6,54
			r_i^{wil}	1,61	1,11	0,30	0,80	1,26	2,04	4,56
3	0,5	15	r_i^{dev}	54,12	12,65	36,26	46,30	51,66	59,37	79,86
			r_i^{pea}	30,69	5,93	17,97	29,08	31,03	32,81	42,38
			r_i^{qu}	16,30	13,10	5,69	6,96	9,71	24,16	47,35
			r_i^{ans}	48,81	16,26	31,67	38,62	41,97	55,41	90,72
			r_i^{wil}	41,13	9,13	22,18	33,95	42,94	46,10	58,35
3	0,5	50	r_i^{dev}	62,30	12,15	35,72	53,76	62,81	71,33	88,02
			r_i^{pea}	45,43	6,45	32,44	40,76	45,55	48,89	61,78
			r_i^{qu}	2,12	1,64	0,20	0,96	1,76	2,86	8,32
			r_i^{ans}	57,93	11,48	33,76	49,88	59,34	67,18	81,83
			r_i^{wil}	57,88	11,41	31,18	50,35	60,00	67,09	80,94
4	0,1	15	r_i^{dev}	4,97	1,89	1,70	3,80	4,35	6,58	8,30
			r_i^{pea}	4,00	1,51	1,62	3,04	3,94	5,07	7,52
			r_i^{qu}	16,32	17,19	2,13	6,95	8,56	18,79	55,67
			r_i^{ans}	17,25	15,74	3,69	7,18	11,45	20,13	52,04
			r_i^{wil}	4,53	1,84	1,31	3,28	4,05	5,88	8,09
4	0,1	50	r_i^{dev}	3,54	2,48	0,44	1,75	2,89	4,58	11,57
			r_i^{pea}	3,02	1,67	0,73	2,09	2,58	3,48	8,52
			r_i^{qu}	2,03	1,32	0,34	1,05	1,80	2,85	6,64
			r_i^{ans}	4,26	2,40	0,43	2,64	3,65	5,36	10,36
			r_i^{wil}	3,38	2,43	0,45	1,59	2,77	4,37	11,35
5	0,1	15	r_i^{dev}	5,13	1,95	2,41	3,66	5,40	6,09	8,39
			r_i^{pea}	4,17	1,42	1,88	3,16	4,30	4,79	6,84
			r_i^{qu}	17,61	22,08	2,18	5,07	8,09	16,11	71,24
			r_i^{ans}	18,57	21,48	2,00	8,47	10,42	16,25	70,94
			r_i^{wil}	4,63	1,83	1,96	3,56	4,66	5,46	7,76
5	0,1	50	r_i^{dev}	3,54	2,16	0,42	2,17	2,81	4,91	8,82
			r_i^{pea}	3,09	1,28	1,33	2,22	2,90	3,74	6,48
			r_i^{qu}	2,11	1,97	0,31	0,92	1,44	2,17	9,68
			r_i^{ans}	4,42	2,70	0,59	2,35	3,83	6,30	13,68
			r_i^{wil}	3,37	2,10	0,38	2,05	2,77	4,74	8,57
6	0,1	15	r_i^{dev}	5,24	2,40	1,56	3,51	4,89	6,45	9,92
			r_i^{pea}	4,48	1,81	1,74	3,34	4,32	5,90	7,30
			r_i^{qu}	18,90	26,01	1,91	4,13	6,60	25,36	103,22
			r_i^{ans}	20,06	25,31	3,05	5,67	7,09	26,77	101,56
			r_i^{wil}	4,79	2,35	1,53	3,04	4,48	6,21	9,49
6	0,1	50	r_i^{dev}	3,46	2,55	0,41	1,86	2,60	4,35	14,83
			r_i^{pea}	2,94	1,80	0,51	1,97	2,58	3,29	11,40
			r_i^{qu}	14,45	90,27	0,14	0,84	1,25	1,91	639,87
			r_i^{ans}	16,65	89,67	0,20	2,34	3,20	5,55	637,79
			r_i^{wil}	3,33	2,54	0,40	1,72	2,50	4,07	14,71
7	0,1	15	r_i^{dev}	5,43	2,67	2,00	3,67	5,08	7,02	10,60
			r_i^{pea}	4,70	2,18	1,90	3,06	4,37	6,34	8,79
			r_i^{qu}	38,85	86,84	1,54	3,23	5,39	9,60	253,78
			r_i^{ans}	39,99	86,84	2,66	4,19	5,53	11,54	253,84
			r_i^{wil}	4,99	2,55	1,80	3,00	4,64	6,47	10,09
7	0,1	50	r_i^{dev}	3,30	2,25	0,37	1,62	2,95	4,45	13,88
			r_i^{pea}	2,74	1,44	0,94	1,76	2,49	3,29	9,80
			r_i^{qu}	4,15	13,73	0,14	0,66	1,03	1,62	95,46
			r_i^{ans}	6,34	13,98	0,38	1,86	3,17	5,37	99,03
			r_i^{wil}	3,13	2,21	0,36	1,45	2,81	4,32	13,66

3.2.2 Modelo gaussiana inversa

Nos 6 primeiros cenários consideramos a seguinte estrutura para o modelo com resposta gaussiana inversa:

$$\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}. \quad (3.19)$$

Diversos cenários foram considerados nos estudos de simulação de MC para avaliar a performance desses resíduos. Nos 4 primeiros cenários consideramos a distribuição uniforme padrão para gerar as variáveis preditoras. No primeiro cenário consideramos $\beta_0 = 3, \beta_1 = 2, \beta_2 = 1$ e $\sigma = 0,02$ resultando em $\mu \in (20,085; 403,429)$. No segundo e terceiro cenários consideramos os mesmos valores dos $\beta_j, j = 0, 1, 2$ do cenário I. Porém variamos o valor de σ , considerando $\sigma = 0,01$ e $\sigma = 0,03$ nos cenários II e III respectivamente. No quarto cenário variamos os valores de μ_i para próximo de zero e mantivemos o mesmo coeficiente de variação do cenário I, ou seja, consideramos $\beta_0 = -3; \beta_1 = 1,5, \beta_2 = 1$, resultando em $\mu \in (0,049; 0,606)$ e $\sigma = 0,5$. No quinto e sexto cenário mudamos as funções de distribuição das variáveis preditoras, considerando $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$ no cenário V e $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim IG(0,4; 2)$ no cenário VI. Em ambos os cenários (cenários V e VI) consideramos $\beta_0 = 3, \beta_1 = 2, \beta_2 = 1$ e $\sigma = 0,02$.

No sétimo cenário alteramos a estrutura do modelo (3.19) onde mudamos a função de ligação log para a função de ligação canônica do modelo gaussiana inversa:

$$\frac{1}{\mu_i^2} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}. \quad (3.20)$$

em que $\beta_0 = 0,000006, \beta_1 = 0,002, \beta_2 = 0,001, \sigma = 0,02$ resultando em $\mu \in (18,23; 408,25)$ e utilizamos a distribuição uniforme padrão para gerar as variáveis preditoras. Na Tabela 3.4 apresentamos as alterações feitas em cada cenário em relação ao cenário I. A estrutura do modelo e os coeficientes considerados nesse trabalho, foram baseados em dados reais. E a escolha da função de ligação log foi adotada, pois essa é a função de ligação mais utilizada no ajuste desses modelos. E no cenário 7, optou-se em mudar a função de ligação para analisar o comportamento dos resíduos ao mudarmos a função de ligação.

Tabela 3.4: Alterações feitas nos cenários em relação ao cenário I.

Cenário	Alterações
II	Diminuímos o valor de σ
III	Aumentamos o valor de σ
IV	Valores de μ próximo de zero
V	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$
VI	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim IG(0,4; 2)$
VII	Mudança na função de ligação: função de ligação invesa

Os resultados dos estudos de simulação para o tamanho de amostra $n = 15$ são apresen-

tados nas tabelas de $C1$ a $C7$ (apresentadas no Apêndice C). Para a obtenção dos resultados calculamos a média, variância, assimetria, curtose e estatística de Anderson Darling (AD) de r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} para as $M = 5000$ réplicas de MC.

No cenário I, a variável Y tem média, μ entre 20,08 e 403,43; desvio padrão entre 1,80 e 162,06, assimetria entre 0,27 e 1,20 e coeficiente de variação entre 0,09 e 0,40. Os resultados da simulação de MC para esse cenário sugerem os resíduos em geral tem média, variância e assimetria próximas da normal padrão. Mas aqueles que tiveram suas médias mais próximas de zero foram os resíduos r_i^{pea} e r_i^{qu} , sendo r_i^{pea} aquele com a média mais próxima de zero. A variância próxima de 1 notamos mais nos resíduos de r_i^{pea} , r_i^{dev} e r_i^{wil} . A assimetria mais próxima de zero notamos no resíduo r_i^{wil} , r_i^{dev} e r_i^{qu} . Para os resíduos apresentados notamos o excesso de curtose, sendo esse não muito próximo de zero, além disso os resíduos forneceram resultados de curtose muito próximos. Na análise da estatística AD notamos que os resíduos r_i^{pea} , r_i^{wil} e r_i^{qu} são aqueles que possuem menor estatística AD, nessa ordem. Os outros 2 resíduos (r_i^{qu} e r_i^{ans}) tem AD consideravelmente superiores principalmente por apresentar variância substancialmente inferior a 1.

No cenário II, diminuimos a variância (Tabela C.2) e obtemos resultados similares ao cenário I. A média para os resíduos r_i^{dev} , r_i^{ans} e r_i^{wil} ficam um pouco mais próxima de zero, a assimetria de todos os resíduos aumenta um pouco para os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} . O excesso de curtose é notado para todos os resíduos e as estatísticas AD diminuem para os resíduos r_i^{pea} , r_i^{dev} , r_i^{ans} e r_i^{wil} , para o r_i^{qu} a estatística AD quase não se altera. Ao aumentarmos a variância, cenário III (Tabela C.3) provoca uma piora nos resultados dos resíduos em geral, exceto no excesso de curtose que em média reduz para todos os resíduos. Ao aumentarmos a variância temos que os resíduos r_i^{dev} , r_i^{wil} , r_i^{ans} tem um leve aumento na média, não sendo tão próximas de zero como nos cenários anteriores. E para todos os resíduos, o aumento da variância provoca uma piora na assimetria dos resíduos, ou seja, nesse cenário todos os resíduos apresentam uma assimetria não tão próxima de zero como no cenário I. O resíduo que possui a menor estatística AD no cenário III é o r_i^{qu} , vale ressaltar ainda que o resíduo r_i^{qu} foi o resíduo que sofreu menos influência na variação da variância. A piora nos resultados com o aumento dos valores de σ é explicada pelo fato que o aumento do valor de σ provoca uma maior assimetria na distribuição desses modelos.

Na Tabela C.4, apresentamos o cenário IV, onde mantemos um coeficiente de variação próximo do cenário I e diminuimos os valores do parâmetro μ para próximo de zero e nas Tabelas C.5 e C.6 (cenário V e VI) mudamos a função geradora das variáveis preditoras. Essas mudanças não parecem afetar substancialmente os resultados da simulação, sendo todos os resultados para esses cenários bem próximos aos resultados apresentados para o cenário I.

No último cenário (Tabela C.5), mudamos a função de ligação do modelo com resposta gaussiana inversa para a função canônica $g(\mu_i) = 1/\mu_i^2$ e mais uma vez os resultados apresentados para esse cenários são similares aos resultados do cenário I.

Apresentamos a comparação das medidas descritivas da estatística de Anderson-Darling

para $n = 15$ e $n = 50$ para todos os cenários (Tabela 3.5), observamos que para amostras de tamanho 15 temos que a distribuição do resíduo de Pearson padronizado é mais próxima da normal padrão do que os demais resíduos exceto no cenário que aumentamos a variância, (cenário 3) onde encontramos o r_i^{qu} como o resíduo cuja a distribuição mais se aproxima da normal padrão. Para $n = 50$, os resultados sugerem que o r_i^{qu} nesses modelos tem distribuição consideravelmente mais próxima da normal padrão do que os demais resíduos.

Portanto, para o modelo com resposta gaussiana inversa o resíduo cuja distribuição mais aproxima da normal padrão é o r_i^{pea} quando o tamanho da amostra é pequena ($n=15$). E para amostra de tamanhos moderados ($n=50$) ou grandes em geral o melhor resíduo é o r_i^{qu} , que é aquele cuja estatística AD apresenta maior redução quando o tamanho de amostra aumenta. Note porém que no cenário 7 para $n = 50$ (Tabela 3.5), o valor máximo de r_i^{qu} é consideravelmente maior que o valor máximo dos outros resíduos. Isso, ocorre porque para uma observação temos um valor ajustado muito maior dos que os demais valores ajustados, assim a variabilidade do resíduo r_i^{qu} para essa observação fica próxima de 0,04, bem afastado da variância da normal padrão. Assim, o resíduo quantílico deve ser usado com cautela em situações em que há observações com valores ajustados muito superiores ou muito inferiores as demais.

3.2.3 Considerações finais sobre os estudos de simulação nos MLGs

Os resultados gerais dos resíduos nos estudos de simulação de Monte Carlo foram semelhantes para ambos os modelos (gama e gaussiana inversa). Esses resultados foram próximos em relação à ordenação dos resíduos (melhores e piores). Porém em relação à estatística de Anderson-Darling, o modelo com resposta gama apresenta valores das estatísticas menores do que no modelo com resposta gaussiana inversa. Uma explicação plausível para isso seria que o modelo com resposta gaussiana inversa apresenta uma assimetria maior na variável resposta do que o modelo com resposta gama.

O resíduo Anscombe e o desvio têm segundo artigos distribuição mais próxima da normal do que o resíduo de Pearson. Os resultados apresentados sugerem que isso seja verdade, pois esses resíduos possuem uma menor assimetria do que a do resíduo de Pearson e possuem curtose semelhantes. Porém, os estudos de simulação através da estatística AD sugerem que as distribuições desses resíduos estão mais distantes da normal padrão do que o resíduo de Pearson pelo fato do primeiro ter variância consideravelmente menor do que 1 e o segundo média não tão próxima de zero quanto os demais resíduos. Segundo Williams (1987), o resíduo r_i^{wil} apresenta uma média ligeiramente diferente zero, variância excedendo a 1, alguma curtose e assimetria desprezível e os estudos de simulação confirmam os resultados apresentados por Williams. Conforme já mencionado, não há na literatura estudos envolvendo as propriedades do resíduo quantílico em modelos lineares generalizados. Porém, os estudos de simulação que realizamos sugerem que, em amostras de tamanho moderado ou grande, a distribuição do resíduo quantílico em modelos com resposta gama e gaussiana inversa está

Tabela 3.5: Comparação da estatística de Anderson-Darling para $n=15$ e $n=50$ no modelo gaussiana inversa.

Cenário	σ	n	Resíduos	Média	Desv. Pad	Mínimo	quantiles			Máximo
							Q1	Q2	Q3	
1	0,02	15	r_i^{dev}	22,76	12,90	5,93	13,95	18,66	28,87	52,03
			r_i^{pea}	14,96	8,12	3,50	8,68	15,68	18,87	33,36
			r_i^{qu}	19,90	29,62	2,47	4,12	5,58	18,01	97,14
			r_i^{ans}	35,49	25,74	11,53	17,91	24,89	42,49	98,19
			r_i^{wil}	18,72	11,32	3,68	11,35	16,10	22,33	45,02
1	0,02	50	r_i^{dev}	25,25	20,46	2,30	10,83	16,80	37,06	93,32
			r_i^{pea}	18,94	14,74	2,80	8,35	12,81	26,71	70,49
			r_i^{qu}	2,45	3,35	0,19	0,98	1,45	2,55	22,58
			r_i^{ans}	25,45	19,50	3,99	11,40	17,45	36,13	91,37
			r_i^{wil}	23,80	19,76	1,85	9,54	16,29	34,85	89,89
2	0,01	15	r_i^{dev}	8,15	4,10	1,71	4,33	8,12	11,26	16,40
			r_i^{pea}	5,98	2,51	1,39	4,09	5,89	7,67	10,24
			r_i^{qu}	21,22	33,13	2,49	4,54	5,96	19,25	101,61
			r_i^{ans}	24,98	30,87	5,99	9,10	12,23	21,70	101,44
			r_i^{wil}	7,18	3,69	1,20	3,75	7,43	9,87	14,44
2	0,01	50	r_i^{dev}	7,43	6,23	0,68	2,84	5,39	10,42	26,38
			r_i^{pea}	5,78	4,26	0,38	2,60	4,02	7,91	18,19
			r_i^{qu}	2,68	3,77	0,17	0,78	1,43	3,19	25,11
			r_i^{ans}	8,61	6,23	0,29	3,80	7,50	11,45	25,80
			r_i^{wil}	7,06	6,04	0,63	2,61	4,89	9,93	25,24
3	0,03	15	r_i^{dev}	45,65	28,81	12,98	25,10	41,86	58,14	107,17
			r_i^{pea}	28,64	18,14	2,89	17,50	24,62	40,88	62,00
			r_i^{qu}	19,78	30,43	2,62	3,92	6,38	16,06	100,97
			r_i^{ans}	54,06	26,81	18,61	32,73	45,84	72,99	103,38
			r_i^{wil}	37,12	25,72	6,04	20,23	28,72	49,22	91,55
3	0,03	50	r_i^{dev}	52,45	35,73	9,89	24,37	38,38	79,75	138,65
			r_i^{pea}	40,54	27,08	7,07	17,98	29,30	61,75	104,69
			r_i^{qu}	2,11	2,44	0,35	0,76	1,26	2,40	14,21
			r_i^{ans}	51,20	33,77	11,89	24,90	37,98	77,64	132,82
			r_i^{wil}	49,42	34,24	8,17	23,07	36,66	75,76	132,70
4	0,5	15	r_i^{dev}	24,90	15,07	8,60	14,93	21,13	28,02	57,83
			r_i^{pea}	15,90	8,45	3,82	11,02	14,51	19,35	32,85
			r_i^{qu}	19,21	25,32	1,99	4,54	6,41	25,21	84,33
			r_i^{ans}	36,35	23,28	8,75	21,42	26,06	49,31	89,99
			r_i^{wil}	20,19	13,22	5,30	11,58	17,26	24,16	49,22
4	0,5	50	r_i^{dev}	27,37	16,65	6,57	14,97	20,41	34,98	69,70
			r_i^{pea}	21,13	12,54	5,79	10,80	17,04	29,13	53,65
			r_i^{qu}	1,97	1,98	0,35	0,85	1,23	2,33	11,80
			r_i^{ans}	27,33	15,65	9,55	15,61	21,79	33,88	67,52
			r_i^{wil}	25,68	15,92	5,66	13,58	19,39	33,00	66,09
5	0,02	15	r_i^{dev}	15,22	4,90	5,55	12,05	15,47	18,17	22,85
			r_i^{pea}	9,45	3,18	3,86	7,65	8,67	12,06	14,85
			r_i^{qu}	17,12	20,36	3,10	5,22	8,71	17,60	74,38
			r_i^{ans}	25,76	18,71	7,07	16,35	21,46	24,12	76,85
			r_i^{wil}	12,02	4,76	4,18	9,25	11,97	14,71	20,28
5	0,02	50	r_i^{dev}	18,03	10,74	6,81	12,65	15,91	19,47	70,96
			r_i^{pea}	13,51	6,70	6,62	9,18	12,19	15,20	45,23
			r_i^{qu}	2,17	2,01	0,24	0,74	1,30	3,15	7,89
			r_i^{ans}	18,38	10,56	6,71	12,88	16,46	20,02	69,54
			r_i^{wil}	16,70	9,66	6,29	11,67	14,20	18,26	63,72
6	0,02	15	r_i^{dev}	19,51	10,79	4,94	12,95	14,73	23,81	44,83
			r_i^{pea}	11,89	3,95	6,15	8,23	13,64	14,85	17,75
			r_i^{qu}	16,78	15,44	1,76	4,84	7,83	24,29	49,59
			r_i^{ans}	28,17	20,45	5,26	12,67	19,00	37,94	75,90
			r_i^{wil}	14,61	6,18	4,23	9,57	13,35	19,70	25,17
6	0,02	50	r_i^{dev}	18,60	8,98	4,67	12,54	18,20	22,22	43,41
			r_i^{pea}	14,38	6,16	5,23	9,96	12,81	17,23	31,94
			r_i^{qu}	2,00	1,69	0,34	0,68	1,49	2,83	8,59
			r_i^{ans}	18,71	8,74	6,40	12,55	17,01	21,71	41,55
			r_i^{wil}	17,29	8,25	3,76	11,96	16,71	20,27	39,84
7	0,02	15	r_i^{dev}	8,27	4,25	2,98	4,70	7,10	11,80	15,54
			r_i^{pea}	6,72	2,25	4,00	5,24	5,52	8,20	10,91
			r_i^{qu}	24,78	48,12	2,87	4,59	5,50	15,22	182,17
			r_i^{ans}	27,91	49,09	3,08	5,94	8,34	21,65	186,83
			r_i^{wil}	7,03	3,41	2,79	4,20	6,19	9,00	13,17
7	0,02	50	r_i^{dev}	7,01	4,18	1,13	4,37	6,13	8,29	25,99
			r_i^{pea}	5,63	2,04	2,52	4,49	5,29	6,61	12,50
			r_i^{qu}	4,83	20,09	0,35	0,79	1,27	2,26	143,04
			r_i^{ans}	10,27	21,79	1,23	4,27	6,48	9,43	158,74
			r_i^{wil}	6,48	3,41	1,06	4,28	5,88	7,61	17,22

mais próxima da normal padrão do que a dos resíduos usualmente usados na análise de diagnóstico desses modelos.

3.3 Análise de diagnóstico em modelos inflacionados ZAGA e ZAIG

Na análise de diagnóstico é necessário avaliar diferentes aspectos do ajuste do modelo, tais como investigar se existe problema na função de ligação adotada, ou na distribuição da variável resposta adotada ou ainda se existem observações discrepantes e/ou influentes. Na próxima seção abordaremos técnicas para identificação de observações discrepantes em modelos de regressão ZAGA e ZAIG.

3.3.1 Identificação de observações discrepantes - componente contínuo

Nos modelos inflacionados no zero, a análise de diagnóstico possui algumas peculiaridades, pois em geral temos dois valores ajustados para cada observação. Um é a estimativa do valor esperado dado que a observação assumiu valores diferentes de zero e o outro é a estimativa da probabilidade da observação assumir o valor zero (Pereira, Botter e Sandoval, 2014). Para as observações diferentes de zero temos interesse nas duas estimativas, mas nas observações iguais a zero nosso interesse recai apenas na probabilidade da observação assumir o valor zero. Com isso, dependendo do valor que a observação assume, temos um determinado número de estimativas de interesse e isso dificulta a escolha de um único resíduo para a análise de diagnóstico desse modelo inflacionado. Além disso, neste trabalho nosso interesse recai nos modelos contínuos inflacionados no zero e nesses modelos os parâmetros da parte discreta são ortogonais e separáveis aos parâmetros da parte contínua do modelo, sugerindo assim uma análise de diagnóstico separada para cada parte do modelo (contínua e discreta).

Uma alternativa seria fazer a análise de diagnóstico utilizando um único resíduo e um candidato natural para isso seria o resíduo quantílico aleatorizado proposto por Dunn e Smyth (1996).

Definição 12. *O resíduo quantílico aleatorizado para os modelos inflacionados no zero são definidos como:*

$$q_i = \begin{cases} \Phi^{-1}(F(y_i; \hat{\mu}_i, \hat{\sigma}_i)), & \text{se } y_i > 0 \\ \Phi^{-1}(u_i), & \text{se } y_i = 0, \end{cases} \quad (3.21)$$

em que $\Phi()$ é a função de distribuição acumulada da normal padrão, u_i é uma variável aleatória com distribuição uniforme $U(0, \hat{\alpha}_i)$ e $\hat{\alpha}_i = P(\widehat{y_i = 0})$.

A aleatorização do resíduo quantílico tem a finalidade de produzir resíduos normais e contínuos. Porém, Pereira, Botter e Sandoval (2014) observaram que o resíduo quantílico

não consideraria discrepantes observações que assumem valores próximos de zero e tem alta ou moderada probabilidade de assumir valor zero. Como mencionado anteriormente isso não é razoável em muitas situações. Por exemplo, uma instituição financeira têm interesse em modelar o volume total de investimento de cada cliente e para isso utiliza um modelo de regressão inflacionado no zero. Se um cliente i tem altíssima renda é esperado que $\hat{\mu}_i$ seja razoavelmente grande. Assim se y_i apresentar valor baixo é razoável deduzir que essa observação i é uma observação discrepante. Entretanto se o valor de $\hat{\alpha}_i$ não for baixo (não for inferior a 2,5%, por exemplo), o resíduo quantílico aleatorizado dessa observação nunca seria grande, em módulo, não importando qual valor observado para $y_i > 0$. Nessa situação citada, é plausível presumir que $\hat{\alpha}_i$ não seja tão baixo, pois esse cliente i pode preferir concentrar seus investimentos em outras instituições financeiras. Suponha que $\hat{\alpha}_i = 0,1$, $\hat{\sigma} = 0,6$, $\hat{\mu}_i = 120000$ e $y \sim ZAGA$. Então,

$$P(0 < y < 1000) = 0,000005.$$

Se $y_i = 500$, usando q_i temos:

$$P(\widehat{y_i} < 500) = P(\widehat{y_i} = 0) + P(0 < \widehat{y_i} < 500) \approx 0,1.$$

Assim $q_i = \Phi^{-1}(0,1) = -1,28$. Logo a observação não seria considerada discrepante.

Portanto, nos modelos inflacionados no zero é mais interessante definir um resíduo de forma diferente para cada componente. A análise de diagnóstico do componente discreto do modelo pode ser feita utilizando os resíduos definidos para os modelos de regressão com respostas binárias (Hosmer Jr e Lemeshow, 2004). Discutiremos esse assunto na Seção 3.5.1. Para o componente contínuo poderíamos utilizar os resíduos definidos para os modelos não inflacionados. Por exemplo, para os modelos inflacionados no zero com resposta gama ou gaussiana inversa, poderíamos utilizar os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} e r_i^{wil} . Porém esses resíduos não levam em consideração uma estimativa para a probabilidade da observação assumir valores diferentes zero. Suponha por exemplo que as observações i e j em um modelo inflacionado no zero sejam positivas e seus resíduos são respectivamente $r_i = r_j = 2,5$, sendo r qualquer resíduo presente na literatura para esses modelos. Considerando essa abordagem na análise de diagnóstico, essas duas observações serão consideradas igualmente discrepantes independente do valor de $\hat{\alpha}_j$, que podem ser bem diferentes. A seguir, iremos definir duas classes de resíduos para avaliar o componente contínuo do modelo e que consideram a estimativa da probabilidade da observação assumir valor diferente de zero. Analisaremos as propriedades desses resíduos através de estudos de simulação de MC.

Ospina e Ferrari (2012) definiram um resíduo para os modelos com resposta beta inflacionada, a partir do algoritmo scoring de Fisher utilizado na obtenção dos estimadores de máxima verossimilhança do modelo, que é uma função da estimativa da probabilidade da observação assumir o valor diferente de zero. Baseado nesse resíduo definimos a seguinte

classe de resíduos inflacionados no zero.

Definição 13. *Considere r_i qualquer resíduo para um modelo de regressão com resposta contínua e $(1 - \hat{\alpha}_i)$ a estimativa da probabilidade da observação i assumir valor diferente de zero então definimos r_i^{**} como:*

$$r_i^{**} = \frac{r_i}{\sqrt{(1 - \hat{\alpha}_i)}}. \quad (3.22)$$

Note que r_i^{**} é uma classe de resíduos, pois para cada diferente r_i , temos um diferente resíduo.

Inspirados no resíduo quantílico r_i^{qu} , [Pereira, Botter e Sandoval \(2014\)](#) propuseram uma classe de resíduos para modelos inflacionados no zero que não apresenta os problemas apresentados pelo resíduo quantílico aleatorizado. A ideia para essa outra classe de resíduos é conseguir introduzir num mesmo resíduo, o resíduo r_i e a estimativa $\hat{\alpha}_i$. Porém, para conseguirmos comparar essas medidas, é necessário que ambas possam assumir os mesmos valores no intervalo real. Nesse caso, isso não ocorre, pois $r_i \in (0, \infty)$ e $\hat{\alpha}_i \in [0, 1]$. Para resolver esse problema, pode-se realizar uma transformação em r_i , para o mesmo pertencer ao intervalo $[0, 1]$, ou seja, aplica-se $\Phi(r_i)$, em que $\Phi(\cdot)$ é a função acumulada da distribuição normal. Desta forma temos $\Phi(r_i)$ e $\hat{\alpha}_i$, na mesma escala e as duas medidas são passíveis de comparação. Para gerar resíduos que têm distribuição próxima da normal padrão foi aplicada uma nova transformação em uma função de $\Phi(r_i)$ e $\hat{\alpha}_i$. Aplicamos $\Phi(\cdot)^{-1}$, mesma transformação usada em r_i^{qu} , para assim termos resíduos que tenham distribuições mais próximas da normal padrão. Esse resíduo é definido como:

$$r_i^* = \begin{cases} \Phi^{-1}[(1 - \Phi(|r_i|))(1 - \hat{\alpha}_i)], & \text{se } r_i < 0, \\ \Phi^{-1}[1 - (1 - \Phi(|r_i|))(1 - \hat{\alpha}_i)] & \text{se } r_i > 0. \end{cases} \quad (3.23)$$

A partir de simples manipulações algébricas, temos que: se $r_i < 0$, então

$$\begin{aligned} \Phi^{-1}[(1 - \Phi(|r_i|))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[(1 - \Phi(-r_i))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[(1 - (1 - \Phi(r_i)))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[(1 - 1 + \Phi(r_i))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[\Phi(r_i)(1 - \hat{\alpha}_i)]. \end{aligned}$$

E se $r_i > 0$, temos:

$$\begin{aligned} \Phi^{-1}[1 - (1 - \Phi(|r_i|))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[1 - (1 - \Phi(r_i))(1 - \hat{\alpha}_i)] &= \\ \Phi^{-1}[1 - (1 - \hat{\alpha}_i - \Phi(r_i) + \Phi(r_i)\hat{\alpha}_i)] &= \\ \Phi^{-1}[1 - 1 + \hat{\alpha}_i + \Phi(r_i) - \Phi(r_i)\hat{\alpha}_i] &= \\ \Phi^{-1}[\hat{\alpha}_i + \Phi(r_i)(1 - \hat{\alpha}_i)]. \end{aligned}$$

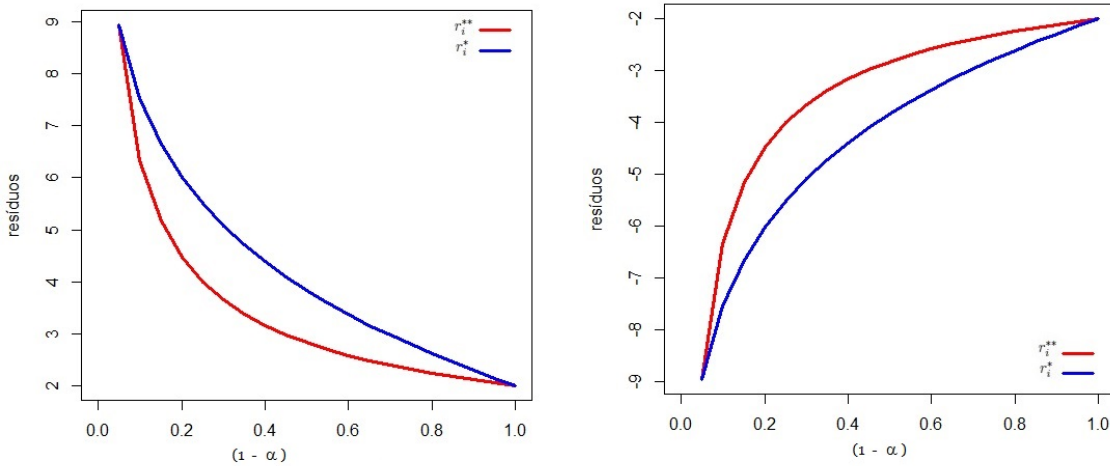
Então reescrevendo a equação (3.23), definimos r_i^* como:

Definição 14. A classe dos resíduos r_i^* é definida como

$$r_i^* = \begin{cases} \Phi^{-1}[\Phi(r_i)(1 - \hat{\alpha}_i)], & \text{se } r_i < 0, \\ \Phi^{-1}[\hat{\alpha}_i + \Phi(r_i)(1 - \hat{\alpha}_i)], & \text{se } r_i > 0. \end{cases} \quad (3.24)$$

Note que (3.24) também define uma classe de resíduos para observações positivas dos modelos de regressão inflacionados com resposta contínua, pois podemos usar como r_i qualquer resíduo para um modelo de regressão com resposta contínua.

Na Figura 3.1, temos os gráficos dos resíduos r_i^* e r_i^{**} , considerando diferentes valores de α e $r_i = 2$ (Figura 3.1(a)) e $r_i = -2$ (Figura 3.1(b)). Notamos que quando $(1 - \alpha_i)$ tende a 1, ou seja, $P(y_i > 0)$ tende a 1, tanto r_i^{**} quanto r_i^* tende ao resíduo r_i . E quanto menor $(1 - \alpha_i)$, ou seja, quando maior for a probabilidade da observação assumir o valor zero, maiores são, em módulo, os valores de r_i^* e r_i^{**} . Note ainda que fixado $1 - \alpha_i$, r_i^* é maior, em módulo, que r_i^{**} . A diferença entre os 2 resíduos é considerável para valores de $(1 - \alpha_i)$ nem tão próximos de zero e nem tão próximos de 1. A seguir apresentamos e demonstramos dois resultados para a classe de resíduos r_i^* .



(a) Gráficos dos resíduos r_i^{**} e r_i^* , considerando $r_i = 2$. (b) Gráficos dos resíduos r_i^{**} e r_i^* , considerando $r_i = -2$.

Figura 3.1: Gráfico dos resíduos para diferentes valores de α .

Resultado 1. Se α_i for conhecido para toda observação i e r_i tiver distribuição normal padrão e se na equação (3.24) considerarmos o valor real de α_i ao invés da estimativa $\hat{\alpha}_i$, temos que para todo $k > \Phi^{-1}(1 - 0,5(1 - \alpha_i)) = \Phi^{-1}(0,5 + 0,5\alpha_i)$,

$$P(r_i^* < -k) = P(r_i^* > k) = 1 - \Phi(k). \quad (3.25)$$

Demonstração.

$$\begin{aligned}
P(r_i^* < -k) &= \\
P(\Phi(r_i)(1 - \alpha_i) < \Phi(-k)) &= \\
P(r_i < \Phi^{-1}\left(\frac{\Phi(-k)}{(1-\alpha_i)}\right)) &= \\
\Phi\left[\Phi^{-1}\left(\frac{\Phi(-k)}{(1-\alpha_i)}\right)\right] (1 - \alpha_i) &= \\
\left[\frac{\Phi(-k)}{(1-\alpha_i)}(1 - \alpha_i)\right] &= 1 - \Phi(k).
\end{aligned} \tag{3.26}$$

□

A demonstração para a cauda superior é semelhante. O resíduo r_i^* é interessante, pois não importa o valor da variável preditora, se r_i possui exatamente distribuição normal padrão e α_i é conhecido, então a probabilidade de ocorrência de valores de r_i^* superiores a uma constante $k > \Phi^{-1}(0,5 + 0,5\alpha_i)$ é igual a probabilidade de uma variável com distribuição normal padrão assumir valores superiores a k . Assim, nas condições colocadas, r_i^* comporta-se como se tivesse distribuição normal. Porém, em geral, não conhecemos os valores desses parâmetros e r_i não tem distribuição exatamente normal. Assim, [Pereira, Botter e Sandoval \(2014\)](#) utilizam os estudos de simulação de MC para estudar o comportamento do resíduo r_i^* e fizeram isso para o modelo de regressão beta inflacionada no zero. Os autores chegaram a conclusão que o resíduo proposto apresenta boas propriedades e resultados melhores do que o resíduo proposto por [Ospina e Ferrari \(2012\)](#). A seguir apresentamos outro resultado para essa classe de resíduo.

Resultado 2. Se $r_i > 0$ e r_i é o resíduo q_i , temos:

$$r_i^* = q_i. \tag{3.27}$$

Demonstração.

$$\begin{aligned}
r_i^* &= \Phi^{-1}\left[\hat{\alpha}_i + \Phi\left[\Phi^{-1}(P(Y_i \leq \widehat{y_i} | Y_i > 0))\right] (1 - \hat{\alpha}_i)\right] \\
&= \Phi^{-1}\left[\hat{\alpha}_i + P(Y_i \leq \widehat{y_i} | Y_i > 0)(1 - \hat{\alpha}_i)\right] \\
&= \Phi^{-1}\left[\hat{\alpha}_i + \frac{P(0 < Y_i \leq y_i)}{P(y_i > 0)}(1 - \hat{\alpha}_i)\right] \\
&= \Phi^{-1}\left[\hat{\alpha}_i + \frac{P(0 < Y_i \leq y_i)}{(1 - \hat{\alpha}_i)}(1 - \hat{\alpha}_i)\right] \\
&= \Phi^{-1}\left[P(\widehat{y_i = 0})_i + P(0 < \widehat{Y_i} \leq y_i)\right] \\
&= \Phi^{-1}\left[P(\widehat{Y_i} \leq y_i)\right] \\
&= \Phi^{-1}\left[F(y_i, \hat{\mu}_i, \hat{\sigma}_i)\right] = q_i.
\end{aligned} \tag{3.28}$$

□

Isso significa que, se considerarmos o resíduo quantílico como r_i e se $r_i > 0$, então r_i^* é igual ao resíduo quantílico aleatorizado. Assim, r_i^* pode ser visto como uma correção do resíduo quantílico aleatorizado quando o valor de r_i é negativo.

Os resíduos mais utilizados nos modelos de regressão gama e gaussiana inversa não são os mesmos do que os resíduos nos modelos de regressão beta. Por esta razão, é importante avaliar essas classes de resíduos nos modelos com resposta gama e gaussiana inversa inflacionados no zero, usando como r_i os resíduos comuns em modelos lineares generalizados como r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} e r_i^{wil} definidos na seção anterior.

3.4 Estudos de simulação para os modelos inflacionados no zero

Utilizamos estudos de simulação de Monte Carlo para analisarmos e compararmos duas classes de resíduos para modelos inflacionados no zero: a classe proposta a partir do resíduo proposto por [Ospina e Ferrari \(2012\)](#), r_i^{**} e a classe proposta por [Pereira, Botter e Sandoval \(2014\)](#), r_i^* . Foram utilizados diversos cenários para analisar a performance desses resíduos e para todos os cenários os resíduos r_i^* e r_i^{**} foram calculados usando os resíduos mais utilizados nesses modelos, ou seja, usamos o resíduo desvio (r_i^{dev}), resíduo de pearson (r_i^{pea}), resíduo quantílico (r_i^{qu}), o resíduo anscombe (r_i^{asc}) e resíduo william (r_i^{wil}) como r_i . Os estudos de simulação foram realizados utilizando o pacote "gamlss" do software estatístico R ([Stasinopoulos et al., 2014](#)) e para cada cenário consideramos $M = 5000$ réplicas de MC. Neste trabalho consideramos os modelos de regressão ZAGA e ZAIG para o estudo de MC.

Após o ajuste, para cada resíduo e para cada uma das n unidades amostrais consideradas nesse trabalho, calculou-se entre as $M = 5000$ réplicas de MC, a porcentagem dos resíduos que foram inferiores -3 , -2 e -1 e superiores a 1 , 2 e 3 . Por exemplo, para a primeira observação das 5000 réplicas de MC calculou-se r_i^* (ou r_i^{**} , após o calculo contamos entre as 5000 réplicas de MC ($\#_j r_{1j}$) aqueles resíduos maiores que 3. Isso foi feito até obter a contagem dos resíduos para todos os intervalos (3,2,1,-1,-2,-3). Na Tabela 3.6, apresentamos a estrutura das tabelas apresentadas nos apêndices D e E.

Neste trabalho consideramos como um bom resíduo aquele que possui essas porcentagens próximas dos respectivos valores teóricos da distribuição normal padrão. Temos interesse em resíduos que possuem propriedades próximas da normal padrão, pois assim teremos resíduos com propriedades conhecidas e teremos gráficos de resíduos que serão interpretados mais facilmente. Por exemplo, iremos considerar r_i^* como um bom resíduo, se para todas as n observações, a porcentagem dos resíduos superiores a 2 esteja próxima de $1 - \Phi(2) = 2.28\%$. Se $y_i = 0$, r_i^* não é definido. Nesse caso, assumimos que r_i^* não é maior que k e nem menor que $-k$, para qualquer $k \in R^+$. O mesmo procedimento é usado para r_i^{**} . Como para diversas observações os resíduos em estudo não estão definidos, a abordagem utilizadas nos estudos de simulação dos MLGs não se aplica aqui. Por isso, optamos em analisar os resultados a partir da Tabela 3.6, em que apresentamos o mínimo, o máximo e os quartis das porcentagens de resíduos superiores e inferiores a determinados valores. Nas próximas seções iremos apresentar os resultados dos estudos de simulação desses resíduos para os modelos

Tabela 3.6: Estrutura das tabelas apresentadas nos apêndices *D* e *E*

replica \ resíduo	r_1^*	r_2^*	r_3^*	...	r_{15}^*
1	r_{11}^*	r_{12}^*	r_{13}^*	...	r_{115}^*
2	r_{21}^*	r_{22}^*	r_{23}^*	...	r_{215}^*
3	r_{31}^*	r_{32}^*	r_{33}^*	...	r_{315}^*
4	r_{41}^*	r_{42}^*	r_{43}^*	...	r_{415}^*
⋮	⋮	⋮	⋮	...	⋮
4999	r_{49991}^*	r_{49992}^*	r_{49993}^*	...	r_{499915}^*
5000	r_{50001}^*	r_{50002}^*	r_{50003}^*	...	r_{500015}^*
$\#_j r_{ij} > 3/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$
$\#_j r_{ij} > 2/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$
$\#_j r_{ij} > 1/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$
$\#_j r_{ij} < -1/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$
$\#_j r_{ij} < -2/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$
$\#_j r_{ij} < -3/5000^*$	$\#_j r_{1j}/5000$	$\#_j r_{2j}/5000$	$\#_j r_{3j}/5000$...	$\#_j r_{15j}/5000$

* Para algumas observações i temos que y_i assume valor zero. Assim os resíduos r_i^* e r_i^{**} não estão definidos e por isso consideramos que r_i^* e r_i^{**} , para qualquer k positivo, não são nem superiores a k e nem inferiores a $-k$.

de regressão ZAGA e ZAIG.

3.4.1 Estudos de simulação para o modelo ZAGA

No estudo de simulação utilizamos o modelo com resposta gama inflacionado no zero e consideramos os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} , r_i^{wil} como r_i . Foram considerados diversos cenários e em cada cenário utilizou-se $M = 5000$ réplicas de MC.

No cenário 1, consideramos a seguinte estrutura para modelo ZAGA:

$$\begin{aligned}
 \log\{\mu_i\} &= \beta_{11} + \beta_{21}x_{i2} + \beta_{31}x_{i3} \\
 \log\{\sigma_i\} &= \beta_{12} \\
 \log\left\{\frac{\alpha_i}{1-\alpha_i}\right\} &= \beta_{13} + \beta_{23}x_{i2} + \beta_{33}x_{i3}.
 \end{aligned} \tag{3.29}$$

As variáveis preditoras utilizadas no cenário 1 foram geradas independentemente a partir de variáveis aleatórias com distribuição uniforme no intervalo $(0; 1)$ e mantidas constantes em todas as $M = 5000$ replicações. Consideramos uma amostra de tamanho $n = 100$ e utilizou-se os seguintes coeficientes nesse cenário, $\beta_{11} = 5,0$, $\beta_{21} = -2,0$, $\beta_{31} = 1$, $\beta_{12} = -2,3$, $\beta_{13} = -0,5$, $\beta_{23} = 1,0$ e $\beta_{33} = -0,5$. Estabelecendo esses valores dos coeficientes, no cenário 1 temos que os intervalos de variação dos parâmetros associados a distribuição gama inflacionada no zero resultam em $\mu \in (20,9; 403,4)$, $\sigma = 0,1$ e $\alpha \in (0,27; 0,62)$. Os resultados são apresentados na Tabela D.1.

Nota-se nesse cenário, que os resíduos r_i^* são os resíduos que apresentam os valores mais próximos dos valores teóricos da normal padrão, enquanto os resíduos r_i^{**} apresentam

valores longe dos valores desejáveis. Entre todos os resíduos r_i^* , o resíduo r_i^{*qu2} é aquele que apresentam resultados mais próximos dos valores teóricos da normal padrão. O resíduo r_i^{*qu} não somente é melhor que os outros resíduos, mas é aquele que possui os mínimos e máximos mais próximos dos valores teóricos. Por exemplo, na Tabela D.1, se considerarmos o intervalo dos resíduos maiores que 2, o valor teórico da normal padrão é de 2,28% e a porcentagem observada dos r_i^{*qu} varia de 1,74 a 2,92.

Diversos outros cenários foram considerados nos estudos de simulação, variando-se, em relação ao cenário 1, ou o tamanho amostral, ou o valor dos parâmetros, ou as variáveis preditoras, ou a função de ligação, ou a estrutura do modelo para o parâmetro σ . Na Tabela 3.7, apresentamos as alterações feitas nos cenários em relação ao cenário 1. No cenário 2 variamos os valores de α , sendo que no cenário 2a diminuimos o valor do parâmetro α , considerando os coeficientes iguais a; $\beta_{13} = -1,5$, $\beta_{23} = 1,0$ e $\beta_{33} = -0,5$ resultando em $\alpha \in (0,11; 0,38)$. No cenário 2b aumentamos o valor do parâmetro α , obtendo-se $\alpha \in (0,50; 0,82)$, considerando como coeficientes β_{13}, β_{23} e β_{33} iguais a 0,5, 1 e $-0,5$, respectivamente. Ao diminuirmos o valor do parâmetro α (Tabela D.2), temos uma ligeira piora nos resultados dos resíduos r_i^* e uma melhora nos resíduos r_i^{**} em relação ao cenário 1. Ou seja, em relação ao cenário 1, temos que no cenário 2a a média da porcentagem observada dos resíduos r_i^* não é tão próxima do valor teórico da normal padrão e a amplitude dos intervalos é um pouco maior quando comparada ao cenário 1. Já os resíduos r_i^{**} têm uma melhora na média da porcentagem observada dos resíduos, ficando mais próxima da normal padrão e a amplitude dos intervalos um pouco menor quando comparada ao cenário 1. Porém, apesar da melhora dos resíduos r_i^{**} , temos que os resíduos r_i^* ainda apresentam a porcentagem observada dos resíduos mais próxima dos valores teóricos da normal padrão. E entre todos os r_i^* , r_i^{*qu} é aquele que apresenta os melhores resultados.

Tabela 3.7: Alterações feitas nos cenários em relação ao cenário 1

Cenário	Alterações
2a	Diminuimos o valor de α
2b	Aumentamos o valor de α
3	Valores de μ próximo de zero
4a	Diminuimos o valor de σ
4b	Aumentamos o valor de σ
5	Mudança no modelo para σ
6a	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim IG(0,4; 2)$
6b	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$
7a	Mudança na função de ligação para o parâmetro μ : função de ligação inversa
7b	Mudança na função de ligação para o parâmetro α : função de ligação proibido
8a	Mudança no tamanho da amostra: n=200
8b	Mudança no tamanho da amostra: n=50

Quando aumentamos o valor do parâmetro α (cenário 2b, Tabela D.3), os resíduos r_i^{**}

² r_i^{*qu} é o resíduo r_i^* considerando o resíduo r_i^{qu} como r_i .

pioram bastante tendo valores da porcentagem observada dos resíduos bem afastadas do valor teórico. Os resíduos r_i^* , têm uma piora nos resultados também, mas mantêm-se num patamar aceitável e bem melhor do que os resíduos r_i^{**} . Ou seja, apesar da piora os resíduos r_i^* quando comparado aos resultados do cenário 1, os resíduos r_i^* apresentam a porcentagem observada dos resíduos próximos dos valores teóricos. E novamente o resíduo r_i^{*qu} foi aquele que apresentou os melhores resultados entre todos os r_i^* .

No próximo cenário (cenário 3), diminuimos o intervalo do parâmetro μ para próximo de zero, ou seja, consideramos $\beta_{11} = 5,0$, $\beta_{21} = -2,0$, $\beta_{31} = 1$ que resultou em $\mu \in (0,48; 6,96)^3$. Os resultados apresentados na Tabela D.4, sugerem que a performance r_i^{*qu} continua consideravelmente melhor do que a dos demais resíduos considerados e não muda de forma significativa em relação ao observado no cenário 1.

No cenário 4, variamos o valor do parâmetro σ , as Tabelas D.5 e D.6 apresentam os resultados dos cenários 4a e 4b, respectivamente. No primeiro diminuimos o valor de σ , considerando $\beta_{12} = -3,0$ que resultou em $\sigma = 0,05$ e no segundo aumentamos o valor de σ considerando $\beta_{12} = -0,69$ que resultou em $\sigma = 0,5$. A mudança no parâmetro σ , parece não afetar os resultados dos estudos, sendo muito próximos aos resultados apresentados no cenário 1, ou seja, a performance do resíduo r_i^{*qu} é melhor do que os demais considerados.

A seguir apresentamos os resultados do cenário 5 (Tabela D.7), em que propomos um modelo diferente para σ . Aqui consideramos a seguinte estrutura para o modelo ZAGA:

$$\log\{\mu_i\} = \beta_{11} + \beta_{21}x_{i2} + \beta_{31}x_{i3} \quad (3.30)$$

$$\log\{\sigma_i\} = \beta_{12} + \beta_{22}x_{i2} + \beta_{32}x_{i3} \quad (3.31)$$

$$\log\left\{\frac{\alpha_i}{1 - \alpha_i}\right\} = \beta_{13} + \beta_{23}x_{i2} + \beta_{33}x_{i3}. \quad (3.32)$$

Nota-se que ao mudarmos o modelo para o parâmetro σ , houve uma piora em todos os resíduos e o resíduo r_i^{*qu} foi o que apresentou os valores mais próximos da normal padrão e o que sofreu menos influência nessa mudança.

No próximo cenário (cenário 6), mudamos a distribuição das variáveis preditoras. No cenário 6a (Tabela D.8) consideramos $x_{i1} \sim NO(0,5; 0,25^2)$ e $x_{i2} \sim IG(0,4; 2)$ e no cenário 6b (Tabela D.9) consideramos $x_{i1} \sim NO(0,5; 0,25^2)$ e $x_{i2} \sim GA(0,4; 1)$. A mudança na função de distribuição das variáveis preditoras não afeta os resultados relativos dos resíduos, sendo esses próximos dos apresentados no cenário 1. Porém a mudança parece aumentar a amplitude dos intervalos.

A Tabela D.10 apresenta os resultados do cenário 7a, onde alteramos a função de ligação do parâmetro μ da função ligação log para a função de ligação identidade. Notamos que a média e mediana são próximas dos valores teóricos, porém observamos um pequeno aumento na amplitude dos intervalos em relação aos resultados do cenário 1. Entre todos os resíduos

³Podemos considerar 6,96 próximo de zero, pois no modelo ZAGA o suporte da distribuição está definido no intervalo $[0, \infty)$

r_i^{*qu} continua sendo o resíduo que apresenta propriedades próximas da distribuição normal padrão. E na Tabela D.11, cenário 7b, alteramos a função de ligação do parâmetro α da logito para a função de ligação probito. Tal mudança não afeta significativamente os resultados, onde r_i^{*qu} continua sendo o melhor resíduo entre todos e os resultados são bem próximos dos resultados do cenário 1.

No cenário 8 variou-se o tamanho amostral. Em um modelo de regressão inflacionado no zero com componente contínuo, deve haver necessariamente no mínimo algumas observações iguais a zero e algumas observações positivas. Entretanto, em estudos de simulação, se considerar tamanhos amostrais muito pequenos pode ocorrer de, em alguma replicação, não haver nenhuma observação igual a zero ou nenhuma observação positiva. Por isso, aqui iremos considerar apenas amostras de tamanho pelo menos moderado. No cenário 8 manteve-se assim os valores dos parâmetros e considerou-se 2 diferentes tamanhos amostrais: $n = 200$ (cenário 8b, Tabela D.12) e $n = 50$ (cenário 8b, Tabela D.13). Pode-se notar que, em relação ao tamanho amostral igual a 100, a performance de r_i^{*qu} piora quando $n = 50$, mas a performance continua sendo aceitável e bem melhor que a dos demais resíduos. Já para $n = 200$, r_i^{*qu} continua sendo o melhor resíduo, mas sua performance não melhora muito em relação ao cenário 1.

Resumindo os resultados dos cenários considerados nos estudos de simulação, tem-se que r_i^{*qu} apresenta boa performance para a realização de análise de diagnóstico do componente contínuo de modelos de regressão gama inflacionados no zero e superior à observada para os demais resíduos propostos. Assim, os resultados dos estudos de simulação sugerem que r_i^{*qu} é adequado para a realização de análise de diagnóstico do componente contínuo de modelos de regressão gama inflacionados no zero.

3.4.2 Estudos de simulação para o modelo ZAIG

Para analisar o desempenho dos resíduos r_i^* e r_i^{**} nos estudos de simulação de MC, utilizou-se o modelo com resposta gaussiana inversa inflacionada no zero e considerou-se os resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{ans} e r_i^{wil} como os resíduos r_i . Foram considerados diversos cenários nos estudos e em cada cenário utilizou-se $M = 5000$ réplicas de MC.

No cenário 1, consideramos a estrutura do modelo ZAIG apresentada abaixo:

$$\log\{\mu_i\} = \beta_{11} + \beta_{21}x_{i2} + \beta_{31}x_{i3} \quad (3.33)$$

$$\log\{\sigma_i\} = \beta_{12} \quad (3.34)$$

$$\log\left\{\frac{\alpha_i}{1 - \alpha_i}\right\} = \beta_{13} + \beta_{23}x_{i2} + \beta_{33}x_{i3}. \quad (3.35)$$

No cenário 1 as variáveis preditoras utilizadas foram geradas independentemente, a partir de variáveis aleatórias com distribuição uniforme padrão, $x_i \sim U(0; 1)$ e foram mantidas constantes em todas as $M = 5000$ replicações. Uma amostra de tamanho $n = 100$ foi considerada e utilizou-se os seguintes coeficientes nesse cenário, $\beta_{11} = 5,0$, $\beta_{21} = -2,0$, $\beta_{31} =$

1, $\beta_{12} = -3,91$, $\beta_{13} = -0,5$, $\beta_{23} = 1,0$ e $\beta_{33} = -0,5$. Estabelecido esses critérios e valores dos coeficientes, no cenário 1 temos que os intervalos dos parâmetros associados à distribuição gaussiana inversa inflacionada no zero são $\mu \in (20,9; 403,4)$, $\sigma = 0.02$ e $\alpha \in (0,27; 0,62)$. A Tabela E.1, fornece os resultados do cenário 1.

Analisando os resultados da simulação para o cenário 1, temos que os resíduos r_i^{**} apresentam valores longe dos valores esperados, ou seja, os valores dos resíduos r_i^{**} estão longe dos valores teóricos da normal padrão, enquanto os resíduos r_i^* apresenta os valores mais próximos dos valores teóricos da normal padrão. Entre todos os resíduos r_i^* , o resíduo r_i^{*qu} é aquele que apresentam resultados mais próximos dos valores teóricos da normal padrão. A média da porcentagem observada do resíduo r_i^{*qu} é bem próxima do valor teórico da normal padrão para todos os intervalos. A amplitude dos intervalos do resíduo r_i^{*qu} é a menor entre todos os resíduos, sugerindo que a porcentagem observada do resíduo r_i^{*qu} é bem próxima dos valores teóricos da distribuição normal padrão.

Consideramos vários outros cenários nos estudos de simulação, onde variou-se, em relação ao cenário 1, ou o tamanho amostral, ou o valor dos parâmetros, ou as variáveis preditoras, ou a função de ligação, ou a estrutura do modelo para o parâmetro σ . Na Tabela 3.8, temos um resumo das alterações feita em cada cenário em relação ao cenário1. No cenário 2 variamos os valores do parâmetro α . Diminuímos o valor de α no cenário 2a, consideramos os coeficientes β_{13}, β_{23} e β_{33} iguais a $-1,5, 1,0$ e $-0,5$ respectivamente, resultando em $\alpha \in (0,12; 0,38)$. E aumentamos o valor de α no cenário 2b, para $\alpha \in (0,50; 0,82)$, considerando os coeficientes iguais a $\beta_{13} = 0,5, 0, \beta_{23} = 1,0$ e $\beta_{33} = -0,5$. Ao diminuirmos o valor do parâmetro α (Tabela E.2), os resíduos r_i^{**} têm uma melhora e os resíduos r_i^* tem uma ligeira piora nos resultados em relação ao cenário 1. Ou seja, as medidas descritivas da porcentagem observada para os resíduos r_i^{**} ficaram mais próximas dos valores teóricos da normal padrão no cenário 2a do que no cenário 1. Já as medidas descritivas da porcentagem observada para os resíduos r_i^* , no cenário 2a, não ficaram tão próximos dos valores teóricos da normal padrão quando comparadas aos resultados do cenário 1. Porém, apesar da melhora dos resíduos r_i^{**} , os resíduos r_i^* continuam melhores que r_i^{**} e r_i^{*qu} é aquele que apresenta os melhores resultados entre todos.

O aumento do valor do parâmetro α (cenário 2b, Tabela E.3), provoca uma piora em todos os resíduos. Os resíduos r_i^{**} pioram bastantes tendo valores da porcentagem observada dos resíduos bem afastadas do valor teórico. Os resíduos r_i^* têm uma piora nos resultados também, mas mantém-se num patamar aceitável e bem melhor do que os resíduos r_i^{**} . Ou seja, apesar da piora quando comparada aos resultados do cenário 1, os resíduos r_i^* apresentam medidas descritivas próximos do valor teórico da normal padrão. E novamente r_i^{*qu} foi aquele que apresentou os melhores resultados entre todos os resíduos.

No cenário 3, alteramos o intervalo do parâmetro μ para próximo de zero, ou seja, consideramos $\beta_{11} = 1, \beta_{21} = -2,0, \beta_{31} = 1$ que resultou em $\mu \in (0,37; 7,39)^4$. Os resultados

⁴A distribuição ZAIG têm o suporte definido em $[0, \infty)$. Assim 7,39 foi considerado um valor próximo de zero, pois nessa distribuição espera-se valores bem altos.

Tabela 3.8: Alterações feitas nos cenários em relação ao cenário 1

Cenário	Alterações
2a	Diminuímos o valor de α
2b	Aumentamos o valor de α
3	Valores de μ próximo de zero
4a	Diminuímos o valor de σ
4b	Aumentamos o valor de σ
5	Mudança no modelo para σ
6a	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim IG(0,4; 2)$
6b	Mudança na variáveis preditoras: $x_1 \sim N(0,5; 0,25^2)$ e $x_2 \sim GA(0,4; 1)$
7a	Mudança na função de ligação para o parâmetro μ : função de ligação invesa
7b	Mudança na função de ligação para o parâmetro α : função de ligação proibito
8a	Mudança no tamanho da amostra: n=200
8b	Mudança no tamanho da amostra: n=50

(Tabela E.4), sugerem que o desempenho de r_i^{*qu} continua consideravelmente melhor do que a dos demais resíduos considerados e não temos alterações significativa em relação aos resultados observado no cenário 1.

No próximo cenário (cenário 4), variamos o valor do parâmetro σ , sendo que os resultados dos cenários 4a e 4b estão apresentados nas tabelas E.5 e E.6 respectivamente. No primeiro diminuímos o valor de σ , considerando $\beta_{12} = -4,60$ que resultou em $\sigma = 0,01$. No segundo aumentamos o valor de σ considerando $\beta_{12} = -3,50$ que resultou em $\sigma = 0,03$. A mudança no parâmetro σ , parece não afetar os resultados dos estudos, sendo muito próximos aos resultados apresentados no cenário 1, ou seja, a performance do resíduo r_i^{*qu} é melhor do que os demais considerados.

A seguir (cenário 5, Tabela E.7), propomos um modelo diferente para σ , considerando a seguinte estrutura para o modelo ZAIG:

$$\log\{\mu_i\} = \beta_{11} + \beta_{21}x_{i2} + \beta_{31}x_{i3} \quad (3.36)$$

$$\log\{\sigma_i\} = \beta_{12} + \beta_{22}x_{i2} + \beta_{32}x_{i3} \quad (3.37)$$

$$\log\left\{\frac{\alpha_i}{1 - \alpha_i}\right\} = \beta_{13} + \beta_{23}x_{i2} + \beta_{33}x_{i3}. \quad (3.38)$$

Ao alterarmos a estrutura do modelo para o parâmetro σ , todos os resíduos pioraram. O resíduo r_i^{*qu} foi o que sofreu menos influência nessa mudança, seus intervalos continuam em um patamar aceitável e com os valores mais próximos da normal padrão.

No próximo cenário (cenário 6), mudamos a distribuição que gera as variáveis preditoras. Consideramos $x_{i1} \sim NO(0,5; 0,25^2)$ e $x_{i2} \sim IG(0,4; 2)$ (cenário 6a, Tabela E.8) e $x_{i1} \sim NO(0,5; 0,25^2)$ e $x_{i2} \sim GA(0,4; 1)$ (cenário 6b, Tabela E.9). Os resultados sugerem que a mudança na distribuição das variáveis preditoras não afeta os resíduos, sendo os resultados muitos próximos dos apresentados no cenário 1.

A mudança na função de ligação logito para a função de ligação proibito, relativa ao

parâmetro α (cenário 7b, Tabela E.11), não afeta significativamente os resultados, em que r_i^{*qu} continua sendo o melhor resíduo entre todos e os resultados são bem próximos dos resultados do cenário 1. Já, na mudança da função de ligação log para a função de ligação identidade, cenário 7a (Tabela E.10), temos um aumento na amplitude dos intervalos. Por exemplo, no cenário 1 (Tabela E.1), a porcentagem dos resíduos menores que < -2 variam entre 1,52 a 2,90 (mínimo e máximo respectivamente), enquanto que no cenário 7a, temos uma variação de 0,48 a 2,72. Porém essa variação está em um patamar aceitável e para o cenário 7a, entre todos os resíduos, o resíduo r_i^{*qu} continua sendo o melhor entre todos os outros resíduos.

As tabelas E.12 e E.13 apresentam os resultados dos cenários 8a e 8b respectivamente. Nesses cenários, variou-se o tamanho amostral considerando-se apenas amostras de tamanho pelo menos moderado ($n=50$)⁵ e manteve-se assim os valores dos parâmetros definidos no cenário 1. Consideraram-se diferentes tamanhos amostrais: $n = 200$ (cenário 8a, Tabela E.12) e $n = 50$ (cenário 8b, Tabela E.13). A performance de r_i^{*qu} piora quando $n = 50$, mas é aceitável e bem melhor que a dos demais resíduos. Já para $n = 200$, r_i^{*qu} continua sendo o melhor resíduo, mas sua performance não melhora muito em relação ao cenário 1.

Portanto, os resultados dos cenários considerados nos estudos de simulação, sugere que r_i^{*qu} apresenta boa performance para a realização de análise de diagnóstico do componente contínuo de modelos de regressão gaussiana inversa inflacionados no zero e é superior à observada para os demais resíduos propostos. Assim, os resultados dos estudos de simulação sugerem que r_i^{*qu} é adequado para a realização de análise de diagnóstico do componente contínuo de modelos de regressão gaussiana inversa inflacionados no zero. Vale ressaltar que os resultados da performance dos resíduos foram muito próximas nas duas distribuições ZAGA e ZAIG, sugerindo que, entre essas duas, a distribuição da variável resposta não influencia na performance dos resíduos. Observa-se ainda que nos estudos de simulação com resposta ZAIG tivemos um aumento da amplitude de cada intervalo, quando comparado à amplitude dos intervalos dos estudos de simulação dos modelos com resposta ZAGA. Por exemplo, no cenário 1 do modelo ZAIG (Tabela E.1), para a porcentagem dos resíduos r_i^{*qu} maiores que < -1 temos que o intervalo varia de 9,68 a 15,26 (mínimo e máximo respectivamente) e no cenário 1 do modelo ZAGA temos uma variação de 14,44 a 17,78. Provavelmente isso se dá pelo fato da distribuição gaussiana inversa ser mais assimétrica que a distribuição gama.

3.4.3 Estudos de simulação para o caso do ajuste do modelo incorreto

Nesta seção temos interesse em analisar as propriedades dos resíduos propostos na Seção 3.3.1 quando ajustamos uma distribuição errada para a variável resposta. Para isso utilizamos

⁵Nos modelos inflacionados no zero temos que considerar uma amostra de tamanho moderado, pois se pegarmos uma amostra pequena pode ocorrer da amostra não ter observação com valor zero.

simulação de Monte Carlo e calculamos os resíduos r_i^* e r_i^{**} usando o resíduo desvio (r_i^{dev}), resíduo de pearson (r_i^{pea}), resíduo quantílico (r_i^{qu}), o resíduo anscombe (r_i^{asc}) e resíduo william (r_i^{wil}) como r_i . Os estudos de simulação foram realizados utilizando o pacote "gamlss" do software estatístico R (Stasinopoulos et al., 2014) e para cada cenário consideramos $M = 5000$ réplicas de MC. No primeiro cenário, Tabela F.1, geramos variáveis respostas com distribuição ZAIG e ajustamos o modelo de regressão ZAGA. E no cenário II, Tabela F.2, geramos a variável resposta com distribuição ZAGA e ajustamos o modelo ZAIG.

Em ambos os cenários propostos foram usadas as seguintes estruturas para os modelos:

$$\log\{\mu_i\} = \beta_{11} + \beta_{21}x_{i2} + \beta_{31}x_{i3} \quad (3.39)$$

$$\log\{\sigma_i\} = \beta_{12} \quad (3.40)$$

$$\log\left\{\frac{\alpha_i}{1 - \alpha_i}\right\} = \beta_{13} + \beta_{23}x_{i2} + \beta_{33}x_{i3}. \quad (3.41)$$

As variáveis preditoras utilizadas nos cenários foram geradas independentemente a partir de variáveis aleatórias com distribuição uniforme no intervalo $(0; 1)$ e mantidas constantes em todas as $M = 5000$ réplicas de MC. Nos cenários I e II (Tabela F.1 e F.2), respectivamente), consideramos $\beta_{11} = 5,0$, $\beta_{21} = -2,0$, $\beta_{31} = 1$, $\beta_{12} = -3,91$, $\beta_{13} = -0,5$, $\beta_{23} = 1,0$ e $\beta_{33} = -0,5$. Para o cenário I, consideramos $\sigma = 0,02$ e para o cenário II temos $\sigma = 0,1$. Em ambos os cenários, notamos que a porcentagem mediana observada dos resíduos estão bem afastadas dos seus respectivos valores teóricos. Para o resíduo r_i^{*qu} , por exemplo, no cenário 1, quando o ajuste do modelo estava correto, a porcentagem mediana observada dos resíduos inferiores a 2 era 2,32% no modelo ZAGA (Tabela D.1) e 2,28% no modelo ZAIG (Tabela E.1), ambos bem próximos do valor teórico de 2,28%. Já, no ajuste do modelo incorreto, a porcentagem mediana observada de resíduos inferiores a 2 é de 2,45% quando ajustamos o modelo ZAIG e a distribuição correta é a ZAGA (Tabela F.1). E é de 0,98% quando a distribuição correta é a ZAIG e ajustamos o modelo ZAGA (Tabela F.2). Também nota-se que a amplitude em todos os intervalos aumenta consideravelmente quando ajustamos o modelo incorreto, ou seja, se a distribuição não está adequada a porcentagem observada dos resíduos se afastam dos valores teóricos da normal padrão. Com isso concluímos que mesmo os resíduos r_i^* só é adequado para detectar observações discrepantes se a distribuição da variável resposta é a correta. Dessa forma, é importante avaliarmos o ajuste do modelo quanto à adequação da distribuição de y_i e isso pode ser feito, por exemplo, a partir de gráficos do tipo envelope que serão discutidos na próxima seção (Seção 3.5.3).

3.5 Outros aspectos da análise de diagnóstico

Como dito anteriormente, na análise de diagnóstico diferentes aspectos devem ser avaliados. Podemos investigar se existem problemas na função de ligação adotada, ou na distribuição da variável resposta adotada, ou ainda se existem observações discrepantes ou influentes.

A seguir discutiremos algumas técnicas para a análise de diagnóstico em modelos ZAGA e ZAIG.

O vetor de parâmetros γ é separável e ortogonal a β e δ . Além disso γ está relacionado apenas com o componente discreto e β e δ estão relacionados com o componente contínuo. Portanto, com exceção da identificação de observações positivas discrepantes discutidas anteriormente (Seção 3.3.1), é razoável realizar duas análises de diagnóstico separadas nos modelos de regressão ZAGA e ZAIG: uma para o componente contínuo e outra para o componente discreto. Assim, podemos utilizar técnicas de diagnóstico usadas nos modelos com resposta gama e gaussiana inversa para o componente contínuo e técnicas de diagnóstico utilizadas nos modelos com resposta binária para o componente discreto. Nas próximas seções definiremos resíduos para o componente discreto (Seção 3.5.1), técnicas de identificação de pontos influentes para o componente contínuo e discreto (Seção 3.5.2) e avaliação do ajuste dos modelos ZAGA e ZAIG (Seção 3.5.3).

3.5.1 Observações discrepantes - componente discreto

Na Seção 3.3.1, definimos resíduos para a identificação de observações positivas discrepantes nos modelos com resposta ZAGA e ZAIG. Nessa seção discutimos os resíduos para o componente discreto dos modelos ZAGA e ZAIG. Na análise de diagnóstico do componente discreto, iremos utilizar os resíduos definidos para o componente contínuo (r_i^{dev} , r_i^{pea} , r_i^{wil}) e apresentaremos a expressão desses resíduos para o componente discreto, exceto o resíduo quantílico e anscombe⁶. Esses resíduos irão ajudar na identificação de observações discrepantes em relação ao componente discreto do modelo.

Em relação ao componente discreto dos modelos, a variável resposta dos modelos ZAGA e ZAIG pode ser vista como uma variável binária que assume para a observação i o valor 0 se $y_i = 0$ e 1 se $y_i > 0$. Assim, podemos utilizar os resíduos de modelos lineares generalizados com as expressões obtidas para os casos de resposta binárias. A partir da equação (3.2), obtemos que o componente do desvio para o componente discreto do modelo é dado por:

$$d(y_i, \hat{\alpha}_i) = \begin{cases} -|2\log(\hat{\alpha}_i)|^{1/2}, & \text{se } y = 0, \\ +|2\log(1 - \hat{\alpha}_i)|^{1/2}, & \text{se } y > 0. \end{cases} \quad (3.42)$$

Definição 15. *Utilizando o componente do desvio definido na equação (3.2), definimos o resíduo desvio padronizado como*

$$r_i^{dev(0)} = \frac{d(y_i, \hat{\alpha}_i)}{\sqrt{1 - \hat{h}_{ii}^0}}, \quad (3.43)$$

⁶O resíduo quantílico, r_i^{qu} , não está definido para o componente discreto do modelo e o resíduo de anscombe não tem uma forma fechada para o caso discreto. O resíduo quantílico aleatorizado é uma extensão do resíduo quantílico e está definido para o componente discreto. Porém, o resíduo quantílico aleatorizado não é útil para a identificação de observação discrepantes, sendo adequado apenas para avaliar o ajuste do modelo.

em que $\hat{h}_{ii}^0 = \hat{\alpha}_i(1 - \hat{\alpha}_i)z_i^\top(Z^\top \hat{V}^0 Z)^{-1}z_i$, $Z = (Z_1, Z_2, \dots, Z_n)$, $V^0 = \text{diag}(V_1^0, V_2^0, \dots, V_n^0)$, tal que $V_i^0 = (\alpha_i(1 - \alpha_i))$.

Definição 16. *Definimos o resíduo de Pearson para o componente discreto como*

$$r_i^{\text{pea}(0)} = \frac{(I_{(1)}(y_i) - (1 - \hat{\alpha}_i))}{\sqrt{\hat{V}_i^0(1 - \hat{h}_{ii}^0)}}, \quad (3.44)$$

em que $V_i^0 = (\alpha_i(1 - \alpha_i))$ e

$$I_{(1)}(y_i) = \begin{cases} 1, & \text{se } y > 0, \\ 0, & \text{se } y = 0. \end{cases} \quad (3.45)$$

Assim como para o componente contínuo, o resíduo de Williams para o componente discreto também é uma ponderação dos resíduos de Pearson e desvio.

Definição 17. *Definimos o resíduo de Williams para o componente discreto como*

$$r_{\text{wil}}^{(0)} = \text{sign}(I_{(1)}(y_i) - (1 - \hat{\alpha}_i))\{(1 - \hat{h}_{ii}^0)(r_{\text{dev}}^{(0)})^2 + \hat{h}_{ii}^0(r_{\text{pea}}^{(0)})^2\}^{\frac{1}{2}}. \quad (3.46)$$

3.5.2 Pontos influentes

Na análise de pontos influentes temos interesse em analisar se uma observação está interferindo desproporcionalmente nas estimativas dos parâmetros modelo. Observação influente é aquela observação que quando retirada do modelo provoca mudanças substanciais nas estimativas dos parâmetros do modelo (Cordeiro e Demétrio, 2008). Assim, o método mais utilizado para detectar quais observações são influentes é a exclusão de observações. Ou seja, para verificar a influência de uma observação, retiramos essa observação do modelo e verificamos as mudanças nas estimativas dos parâmetros do modelo. A medida mais utilizada para isso é a distância de Cook (Cook, 1977), medida que definimos a seguir.

A distância de Cook é uma estatística utilizada para medir o afastamento das estimativas dos parâmetros resultante da exclusão de uma observação (Cordeiro e Demétrio, 2008). Para os modelos lineares generalizados a medida de Cook é definida a seguir (Paula, 2013).

Definição 18. *A distância aproximada de Cook para a i -ésima observação é definida por:*

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_{(i)})^\top (X^\top \hat{W} X)(\hat{\beta} - \hat{\beta}_{(i)})}{\hat{\sigma}^2}, \quad (3.47)$$

em que $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p)^\top$, $p < n$, é um vetor de parâmetros estimados e $\hat{W} = \text{diag}\{\hat{w}_1, \dots, \hat{w}_n\}$ é a matriz de pesos estimados, sendo o peso \hat{w}_i definido como $\hat{w}_i = \frac{1}{V_i} \left(\frac{d\mu_i}{d\eta_i}\right)^2$ e V_i é a função de variância. No caso dos modelos lineares generalizados, $\hat{\beta}_i$ não tem forma fechada. A partir de uma aproximação de $\hat{\beta}_i$, (Cook e Weisberg, 1982), obtém-se que:

$$D_i \cong \left\{ \frac{\hat{h}_{ii}}{1 - \hat{h}_{ii}} \right\} (r_i^{pea})^2 \quad (3.48)$$

Nos modelos inflacionados no zero sugerimos fazer a análise de observações influentes separando o componente discreto do componente contínuo, pois ambos os componentes são ortogonais. Além disso, na análise de pontos influentes estamos interessados em analisar a influência de uma observação nas estimativas de máxima verossimilhança dos parâmetros. Assim podemos ter uma observação influenciando a estimativa de um componente, mas não do outro componente. Assim, definiremos a distância de Cook para o componente contínuo e para o componente discreto dos modelos ZAGA e ZAIG. A partir da equação (3.6), a distância de Cook para o componente contínuo do modelo é definida como:

$$D_i \cong \frac{\hat{h}_{ii}}{(1 - \hat{h}_{ii})} \left(\frac{(y_i - \hat{\mu}_i)}{\hat{\sigma}_i \sqrt{\hat{V}_i(1 - \hat{h}_{ii})}} \right)^2. \quad (3.49)$$

Assim para o modelo com resposta ZAGA temos que a distância de Cook para o componente contínuo do modelo é dado por:

$$D_i \cong \frac{\hat{h}_{ii}}{\hat{\sigma}_i^2(1 - \hat{h}_{ii})^2} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i^2}. \quad (3.50)$$

E no modelo com resposta ZAIG definimos a distância de Cook para o componente contínuo do modelo como:

$$D_i \cong \frac{\hat{h}_{ii}}{\hat{\sigma}_i^2(1 - \hat{h}_{ii})^2} \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i^3}. \quad (3.51)$$

E para o componente discreto, para ambos os modelos, definimos a distância aproximada de Cook como:

$$D_i^{(0)} \cong \frac{\hat{h}_{ii}^0}{(1 - \hat{h}_{ii}^0)^2} \frac{(I_{(1)}(y_i) - (1 - \hat{\alpha}_i))^2}{\hat{\alpha}_i(1 - \hat{\alpha}_i)}. \quad (3.52)$$

Consideramos uma observação influente quando D_i é relativamente alto quando comparada aos D_i das outras observações. Na próxima seção, discutiremos os gráficos de diagnóstico utilizados na análise de diagnóstico dos modelos ZAGA e ZAIG.

3.5.3 Gráficos de diagnóstico

Na análise de diagnóstico de modelos de regressão é comum a utilização de gráficos para uma melhor visualização dos resultados. Como dito anteriormente devido à ortogonalidade do parâmetro γ com os parâmetros β e δ , é sugerida uma análise de diagnóstico separada para cada componente do modelo (componente contínuo e discreto). Assim para o componente contínuo dos modelos ZAGA e ZAIG sugerimos os seguintes gráficos para análise de diagnóstico:

1. gráfico de dispersão entre r_i^{*qu} e $\hat{\mu}_i$ para identificação de observações discrepantes (*outliers*).

De acordo com os estudos de simulação realizados, o resíduo r_i^{*qu} foi aquele que apresentou melhor performance para a realização da análise de diagnóstico para o componente contínuo. Por isso, optamos por utilizar o resíduo r_i^{*qu} para a identificação de pontos "outliers".

2. gráfico de dispersão entre o resíduo r_i e $\hat{\mu}_i$ para avaliação do ajuste do modelo.

Em gráficos de dispersão entre o resíduo r_i e $\hat{\mu}_i$, quando o modelo está adequado os pontos ficarão distribuídos sem nenhum padrão, formando uma "nuvem" de pontos. Optamos por usar os resíduos r_i , no lugar dos resíduos r_i^* , pois gráfico dos resíduos r_i^* podem apresentar padrão mesmo que o modelo esteja bem ajustado, já que para valores de $\hat{\alpha}_i$ moderados ou grandes, não é possível a observação de resíduos próximos de zero. Aqui optamos por utilizar o resíduo r_i^{qu} , se o tamanho da amostra for pelo menos moderado, pois nos estudos de simulação apresentadas na Seção 3.2 foi este resíduo que obteve distribuição mais próximo da normal padrão para tamanho amostral igual a 50.

3. gráfico de D_i em função da ordem da observação é recomendado para a identificação de observações influentes.
4. gráfico de probabilidade normal com envelope simulado para r_i (sendo r_i qualquer resíduo utilizado nos modelos com resposta gama ou gaussiana inversa definidos na Seção 3.1).

Pelos mesmos motivos discutidos no segundo gráfico, sugerimos a utilização do resíduo r_i^{qu} para a realização do gráfico de probabilidade normal.

Analisar se o gráfico normal de probabilidade sugere um ajuste inadequado, pode ser difícil. Por isso [Atkinson \(1985\)](#) propôs a construção de bandas de confianças, geradas através de simulações, chamadas de "envelope". Para a construção do gráfico de envelope temos que considerar apenas as observações em que $y_i \in (0, \infty)$, já que estamos avaliando o ajuste do componente contínuo do modelo. Assim obtemos o gráfico de probabilidade normal com envelope simulado para r_i no modelo ZAIG⁷ a partir do seguinte procedimento:

I Gerar uma amostra da variável resposta considerando se a observação $y_i > 0$, então gerar $y_i^* \sim IG(\hat{\mu}_i, \hat{\sigma}_i)$ e se $y_i = 0$ então $y_i^* = 0$.

II Ajuste o modelo ZAIG para a amostra simulada, $(y^*) = (y_1^*, y_2^*, \dots, y_n^*)$ e obtenha o valor absoluto de cada um dos r_i para as observações em que $y_i > 0$ e ordene-as em ordem crescente.

III Repita o passo I e II k vezes. ([Atkinson \(1985\)](#) sugere utilizar $k = 19$.)

⁷ O gráfico de probabilidade meio normal para o modelo ZAGA segue o mesmo procedimento.

- IV Considerando $n^* = n - \sum_{i=1}^n I_0(y_i)$, conjunto de estatística de ordem que contém as k observações cada, obtenha para cada conjunto sua média, mínimo e seu máximo.
- V Represente no gráfico esses valores e os valores absolutos dos resíduos r_i ordenados no eixo vertical do gráfico e os escores normais dados por $\Phi^{-1}(\frac{i-3/8}{n+1/4})$ no eixo horizontal, em que Φ é a função de distribuição normal padrão.

Na análise de diagnóstico do modelo para o componente discreto sugerimos os seguintes gráficos:

- diagrama de dispersão entre $r_i^{(0)}$ ⁸ e $\hat{\alpha}_i$ para a identificação de observações discrepantes.
- gráfico de $D_i^{(0)}$ em função da ordem da observação é recomendado para a identificação de observações influentes.
- gráfico de probabilidade normal com envelope simulado para $r_i^{(0)}$ para análise do ajuste global do modelo.

Para a construção do gráfico de envelope do componente discreto, consideramos $y.disc_i$ igual a 1 se $y_i > 0$ e $y.disc_i = 0$ se $y_i = 0$, como nossa variável resposta do componente discreto do modelo. Assim, observamos que a nossa nova variável resposta tem resposta binária. Com isso, obtemos o gráfico de probabilidade normal com envelope simulado para r_i utilizando o mesmo procedimento do gráfico de envelope para modelo com resposta binária. A seguir descrevemos os passos para o procedimento para a construção do gráfico de probabilidade normal com envelope simulado.

- I Gerar n observações com distribuição binomial, $y.disc_i^* \sim Binomial(1, 1 - \hat{\alpha}_i)$.
- II Ajuste o modelo Binomial para a amostra simulada, $(y^*) = (y_1^*, y_2^*, \dots, y_n^*)$ e obtenha o valor absoluto de cada um dos $r_i^{(0)}$ e ordene-as cada grupo de n resíduos em ordem crescente.
- III Repita o passo I e II k vezes.
- IV Considerando n , o conjunto de estatística de ordem que contém as k observações cada, obtenha para cada conjunto sua média, mínimo e seu máximo.
- V Represente no gráfico esses valores e os valores absolutos dos resíduos r_i ordenados no eixo vertical do gráfico e os escores normais dados por $\Phi^{-1}(\frac{i-3/8}{n+1/4})$ no eixo horizontal, em que Φ é a função de distribuição normal padrão.

Para cada observação i os resíduos $r_i^{(0)}$ podem assumir apenas 2 valores, um se $y_i = 0$ e o outro se $y_i > 0$. Assim, a distribuição desses resíduos não é aproximadamente normal. Logo nos gráficos de dispersão entre $r_i^{(0)}$ e $\hat{\alpha}_i$ não é esperado que os pontos se distribuam sem

⁸Os resíduos $r_i^{(0)}$ podem ser qualquer um dos resíduos definidos para o componente discreto (Seção 3.5.1).

nenhum padrão, como é observado em bons resíduos para o componente contínuo quando o modelo está bem ajustado. Assim, para o componente discreto, esse gráfico não traz muita informação a respeito da qualidade do ajuste como na análise dos resíduos do componente contínuo nos modelos com resposta ZAGA e ZAIG (Pereira et al., 2012). Assim para a análise de diagnóstico para o componente discreto sugerimos um teste de qualidade de ajuste, que definiremos na próxima seção.

3.5.4 Teste de qualidade de ajuste para o componente discreto

Alguns testes de ajuste existentes na literatura podem ser utilizados na análise do componente discreto, entre eles: teste de Royston (Royston, 1992), teste qui-quadrado de Pearson (McCullagh, 1985), teste de Hosmer-Lemeshow (Hosmer Jr e Lemeshow, 2004) e teste de Stukel (Stukel, 1988). Hosmer et al. (1997) através de estudos de simulação estudaram testes de ajuste para o modelo logístico e concluíram que o teste de Stukel é mais poderoso para detectar falta de ajuste do modelo. Por isso, no próximo capítulo, optamos por utilizar o teste de Stukel para a análise do ajuste da parte discreta do modelo. A seguir descreveremos o teste de Stukel com mais detalhes.

O teste de Stukel, foi proposto por Stukel (1988) e é baseado na comparação entre o modelo para respostas binárias ajustado com um modelo mais geral. O modelo mais geral tem a adição de dois parâmetros. Cada parâmetro adicionado influencia o comportamento de uma das caudas da curva de probabilidade (Liu, 2007).

Considere $\hat{\alpha}_i$ e $\hat{\zeta}_i$ as estimativas de máxima verossimilhança de α_i e $\zeta_i = \mathbf{z}_i^\top \boldsymbol{\gamma}$, para $i = 1, 2, \dots, n$, respectivamente. As duas novas covariáveis são definidas como:

$$z_{1i} = 0,5\hat{\zeta}_i^2 I(\hat{\alpha}_i > 0,5) \quad (3.53)$$

$$z_{2i} = -0,5\hat{\zeta}_i^2 I(\hat{\alpha}_i < 0,5). \quad (3.54)$$

A generalização do modelo para resposta binária é obtida pela adição de duas covariáveis z_{1i} e z_{2i} da seguinte maneira:

$$g(\alpha_i) = \zeta_i + \gamma_{1i}z_{1i} + \gamma_{2i}z_{2i}. \quad (3.55)$$

A hipótese nula do teste é expressada como:

$$H_0 : \gamma_{1i} = \gamma_{2i} = 0, \quad (3.56)$$

Observe que a hipótese nula corresponde a hipótese de adequação do ajuste modelo.

Stukel (1988) em seu artigo propôs o teste score para o cálculo da estatística, mas Hosmer et al. (1991), sugere que a análise do teste de Stukel pode ser feito usando o teste score ou o teste da razão de verossimilhança. Por ser mais fácil a implementação, optamos em utilizar o teste da razão de verossimilhança. Então, seja $l_1(Z)$ o logaritmo da função

de verossimilhança do modelo sem as covariáveis z_{1i} e z_{2i} e $l_2(Z)$ o logaritmo da função de verossimilhança do modelo geral (equação (3.55)). A estatística do teste é dada por:

$$ST = 2(l_2(Z) - l_1(Z)). \quad (3.57)$$

Sob a hipótese nula, a estatística ST tem distribuição assintótica qui-quadrado com dois graus de liberdade. A não rejeição da hipótese nula implica que não temos evidências de que exista problemas com o ajuste do componente discreto do modelo ZAGA ou ZAIG.

No próximo capítulo apresentaremos uma aplicação de um ajuste no modelo ZAIG juntamente com sua análise de diagnóstico.

Capítulo 4

Aplicação

Neste capítulo apresentamos uma aplicação do modelo ZAIG para dados reais de fundos de investimentos no Brasil. Na Seção 4.1, descrevemos os dados e apresentamos uma análise descritiva. Na Seção 4.2, ajustamos o modelo ZAIG, selecionamos um modelo final, realizamos uma análise de diagnóstico e para finalizar apresentamos as conclusões obtidas. E na Seção 4.3 apresentamos alguns comentários relevantes sobre a identificação de observações positivas discrepantes.

4.1 Descrição dos dados

Os fundos de investimentos são tipos de investimentos coletivos, em que existe um administrador do fundo, geralmente uma instituição financeira, que define os objetivos, políticas de investimentos e outras regras gerais para regulamentar o fundo. E através de recursos financeiros de diversos investidores, os administradores dos fundos investem em diferentes tipos de ativos (ações, moedas estrangeiras, investimentos estrangeiros entre outros). Existem diferentes tipos de fundos: cambiais, renda fixa, referenciados DI, entre outros. Aqui utilizamos apenas os fundos referenciados DI para a nossa análise (Pereira, Novaes, Dias et al., 2014).

Os fundos referenciados DI¹ são um tipo de investimento considerado conservador e são fundos que buscam proteger o patrimônio do investidor. Esses fundos são atrelados à taxa de juros interbancária, o chamado CDI (Certificado de Depósito Interbancário), que por sua vez acompanha a variação da taxa básica de juros da economia brasileira, a Selic. Se a Selic aumenta, o rendimento do fundo DI sobe; se a Selic cai, o mesmo ocorre com o fundo DI (Fonseca et al., 2007). Para a nossa análise de dados, utilizamos os fundos referenciados DI existentes em junho de 2013. Os dados foram fornecidos pela empresa AMBIMA².

O objetivo da nossa análise é modelar a taxa de crescimento dos fundos de investimentos do tipo referenciados DI. Para isso, definimos como variável resposta taxa de crescimento

¹DI - Depósito interbancários, pois esses fundos tendem a acompanhar o certificado de depósito. Esses certificados representam a taxa de juros aplicada nos empréstimos interbancários.

²www.portalambima.com.br

entre o patrimônio líquido (pl) de junho de 2015 e o patrimônio líquido (pl) de junho de 2013 para cada fundo DI ($y = (pl.junho2015)/(pl.junho2013)$). Portanto nossa variável resposta vai assumir valor 0 quando o fundo DI por algum motivo não estava ativo em junho de 2015 e valor entre $(0, \infty)$ para aqueles fundos que estavam ativos em junho de 2015. Como covariáveis utilizamos as variáveis instituições financeiras, tempo do fundo (data final³ - data de criação do fundo) e a taxa de crescimento do fundo no semestre anterior a data final ($(pl$ de junho de 2013)/(pl de dezembro de 2012), sendo que esta última denominamos como variável razão. Consideramos aqui todos os fundos de investimentos referenciados DI existentes em junho de 2013 e que tinham pelo menos 1 ano de existência. Retiramos algumas observações da nossa análise, entre elas aqueles fundos que não tinham informações da pl-dezembro 2012 (2 fundos tiveram essa característica). Também foram retiradas aquelas observações que tinham pl-junho 2013 menores que R\$5.000.000,00 (fundos muito pequenos). Também foram retirados os fundos de instituições financeiras que tinham menos do que 5 fundos referenciados DI. Além disso, para aquelas observações que tinham pl de junho de 2015 menores que R\$5.000.000,00, consideramos o valor da variável resposta como sendo igual a 0. A justificativa para isso é que fundos que têm patrimônio líquido muito pequeno são considerados pela instituição financeira como inativos. São aqueles fundos para os quais a instituição não recebe novos investimentos e que a instituição gostaria de fechar. Porém sem o consentimento dos investidores o banco não pode fechar o fundo, mas o mesmo recebe pouquíssima atenção da instituição financeira. Note que considerar o pl de junho de 2015 igual a zero para fundos com pl bem pequeno nessa data, não impede a ocorrência de valores bem próximos de 0 para a variável resposta, pois o pl de muitos fundos é superior a um bilhão de reais em junho de 2013 e um decréscimo para valores pouco superior a 5 milhões resulta em um valor bem baixo da variável resposta.

Após os filtros mencionados a população em estudo tem tamanho $n = 344$. A seguir apresentamos uma análise descritiva dessas variáveis. É importante ressaltar que neste trabalho estamos considerando a população toda na análise dos dados e não uma amostra da população.

A variável resposta é a taxa de crescimento do fundo DI, conforme definido anteriormente. Temos que das 344 observações existentes na população final, 44 são zeros. Ou seja, 12,8% dos fundos DI foram fechados ou tinham pl inferior a 5 milhões em junho de 2015. Na Tabela 4.1, apresentamos um resumo das medidas descritivas da variável resposta Y , em que consideramos apenas as observações diferentes de zero. Pode-se observar na Tabela 4.1 que a variável resposta tem um crescimento médio de 1,697 e uma variação de $[0,0445;45,020]$ e um desvio padrão de 2,972. Na Figura 4.1, temos o histograma da variável resposta Y considerando apenas as observações diferentes de zeros. No gráfico dessa figura, observamos que a variável resposta tem distribuição assimétrica a direita e uma alta concentração dos valores da variável resposta entre $(0,2)$.

³A data final definimos 30/06/2013, pois é a data utilizada como base das nossas covariáveis.

Tabela 4.1: Medidas descritivas da variável resposta Y considerando apenas as observações diferentes de zero.

Variável	Minimo	Q1	Mediana	Média	Q3	Máximo	Desv. Pad
Y	0,0445	0,7611	1,0420	1,6970	1,6720	45,0200	2,9722

Na Tabela 4.2 apresentamos as medidas descritivas das covariáveis utilizadas na análise. Observamos que na população a taxa de crescimento entre dezembro de 2012 a junho de 2013 (variável razão), foi em média de 1,26, a menor taxa de crescimento foi de 0,05 e a maior de 31,97, com desvio padrão de 1,81. Já o patrimônio líquido desses fundos em junho de 2013 foi em média 1153,10 milhões, sendo o menor patrimônio líquido de 5,18 milhões e o maior de 23050,00 milhões, tendo um desvio padrão de 2738,39 milhões. E a variável tempo do fundo tem média 9,4 anos, tendo o fundo mais novo 1,13 anos e o mais antigo 22,88 anos, com um desvio padrão de 5,18 anos. Através de tabelas e gráficos, não apresentados aqui, estudamos o relacionamento das covariáveis com a variável resposta para a inserção das mesmas no ajuste do modelo da melhor maneira possível. Assim, em alguns casos, categorizamos as covariáveis e em outros casos tomamos o logaritmo da variável, como na variável razão. Na Tabela 4.3, apresentamos as categorizações feitas nas covariáveis inseridas no modelo. Note na tabela que uma mesma covariável foi categorizada de forma diferente em cada um dos submodelos, com o objetivo da obtenção de um melhor ajuste.

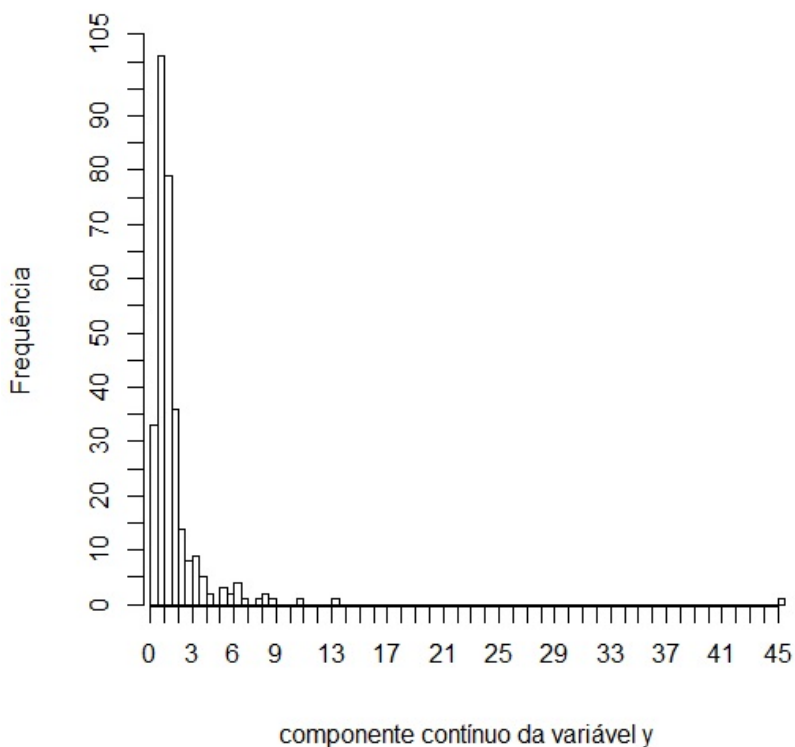


Figura 4.1: Histograma da variável resposta - observações diferentes de zero.

Analisando a Tabela 4.3, temos a primeira variável banco- μ , em que temos uma categorização da variável instituição financeira. No grupo I temos os bancos que possuem baixa taxa de crescimento (14,83% das instituições), grupo II taxa de crescimento mediano (29% das instituições) e grupo III, instituições com as taxas de crescimentos maiores (56% das instituições). Na categorização da variável banco- α , temos no grupo I os bancos cujos fundos têm baixa probabilidade de assumir zero na variável resposta e no grupo II estão os bancos cujos fundos têm probabilidade de assumir zero na variável resposta moderada, em torno de 20%.

Tabela 4.2: *Medidas descritivas das covariáveis.*

	Minimo	Q1	Mediana	Média	Q3	Máximo	Desv. Pad
razão	0,050	0,900	1,015	1,260	1,152	31,970	1,810
pl/milhão	5,18	56,92	223,04	1153,10	939,700	23050,00	2738,385
tempo	1,129	4,764	8,468	9,379	13,630	22,880	5,182

Tabela 4.3: *Tabela de frequências das covariáveis categorizadas.*

Variável	Categorias	Freq. Abs	Freq. Rel (%)
banco- μ	I	51	14,83
	II	99	28,78
	III	194	56,40
banco- α	I	289	84,01
	II	55	15,99
tempo- μ	I-<6	126	36,63
	II-<15	157	45,64
	III-> 15	61	17,73
tempo- α	I- \leq 3,65	49	14,24
	II-> 3,65	295	85,76
tempo- σ	I-< 6	126	36,63
	II-> 6	218	63,37
pl- α	I-< 57	86	25,00
	II-< 930	171	49,71
	III-> 930	87	25,29
razao- α	I- < 0,5	6	1,74
	II- \leq 1	154	44,77
	III-> 1	184	53,49

Ainda analisando a Tabela 4.3, temos a categorização da variável tempo, tal que, a variável tempo- μ vale I para fundos com menos de 6 anos (36,6%), II fundos entre 6 e 15 anos (45,6%) e III fundos com mais de 15 anos (17,7%). E a variável tempo- α , 14% estão no grupo I (fundos com menos de 3,65 anos) e 86% estão no grupo II (fundos com mais de 3,65 anos). A variável tempo- σ esta categorizada em 2 níveis, I (fundos com menos de 6 anos) que corresponde a 37% da população e nível II (fundos com mais de 6 anos) que representa 63% da população. As covariáveis pl- α e razão- α estão categorizadas em 3 níveis. Temos que 25% da população apresentam um patrimônio líquido de até 57 milhões (nível I), 50% patrimônio

entre 57 e 930 milhões (nível II) e 25% patrimônio maior que 930 milhões (nível III). E na última covariável temos que apenas 1,74% da população tem uma taxa de crescimento de 6 meses de menos que 0,5; 44,77% dos fundos apresentam uma taxa de crescimento de 6 meses menores que 1 e 53,49% uma taxa acima de 1. A partir das variáveis definidas nesta seção, na seção a seguir (Seção 4.2), apresentamos o ajuste e análise de diagnóstico do modelo ZAIG para os fundos DI.

4.2 Ajuste e análise de diagnóstico

O modelo considerado para o ajuste da variável taxa de crescimento do fundo DI foi o ZAIG. Utilizou-se o modelo de regressão ZAIG e não o ZAGA, pois a análise de diagnóstico e o AIC sugeriram um ajuste consideravelmente superior com o primeiro modelo. Utilizou-se o pacote "gamlss" para o ajuste do mesmo e as covariáveis utilizadas no ajuste foram as variáveis apresentadas na seção anterior (Seção 4.1). Por estarmos trabalhando com a população e não com uma amostra, os resultados dos testes de hipóteses devem ser vistos com cautela. Por isso, a inclusão ou não da variável no modelo se deu através de uma seleção de modelos, conforme descrito na Seção 2.5. Portanto, utilizando o critério AIC, obteve-se um modelo final para cada um dos parâmetros da distribuição ZAIG e realizou-se uma análise de diagnóstico conforme discutido no Capítulo 3.

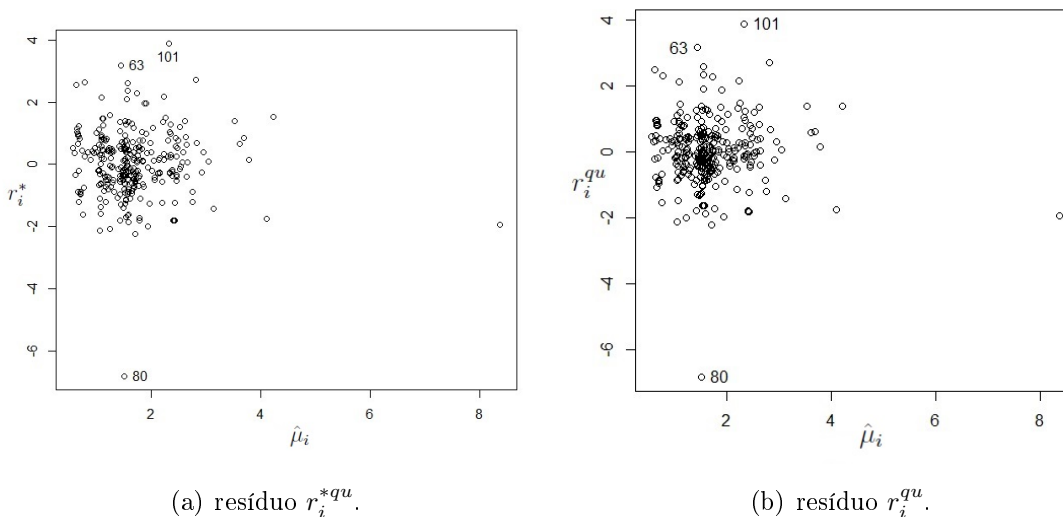


Figura 4.2: Análise de diagnóstico componente contínuo do modelo ZAIG.

Na Figura 4.2, apresentamos a análise de diagnóstico do componente contínuo do modelo ZAIG. Do lado esquerdo temos o gráfico entre $\hat{\mu}$ e os resíduos r_i^{*qu} (Figura 4.2(a)), discutido na Seção 3.3.1. Observa-se que 3 observações destacam-se com resíduos em módulo maiores que 3, sendo elas as observações 63, 80 e 101. Por serem possíveis observações discrepantes, analisamos posteriormente como o modelo se comporta com a retirada dessas observações. O

gráfico entre r_i^{qu} e $\hat{\mu}$ (Figura 4.2(b)), utilizamos para uma avaliação do ajuste do modelo. Fora as observações discrepantes, aparentemente o componente contínuo do modelo selecionado está adequado para os dados.

Na Figura 4.3, temos os gráficos entre $\hat{\alpha}$ e $r_i^{(0)}$, sendo $r_i^{(0)}$ qualquer um dos resíduos para o componente discreto apresentados na Seção 3.5.1. Esses gráficos auxiliam na identificação de pontos "outliers". Nos gráficos entre resíduos $r_{dev}^{(0)}$ e $r_{wil}^{(0)}$ pelo parâmetro $\hat{\alpha}$ (Figuras 4.3(a) e 4.3(c)), destaca-se a observação 89. Já no gráfico entre o resíduo $r_{pea}^{(0)}$ e $\hat{\alpha}$ (Figura 4.3(b)) além da observação 89, mais observações aparecem como aberrantes. Entre essas observações destacamos a observação 2, 87, 191 e 249. Posteriormente iremos analisar como o modelo se comporta com a retirada dessas observações.

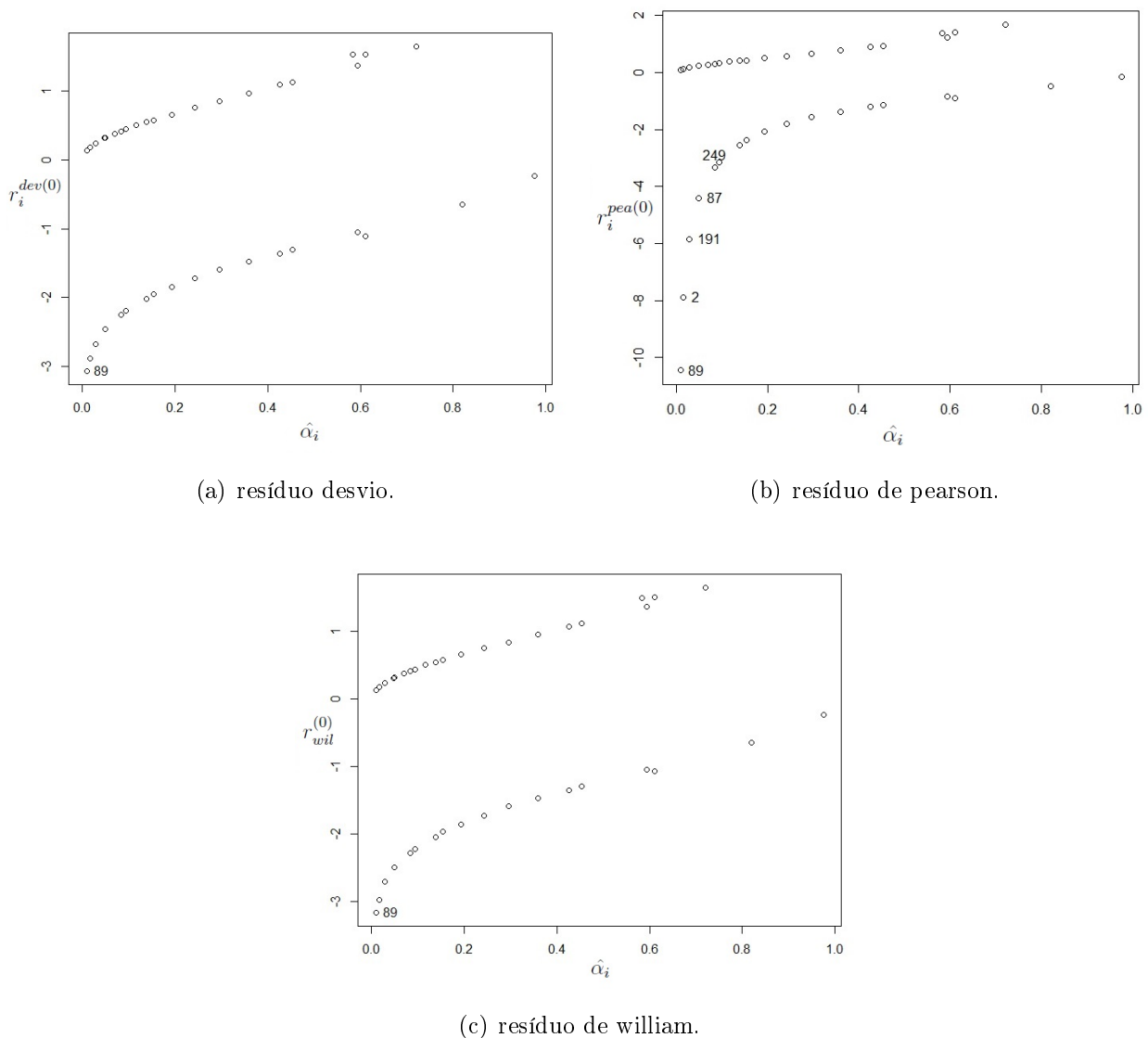


Figura 4.3: Análise de diagnóstico componente discreto do modelo ZAIG.

Uma parte importante na análise de diagnóstico é a análise de observações influentes. A Figura 4.4, apresenta a distância de Cook para o componente contínuo (Figura 4.4(a)) e para

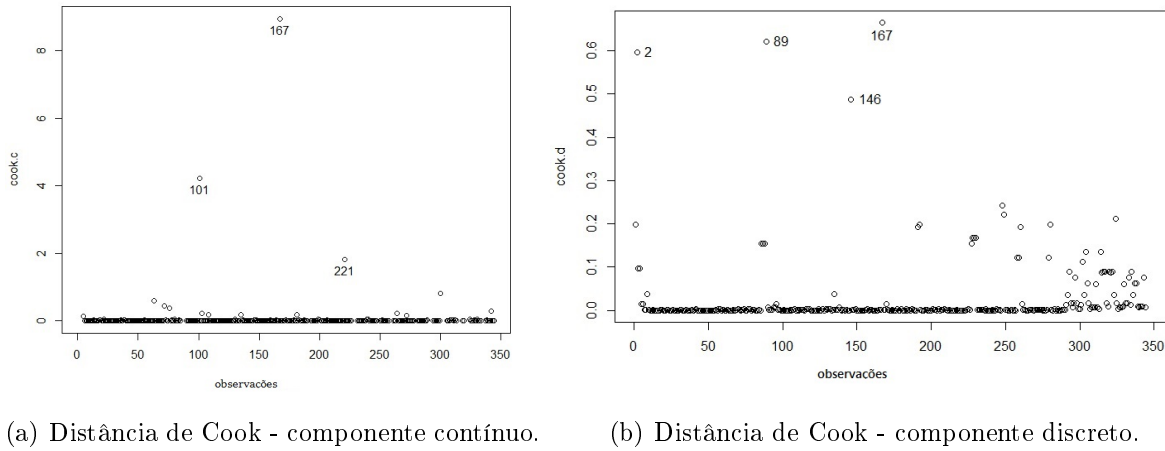
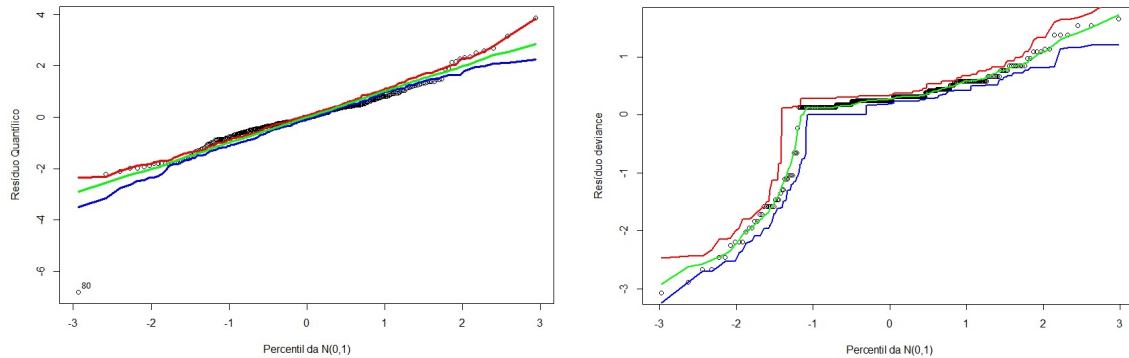


Figura 4.4: *Análise de diagnóstico distância de Cook.*

o componente discreto (Figura 4.4(b)), como definidos na Seção 3.5.2. Observamos que para o componente contínuo as observações que se destacam como influentes são a observação 101, 167 e 221. E para o componente discreto (Figura 4.4(b)), as observações que se destacam são as observações 2, 89, 146 e 167.

Para finalizar a análise de diagnóstico do modelo apresentamos os gráficos de envelope (Figura 4.5) para o componente contínuo e discreto. No envelope para o componente contínuo (Figura 4.5(a)), com exceção da observação 80, todas as observações estão dentro ou muito próximos dos limites determinados pelo envelope. E o envelope para o componente discreto (Figura 4.5(b)), sugere que o ajuste para o componente discreto do modelo está adequado, não apresentando nenhum ponto fora do envelope. Além disso, o valor-p do teste de Stukel foi de 0,63, confirmando assim o ajuste adequado do componente discreto do modelo.

A Tabela 4.4 apresenta os resultados do modelo ZAIG final para a taxa de crescimento dos fundos referenciados DI. Observamos que todas as variáveis incluídas no modelo final são significantes no modelo (valor-p menor que 5%). Observando o sinal das estimativas dos parâmetros conclui-se que para a variável banco- μ , a mudança de uma categoria para outra categoria causa um aumento na média dos valores positivos da variável resposta, taxa de crescimento dos fundos DI. Ou seja, bancos que estão na categoria III têm uma maior taxa média de crescimento dos fundos DI do que bancos que pertencem à categoria II e I considerando apenas os valores positivos. E bancos que pertencem à categoria II tem uma maior taxa média de crescimento dos fundos DI do que os bancos na categoria I. Isso faz sentido, uma vez que na categorização da variável banco agrupamos os mesmo de acordo com sua taxa média de crescimento, ou seja, na categoria III temos os bancos com maiores taxas médias de crescimento. Analisando a variável tempo- μ , temos que a mudança de categoria leva a uma redução da taxa média de crescimento dos fundos DI para aqueles fundos que se mantiveram ativos. Por exemplo, bancos que possuem fundos com mais de 6 anos têm uma taxa média de crescimentos dos fundos DI, menor que os bancos que possuem fundos



(a) Envelope componente contínuo - resíduo quantílico. (b) Envelope componente discreto - resíduo desvio.

Figura 4.5: Gráfico envelope.

com menos de 6 anos. Isso é razoável, uma vez que fundos mais antigos tendem a ser muito grandes e assim o crescimento relativo tende a ser menor. A última variável para o modelo μ é a variável $\log(\text{razão})$, observamos que um aumento da variável $\log(\text{razão})$ aumenta o valor da taxa média de crescimento dos fundos DI que se mantiveram ativos. No modelo para σ , nota-se que quanto maior o logaritmo da variável razão, menor é o valor do parâmetro de dispersão. E em relação à variável tempo- σ , fundos com mais de 6 anos apresentam valores de σ menores do que aqueles fundos com menos de 6 anos. Para o parâmetro σ não temos uma interpretação direta, pois a variância da ZAIG é função de σ e μ .

Tabela 4.4: Modelo final para a variável taxa de crescimento do fundo referenciados DI existente em junho de 2013 a junho de 2015.

Equação	Variável	Estimativas	Erro Padrão	Valor t	p-valor	exp()
μ	intercepto	0,03	0,15	0,17	0,86	1,03
	banco- μ -II	0,51	0,13	4,01	0,00	1,67
	banco- μ -III	0,83	0,12	6,97	0,00	2,29
	tempo- μ -II	-0,44	0,13	-3,49	0,00	0,64
	tempo- μ -III	-0,32	0,16	-1,99	0,05	0,73
	$\log(\text{razao})$	0,37	0,10	3,73	0,00	1,44
σ	intercepto	-0,17	0,07	-2,32	0,02	0,84
	$\log(\text{razao})$	-0,24	0,07	-3,36	0,00	0,78
	tempo- σ -II	-0,32	0,09	-3,46	0,00	0,73
α	(Intercept)	1,58	0,98	1,61	0,11	4,87
	banco- α -II	2,09	0,42	5,00	0,00	8,10
	tempo- α -II	-1,13	0,47	-2,40	0,02	0,32
	pl- α -II	-1,25	0,42	-3,01	0,00	0,29
	pl- α -III	-2,41	0,79	-3,06	0,00	0,09
	razao- α -II	-2,16	0,96	-2,25	0,02	0,12
	razao- α -III	-2,73	0,98	-2,78	0,01	0,07

Continuamos a análise da Tabela 4.4, observando o sinal das estimativas do modelo para

o parâmetro α . Temos que para a variável banco- α , bancos que pertencem à categoria II têm uma probabilidade de ter um fundo de investimento encerrado maiores que aqueles bancos que pertencem a categoria I. Isso é razoável, visto que na categoria I temos os fundos com proporção amostral de zeros menores que 20%. Na variável pl- α , fundos com o pl > 930 milhões tendem a ter uma menor probabilidade de serem encerrados do que aqueles fundos com pl < 930 milhões. Já na variável razão- α , temos que o aumento na taxa de crescimento do fundo no semestre anterior a data final (razão) diminui a probabilidade do fundo ser encerrado.

No ajuste do modelo ZAIG foram usadas as funções de ligação logarítmica e logito. Assim a relação entre as variáveis presentes no modelo ZAIG e a taxa de crescimento dos fundos de referenciados DI pode ser melhor interpretada a partir de exponenciais das estimativas dos parâmetros (apresentadas na última coluna da Tabela 4.4). Assim, estima-se que, por exemplo, fundos que pertencem a bancos da categoria 3 e se mantêm ativos têm uma taxa de crescimento médio em 2 anos, $129\%((2,29 - 1) \times 100\%)$ maior que fundos que pertencem a bancos da categoria 1 (mantendo-se as outras variáveis fixas). Estima-se também que a chance de um fundo se tornar inativo em um período de 2 anos é $710\%((8,10 - 1) \times 100\%)$ maior em fundos dos bancos que pertencem a bancos do categoria 2 do que em bancos da categoria 1, mantidas as demais variáveis constantes. Em relação às estimativas do submodelo para σ , a interpretação das estimativas não é tão direta, pois a variância na ZAIG depende de μ e σ . A interpretação das demais estimativas pode ser feita de forma equivalente.

Tabela 4.5: Tabela da variação percentual das estimativas dos parâmetros com a retirada das observações.

Equação	Variável	Estimativas	% de mudança sem a obs.					
			80	101	167	2	89	146
μ	intercepto	0,0257	234,16	-500,89*	-19,84	0,00	0,00	-22,18
	banco- μ -II	0,5111	-3,05	1,69	-3,43	0,00	0,00	-0,22
	banco- μ -III	0,8272	-2,37	-4,03	-4,51	0,00	0,00	0,33
	tempo- μ -II	-0,4445	9,75	-31,94	-5,83	0,00	0,00	-1,01
	tempo- μ -III	-0,3168	18,70	-47,63	8,31	0,00	0,00	-2,14
	log(razao)	0,3670	-11,01	-5,77	37,33	0,00	0,00	0,81
σ	intercepto	-0,1716	16,10	34,36	1,93	0,00	0,00	-0,93
	log(razao)	-0,2424	-53,36	-1,27	-9,76	0,00	0,00	3,14
	tempo- σ -II	-0,3177	37,22	-18,81	1,75	0,00	0,00	0,74
α	(Intercepto)	1,5838	-0,17	-0,17	50,66	1,38	3,39	42,69
	banco- α -II	2,0913	-0,21	-0,21	-0,28	4,02	4,26	-0,06
	tempo- α -II	-1,1344	-0,19	-0,19	12,17	2,11	6,07	-3,03
	pl- α -II	-1,2499	-0,42	-0,42	-8,80	1,92	0,54	2,41
	pl- α -III	-2,4115	-0,03	-0,03	-2,92	30,52	29,80	0,77
	razao- α -II	-2,1608	-0,17	-0,17	34,07	2,62	-0,49	31,90
	razao- α -III	-2,7271	-0,01	-0,01	28,35	-0,73	3,40	25,04

* não faz sentido o cálculo da variação percentual quando temos uma mudança de sinal nas estimativas, como houve no caso desse intercepto para o qual temos uma mudança de 0,0257 para -0,103.

Na Tabela 4.5, apresentamos o cálculo da variação percentual das estimativas dos parâmetros quando retiramos observações que se mostraram discrepantes ou/e influentes. A variação percentual das estimativas dos parâmetros das observações 80 e 101 são altas para os submodelos μ e σ , mas não são altas para o submodelo α . Isso ocorre, pois ambas as observações foram identificadas como discrepantes no componente contínuo do modelo, mas não no componente discreto. Por isso que sua exclusão não causa grandes variações percentual das estimativas dos parâmetros para o submodelo α . A observação 101, além de discrepante foi identificada como influente. Por esse motivo, a variação percentual das estimativas é maior com a retirada dessa observação. A observação 167 foi identificada como influente no componente contínuo e discreto e por isso é a única observação que altera consideravelmente as estimativas tanto no submodelo para μ como no submodelo para α . Já, as observações 2 e 89 são influentes apenas no componente discreto e observamos que a sua exclusão não causa nenhuma alteração nas estimativas nos submodelos para μ e σ . Isso ocorre porque essas observações têm valor da variável resposta igual a 0 e em modelos contínuos inflacionados no zero, observações com essas características não influenciam as estimativas dos parâmetros dos submodelos para μ e σ . A observação 80 não foi identificada como influente, mas apenas como discrepante. Porém, a Tabela 4.5 sugere que ela é bastante influente nas estimativas do modelo para σ . Ela não tem alto valor no componente contínuo da distância de Cook porque esta medida está medindo apenas a influência no submodelo para μ . Apesar de termos uma amostra de tamanho 344, a exclusão de cada uma dessas 6 observações leva a uma alteração de mais de 29% em pelo menos uma das estimativas. Isso reforça a importância da identificação de observações discrepantes e influentes. Embora essas 6 observações causem considerável alteração nas estimativas dos parâmetros do modelo, sua exclusão não altera os sinais das estimativas dos parâmetros (com exceção de uma única estimativa do intercepto que tem valor muito próximo de 0 no modelo final). Além disso, com exceção também do intercepto para μ , as variações percentuais embora consideráveis, não são extremamente altas. Isso, aliado à análise de diagnóstico efetuada, sugere que o modelo com todas as observações é adequado para o estudo da taxa de crescimento de fundos referenciados DI.

4.3 Discussão sobre a identificação de observações positivas discrepantes.

Como dito anteriormente (Seção 3.5.2), se utilizarmos, na análise de diagnóstico dos modelos contínuos inflacionados no zero, resíduos existentes na literatura para identificação de observação discrepantes, teremos alguns problemas. Nesta seção discutimos alguns pontos relevantes sobre esses resíduos.

Nos modelos inflacionados no zero é comum utilizar o resíduo quantílico aleatorizado, q_i , para a identificação de pontos outliers. Mas como discutido na Seção 3.3.1, temos que o resíduo q_i não consegue captar como outliers observações que assumem valores próximos de

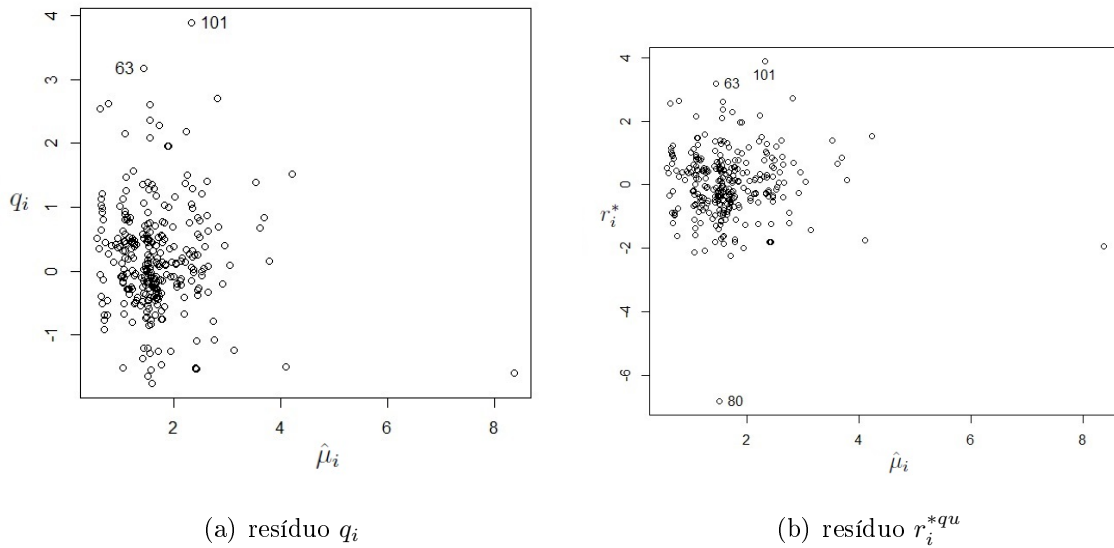


Figura 4.6: Resíduos para o componente contínuo do modelo ZAIG.

zero e têm alta ou moderada probabilidade de assumir o valor zero. Na Figura 4.6, temos um exemplo desse problema apresentado pelo resíduo q_i . A observação 80 tem $\hat{\alpha} = 0,049$. Observamos que essa observação não aparece como discrepante no gráfico de q_i versus $\hat{\mu}$ (Figura 4.6(a)), mas no gráfico de r_i^{*qu} versus $\hat{\mu}$ (Figura 4.6(b)), essa observação é considerada discrepante. Analisando a observação 80, notamos que realmente ela é discrepante, pois para a taxa de crescimento do fundo referenciados DI ela assume o valor de 0,045, na variável banco- μ ela pertence a categoria III, na variável tempo- μ ela pertence a categoria III e o valor da taxa de crescimento do último semestre foi de 0,71. Ela é considerada discrepante, pois na variável banco- μ ela pertence a uma categoria em que estão os bancos com maiores taxas de crescimentos dos fundos referenciados DI e observamos que a taxa de crescimento desse fundo é baixa. Assim o valor estimado para a taxa de crescimento dos fundos referenciados DI para a observação 80 foi de 1,51. É interessante notar ainda que essa observação é bastante influente no submodelo para σ conforme mencionado na Seção 4.2 e isso não seria notado se utilizássemos apenas a distância de Cook e o resíduo quantílico aleatorizado na análise de diagnóstico.

Na Seção 3.3.1 outro ponto discutido foi a utilização dos resíduos r_i^{dev} , r_i^{pea} , r_i^{qu} , r_i^{wil} e r_i^{ans} na análise de observações discrepantes. Porém, como foi dito, se utilizarmos esses resíduos podemos ter observações com o mesmo resíduo r_i não importando o valor de $\hat{\alpha}_i$. Por exemplo, se considerarmos a observação 167 do conjunto de dados analisado na Seção 4.2, temos que as variáveis para essa observação assumem os valores apresentados na Tabela 4.6. O resíduo r_i^{qu} para essa observação é igual a $r_i^{qu} = 2,32$. Suponha, por exemplo, que nesse conjunto de dados tivesse mais duas observações: # e *, com as características apresentadas na Tabela 4.7. Observa-se que os valores das observações # e * para as variáveis do modelo para o parâmetro μ e σ são iguais aos valores dessas variáveis na observação 167, mas apresentam

Tabela 4.6: *Valores das variáveis para a observação 167*

Variável	Valor
y[167]	5,93
banco- μ [167]	3
banco- α [167]	1
tempo- μ [167]	1
tempo- α [167]	1
tempo- σ [167]	2
pl- α [167]	2
razao- α [167]	1
log(razao)[167]	-2,99

diferença em relação ao modelo para o parâmetro α . Se utilizarmos o resíduo r_i^{qu} teremos sempre o mesmo valor de resíduo, ou seja, $r_i^{qu} = 2,32$. Agora utilizando o resíduo r_i^{*qu} teríamos resíduos iguais a 2,63, 2,922 e 2,369 para as observações 167, # e * respectivamente. Assim ao utilizar o r_i^{qu} não teríamos essa variação dos resíduos e já com o resíduo r_i^{*qu} temos essa variação dos resíduos, conseguindo detectar melhor quando temos observações discrepantes. Como na observação # para qual temos um resíduo consideravelmente mais alto quando comparado a observação *.

Tabela 4.7: *Valores das variáveis para a observação # e **

Variável	Valor	Variável	Valor
banco- μ #	3	banco- μ *	3
banco- α #	1	banco- α *	1
tempo- μ #	1	tempo- μ *	1
tempo- α #	1	tempo- α *	2
tempo- σ #	2	tempo- σ *	2
pl- α #	1	pl- α *	3
razao- α #	1	razao- α *	1
log(razao)#	-2,99	log(razao)*	-2,99

Capítulo 5

Conclusões e trabalhos futuros

O trabalho teve como objetivo principal estudar as propriedades de duas classes de resíduos nos modelos de regressão com resposta gama inflacionados e gaussiana inversa inflacionados ambos no valor zero. Tais modelos são utilizados em diversas áreas como financeira, saúde pública, medicina e seguro. Para alcançar tal objetivo, definimos primeiramente as distribuições gama inflacionada no zero e gaussiana inversa inflacionada no zero. Em seguida, discutimos os modelos de regressão gama inflacionada no zero e gaussiana inversa inflacionada no zero. Para esses modelos, obtivemos o vetor escore e a matriz de informação de Fisher e discutimos estimação, intervalos de confiança, teste de hipóteses e seleção de variáveis.

No Capítulo 3, discutimos a análise de diagnóstico desses modelos. Em um primeiro momento, definimos os modelos lineares generalizados e os resíduos utilizados em tais modelos (deviance, Pearson, quantílico, Anscombe e Williams). Por meio de estudos de simulação de Monte Carlo examinamos as propriedades desses resíduos nos modelos de regressão com resposta gama e gaussiana inversa. Concluímos que o resíduo que tem distribuição mais próxima da normal padrão é o resíduo de Pearson para amostras pequenas e o resíduo quantílico para amostras moderadas ou grandes. O conhecimento das propriedades desses resíduos nos modelos gama e gaussiana inversa foram importantes, pois esses mesmos resíduos são utilizados na definição de novos resíduos nos modelos de regressão inflacionados no zero.

Discutimos também os resíduos para os modelos de regressão gama e gaussiana inversa inflacionados no zero. Apresentamos duas classes de resíduos para realizar análise de diagnóstico do componente contínuo desses modelos e estudamos suas propriedades por meio de estudos de simulação de Monte Carlo. Os resultados dos estudos de simulação sugerem que o resíduo de uma das classes que é função do resíduo quantílico possui boas propriedades.

Além disso, discutimos outros aspectos da análise de diagnóstico tais como, análise de observações discrepantes no componente discreto, observações influentes nos modelos inflacionados no zero, gráficos de diagnósticos para auxiliar na análise de diagnóstico e para finalizar discutimos um teste de qualidade do ajuste para o componente discreto.

Para completar nosso trabalho, aplicamos a metodologia apresentada em um banco de

dados reais. Isso foi útil, pois conseguimos exemplificar toda a parte teórica, além de conseguir mostrar os problemas apresentados pelos resíduos que já existem na literatura para os modelos contínuos inflacionados no zero.

5.1 Próximos passos

Como trabalhos futuros temos algumas propostas:

- Desenvolvimento de um pacote no software R para a análise de diagnóstico em modelos ZAGA e ZAIG.
- Estudo das propriedades das classes de resíduos estudadas para outros modelos de regressão contínuos inflacionados no zero.
- Desenvolvimento de análise de influência local em modelos ZAGA e ZAIG. Essa análise consiste em verificar se existem observações que causam efeitos desproporcionais nos resultados da análise inferencial, quando pequenas perturbações são incorporadas no modelo e/ou nos dados do estudo.

Apêndice A

Demonstrações dos resultados do Capítulo (2)

A.1 Modelo ZAGA

A.1.1 Obtenção da função escore para o modelo ZAGA

A função escore é obtida derivando o logaritmo da função de verossimilhança (Equação (2.11)) em relação a cada parâmetro. Assim, para obtermos o termo da função escore, associado ao parâmetro γ_j , derivamos a equação (2.11) em relação ao parâmetro γ_j . Com isso temos:

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \gamma_j} &= \frac{\partial l_1(\gamma)}{\partial \gamma_j} = \sum_{i=1}^n \frac{\partial l_{i1}(\alpha_i)}{\partial \gamma_j} = \sum_{i=1}^n \left\{ \frac{\partial l_{i1}(\alpha_i)}{\alpha_i} \frac{\partial \alpha_i}{\partial \zeta_i} \frac{\partial \zeta_i}{\partial \gamma_j} \right\} = \\ & \sum_{i=1}^n \left\{ \left[\frac{I_{\{0\}}(y_i)}{\alpha_i} - \frac{(1-I_{\{0\}}(y_i))}{(1-\alpha_i)} \right] \frac{\partial \gamma_i}{\partial \zeta_i} z_{ij} \right\} = \sum_{i=1}^n \left\{ \left[\frac{I_{\{0\}}(y_i) - \alpha_i}{\alpha_i(1-\alpha_i)} \right] \frac{\partial \gamma_i}{\partial \zeta_i} z_{ij} \right\} \end{aligned} \quad (\text{A.1})$$

O termo de posição j do vetor que é o segundo termo de $U(\theta)$ (Equação (2.13)) é dado por:

$$\frac{\partial l(\theta)}{\partial \beta_j} = \frac{\partial l_{2i}(\beta, \delta)}{\partial \beta_j} = \sum_{\substack{i: \\ y \neq 0}} \left\{ \frac{\partial l_{i2}(\mu_i, \sigma_i)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} \right\} = \sum_{\substack{i: \\ y \neq 0}} \left\{ \left[\frac{(y_i - \mu_i)}{\mu_i^2 \sigma^2} \right] \frac{\partial \mu_i}{\partial \eta_{i1}} x_{ij} \right\} \quad (\text{A.2})$$

E o termo de posição j do vetor que é último termo de $U(\theta)$ (Equação (2.13)) é dado por:

$$\begin{aligned} \frac{\partial l(\theta)}{\partial \delta_j} &= \frac{\partial l_{2i}(\beta, \delta)}{\partial \delta_j} = \sum_{\substack{i: \\ y \neq 0}} \left\{ \frac{\partial l_{i2}(\mu_i, \sigma_i)}{\partial \sigma_j} \frac{\sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} \right\} = \\ & \sum_{\substack{i: \\ y \neq 0}} \left\{ \left[\frac{2}{\sigma_i^3} \left(\psi\left(\frac{1}{\sigma_i^2}\right) - \log(y_i) + \log(\sigma_i^2) + \log(\mu_i) - 1 + y_i/\mu \right) \right] \frac{\partial \mu_i}{\partial \eta_{i2}} v_{ij} \right\} = \\ & \sum_{\substack{i: \\ y \neq 0}} \left\{ \left[\frac{2}{\sigma_i^3} \left(\psi\left(\frac{1}{\sigma_i^2}\right) + \log\left(\frac{\mu_i}{y_i}\right) + \frac{(y_i - \mu_i)}{\mu_i} + \log(\sigma_i^2) \right) \right] \frac{\partial \mu_i}{\partial \eta_{i2}} v_{ij} \right\} = \\ & \sum_{\substack{i: \\ y \neq 0}} \left\{ \left[\frac{2}{\sigma_i^3} (m_i - \sigma_i^*) \right] \frac{\partial \mu_i}{\partial \eta_{i2}} v_{ij} \right\} \end{aligned} \quad (\text{A.3})$$

em que $m_i = \psi\left(\frac{1}{\sigma_i^2}\right)\log\left(\frac{\mu_i}{y_i}\right) + \frac{(y_i - \mu_i)}{\mu_i}$, tal que $\psi\left(\frac{1}{\sigma_i^2}\right) = \frac{\Gamma'\left(\frac{1}{\sigma_i^2}\right)}{\Gamma\left(\frac{1}{\sigma_i^2}\right)}$ e $\sigma_i^* = -\log(\sigma_i^2)$.

Podemos reescrever os vetores

$$\begin{aligned} U_\gamma(\alpha) &= \left(\frac{\partial l(\theta)}{\gamma_1}, \dots, \frac{\partial l(\theta)}{\partial \gamma_{p_1}} \right) \\ U_\beta(\mu) &= \left(\frac{\partial l(\theta)}{\beta_1}, \dots, \frac{\partial l(\theta)}{\partial \beta_{p_2}} \right) \\ U_\delta(\sigma) &= \left(\frac{\partial l(\theta)}{\delta_1}, \dots, \frac{\partial l(\theta)}{\partial \delta_{p_3}} \right) \end{aligned} \quad (\text{A.4})$$

na forma matricial como apresentado na equação (2.14).

A.1.2 Obtenção da matriz de informação de Fisher - modelo ZAGA

A seguir apresentaremos as segundas derivadas do logaritmo da função de verossimilhança em relação a cada um dos parâmetros. Iremos em alguns casos denotar $l_{i1}(\alpha)$ e $l_{i2}(\mu_i, \sigma_i)$ apenas como l_{i1} e l_{i2} respectivamente. Utilizando as derivadas apresentadas na seção anterior temos:

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \gamma^2} &= \frac{\partial^2 l_1(\alpha)}{\partial \gamma_j \gamma_l} = \sum_{i=1}^n \frac{\partial^2 l_{i1}(\alpha)}{\partial \gamma_j \gamma_l} = \sum_{i=1}^n \left[\left(\frac{\partial^2 l_{i1}}{\partial \alpha_i^2} \frac{\partial \alpha_i}{\partial \zeta_i} \frac{\partial \zeta_i}{\partial \gamma_j} \right) \frac{\partial \alpha_i}{\partial \zeta_i} \frac{\partial \zeta_i}{\partial \gamma_j} + \frac{\partial l_{i1}}{\partial \alpha_i} \frac{\partial^2 \alpha_i}{\partial \zeta^2} \frac{\partial \zeta_i}{\partial \gamma_l} \frac{\partial \zeta_i}{\partial \gamma_j} \right] = \\ &= \sum_{i=1}^n \left\{ \left[\left(\frac{-I_{(0)}(y_i)}{\alpha_i^2} - \frac{I_{(0,\infty)}(y_i)}{(1-\alpha_i)^2} \right) \left(\frac{\partial \alpha_i}{\partial \zeta_i} \right)^2 + \left(\frac{I_{(0)}(y_i)}{\alpha_i} - \frac{I_{(0,\infty)}(y_i)}{(1-\alpha_i)} \right) \frac{\partial^2 \alpha_i}{\partial \zeta^2} \right] z_{ij} z_{il} \right\} \end{aligned} \quad (\text{A.5})$$

Temos que

$$\begin{aligned} E \left(-\frac{I_{(0)}(y_i)}{\alpha_i^2} - \frac{I_{(0,\infty)}(y_i)}{(1-\alpha_i)^2} \right) &= -\frac{1}{\alpha_i(1-\alpha_i)} \\ E \left(\frac{I_{(0)}(y_i)}{\alpha_i} - \frac{I_{(0,\infty)}(y_i)}{(1-\alpha_i)} \right) &= 0 \end{aligned} \quad (\text{A.6})$$

Portanto

$$-E \left(\sum_{i=1}^n \frac{\partial^2 l_{i1}(\alpha)}{\partial \gamma_j \gamma_l} \right) = \sum_{i=1}^n \left\{ \frac{1}{\alpha_i(1-\alpha_i)} \left(\frac{\partial \alpha_i}{\partial \zeta_i} \right)^2 z_{ij} z_{il} \right\} \quad (\text{A.7})$$

Para a obtenção dos demais termos, utilizaremos a proposição seguir. A demonstração de um proposição semelhante a essa pode ser vista em Ospina (2008).

Proposição 2. *Sejam y_1, y_2, \dots, y_n variáveis aleatórias independente, tal que, y_i tem distribuição inflacionada no zero, como definida na equação (2.1). Se F é uma função contínua então:*

$$E \left(\sum_{i:y \neq 0} F(y_i) \right) = \sum_{i=1}^n (1 - \alpha_i) E(F(y_i) | I_{(0,\infty)}(y_i = 1)) \quad (\text{A.8})$$

Utilizando as derivadas apresentadas na seção anterior temos:

$$\begin{aligned}
 \frac{\partial^2 l(\theta)}{\partial \beta^2} &= \frac{\partial^2 l_2(\mu, \sigma)}{\partial \beta_j \beta_l} = \sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \beta_j \beta_l} \\
 &= \sum_{i:y \neq 0} \left[\left(\frac{\partial^2 l_{i2}}{\partial \mu_i^2} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} \right) \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} + \frac{\partial l_{i2}}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \frac{\partial \eta_{i1}}{\partial \beta_l} \frac{\partial \eta_{ij}}{\partial \beta_j} \right] \\
 &= \sum_{i:y \neq 0} \left\{ \left[\left(-\frac{2y_i - \mu_i}{\mu_i^3 \sigma_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 + \left(\frac{y_i - \mu_i}{\mu_i^2 \sigma_i^2} \right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \right] x_{ij} x_{il} \right\}
 \end{aligned} \tag{A.9}$$

Temos que:

$$\begin{aligned}
 E\left(\sum_{i:y \neq 0} -\frac{2y_i - \mu_i}{\mu_i^3 \sigma_i^2}\right) &= \left(-\frac{1}{\mu_i^2 \sigma_i^2}\right) \\
 E\left(\sum_{i:y \neq 0} \frac{y_i - \mu_i}{\mu_i^2 \sigma_i^2}\right) &= 0
 \end{aligned} \tag{A.10}$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned}
 -E\left(\sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu, \sigma)}{\partial \beta_j \beta_l}\right) &= -E\left(\sum_{i:y \neq 0} \left\{ \left[\left(-\frac{2y_i - \mu_i}{\mu_i^3 \sigma_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 + \left(\frac{y_i - \mu_i}{\mu_i^2 \sigma_i^2} \right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \right] x_{ij} x_{il} \right\}\right) \\
 &= -\sum_{i=1}^n (1 - \alpha_i) \left\{ E\left(-\frac{2y_i - \mu_i}{\mu_i^3 \sigma_i^2}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} + E\left(\frac{y_i - \mu_i}{\mu_i^2 \sigma_i^2}\right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} x_{ij} x_{il} \right\} \\
 &= \sum_{i=1}^n (1 - \alpha_i) \left\{ \left(\frac{1}{\mu_i^2 \sigma_i^2}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} \right\}
 \end{aligned} \tag{A.11}$$

De forma semelhante temos:

$$\begin{aligned}
 \frac{\partial^2 l(\theta)}{\partial \delta^2} &= \frac{\partial^2 l_2(\mu, \sigma)}{\partial \delta_j \delta_l} = \sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \delta_j \delta_l} \\
 &= \sum_{i:y \neq 0} \left[\left(\frac{\partial^2 l_{i2}}{\partial \sigma_i \sigma_i} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} \right) \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} + \frac{\partial l_{i2}}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2} \frac{\partial \eta_{i2}}{\partial \delta_l} \frac{\partial \delta_{ij}}{\partial \delta_j} \right] \\
 &= \sum_{i:y \neq 0} \left\{ \left[\left(\frac{-6}{\sigma_i^4} \left[\psi\left(\frac{1}{\sigma_i^2}\right) + \log\left(\frac{\mu_i}{y_i}\right) + \left(\frac{y_i - \mu_i}{\mu_i}\right) + \log(\sigma_i^2) \right] + \frac{4}{\sigma_i^6} \left[-\psi'\left(\frac{1}{\sigma_i^2}\right) + \sigma_i^2 \right] \right) \left(\frac{\partial \sigma_i}{\partial \eta_{i2}}\right)^2 \right. \right. \\
 &\quad \left. \left. + \frac{2}{\sigma_i^3} \left(\psi\left(\frac{1}{\sigma_i^2}\right) - \log(y_i) + \log(\sigma_i^2) + \log(\mu_i) + \frac{y_i - \mu_i}{\mu_i} \right) \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2} \right] v_{ij} v_{il} \right\} \\
 &= \sum_{i:y \neq 0} \left\{ \left[a_i \left(\frac{\partial \sigma_i}{\partial \eta_{i2}}\right)^2 + b_i \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2} \right] v_{ij} v_{il} \right\}
 \end{aligned} \tag{A.12}$$

$$\begin{aligned}
 \text{em que } a_i &= \frac{-6}{\sigma_i^4} \left[\psi\left(\frac{1}{\sigma_i^2}\right) + \log\left(\frac{\mu_i}{y_i}\right) + \left(\frac{y_i - \mu_i}{\mu_i}\right) + \log(\sigma_i^2) \right] + \frac{4}{\sigma_i^6} \left[-\psi'\left(\frac{1}{\sigma_i^2}\right) + \sigma_i^2 \right] \text{ e} \\
 b_i &= \frac{2}{\sigma_i^3} \left(\psi\left(\frac{1}{\sigma_i^2}\right) - \log(y_i) + \log(\sigma_i^2) + \log(\mu_i) + \frac{y_i - \mu_i}{\mu_i} \right) \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2}
 \end{aligned}$$

Temos que:

$$\begin{aligned}
 E(\log(y)) &= \psi\left(\frac{1}{\sigma_i^2}\right) - \ln\left(\frac{1}{\mu_i \sigma_i^2}\right) \text{ (Maiores informações em Johnson et al. (1994))} \\
 E\left(\sum_{i:y \neq 0} a\right) &= \frac{4}{\sigma_i^6} \left(-\psi'\left(\frac{1}{\sigma_i^2}\right) + \sigma_i^2 \right) \\
 E\left(\sum_{i:y \neq 0} \frac{y_i - \mu_i}{\mu_i^3 \sigma_i^2}\right) &= 0
 \end{aligned} \tag{A.13}$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned}
 -E \left(\sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu, \sigma)}{\partial \delta_j \partial \delta_i} \right) &= -E \left(\sum_{i:y \neq 0} \left\{ \left[(a) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 + (b) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \right] x_{ij} x_{il} \right\} \right) \\
 &= -\sum_{i=1}^n (1 - \alpha_i) \left\{ E(a_i) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 x_{ij} x_{il} + E(b_i) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} x_{ij} x_{il} \right\} \\
 &= \sum_{i=1}^n (1 - \alpha_i) \left\{ \frac{4}{\sigma_i^6} \left(\psi' \left(\frac{1}{\sigma_i^2} \right) - \sigma_i^2 \right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 x_{ij} x_{il} \right\}
 \end{aligned} \tag{A.14}$$

De forma análoga

$$\frac{\partial^2 l(\theta)}{\partial \beta_j \partial \delta_i} = \frac{\partial^2 l_2(\mu, \sigma)}{\partial \beta_j \partial \delta_i} = \sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \beta_j \partial \delta_i} = \sum_{i:y \neq 0} \left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i1}}{\partial \beta_i} \frac{\partial \eta_{i2}}{\partial \delta_j} \right) \tag{A.15}$$

Temos que:

$$\frac{\partial^2 l(\theta)}{\partial \beta_j \partial \delta_i} = -\frac{2(y_i - \mu_i)}{\mu_i^2 \sigma_i^3} \tag{A.16}$$

E em decorrência da ortogonalidade dos parâmetro μ_i e σ_i temos que

$$E \left(\sum_{i:y \neq 0} -\frac{2(y_i - \mu_i)}{\mu_i^2 \sigma_i^3} \right) = 0 \tag{A.17}$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned}
 -E \left(\sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \beta_j \partial \delta_i} \right) &= -E \left(\sum_{i:y \neq 0} \left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i1}}{\partial \beta_i} \frac{\partial \eta_{i2}}{\partial \delta_j} \right) \right) \\
 &= \sum_{i=1}^n (1 - \alpha_i) - E \left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} x_{ij} v_{il} \mid I_{(0, \infty)}(y_i) = 1 \right) = 0
 \end{aligned} \tag{A.18}$$

A.2 Modelo ZAIG

A.2.1 Obtenção da função escore para o modelo ZAIG

Os termos da equação (2.22), como dito anteriormente, são obtidos derivando o logaritmo da função de verossimilhança (equação (2.21)) em relação aos parâmetros. O primeiro termo da função escore, associado ao parâmetro γ_j , para o modelo ZAIG é obtido de forma similar ao vetor U_γ do modelo ZAGA (Seção (A.1.1)) e assim omitiremos os resultados desse termo. O segundo termo da equação (2.22) é dado por:

$$U_\beta = \frac{\partial l(\theta)}{\partial \beta_j} = \frac{\partial l_{2i}(\beta, \delta)}{\partial \beta_j} = \sum_{i:y \neq 0} \left\{ \frac{\partial l_{i2}(\mu_i, \sigma_i)}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} \right\} = \sum_{i:y \neq 0} \left\{ \left[\frac{(y_i - \mu_i)}{\mu_i^3 \sigma_i^2} \right] \frac{\partial \mu_i}{\partial \eta_{i1}} x_{ij} \right\} \tag{A.19}$$

Da mesma maneira da forma anteriores, o último termo da equação (2.22) é dado por:

$$\begin{aligned}
 U_\delta &= \frac{\partial l(\theta)}{\partial \delta_j} = \frac{\partial l_{2i}(\beta, \delta)}{\partial \delta_j} = \sum_{i:y \neq 0} \left\{ \frac{\partial l_{i2}(\mu_i, \sigma_i)}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} \right\} = \sum_{i:y \neq 0} \left\{ \left[\frac{(y_i - \mu_i)^2}{(y_i \mu_i^2 \sigma_i^2)} - \frac{1}{\sigma_i} \right] \frac{\partial \sigma_i}{\partial \eta_{i2}} v_{ij} \right\} = \\
 &\sum_{i:y \neq 0} \left\{ \left[\frac{1}{\sigma_i^3} (s_i - \sigma_i^2) \right] \frac{\partial \sigma_i}{\partial \eta_{i2}} v_{ij} \right\}
 \end{aligned} \tag{A.20}$$

em que $s_i = \frac{(y_i - \mu_i)^2}{y_i \mu_i^2}$.

Os vetores

$$\begin{aligned} U_\gamma(\alpha) &= \left(\frac{\partial l(\theta)}{\gamma_1}, \dots, \frac{\partial l(\theta)}{\partial \gamma_{p_1}} \right) \\ U_\beta(\mu) &= \left(\frac{\partial l(\theta)}{\beta_1}, \dots, \frac{\partial l(\theta)}{\partial \beta_{p_2}} \right) \\ U_\delta(\sigma) &= \left(\frac{\partial l(\theta)}{\delta_1}, \dots, \frac{\partial l(\theta)}{\partial \delta_{p_3}} \right) \end{aligned} \quad (\text{A.21})$$

são apresentados na forma matricial na equação (2.23).

A.2.2 Obtenção da matriz de informação de Fisher - modelo ZAIG

Iremos encontrar a matriz de informação de Fisher para o modelo ZAIG de forma similar de como obtemos a matriz de informação de Fisher para o modelo ZAGA. Omitimos os resultados da informação de Fisher para o parâmetro α_i , por ser obtido da mesma maneira para o modelo ZAGA. A seguir apresentaremos as segundas derivadas do logarítmico da função de verossimilhança e os termos da matriz de informação de Fisher.

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \beta^2} &= \frac{\partial^2 l_2(\mu, \sigma)}{\partial \beta_j \beta_l} \\ &= \sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \beta_j \beta_l} = \sum_{i:y \neq 0} \left[\left(\frac{\partial^2 l_{i2}}{\partial \mu_i^2} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} \right) \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \eta_{i1}}{\partial \beta_j} + \frac{\partial l_{i2}}{\partial \mu_i} \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \frac{\partial \eta_{i1}}{\partial \beta_l} \frac{\partial \eta_{ij}}{\partial \beta_j} \right] \\ &= \sum_{i:y \neq 0} \left\{ \left[\left(\frac{2\mu_i - 3y_i}{\mu_i^4 \sigma_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 + \left(\frac{y_i - \mu_i}{\mu_i^3 \sigma_i^2} \right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \right] x_{ij} x_{il} \right\} \end{aligned} \quad (\text{A.22})$$

Temos que:

$$\begin{aligned} E\left(\sum_{i:y \neq 0} \frac{2\mu_i - 3y_i}{\mu_i^4 \sigma_i^2}\right) &= -\frac{1}{\mu_i^3 \sigma_i^2} \\ E\left(\sum_{i:y \neq 0} \frac{y_i - \mu_i}{\mu_i^3 \sigma_i^2}\right) &= 0 \end{aligned} \quad (\text{A.23})$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned} -E\left(\sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu, \sigma)}{\partial \beta_j \beta_l}\right) &= -E\left(\sum_{i:y \neq 0} \left\{ \left[\left(\frac{2\mu_i - 3y_i}{\mu_i^4 \sigma_i^2} \right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}} \right)^2 + \left(\frac{y_i - \mu_i}{\mu_i^3 \sigma_i^2} \right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} \right] x_{ij} x_{il} \right\}\right) \\ &= -\sum_{i=1}^n (1 - \alpha_i) \left\{ E\left(\frac{2\mu_i - 3y_i}{\mu_i^4 \sigma_i^2}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} + E\left(\frac{y_i - \mu_i}{\mu_i^3 \sigma_i^2}\right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} x_{ij} x_{il} \right\} \\ &= \sum_{i=1}^n (1 - \alpha_i) \left\{ \left(\frac{1}{\mu_i^3 \sigma_i^2}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} \right\} \end{aligned} \quad (\text{A.24})$$

De forma semelhante temos:

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \delta^2} &= \frac{\partial^2 l_2(\mu, \sigma)}{\partial \delta_j \delta_l} = \sum_{i:y \neq 0} \frac{\partial^2 l_{i2}(\mu_i, \sigma_i)}{\partial \delta_j \delta_l} \\ &= \sum_{i:y \neq 0} \left[\left(\frac{\partial^2 l_{i2}}{\partial \sigma_i \sigma_i} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} \right) \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i2}}{\partial \delta_j} + \frac{\partial l_{i2}}{\partial \sigma_i} \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2} \frac{\partial \eta_{i2}}{\partial \delta_l} \frac{\partial \delta_{ij}}{\partial \delta_j} \right] \\ &= \sum_{i:y \neq 0} \left\{ \left[\left(\frac{1}{\sigma_i^2} - \frac{3(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^4} \right) \left(\frac{\partial \sigma_i}{\partial \eta_{i2}} \right)^2 + \left(-\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^3} \right) \frac{\partial^2 \sigma_i}{\partial \eta_{i2}^2} \right] v_{ij} v_{il} \right\} \end{aligned} \quad (\text{A.25})$$

Temos que:

$$\begin{aligned} E\left(\sum_{i:y \neq 0} \left(\frac{1}{\sigma_i^2} - \frac{3(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^4}\right)\right) &= \left(-\frac{2}{\sigma_i^2}\right) \\ E\left(\sum_{i:y \neq 0} \left(-\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^3}\right)\right) &= 0 \end{aligned} \quad (\text{A.26})$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned} &-E\left(\sum_{i:y \neq 0} \frac{\partial^2 l_2(\mu, \sigma)}{\partial \delta_j \partial \delta_l}\right) = \\ &-E\left(\sum_{i:y \neq 0} \left\{ \left[\left(\frac{1}{\sigma_i^2} - \frac{3(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^4}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 + \left(-\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^3}\right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2}\right] x_{ij} x_{il} \right\}\right) \\ &= -\sum_{i=1}^n (1 - \alpha_i) \left\{ E\left(\frac{1}{\sigma_i^2} - \frac{3(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^4}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} + E\left(-\frac{1}{\sigma_i} + \frac{(y_i - \mu_i)^2}{y_i \mu_i^2 \sigma_i^3}\right) \frac{\partial^2 \mu_i}{\partial \eta_{i1}^2} x_{ij} x_{il} \right\} \\ &= \sum_{i=1}^n (1 - \alpha_i) \left\{ \left(-\frac{2}{\sigma_i^2}\right) \left(\frac{\partial \mu_i}{\partial \eta_{i1}}\right)^2 x_{ij} x_{il} \right\} \end{aligned} \quad (\text{A.27})$$

De forma análoga

$$\frac{\partial^2 l(\theta)}{\partial \beta_j \partial \delta_l} = \frac{\partial^2 l_2(\mu, \sigma)}{\partial \beta_j \partial \delta_l} = \sum_{i:y \neq 0} \frac{\partial^2 l_2(\mu_i, \sigma_i)}{\partial \beta_j \partial \delta_l} = \sum_{i:y \neq 0} \left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i1}}{\partial \beta_i} \frac{\partial \eta_{i2}}{\partial \delta_j} \right) \quad (\text{A.28})$$

Temos que:

$$\frac{\partial^2 l(\theta)}{\partial \beta_j \partial \delta_l} = -\frac{2(y_i - \mu_i)}{\mu_i^3 \sigma_i^3} \quad (\text{A.29})$$

E em decorrência da ortogonalidade dos parâmetro μ_i e σ temos que

$$E\left(\sum_{i:y \neq 0} -\frac{2(y_i - \mu_i)}{\mu_i^3 \sigma_i^3}\right) = 0 \quad (\text{A.30})$$

então, utilizando o resultado da proposição (2), temos

$$\begin{aligned} &-E\left(\sum_{i:y \neq 0} \frac{\partial^2 l_2(\mu_i, \sigma_i)}{\partial \beta_j \partial \delta_l}\right) = -E\left(\sum_{i:y \neq 0} \left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} \frac{\partial \eta_{i1}}{\partial \beta_i} \frac{\partial \eta_{i2}}{\partial \delta_j}\right)\right) \\ &= \sum_{i=1}^n (1 - \alpha_i) - E\left(\frac{\partial^2 l_{i2}}{\partial \mu_i \sigma_i} \frac{\partial \mu_i}{\partial \eta_{i1}} \frac{\partial \sigma_i}{\partial \eta_{i2}} x_{ij} v_{il} | I(0, \infty)(y_i) = 1\right) = 0 \end{aligned} \quad (\text{A.31})$$

Apêndice B

Tabelas dos estudos de simulação no
modelo gama

Tabela B.1: Resultados da simulação para o modelo gama - cenário I.

n	Média			Desvio			Pearson			Quantílico						
	média	var	AD	assim	curtose	AD	média	var	AD	assim	curtose	AD				
1	67,74	-0,05	1,03	-0,05	-0,40	7,60	-0,02	1,03	0,09	-0,41	5,56	-0,01	0,85	-0,05	-0,40	5,71
2	100,31	-0,01	0,98	-0,03	-0,46	2,75	0,02	0,98	0,11	-0,45	3,51	0,02	0,84	-0,03	-0,46	7,70
3	215,70	-0,05	1,02	-0,01	-0,45	8,82	-0,02	1,02	0,13	-0,44	6,58	-0,01	0,83	-0,01	-0,45	7,73
4	115,59	-0,03	0,99	-0,01	-0,45	5,35	0,00	0,98	0,14	-0,44	5,36	0,00	0,88	-0,01	-0,44	4,01
5	131,63	-0,03	0,97	-0,09	-0,42	3,82	0,00	0,96	0,05	-0,43	2,36	0,00	0,88	-0,09	-0,42	4,29
6	200,68	-0,01	0,98	-0,09	-0,34	1,38	0,02	0,97	0,06	-0,37	1,88	0,02	0,80	-0,09	-0,34	13,26
7	33,84	-0,02	1,00	-0,10	-0,39	2,72	0,01	1,00	0,03	-0,42	2,11	0,02	0,70	-0,10	-0,39	33,22
8	80,06	-0,02	0,99	-0,03	-0,37	2,63	0,00	0,99	0,11	-0,37	2,48	0,01	0,83	-0,03	-0,37	7,54
9	48,23	-0,02	0,99	-0,06	-0,43	3,08	0,00	0,99	0,07	-0,44	2,84	0,01	0,65	-0,06	-0,43	45,32
10	52,03	-0,03	0,99	-0,04	-0,37	3,21	0,00	0,98	0,09	-0,36	1,86	0,01	0,72	-0,04	-0,37	25,82
11	48,73	-0,04	1,02	-0,06	-0,42	6,63	-0,01	1,01	0,08	-0,44	4,85	0,00	0,85	-0,06	-0,42	5,35
12	114,40	-0,03	1,03	-0,01	-0,43	6,16	0,00	1,03	0,11	-0,43	5,60	0,01	0,64	-0,01	-0,43	49,03
13	194,27	-0,05	1,03	-0,02	-0,40	8,51	-0,02	1,02	0,13	-0,40	6,92	-0,01	0,85	-0,02	-0,40	5,93
14	103,73	-0,03	1,00	-0,02	-0,43	4,59	0,00	1,00	0,13	-0,40	3,50	0,00	0,93	-0,02	-0,43	1,57
15	208,92	-0,02	1,04	-0,05	-0,47	6,09	0,01	1,04	0,09	-0,46	6,06	0,02	0,79	-0,05	-0,47	12,50
Média	114,39	-0,03	1,00	-0,04	-0,42	4,89	0,00	1,00	0,09	-0,42	4,10	0,01	0,80	-0,04	-0,42	15,26
Desv.Pad	63,28	0,01	0,02	0,03	0,04	2,33	0,01	0,02	0,03	0,03	1,82	0,01	0,09	0,03	0,04	15,52
N	Média			Anscombe			Williams									
	média	var	AD	assim	curtose	AD	média	var	AD	assim	curtose	AD				
1	67,74	-0,04	0,85	-0,05	-0,40	9,19	-0,04	1,03	-0,02	-0,40	6,93	-0,01	0,98	-0,01	-0,46	2,60
2	100,31	-0,01	0,84	-0,03	-0,46	6,47	-0,04	1,02	0,02	-0,45	8,07	-0,03	0,99	0,01	-0,45	5,13
3	215,70	-0,04	0,83	-0,01	-0,45	11,78	-0,03	0,97	-0,08	-0,42	3,51	-0,01	0,98	-0,06	-0,34	1,18
4	115,59	-0,03	0,88	-0,01	-0,45	6,27	-0,01	1,00	-0,06	-0,40	2,16	-0,02	0,99	-0,01	-0,37	2,32
5	131,63	-0,03	0,88	-0,09	-0,42	5,32	-0,01	0,99	-0,02	-0,44	2,62	-0,02	0,99	-0,02	-0,44	2,62
6	200,68	-0,01	0,80	-0,09	-0,34	11,45	-0,01	0,99	-0,01	-0,37	2,50	-0,02	0,99	-0,01	-0,37	2,50
7	33,84	-0,02	0,70	-0,10	-0,39	32,41	-0,02	0,99	-0,01	-0,43	6,05	-0,03	1,02	-0,04	-0,43	6,05
8	80,06	-0,02	0,83	-0,03	-0,37	8,24	-0,02	0,99	-0,01	-0,43	5,57	-0,04	1,03	0,03	-0,43	5,57
9	48,23	-0,02	0,65	-0,06	-0,43	45,14	-0,02	0,99	-0,01	-0,41	7,94	-0,04	1,02	0,01	-0,41	7,94
10	52,03	-0,03	0,72	-0,04	-0,38	26,67	-0,03	1,02	-0,04	-0,43	4,36	-0,03	1,00	-0,01	-0,43	4,36
11	48,73	-0,04	0,85	-0,06	-0,43	7,87	-0,02	1,03	0,03	-0,43	5,73	-0,01	1,04	-0,01	-0,47	5,73
12	114,40	-0,02	0,64	-0,01	-0,43	49,88	-0,02	1,03	0,03	-0,43	4,36	-0,01	1,04	-0,01	-0,47	5,73
13	194,27	-0,04	0,85	-0,02	-0,40	9,88	-0,04	1,02	0,01	-0,41	7,94	-0,03	1,00	-0,01	-0,43	4,36
14	103,73	-0,03	0,93	-0,02	-0,43	3,69	-0,03	1,00	-0,01	-0,43	4,36	-0,01	1,04	-0,01	-0,47	5,73
15	208,92	-0,01	0,79	-0,05	-0,47	11,69	-0,01	1,04	0,09	-0,46	6,06	0,02	0,79	-0,05	-0,47	12,50
Média	114,39	-0,03	0,80	-0,04	-0,42	16,39	-0,02	1,00	-0,02	-0,42	4,45	0,01	0,02	0,03	0,04	2,23
Desv.Pad	63,28	0,01	0,09	0,03	0,04	14,87	0,01	0,02	0,03	0,04	1,87	0,01	0,02	0,03	0,04	2,23

Tabela B.2: Resultados da simulação para o modelo gama - cenário II.

n	Desvio			Pearson			Quantilico				
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	67,74	-0,03	1,03	-0,03	-0,40	4,83	-0,01	1,02	0,05	-0,40	3,95
2	100,31	-0,01	1,02	-0,04	-0,47	4,43	0,00	1,02	0,03	-0,47	4,28
3	215,70	-0,03	0,96	0,02	-0,38	4,60	-0,02	0,96	0,09	-0,38	4,03
4	115,59	-0,02	1,01	-0,02	-0,45	3,97	0,00	1,01	0,05	-0,45	3,61
5	131,63	-0,02	0,96	0,00	-0,32	1,19	0,00	0,95	0,08	-0,32	1,18
6	200,68	0,00	0,98	-0,02	-0,41	2,13	0,01	0,98	0,05	-0,40	2,49
7	33,84	-0,02	1,00	0,00	-0,41	3,66	-0,01	1,00	0,06	-0,41	3,54
8	80,06	0,01	1,04	0,00	-0,38	3,59	0,02	1,04	0,08	-0,38	4,70
9	48,23	-0,02	1,04	-0,06	-0,47	6,40	0,00	1,04	0,00	-0,48	5,91
10	52,03	-0,03	1,04	-0,04	-0,54	10,98	-0,02	1,03	0,02	-0,55	9,79
11	48,73	0,00	1,01	0,01	-0,44	3,35	0,01	1,01	0,08	-0,43	3,76
12	114,40	0,00	0,96	0,01	-0,43	2,22	0,01	0,96	0,07	-0,43	2,61
13	194,27	-0,01	0,99	-0,02	-0,43	2,73	0,00	0,99	0,05	-0,42	2,57
14	103,73	0,00	0,98	-0,02	-0,40	1,78	0,01	0,99	0,06	-0,40	2,38
15	208,92	-0,03	1,00	0,00	-0,44	4,85	-0,01	1,00	0,07	-0,44	4,33
Média	114,39	-0,01	1,00	-0,01	-0,42	4,05	0,00	1,00	0,06	-0,42	3,94
Desv.Pad	63,28	0,01	0,03	0,02	0,05	2,36	0,01	0,03	0,02	0,05	1,98

N	Anscombe			Williams							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	67,74	-0,03	0,84	-0,03	-0,40	7,47	-0,03	1,02	-0,01	-0,40	4,60
2	100,31	-0,01	0,87	-0,04	-0,47	4,55	-0,01	1,02	-0,03	-0,47	4,34
3	215,70	-0,03	0,78	0,02	-0,38	17,45	-0,03	0,96	0,04	-0,38	4,41
4	115,59	-0,02	0,89	-0,02	-0,45	3,29	-0,02	1,01	-0,02	-0,45	3,88
5	131,63	-0,01	0,87	0,00	-0,32	4,75	-0,01	0,96	0,01	-0,32	1,14
6	200,68	0,00	0,81	-0,02	-0,41	10,13	0,00	0,98	-0,01	-0,41	2,13
7	33,84	-0,02	0,70	0,00	-0,41	32,28	-0,02	1,00	0,02	-0,41	3,53
8	80,06	0,01	0,87	0,00	-0,38	3,92	0,01	1,04	0,01	-0,38	3,70
9	48,23	-0,01	0,68	-0,06	-0,47	34,13	-0,01	1,04	-0,04	-0,47	6,14
10	52,03	-0,03	0,76	-0,04	-0,54	18,46	-0,03	1,04	-0,02	-0,55	10,57
11	48,73	0,00	0,85	0,01	-0,44	5,74	0,00	1,01	0,02	-0,44	3,35
12	114,40	0,00	0,59	0,01	-0,43	65,34	0,00	0,96	0,03	-0,43	2,28
13	194,27	-0,01	0,82	-0,02	-0,43	9,17	-0,01	0,99	-0,01	-0,43	2,63
14	103,73	0,00	0,91	-0,02	-0,40	1,86	0,00	0,98	-0,01	-0,40	1,79
15	208,92	-0,02	0,77	0,00	-0,44	17,31	-0,02	1,00	0,02	-0,44	4,64
Média	114,39	-0,01	0,80	-0,01	-0,42	15,72	-0,01	1,00	0,00	-0,43	3,94
Desv.Pad	63,28	0,01	0,09	0,02	0,05	17,03	0,01	0,03	0,02	0,05	2,25

Tabela B.3: Resultados da simulação para o modelo gama - cenário III.

n	Média			Desvio			Pearson			Quantílico		
	média	var	AD	média	var	AD	média	var	AD	média	var	AD
1	67,74	1,06	50,58	-0,21	-0,25	0,00	0,98	0,49	30,55	0,02	0,86	10,62
2	100,31	1,06	43,43	-0,15	-0,39	0,02	1,01	0,51	33,25	0,04	0,89	9,63
3	215,70	1,09	72,07	-0,18	-0,27	-0,02	1,01	0,51	42,38	0,00	0,87	6,08
4	115,59	1,09	51,66	-0,17	-0,37	0,02	1,03	0,53	34,75	0,03	0,96	6,72
5	131,63	1,05	37,80	-0,17	-0,39	0,03	0,99	0,51	31,33	0,04	0,94	9,71
6	200,68	1,08	42,28	-0,20	-0,27	0,02	1,01	0,49	29,06	0,04	0,87	12,06
7	33,84	1,10	50,93	-0,30	-0,12	-0,01	1,00	0,37	20,28	0,03	0,76	30,55
8	80,06	1,09	52,57	-0,21	-0,28	0,01	1,01	0,49	31,03	0,02	0,90	7,21
9	48,23	1,09	71,52	-0,26	-0,24	-0,04	0,98	0,38	30,90	0,02	0,71	35,59
10	52,03	1,10	36,26	-0,32	-0,14	0,02	1,00	0,38	17,97	0,05	0,80	29,27
11	48,73	1,11	54,96	-0,24	-0,32	0,01	1,02	0,45	27,80	0,02	0,92	8,37
12	114,40	1,12	79,86	-0,35	0,03	-0,06	0,98	0,32	32,38	0,01	0,68	47,35
13	194,27	1,08	60,18	-0,21	-0,30	-0,01	1,00	0,50	32,36	0,01	0,89	6,60
14	103,73	1,06	49,16	-0,14	-0,43	0,02	1,01	0,55	37,23	0,02	0,97	5,69
15	208,92	1,04	58,56	-0,21	-0,29	-0,02	0,95	0,44	29,09	0,02	0,78	19,06
Média	114,39	1,08	54,12	-0,22	-0,27	0,00	1,00	0,46	30,69	0,03	0,85	16,30
Desv.Pad	63,28	0,02	12,65	0,06	0,12	0,02	0,02	0,07	5,93	0,01	0,09	13,10
n	Média			Anscombe			Williams					
	média	var	AD	média	var	AD	média	var	AD			
1	67,74	0,86	44,25	-0,14	-0,30	-0,12	1,05	-0,09	39,50	-0,32	0,32	
2	100,31	0,89	35,32	-0,13	-0,43	-0,11	1,05	-0,05	35,33	-0,42	0,42	
3	215,70	0,87	58,29	-0,16	-0,32	-0,15	1,08	-0,05	58,35	-0,33	0,33	
4	115,59	0,95	40,44	-0,16	-0,41	-0,13	1,08	-0,10	44,19	-0,40	0,40	
5	131,63	0,94	31,67	-0,15	-0,42	-0,11	1,04	-0,11	32,57	-0,41	0,41	
6	200,68	0,87	35,17	-0,19	-0,32	-0,11	1,07	-0,08	32,25	-0,33	0,33	
7	33,84	0,76	53,50	-0,28	-0,19	-0,11	1,07	-0,10	32,13	-0,28	0,28	
8	80,06	0,90	41,79	-0,19	-0,33	-0,13	1,08	-0,10	41,97	-0,34	0,34	
9	48,23	0,71	75,37	-0,24	-0,30	-0,13	1,06	-0,05	47,96	-0,38	0,38	
10	52,03	0,80	36,79	-0,30	-0,20	-0,09	1,08	-0,14	22,18	-0,29	0,29	
11	48,73	0,92	40,95	-0,22	-0,36	-0,13	1,10	-0,13	43,18	-0,38	0,38	
12	114,40	0,68	90,72	-0,33	-0,05	-0,14	1,07	-0,10	52,21	-0,23	0,23	
13	194,27	0,89	48,57	-0,19	-0,35	-0,14	1,07	-0,09	47,79	-0,35	0,35	
14	103,73	0,97	41,97	-0,12	-0,47	-0,13	1,06	-0,09	44,41	-0,45	0,45	
15	208,92	0,78	57,31	-0,20	-0,34	-0,13	1,02	-0,07	42,94	-0,37	0,37	
Média	114,39	0,85	48,81	-0,20	-0,32	-0,12	1,07	-0,09	41,13	-0,35	0,35	
Desv.Pad	63,28	0,01	16,26	0,06	0,11	0,02	0,02	0,03	9,13	0,06	0,06	

Tabela B.4: Resultados da simulação para o modelo gama - cenário IV.

n	Desvio				Pearson				Quantilico							
	Média	média	var	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD	
1	0,15	-0,04	1,02	-0,08	-0,46	6,99	-0,01	1,01	0,05	-0,47	4,80	0,00	0,84	-0,08	-0,45	7,10
2	0,18	-0,05	0,98	-0,01	-0,41	7,41	-0,02	0,98	0,14	-0,39	5,34	-0,01	0,84	-0,01	-0,40	6,80
3	0,35	-0,03	1,00	-0,08	-0,42	4,35	0,00	0,99	0,06	-0,43	2,92	0,01	0,81	-0,08	-0,42	11,07
4	0,22	-0,02	1,01	-0,04	-0,42	3,98	0,01	1,01	0,10	-0,42	4,05	0,01	0,90	-0,04	-0,42	3,14
5	0,23	-0,04	1,01	-0,09	-0,35	4,59	-0,01	1,00	0,06	-0,37	2,53	0,00	0,92	-0,09	-0,35	2,13
6	0,34	-0,04	0,99	-0,06	-0,40	6,17	-0,02	0,98	0,08	-0,41	3,94	-0,01	0,81	-0,06	-0,40	10,02
7	0,08	-0,02	1,00	-0,10	-0,41	3,87	0,00	1,00	0,03	-0,42	2,52	0,01	0,70	-0,10	-0,41	32,76
8	0,17	-0,01	1,01	-0,09	-0,41	3,30	0,02	1,00	0,05	-0,42	3,18	0,02	0,85	-0,09	-0,41	8,56
9	0,12	-0,02	0,97	-0,09	-0,38	1,70	0,01	0,97	0,03	-0,41	1,62	0,02	0,64	-0,09	-0,38	52,04
10	0,10	-0,03	1,01	-0,04	-0,41	5,17	-0,01	1,00	0,09	-0,42	4,11	0,00	0,74	-0,04	-0,41	21,62
11	0,11	-0,04	0,99	-0,05	-0,45	7,57	-0,01	0,99	0,09	-0,45	5,41	-0,01	0,83	-0,05	-0,45	7,62
12	0,18	0,00	1,02	-0,05	-0,44	3,81	0,03	1,02	0,07	-0,45	5,43	0,03	0,63	-0,05	-0,44	55,67
13	0,31	-0,02	1,01	-0,05	-0,39	3,59	0,01	1,01	0,09	-0,40	3,43	0,01	0,83	-0,05	-0,39	7,74
14	0,19	-0,03	1,04	-0,02	-0,44	8,30	0,00	1,03	0,13	-0,43	7,52	0,00	0,96	-0,02	-0,44	2,58
15	0,36	-0,03	1,01	-0,05	-0,40	3,79	0,00	1,01	0,09	-0,41	3,17	0,01	0,77	-0,05	-0,40	15,96
Média	0,21	-0,03	1,00	-0,06	-0,41	4,97	0,00	1,00	0,08	-0,42	4,00	0,01	0,80	-0,06	-0,41	16,32
Desv.Pad	0,10	0,01	0,02	0,03	0,03	1,89	0,01	0,02	0,03	0,03	1,51	0,01	0,10	0,03	0,03	17,19
n	Anscombe				Williams											
	Média	média	var	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD	
1	0,15	-0,04	0,84	-0,08	-0,46	9,33	-0,03	1,02	-0,06	-0,46	6,30	-0,03	1,02	-0,06	-0,46	6,30
2	0,18	-0,05	0,84	-0,01	-0,41	11,45	-0,05	0,98	0,01	-0,41	6,84	-0,05	0,98	0,01	-0,41	6,84
3	0,35	-0,03	0,81	-0,08	-0,42	11,76	-0,02	1,00	-0,06	-0,42	3,76	-0,02	1,00	-0,06	-0,42	3,76
4	0,22	-0,02	0,90	-0,04	-0,42	3,69	-0,02	1,01	-0,03	-0,42	3,77	-0,02	1,01	-0,03	-0,42	3,77
5	0,23	-0,03	0,92	-0,09	-0,35	3,98	-0,03	1,01	-0,08	-0,35	4,22	-0,03	1,01	-0,08	-0,35	4,22
6	0,34	-0,04	0,81	-0,06	-0,41	13,31	-0,04	0,99	-0,03	-0,41	5,47	-0,04	0,99	-0,03	-0,41	5,47
7	0,08	-0,02	0,70	-0,10	-0,41	32,15	-0,01	1,00	-0,06	-0,41	3,08	-0,01	1,00	-0,06	-0,41	3,08
8	0,17	-0,01	0,85	-0,09	-0,41	6,47	-0,01	1,01	-0,07	-0,41	3,00	-0,01	1,01	-0,07	-0,41	3,00
9	0,12	-0,01	0,64	-0,09	-0,38	50,45	-0,01	0,97	-0,05	-0,38	1,31	-0,01	0,97	-0,05	-0,40	1,31
10	0,10	-0,03	0,74	-0,04	-0,42	23,47	-0,03	1,00	-0,01	-0,42	4,52	-0,03	1,00	-0,01	-0,42	4,52
11	0,11	-0,04	0,83	-0,05	-0,45	10,96	-0,04	0,99	-0,03	-0,45	6,93	-0,04	0,99	-0,03	-0,45	6,93
12	0,18	0,00	0,63	-0,05	-0,44	52,04	0,01	1,02	-0,01	-0,44	4,05	0,01	1,02	-0,01	-0,44	4,05
13	0,31	-0,02	0,83	-0,05	-0,39	7,89	-0,02	1,01	-0,02	-0,39	3,27	-0,02	1,01	-0,02	-0,39	3,27
14	0,19	-0,03	0,96	-0,02	-0,44	5,04	-0,03	1,03	-0,01	-0,44	8,09	-0,03	1,03	-0,01	-0,44	8,09
15	0,36	-0,02	0,77	-0,05	-0,40	16,79	-0,02	1,01	-0,02	-0,40	3,29	-0,02	1,01	-0,02	-0,40	3,29
Média	0,21	-0,03	0,80	-0,06	-0,41	17,25	-0,02	1,00	-0,03	-0,42	4,53	-0,02	1,00	-0,03	-0,42	4,53
Desv.Pad	0,10	0,01	0,10	0,03	0,03	15,74	0,02	0,02	0,03	0,03	1,84	0,02	0,02	0,03	0,03	1,84

Tabela B.5: Resultados da simulação para o modelo gama - cenário V.

n	Desvio			Pearson			Quantilico				
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	110,01	-0,03	1,01	-0,07	-0,49	5,71	0,00	1,01	0,07	-0,49	4,14
2	75,66	-0,05	0,98	-0,01	-0,40	7,97	-0,02	0,98	0,14	-0,38	5,69
3	137,96	-0,03	1,00	-0,05	-0,38	4,12	0,00	0,99	0,09	-0,38	2,81
4	107,90	-0,03	1,02	-0,07	-0,48	5,91	0,00	1,02	0,05	-0,48	4,64
5	140,32	-0,05	1,00	-0,11	-0,32	5,40	-0,02	0,99	0,04	-0,36	2,51
6	132,92	-0,05	0,99	-0,05	-0,43	7,47	-0,02	0,98	0,10	-0,43	4,86
7	104,10	-0,02	0,99	-0,09	-0,37	2,41	0,01	0,99	0,05	-0,38	1,88
8	170,48	-0,02	1,01	-0,08	-0,38	2,64	0,01	1,00	0,04	-0,39	2,00
9	76,57	0,00	1,00	-0,06	-0,40	2,45	0,03	1,00	0,08	-0,41	4,30
10	65,33	-0,03	1,01	-0,05	-0,45	5,89	0,00	1,01	0,09	-0,45	4,61
11	57,42	-0,04	1,00	-0,06	-0,41	5,08	-0,01	1,00	0,07	-0,43	3,52
12	113,39	0,00	1,01	-0,06	-0,44	3,51	0,03	1,01	0,09	-0,47	5,74
13	97,61	-0,02	1,00	-0,04	-0,40	3,81	0,01	1,00	0,12	-0,41	4,23
14	61,04	-0,03	1,02	-0,06	-0,44	6,27	-0,01	1,01	0,08	-0,45	4,72
15	135,80	-0,03	1,03	-0,06	-0,49	8,39	0,00	1,03	0,08	-0,50	6,84
Média	105,77	-0,03	1,00	-0,06	-0,42	5,13	0,00	1,00	0,08	-0,43	4,17
Desv.Pad	33,74	0,02	0,01	0,02	0,05	1,95	0,02	0,01	0,03	0,04	1,42
n	Anscombe			Williams							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	110,01	-0,03	0,91	-0,07	-0,49	4,01	-0,03	1,01	-0,06	-0,49	5,35
2	75,66	-0,05	0,85	-0,01	-0,40	10,70	-0,05	0,98	0,01	-0,40	7,43
3	137,96	-0,03	0,83	-0,05	-0,38	9,88	-0,03	1,00	-0,03	-0,38	3,60
4	107,90	-0,02	0,58	-0,07	-0,48	70,94	-0,01	1,02	-0,02	-0,48	5,01
5	140,32	-0,04	0,84	-0,11	-0,32	10,42	-0,04	1,00	-0,08	-0,33	4,66
6	132,92	-0,05	0,87	-0,05	-0,43	8,83	-0,05	0,99	-0,03	-0,43	6,93
7	104,10	-0,01	0,77	-0,09	-0,37	15,93	-0,01	0,99	-0,06	-0,37	1,96
8	170,48	-0,01	0,59	-0,08	-0,38	67,77	-0,01	1,01	-0,03	-0,39	2,02
9	76,57	0,00	0,82	-0,06	-0,40	8,11	0,01	1,00	-0,04	-0,40	2,48
10	65,33	-0,03	0,77	-0,05	-0,45	16,57	-0,03	1,01	-0,01	-0,45	5,24
11	57,42	-0,03	0,73	-0,06	-0,41	25,81	-0,03	1,00	-0,02	-0,42	4,29
12	113,39	0,00	0,93	-0,06	-0,45	2,00	0,00	1,01	-0,05	-0,45	3,51
13	97,61	-0,02	0,93	-0,04	-0,40	2,81	-0,02	1,00	-0,02	-0,40	3,68
14	61,04	-0,03	0,78	-0,06	-0,44	14,98	-0,03	1,02	-0,03	-0,44	5,57
15	135,80	-0,03	0,83	-0,06	-0,49	9,82	-0,03	1,03	-0,03	-0,49	7,76
Média	105,77	-0,03	0,80	-0,06	-0,42	18,57	-0,02	1,00	-0,03	-0,42	4,63
Desv.Pad	33,74	0,01	0,10	0,02	0,05	21,48	0,02	0,01	0,02	0,05	1,83

Tabela B.6: Resultados da simulação para o modelo gama - cenário VI.

n	Média			Desvio			Pearson			Quantilico			
	média	var	AD	assim	curtose	AD	média	var	AD	assim	curtose	AD	
1	52,29	-0,05	1,03	-0,04	-0,46	9,92	-0,02	1,03	0,11	-0,46	7,05	-0,02	0,94
2	520,94	-0,01	0,97	-0,07	-0,39	1,96	0,01	0,97	0,04	-0,40	1,74	0,02	0,53
3	114,98	-0,04	1,00	0,00	-0,43	7,30	-0,01	1,00	0,15	-0,42	5,95	-0,01	0,90
4	45,85	-0,03	0,99	-0,01	-0,43	5,76	0,00	0,99	0,13	-0,43	5,86	0,00	0,88
5	20,61	-0,03	0,99	-0,07	-0,46	4,89	0,00	0,99	0,06	-0,46	3,90	0,01	0,70
6	47,17	0,00	0,98	-0,07	-0,34	1,56	0,03	0,98	0,09	-0,35	3,26	0,03	0,89
7	109,62	-0,03	0,98	-0,06	-0,38	3,09	0,00	0,98	0,07	-0,38	1,79	0,01	0,71
8	251,15	-0,03	0,99	-0,05	-0,38	4,50	-0,01	0,99	0,09	-0,39	3,42	0,00	0,79
9	46,31	-0,03	0,99	-0,05	-0,51	6,26	0,00	0,99	0,09	-0,50	5,30	0,01	0,90
10	358,85	-0,03	0,97	-0,07	-0,40	3,39	0,00	0,97	0,06	-0,41	2,17	0,01	0,68
11	89,26	-0,05	1,02	-0,05	-0,42	9,17	-0,02	1,02	0,08	-0,44	6,25	-0,01	0,77
12	120,24	-0,01	1,06	-0,06	-0,46	6,64	0,02	1,05	0,08	-0,46	7,30	0,02	0,86
13	30,01	-0,03	1,00	-0,02	-0,38	4,81	0,00	1,00	0,13	-0,38	4,32	0,00	0,86
14	24,11	-0,03	1,01	0,00	-0,44	5,76	0,00	1,01	0,13	-0,41	4,87	0,01	0,74
15	57,08	-0,02	1,03	-0,05	-0,42	3,64	0,02	1,03	0,09	-0,42	4,00	0,02	0,88
Média	125,90	-0,03	1,00	-0,04	-0,42	5,24	0,00	1,00	0,09	-0,42	4,48	0,01	0,80
Desv.Pad	143,30	0,01	0,02	0,03	0,04	2,40	0,01	0,02	0,03	0,04	1,81	0,01	0,11
n	Média			Anscombe			Williams			Quantilico			
	média	var	AD	assim	curtose	AD	média	var	AD	assim	curtose	AD	
1	52,29	-0,05	0,94	-0,04	-0,46	6,76	-0,05	1,03	-0,03	-0,46	9,49	-0,05	0,94
2	520,94	-0,01	0,53	-0,07	-0,39	101,56	0,00	0,97	-0,02	-0,40	1,53	0,00	0,53
3	114,98	-0,04	0,90	0,00	-0,44	6,32	-0,04	1,00	0,02	-0,43	6,95	-0,04	0,90
4	45,85	-0,03	0,88	-0,01	-0,43	6,40	-0,03	0,99	0,01	-0,43	5,54	0,00	0,88
5	20,61	-0,02	0,70	-0,07	-0,46	30,58	-0,02	0,99	-0,03	-0,46	4,24	0,01	0,70
6	47,17	0,00	0,89	-0,07	-0,34	3,05	0,00	0,98	-0,05	-0,34	1,53	0,01	0,89
7	109,62	-0,03	0,71	-0,06	-0,38	30,01	-0,02	0,98	-0,03	-0,38	2,37	-0,02	0,71
8	251,15	-0,03	0,79	-0,04	-0,38	14,62	-0,03	0,99	-0,02	-0,39	3,96	0,00	0,79
9	46,31	-0,03	0,90	-0,05	-0,51	5,01	-0,02	0,99	-0,04	-0,51	5,98	0,01	0,90
10	358,85	-0,03	0,68	-0,07	-0,40	38,02	-0,02	0,97	-0,03	-0,40	2,66	0,01	0,68
11	89,26	-0,05	0,77	-0,05	-0,43	19,72	-0,05	1,02	-0,02	-0,43	8,10	-0,01	0,77
12	120,24	-0,01	0,86	-0,06	-0,46	4,85	0,00	1,05	-0,03	-0,46	6,45	0,02	0,86
13	30,01	-0,03	0,86	-0,02	-0,38	7,09	-0,03	1,00	0,00	-0,38	4,48	0,00	0,86
14	24,11	-0,03	0,74	0,00	-0,44	23,54	-0,02	1,01	0,04	-0,44	5,16	0,01	0,74
15	57,08	-0,01	0,88	-0,05	-0,42	3,30	-0,01	1,03	-0,03	-0,42	3,43	0,02	0,88
Média	125,90	-0,03	0,80	-0,04	-0,42	20,06	-0,02	1,00	-0,02	-0,42	4,79	0,01	0,80
Desv.Pad	143,30	0,01	0,11	0,03	0,04	25,31	0,01	0,02	0,02	0,04	2,35	0,01	0,11

Tabela B.7: Resultados da simulação para o modelo gama - cenário VII.

n	Desvio			Pearson			Quantilico				
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	52,49	-0,03	1,00	-0,05	-0,39	3,95	0,00	0,99	0,10	-0,39	3,07
2	31,31	-0,03	1,02	-0,09	-0,40	5,08	0,00	1,02	0,06	-0,41	3,13
3	23,40	-0,06	1,01	-0,02	-0,41	9,99	-0,02	1,01	0,12	-0,41	7,25
4	33,68	-0,04	1,02	-0,05	-0,40	5,63	-0,01	1,01	0,10	-0,40	4,07
5	28,66	-0,02	0,95	-0,03	-0,39	2,00	0,01	0,95	0,11	-0,38	1,90
6	24,56	-0,03	1,00	-0,04	-0,39	3,51	0,00	1,00	0,10	-0,39	3,06
7	95,25	-0,03	0,99	-0,09	-0,39	3,83	-0,02	0,99	0,01	-0,41	2,22
8	45,28	0,01	0,98	-0,04	-0,44	2,60	0,04	0,98	0,10	-0,45	5,92
9	86,42	-0,05	1,00	0,01	-0,36	8,10	-0,03	1,00	0,11	-0,36	6,69
10	48,48	-0,02	0,97	-0,03	-0,38	2,57	0,00	0,97	0,10	-0,37	2,10
11	62,17	0,00	1,05	-0,09	-0,45	6,03	0,03	1,04	0,06	-0,47	7,32
12	26,97	-0,05	1,00	-0,02	-0,38	8,02	-0,02	1,00	0,12	-0,38	5,99
13	23,69	-0,01	1,00	-0,08	-0,49	4,38	0,02	0,99	0,05	-0,50	4,37
14	32,93	-0,03	1,01	-0,01	-0,39	5,08	0,00	1,01	0,14	-0,39	4,60
15	24,96	-0,05	1,04	-0,02	-0,48	10,60	-0,01	1,03	0,12	-0,47	8,79
Média	42,68	-0,03	1,00	-0,04	-0,41	5,43	0,00	1,00	0,09	-0,41	4,70
Desv.Pad	22,86	0,02	0,02	0,03	0,04	2,67	0,02	0,02	0,04	0,04	2,18
n	Anscombe			Williams							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	52,49	-0,03	0,85	-0,05	-0,39	6,89	-0,03	1,00	-0,03	-0,39	3,56
2	31,31	-0,03	0,90	-0,09	-0,40	4,01	-0,03	1,02	-0,07	-0,40	4,64
3	23,40	-0,05	0,89	-0,02	-0,41	9,16	-0,05	1,01	0,00	-0,41	9,43
4	33,68	-0,04	0,92	-0,05	-0,40	4,32	-0,03	1,02	-0,04	-0,40	5,27
5	28,66	-0,02	0,86	-0,03	-0,39	5,53	-0,02	0,95	-0,02	-0,39	1,80
6	24,56	-0,03	0,88	-0,04	-0,39	4,45	-0,02	1,00	-0,02	-0,39	3,23
7	95,25	-0,02	0,35	-0,09	-0,37	253,84	-0,02	0,99	-0,03	-0,40	2,59
8	45,28	0,01	0,86	-0,04	-0,44	5,29	0,01	0,98	-0,02	-0,44	2,78
9	86,42	-0,03	0,35	0,02	-0,34	253,25	-0,04	1,00	0,07	-0,37	6,99
10	48,48	-0,02	0,74	-0,03	-0,39	21,56	-0,02	0,97	0,00	-0,38	2,13
11	62,17	0,00	0,91	-0,09	-0,45	2,66	0,00	1,05	-0,07	-0,45	5,94
12	26,97	-0,05	0,82	-0,02	-0,38	13,92	-0,04	1,00	0,00	-0,38	7,34
13	23,69	-0,01	0,88	-0,08	-0,49	4,07	-0,01	1,00	-0,07	-0,49	4,15
14	32,93	-0,03	0,93	-0,01	-0,39	3,84	-0,03	1,01	0,00	-0,39	4,87
15	24,96	-0,04	0,90	-0,02	-0,48	7,11	-0,04	1,04	0,00	-0,48	10,09
Média	42,68	-0,03	0,80	-0,04	-0,41	39,99	-0,02	1,00	-0,02	-0,41	4,99
Desv.Pad	22,86	0,02	0,19	0,03	0,04	86,84	0,02	0,02	0,04	0,04	2,55

Apêndice C

Tabelas dos estudos de simulação no modelo gaussiana inversa

Tabela C.1: Resultados da simulação para o modelo gaussiana inversa - cenário I.

n	r_i^{dev}				r_i^{pca}				r_i^{qu}							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	67,74	-0,07	1,00	-0,09	-0,38	12,67	0,00	0,99	0,25	-0,36	8,93	0,02	0,83	-0,09	-0,37	8,80
2	100,31	-0,08	1,00	-0,05	-0,47	18,66	0,01	1,00	0,34	-0,39	16,30	0,02	0,87	-0,06	-0,46	5,58
3	215,70	-0,15	1,03	-0,14	-0,30	52,03	-0,02	1,00	0,48	-0,22	33,36	0,00	0,91	-0,14	-0,29	2,64
4	115,59	-0,09	1,04	-0,10	-0,43	22,84	0,01	1,03	0,35	-0,33	16,23	0,02	0,93	-0,10	-0,42	3,74
5	131,63	-0,09	0,99	-0,10	-0,42	18,42	0,01	0,98	0,36	-0,33	15,08	0,03	0,90	-0,11	-0,41	5,53
6	200,68	-0,11	1,02	-0,15	-0,35	26,86	0,02	1,01	0,44	-0,25	21,47	0,03	0,91	-0,15	-0,33	7,15
7	33,84	-0,05	1,03	-0,17	-0,37	7,94	0,00	1,02	0,03	-0,41	3,50	0,02	0,54	-0,17	-0,37	97,14
8	80,06	-0,08	1,03	-0,09	-0,41	16,05	0,00	1,02	0,28	-0,38	11,64	0,01	0,88	-0,09	-0,41	4,80
9	48,23	-0,07	1,02	-0,11	-0,42	13,17	-0,02	1,00	0,13	-0,44	7,05	0,02	0,57	-0,11	-0,42	79,37
10	52,03	-0,04	1,02	-0,17	-0,35	5,93	0,02	1,01	0,12	-0,42	4,23	0,04	0,70	-0,16	-0,35	37,61
11	48,73	-0,07	1,05	-0,11	-0,44	14,74	0,00	1,04	0,18	-0,43	8,43	0,01	0,88	-0,11	-0,44	4,49
12	114,40	-0,12	1,02	-0,18	-0,25	30,87	-0,03	0,98	0,24	-0,36	15,68	0,00	0,73	-0,19	-0,24	26,94
13	194,27	-0,14	1,03	-0,14	-0,38	42,89	-0,01	0,99	0,44	-0,27	24,82	0,00	0,91	-0,14	-0,37	3,21
14	103,73	-0,10	1,00	-0,07	-0,48	23,52	0,00	0,99	0,35	-0,38	16,63	0,01	0,93	-0,07	-0,47	2,47
15	208,92	-0,13	0,98	-0,15	-0,36	34,76	0,00	0,95	0,42	-0,26	21,12	0,02	0,85	-0,16	-0,35	9,08
Média	114,39	-0,09	1,02	-0,12	-0,39	22,76	0,00	1,00	0,29	-0,35	14,96	0,02	0,82	-0,12	-0,38	19,90
Desv.Pad	63,28	0,03	0,02	0,04	0,06	12,90	0,01	0,02	0,13	0,07	8,12	0,01	0,13	0,04	0,06	29,62
n	r_i^{gms}				r_i^{unl}											
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD					
1	67,74	-0,06	0,82	-0,09	-0,39	17,16	-0,06	1,00	-0,03	-0,39	10,24					
2	100,31	-0,07	0,86	-0,05	-0,48	17,88	-0,07	1,00	0,00	-0,47	16,10					
3	215,70	-0,14	0,88	-0,12	-0,35	46,64	-0,13	1,03	-0,06	-0,33	45,02					
4	115,59	-0,09	0,91	-0,10	-0,45	18,45	-0,08	1,04	-0,05	-0,44	19,84					
5	131,63	-0,08	0,88	-0,10	-0,45	17,94	-0,08	0,99	-0,06	-0,43	15,82					
6	200,68	-0,10	0,88	-0,14	-0,39	24,89	-0,09	1,02	-0,07	-0,37	21,86					
7	33,84	-0,03	0,54	-0,17	-0,37	98,19	-0,03	1,03	-0,07	-0,40	4,85					
8	80,06	-0,07	0,87	-0,09	-0,43	14,71	-0,06	1,03	-0,04	-0,42	13,42					
9	48,23	-0,05	0,56	-0,11	-0,42	84,88	-0,04	1,01	-0,01	-0,44	8,99					
10	52,03	-0,03	0,70	-0,16	-0,36	34,53	-0,02	1,02	-0,08	-0,39	3,68					
11	48,73	-0,06	0,87	-0,11	-0,45	11,53	-0,06	1,05	-0,06	-0,45	12,46					
12	114,40	-0,10	0,72	-0,18	-0,28	49,38	-0,09	1,01	-0,06	-0,32	22,80					
13	194,27	-0,13	0,88	-0,13	-0,42	38,34	-0,12	1,02	-0,07	-0,40	36,40					
14	103,73	-0,09	0,91	-0,07	-0,50	20,81	-0,09	1,00	-0,03	-0,48	21,36					
15	208,92	-0,11	0,82	-0,14	-0,40	36,98	-0,11	0,98	-0,07	-0,39	28,01					
Média	114,39	-0,08	0,81	-0,12	-0,41	35,49	-0,08	1,02	-0,05	-0,41	18,72					
Desv.Pad	63,28	0,03	0,12	0,04	0,06	25,74	0,03	0,02	0,02	0,05	11,32					

Tabela C.2: Resultados da simulação para o modelo gaussiana inversa - cenário II.

n	r_i^{dev}				r_i^{pca}				r_i^{ga}							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	67,74	-0,04	1,02	-0,09	-0,45	7,73	-0,01	1,01	0,08	-0,47	5,28	0,00	0,84	-0,09	-0,45	6,84
2	100,31	-0,06	0,98	-0,02	-0,41	10,93	-0,02	0,98	0,19	-0,37	7,87	-0,01	0,85	-0,02	-0,40	5,96
3	215,70	-0,07	1,01	-0,12	-0,41	12,69	0,00	0,99	0,19	-0,39	6,82	0,01	0,88	-0,12	-0,40	4,76
4	115,59	-0,04	1,01	-0,06	-0,42	6,36	0,01	1,00	0,17	-0,40	5,89	0,01	0,90	-0,06	-0,42	3,39
5	131,63	-0,06	1,01	-0,12	-0,34	9,01	-0,01	1,00	0,14	-0,36	4,57	0,00	0,91	-0,12	-0,33	2,49
6	200,68	-0,08	0,99	-0,09	-0,40	16,40	-0,02	0,98	0,22	-0,38	9,67	-0,01	0,87	-0,09	-0,39	4,64
7	33,84	-0,02	1,01	-0,11	-0,43	4,44	0,00	1,00	-0,02	-0,45	3,04	0,01	0,53	-0,11	-0,43	101,61
8	80,06	-0,02	1,01	-0,10	-0,40	3,43	0,02	1,01	0,09	-0,41	3,61	0,03	0,86	-0,10	-0,40	8,26
9	48,23	-0,02	0,97	-0,10	-0,38	1,71	0,01	0,96	0,02	-0,41	1,39	0,02	0,54	-0,10	-0,38	96,84
10	52,03	-0,03	1,00	-0,06	-0,39	4,20	0,00	1,00	0,08	-0,40	3,31	0,01	0,69	-0,06	-0,39	34,60
11	48,73	-0,05	1,00	-0,06	-0,45	8,12	-0,01	1,00	0,09	-0,46	5,79	-0,01	0,84	-0,06	-0,45	6,78
12	114,40	-0,01	1,02	-0,08	-0,43	4,22	0,03	1,02	0,13	-0,45	6,67	0,04	0,72	-0,08	-0,44	30,25
13	194,27	-0,06	1,00	-0,08	-0,36	9,87	0,00	1,00	0,22	-0,36	7,48	0,01	0,88	-0,08	-0,35	4,44
14	103,73	-0,05	1,04	-0,03	-0,44	11,60	0,00	1,03	0,19	-0,41	10,24	0,00	0,95	-0,03	-0,43	2,75
15	208,92	-0,07	1,02	-0,08	-0,40	11,60	0,00	1,01	0,22	-0,39	8,13	0,01	0,87	-0,09	-0,40	4,75
Média	114,39	-0,05	1,01	-0,08	-0,41	8,15	0,00	1,00	0,13	-0,41	5,98	0,01	0,81	-0,08	-0,40	21,22
Desv.Pad	63,28	0,02	0,02	0,03	0,03	4,10	0,01	0,02	0,08	0,04	2,51	0,01	0,13	0,03	0,03	33,13
n	r_i^{ams}				r_i^{unt}											
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD					
1	67,74	-0,04	0,84	-0,09	-0,45	9,90	-0,04	1,02	-0,06	-0,46	6,84					
2	100,31	-0,06	0,84	-0,02	-0,41	14,10	-0,06	0,98	0,01	-0,41	9,92					
3	215,70	-0,07	0,87	-0,12	-0,42	12,94	-0,06	1,00	-0,08	-0,41	10,73					
4	115,59	-0,04	0,89	-0,06	-0,43	6,04	-0,04	1,01	-0,03	-0,42	5,74					
5	131,63	-0,06	0,91	-0,12	-0,34	8,29	-0,05	1,01	-0,10	-0,34	7,98					
6	200,68	-0,08	0,86	-0,08	-0,41	17,62	-0,07	0,99	-0,05	-0,40	14,44					
7	33,84	-0,02	0,53	-0,11	-0,43	101,44	-0,01	1,01	-0,07	-0,44	3,52					
8	80,06	-0,02	0,85	-0,10	-0,41	5,99	-0,01	1,01	-0,07	-0,41	2,97					
9	48,23	-0,01	0,54	-0,10	-0,38	95,43	0,00	0,97	-0,05	-0,39	1,20					
10	52,03	-0,03	0,69	-0,06	-0,39	35,92	-0,02	1,00	-0,01	-0,40	3,48					
11	48,73	-0,04	0,84	-0,06	-0,46	10,38	-0,04	1,00	-0,03	-0,46	7,43					
12	114,40	-0,01	0,72	-0,08	-0,44	25,79	0,00	1,02	-0,02	-0,45	3,98					
13	194,27	-0,06	0,87	-0,08	-0,37	10,58	-0,05	1,00	-0,04	-0,37	8,52					
14	103,73	-0,05	0,95	-0,03	-0,44	7,97	-0,05	1,04	-0,01	-0,44	11,08					
15	208,92	-0,06	0,86	-0,08	-0,41	12,23	-0,06	1,02	-0,04	-0,41	9,81					
Média	114,39	-0,04	0,80	-0,08	-0,41	24,98	-0,04	1,00	-0,04	-0,41	7,18					
Desv.Pad	63,28	0,02	0,13	0,03	0,03	30,87	0,02	0,02	0,03	0,03	3,69					

Tabela C.3: Resultados da simulação para o modelo gaussiana inversa - cenário III.

n	r_i^{dev}				r_i^{pca}				r_i^{qlc}						
	Média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	67,74	1,07	-0,15	-0,34	33,69	-0,01	1,04	0,36	-0,34	20,98	0,01	0,89	-0,15	-0,33	4,15
2	100,31	1,00	-0,13	-0,40	26,94	0,02	0,99	0,46	-0,26	22,31	0,04	0,88	-0,13	-0,38	9,47
3	215,70	1,05	-0,18	-0,33	107,17	-0,02	0,98	0,68	-0,01	62,00	-0,01	0,95	-0,19	-0,32	2,62
4	115,59	1,02	-0,11	-0,37	50,51	0,00	1,00	0,54	-0,21	38,84	0,01	0,92	-0,11	-0,35	3,69
5	131,63	1,00	-0,19	-0,33	48,66	0,00	0,95	0,50	-0,20	29,90	0,02	0,92	-0,20	-0,30	5,38
6	200,68	1,01	-0,28	-0,01	65,76	0,00	0,94	0,60	-0,05	42,91	0,02	0,91	-0,29	0,01	9,09
7	33,84	1,04	-0,28	-0,27	12,98	-0,01	1,00	0,01	-0,43	2,89	0,03	0,55	-0,28	-0,27	100,97
8	80,06	1,03	-0,13	-0,35	27,26	0,01	1,01	0,42	-0,26	21,07	0,03	0,89	-0,13	-0,33	6,38
9	48,23	1,03	-0,19	-0,35	14,28	0,00	1,01	0,17	-0,43	6,65	0,04	0,58	-0,19	-0,34	81,46
10	52,03	1,03	-0,18	-0,34	18,70	0,00	1,00	0,23	-0,36	7,36	0,03	0,71	-0,18	-0,33	33,52
11	48,73	1,06	-0,15	-0,37	23,27	0,00	1,04	0,27	-0,40	14,03	0,01	0,90	-0,16	-0,37	4,73
12	114,40	1,07	-0,18	-0,32	41,86	0,00	1,03	0,43	-0,29	24,62	0,04	0,77	-0,18	-0,30	22,64
13	194,27	1,06	-0,21	-0,17	93,24	-0,02	0,99	0,66	-0,02	58,04	0,00	0,96	-0,21	-0,15	2,82
14	103,73	1,03	-0,11	-0,40	43,16	0,01	1,02	0,53	-0,15	31,37	0,01	0,97	-0,12	-0,37	2,66
15	208,92	1,08	-0,22	-0,35	77,34	0,01	1,02	0,60	-0,15	46,68	0,02	0,95	-0,23	-0,32	7,19
Média	114,39	1,04	-0,18	-0,31	45,65	0,00	1,00	0,43	-0,24	28,64	0,02	0,85	-0,18	-0,30	19,78
Desv.Pad	63,28	0,05	0,03	0,10	28,81	0,01	0,03	0,19	0,14	18,14	0,01	0,14	0,05	0,10	30,43
n	r_i^{ans}				r_i^{unl}										
	Média	var	assim	curtose	AD	média	var	assim	curtose	AD					
1	67,74	0,87	-0,14	-0,37	27,38	-0,10	1,07	-0,06	-0,37	27,53					
2	100,31	0,85	-0,12	-0,44	26,54	-0,09	1,00	-0,05	-0,42	21,43					
3	215,70	0,88	-0,15	-0,42	93,67	-0,20	1,04	-0,08	-0,36	91,55					
4	115,59	0,89	-0,10	-0,42	45,49	-0,13	1,02	-0,04	-0,39	44,37					
5	131,63	0,88	-0,18	-0,39	45,84	-0,14	1,00	-0,13	-0,35	41,93					
6	200,68	0,85	-0,24	-0,16	63,58	-0,16	1,00	-0,17	-0,10	54,06					
7	33,84	0,54	-0,28	-0,28	103,38	-0,04	1,02	-0,15	-0,37	6,04					
8	80,06	0,86	-0,12	-0,38	25,46	-0,09	1,03	-0,05	-0,37	22,08					
9	48,23	0,57	-0,19	-0,35	85,79	-0,04	1,02	-0,03	-0,41	7,62					
10	52,03	0,70	-0,17	-0,36	42,85	-0,06	1,02	-0,05	-0,38	11,41					
11	48,73	0,88	-0,15	-0,39	18,61	-0,08	1,06	-0,09	-0,40	19,03					
12	114,40	0,74	-0,16	-0,36	48,26	-0,10	1,06	0,00	-0,38	28,72					
13	194,27	0,89	-0,17	-0,29	81,59	-0,18	1,05	-0,10	-0,23	80,05					
14	103,73	0,93	-0,11	-0,44	38,08	-0,12	1,03	-0,06	-0,40	38,72					
15	208,92	0,88	-0,19	-0,42	64,40	-0,16	1,07	-0,10	-0,39	62,22					
Média	114,39	0,81	-0,16	-0,36	54,06	-0,11	1,03	-0,08	-0,35	37,12					
Desv.Pad	63,28	0,05	0,12	0,05	26,81	0,05	0,02	0,05	0,08	25,72					

Tabela C.4: Resultados da simulação para o modelo gaussiana inversa - cenário IV.

n	Média				Desvio				Pearson				Quantilico			
	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD	
1	0,15	-0,09	1,01	-0,11	-0,37	21,13	-0,01	1,00	0,29	-0,34	12,85	0,01	0,84	-0,11	-0,36	6,94
2	0,18	-0,09	1,05	-0,18	-0,36	19,71	0,01	1,02	0,26	-0,37	10,37	0,02	0,91	-0,18	-0,35	5,50
3	0,35	-0,16	1,03	-0,17	-0,30	57,83	-0,03	0,98	0,45	-0,25	32,13	-0,01	0,90	-0,17	-0,29	3,30
4	0,22	-0,11	1,04	-0,13	-0,36	30,30	0,00	1,01	0,37	-0,31	18,82	0,00	0,93	-0,14	-0,35	2,45
5	0,23	-0,10	0,96	-0,11	-0,34	20,15	0,01	0,95	0,39	-0,24	15,08	0,02	0,88	-0,11	-0,33	5,78
6	0,34	-0,13	1,02	-0,17	-0,28	36,37	0,00	0,99	0,44	-0,23	23,02	0,02	0,90	-0,17	-0,26	5,34
7	0,08	-0,07	1,01	-0,17	-0,30	11,16	-0,02	0,99	0,07	-0,38	3,82	0,02	0,57	-0,17	-0,30	84,33
8	0,17	-0,05	1,00	-0,09	-0,42	9,01	0,04	1,00	0,32	-0,38	14,51	0,05	0,85	-0,09	-0,41	12,66
9	0,12	-0,08	1,01	-0,08	-0,35	16,83	-0,02	0,99	0,23	-0,36	10,97	0,02	0,61	-0,08	-0,35	62,86
10	0,10	-0,06	0,99	-0,13	-0,41	8,60	0,01	0,98	0,16	-0,42	4,72	0,03	0,68	-0,13	-0,40	39,33
11	0,11	-0,06	1,07	-0,14	-0,40	13,04	0,02	1,05	0,20	-0,46	11,07	0,02	0,90	-0,14	-0,40	6,41
12	0,18	-0,10	1,03	-0,15	-0,31	23,83	-0,02	1,00	0,26	-0,34	12,38	0,02	0,69	-0,15	-0,31	37,76
13	0,31	-0,11	1,02	-0,19	-0,42	25,74	0,02	0,99	0,36	-0,37	16,09	0,03	0,89	-0,19	-0,40	9,76
14	0,19	-0,10	1,03	-0,08	-0,35	24,09	0,01	1,02	0,41	-0,26	19,88	0,01	0,96	-0,08	-0,34	1,99
15	0,36	-0,15	1,06	-0,15	-0,39	55,79	-0,02	1,01	0,45	-0,31	32,85	0,00	0,90	-0,15	-0,38	3,74
Média	0,21	-0,10	1,02	-0,14	-0,36	24,90	0,00	1,00	0,31	-0,33	15,90	0,02	0,83	-0,14	-0,35	19,21
Desv.Pad	0,10	0,03	0,03	0,04	0,05	15,07	0,02	0,02	0,11	0,07	8,45	0,01	0,13	0,04	0,05	25,32
Williams																
n	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	0,15	-0,09	0,83	-0,11	-0,39	22,93	-0,08	1,01	-0,04	-0,39	17,26	-0,08	1,01	-0,04	-0,39	17,26
2	0,18	-0,08	0,89	-0,17	-0,38	16,59	-0,08	1,04	-0,12	-0,38	15,90	-0,08	1,04	-0,12	-0,38	15,90
3	0,35	-0,15	0,87	-0,15	-0,35	51,83	-0,14	1,03	-0,08	-0,33	49,22	-0,14	1,03	-0,08	-0,33	49,22
4	0,22	-0,11	0,91	-0,13	-0,39	26,06	-0,10	1,03	-0,08	-0,38	26,32	-0,10	1,03	-0,08	-0,38	26,32
5	0,23	-0,09	0,86	-0,11	-0,37	22,13	-0,09	0,96	-0,06	-0,35	17,27	-0,09	0,96	-0,06	-0,35	17,27
6	0,34	-0,12	0,86	-0,16	-0,32	34,87	-0,11	1,02	-0,09	-0,31	29,92	-0,11	1,02	-0,09	-0,31	29,92
7	0,08	-0,05	0,56	-0,17	-0,31	89,99	-0,05	1,01	-0,07	-0,35	6,46	-0,05	1,01	-0,07	-0,35	6,46
8	0,17	-0,05	0,84	-0,09	-0,44	11,87	-0,04	1,00	-0,03	-0,43	7,31	-0,04	1,00	-0,03	-0,43	7,31
9	0,12	-0,06	0,60	-0,08	-0,36	73,54	-0,06	1,00	0,04	-0,37	12,20	-0,06	1,00	0,04	-0,37	12,20
10	0,10	-0,05	0,68	-0,13	-0,41	41,54	-0,04	0,99	-0,04	-0,43	5,30	-0,04	0,99	-0,04	-0,43	5,30
11	0,11	-0,05	0,89	-0,14	-0,42	8,75	-0,05	1,06	-0,09	-0,43	10,97	-0,05	1,06	-0,09	-0,43	10,97
12	0,18	-0,08	0,68	-0,14	-0,33	53,85	-0,07	1,02	-0,01	-0,36	16,22	-0,07	1,02	-0,01	-0,36	16,22
13	0,31	-0,10	0,87	-0,18	-0,45	23,86	-0,09	1,02	-0,11	-0,44	20,03	-0,09	1,02	-0,11	-0,44	20,03
14	0,19	-0,09	0,94	-0,07	-0,38	20,72	-0,09	1,03	-0,04	-0,36	21,99	-0,09	1,03	-0,04	-0,36	21,99
15	0,36	-0,14	0,87	-0,14	-0,43	46,79	-0,13	1,05	-0,06	-0,42	46,51	-0,13	1,05	-0,06	-0,42	46,51
Média	0,21	-0,09	0,81	-0,13	-0,38	36,35	-0,08	1,02	-0,06	-0,38	20,19	-0,08	1,02	-0,06	-0,38	20,19
Desv.Pad	0,10	0,03	0,12	0,03	0,04	23,28	0,03	0,03	0,04	0,04	13,22	0,03	0,03	0,04	0,04	13,22

Tabela C.5: Resultados da simulação para o modelo gaussiana inversa - cenário V.

n	Média	r_i^{dev}				r_i^{pre}				r_i^{qu}						
		média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	69,89	-0,09	1,05	-0,08	-0,44	22,10	-0,01	1,03	0,28	-0,40	14,95	-0,01	0,97	-0,08	-0,43	1,76
2	367,12	-0,15	0,99	-0,32	-0,17	44,83	-0,01	0,92	0,38	-0,31	17,75	-0,01	0,70	-0,32	-0,15	49,59
3	96,53	-0,09	1,01	-0,06	-0,41	21,43	0,00	1,01	0,35	-0,32	16,54	0,01	0,91	-0,06	-0,40	2,55
4	68,86	-0,07	1,01	-0,05	-0,41	14,73	0,01	1,01	0,30	-0,39	14,75	0,01	0,91	-0,05	-0,40	3,56
5	46,26	-0,06	1,01	-0,13	-0,43	11,31	-0,01	1,00	0,12	-0,47	6,15	0,02	0,67	-0,13	-0,43	41,02
6	61,77	-0,04	0,99	-0,11	-0,31	4,94	0,03	0,99	0,24	-0,32	7,69	0,04	0,92	-0,11	-0,31	6,74
7	74,20	-0,08	1,00	-0,14	-0,34	13,39	-0,01	0,98	0,20	-0,36	6,27	0,02	0,74	-0,14	-0,34	24,09
8	204,70	-0,12	1,01	-0,23	-0,25	28,94	0,00	0,96	0,32	-0,34	14,07	0,03	0,77	-0,24	-0,23	24,49
9	58,14	-0,07	1,01	-0,09	-0,50	14,91	0,00	1,00	0,22	-0,46	9,87	0,01	0,91	-0,09	-0,49	4,26
10	268,41	-0,14	1,00	-0,27	-0,23	37,62	-0,01	0,93	0,34	-0,36	15,79	0,04	0,73	-0,28	-0,22	35,64
11	66,74	-0,10	1,04	-0,12	-0,40	25,52	-0,03	1,02	0,20	-0,43	13,64	-0,01	0,77	-0,12	-0,40	15,69
12	86,85	-0,06	1,07	-0,12	-0,44	14,10	0,03	1,06	0,27	-0,43	14,22	0,04	0,91	-0,12	-0,44	7,83
13	49,51	-0,07	1,02	-0,08	-0,36	14,66	-0,01	1,01	0,23	-0,37	10,14	0,00	0,86	-0,08	-0,35	5,42
14	51,56	-0,06	1,02	-0,06	-0,46	12,51	0,00	1,01	0,22	-0,40	8,44	0,02	0,74	-0,06	-0,45	22,41
15	57,25	-0,06	1,04	-0,11	-0,41	11,67	0,01	1,03	0,20	-0,41	8,02	0,02	0,86	-0,11	-0,41	6,66
Média	108,52	-0,08	1,02	-0,13	-0,37	19,51	0,00	1,00	0,26	-0,38	11,89	0,02	0,82	-0,13	-0,36	16,78
Desv.Pad	94,96	0,03	0,02	0,08	0,09	10,79	0,02	0,04	0,07	0,05	3,95	0,02	0,10	0,08	0,10	15,44
n	Média	r_i^{gas}				r_i^{unl}										
		média	var	assim	curtose	AD	média	var	assim	curtose	AD					
1	69,89	-0,09	0,96	-0,07	-0,45	17,42	-0,08	1,05	-0,05	-0,44	20,45					
2	367,12	-0,12	0,66	-0,30	-0,25	75,90	-0,11	0,98	-0,11	-0,31	25,17					
3	96,53	-0,08	0,90	-0,06	-0,42	19,00	-0,08	1,01	-0,02	-0,41	19,05					
4	68,86	-0,06	0,90	-0,05	-0,42	13,08	-0,06	1,01	-0,02	-0,42	13,42					
5	46,26	-0,05	0,66	-0,13	-0,44	45,65	-0,04	1,01	-0,04	-0,46	8,03					
6	61,77	-0,04	0,91	-0,11	-0,33	5,26	-0,04	0,99	-0,08	-0,32	4,23					
7	74,20	-0,07	0,73	-0,14	-0,35	33,23	-0,06	0,99	-0,05	-0,36	9,20					
8	204,70	-0,10	0,75	-0,22	-0,29	42,65	-0,09	1,00	-0,10	-0,32	19,15					
9	58,14	-0,06	0,91	-0,09	-0,51	12,26	-0,06	1,01	-0,06	-0,50	13,35					
10	268,41	-0,11	0,70	-0,26	-0,29	59,91	-0,10	0,99	-0,10	-0,33	23,10					
11	66,74	-0,09	0,77	-0,11	-0,41	33,09	-0,08	1,04	-0,03	-0,43	20,25					
12	86,85	-0,05	0,89	-0,12	-0,46	8,58	-0,04	1,07	-0,06	-0,46	11,91					
13	49,51	-0,07	0,85	-0,08	-0,37	15,72	-0,06	1,02	-0,03	-0,37	12,64					
14	51,56	-0,05	0,73	-0,06	-0,47	28,80	-0,05	1,02	0,02	-0,46	9,70					
15	57,25	-0,06	0,85	-0,11	-0,42	11,98	-0,05	1,04	-0,05	-0,42	9,44					
Média	108,52	-0,07	0,81	-0,13	-0,39	28,17	-0,07	1,02	-0,05	-0,40	14,61					
Desv.Pad	94,96	0,02	0,10	0,07	0,08	20,45	0,02	0,02	0,03	0,06	6,18					

Tabela C.6: Resultados da simulação para o modelo inversa gaussiana - cenário VI.

n	r_i^{dev}				r_i^{pca}				r_i^{qu}							
	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD	
1	82,24	-0,08	1,02	-0,12	-0,47	16,77	0,00	1,00	0,25	-0,44	9,95	0,01	0,92	-0,12	-0,46	3,81
2	72,55	-0,10	0,99	-0,06	-0,38	22,85	-0,02	0,98	0,30	-0,31	14,85	-0,01	0,87	-0,06	-0,38	3,97
3	89,05	-0,09	1,01	-0,13	-0,34	17,47	0,00	0,99	0,28	-0,33	10,23	0,01	0,85	-0,13	-0,34	6,92
4	98,21	-0,08	1,04	-0,17	-0,43	17,99	-0,01	1,01	0,18	-0,46	7,94	0,03	0,64	-0,17	-0,42	53,24
5	76,97	-0,09	1,01	-0,18	-0,25	18,36	-0,02	0,98	0,20	-0,32	7,94	0,00	0,86	-0,18	-0,25	6,47
6	78,94	-0,10	1,00	-0,10	-0,41	22,61	-0,02	0,98	0,28	-0,36	13,33	-0,01	0,89	-0,10	-0,40	3,35
7	89,14	-0,07	1,01	-0,18	-0,31	10,80	0,01	0,98	0,21	-0,35	4,79	0,03	0,79	-0,18	-0,31	18,55
8	72,86	-0,07	1,02	-0,18	-0,32	10,07	0,00	1,00	0,13	-0,39	3,86	0,03	0,59	-0,18	-0,31	74,38
9	62,58	-0,04	1,02	-0,12	-0,37	5,55	0,03	1,01	0,20	-0,40	7,36	0,04	0,83	-0,12	-0,36	13,90
10	60,75	-0,07	1,02	-0,11	-0,43	14,42	-0,01	1,01	0,20	-0,43	8,57	0,01	0,77	-0,11	-0,42	16,66
11	61,97	-0,08	1,02	-0,13	-0,38	13,99	-0,01	1,00	0,17	-0,42	7,07	0,01	0,73	-0,14	-0,38	24,29
12	77,74	-0,05	1,02	-0,11	-0,41	9,29	0,03	1,02	0,27	-0,43	13,11	0,04	0,95	-0,11	-0,40	6,48
13	76,45	-0,07	1,01	-0,09	-0,36	13,30	0,01	1,01	0,30	-0,35	12,42	0,02	0,95	-0,09	-0,35	3,10
14	63,53	-0,08	1,03	-0,13	-0,41	15,47	-0,01	1,01	0,19	-0,44	8,67	0,01	0,80	-0,13	-0,41	12,94
15	72,56	-0,08	1,04	-0,13	-0,47	19,38	0,00	1,02	0,20	-0,48	11,71	0,01	0,83	-0,13	-0,46	8,71
Média	75,70	-0,08	1,02	-0,13	-0,38	15,22	0,00	1,00	0,22	-0,39	9,45	0,02	0,82	-0,13	-0,38	17,12
Desv.Pad	10,97	0,02	0,01	0,03	0,06	4,90	0,02	0,01	0,05	0,05	3,18	0,02	0,10	0,04	0,06	20,36
n	r_i^{ans}				r_i^{ans}				r_i^{ans}							
	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD	
1	82,24	-0,08	0,91	-0,12	-0,48	13,95	-0,07	1,02	-0,08	-0,47	14,54	-0,07	1,02	-0,08	-0,47	14,54
2	72,55	-0,09	0,86	-0,06	-0,40	23,27	-0,09	0,99	-0,02	-0,39	20,28	-0,09	0,99	-0,02	-0,39	20,28
3	89,05	-0,08	0,84	-0,12	-0,36	19,92	-0,07	1,01	-0,06	-0,36	14,03	-0,07	1,01	-0,06	-0,36	14,03
4	98,21	-0,06	0,63	-0,17	-0,44	59,83	-0,05	1,03	-0,04	-0,47	10,92	-0,05	1,03	-0,04	-0,47	10,92
5	76,97	-0,09	0,85	-0,18	-0,27	20,83	-0,08	1,01	-0,12	-0,28	14,88	-0,08	1,01	-0,12	-0,28	14,88
6	78,94	-0,09	0,88	-0,10	-0,42	21,92	-0,09	1,00	-0,06	-0,42	19,96	-0,09	1,00	-0,06	-0,42	19,96
7	89,14	-0,06	0,78	-0,18	-0,33	22,45	-0,05	1,00	-0,09	-0,35	6,87	-0,05	1,00	-0,09	-0,35	6,87
8	72,86	-0,05	0,59	-0,18	-0,33	76,85	-0,04	1,01	-0,05	-0,37	5,20	-0,04	1,01	-0,05	-0,37	5,20
9	62,58	-0,03	0,82	-0,12	-0,38	10,81	-0,03	1,02	-0,06	-0,39	4,18	-0,03	1,02	-0,06	-0,39	4,18
10	60,75	-0,06	0,76	-0,11	-0,44	24,97	-0,06	1,02	-0,03	-0,44	11,07	-0,06	1,02	-0,03	-0,44	11,07
11	61,97	-0,06	0,73	-0,13	-0,39	32,90	-0,06	1,01	-0,05	-0,41	10,07	-0,06	1,01	-0,05	-0,41	10,07
12	77,74	-0,05	0,94	-0,10	-0,42	7,07	-0,05	1,02	-0,08	-0,42	8,43	-0,05	1,02	-0,08	-0,42	8,43
13	76,45	-0,07	0,93	-0,09	-0,38	11,40	-0,06	1,01	-0,06	-0,37	12,16	-0,06	1,01	-0,06	-0,37	12,16
14	63,53	-0,07	0,79	-0,12	-0,42	21,46	-0,06	1,03	-0,06	-0,43	11,97	-0,06	1,03	-0,06	-0,43	11,97
15	72,56	-0,07	0,82	-0,12	-0,48	18,75	-0,07	1,04	-0,06	-0,49	15,73	-0,07	1,04	-0,06	-0,49	15,73
Média	75,70	-0,07	0,81	-0,13	-0,40	25,76	-0,06	1,01	-0,06	-0,40	12,02	-0,06	1,01	-0,06	-0,40	12,02
Desv.Pad	10,97	0,02	0,10	0,03	0,06	18,71	0,02	0,01	0,03	0,05	4,76	0,02	0,01	0,03	0,05	4,76

Tabela C.7: Resultados da simulação para o modelo gaussiana inversa - cenário VII.

n	r_i^{dev}				r_i^{pca}				r_i^{ga}							
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD	média	var	assim	curtose	AD
1	28,61	-0,05	1,03	-0,08	-0,39	7,94	0,00	1,02	0,15	-0,39	5,31	0,01	0,87	-0,08	-0,39	4,60
2	24,89	-0,04	1,02	-0,07	-0,43	6,64	0,01	1,01	0,14	-0,44	5,95	0,01	0,89	-0,07	-0,43	3,88
3	20,50	-0,07	0,97	0,00	-0,40	13,57	-0,03	0,97	0,19	-0,39	10,34	-0,02	0,84	0,00	-0,41	7,79
4	23,86	-0,04	1,00	-0,06	-0,43	6,78	0,00	1,00	0,15	-0,43	5,26	0,01	0,90	-0,06	-0,43	2,87
5	23,03	-0,04	0,96	-0,03	-0,32	4,54	0,00	0,95	0,19	-0,30	4,15	0,01	0,87	-0,03	-0,31	4,59
6	20,82	-0,04	0,99	-0,05	-0,42	4,87	0,01	0,99	0,15	-0,39	4,00	0,01	0,86	-0,05	-0,42	5,36
7	43,53	-0,07	1,01	-0,12	-0,36	11,33	-0,03	0,99	0,08	-0,41	5,52	0,02	0,42	-0,12	-0,31	182,17
8	26,83	-0,02	1,04	-0,04	-0,38	4,17	0,03	1,04	0,18	-0,37	7,56	0,03	0,90	-0,04	-0,38	5,79
9	33,67	-0,07	1,04	-0,13	-0,43	15,23	-0,03	1,03	0,06	-0,47	8,46	0,00	0,56	-0,13	-0,40	83,33
10	32,31	-0,06	1,03	-0,10	-0,52	15,54	-0,01	1,02	0,10	-0,55	10,91	0,01	0,75	-0,10	-0,52	20,52
11	33,48	-0,01	1,01	-0,06	-0,44	3,68	0,04	1,02	0,18	-0,41	7,94	0,05	0,87	-0,06	-0,44	9,93
12	23,93	-0,05	0,98	-0,04	-0,43	7,43	-0,01	0,98	0,15	-0,42	5,42	0,01	0,71	-0,04	-0,44	26,90
13	20,96	-0,05	0,99	-0,03	-0,42	7,10	-0,01	0,99	0,16	-0,40	5,18	0,00	0,86	-0,03	-0,41	4,59
14	24,63	-0,02	0,99	-0,05	-0,37	2,98	0,02	0,99	0,17	-0,37	5,21	0,03	0,92	-0,05	-0,37	3,90
15	20,64	-0,06	1,01	-0,01	-0,44	12,26	-0,02	1,01	0,17	-0,42	9,58	-0,01	0,86	-0,01	-0,44	5,50
Média	26,78	-0,05	1,00	-0,06	-0,41	8,27	0,00	1,00	0,15	-0,41	6,72	0,01	0,80	-0,06	-0,41	24,78
Desv.Pad	6,50	0,02	0,03	0,04	0,05	4,25	0,02	0,02	0,04	0,06	2,25	0,02	0,14	0,04	0,05	48,12
n	r_i^{ans}				r_i^{wsl}											
	Média	média	var	assim	curtose	AD	média	var	assim	curtose	AD					
1	28,61	-0,04	0,86	-0,08	-0,39	8,34	-0,04	1,03	-0,04	-0,39	6,79					
2	24,89	-0,04	0,88	-0,07	-0,44	6,01	-0,03	1,02	-0,05	-0,44	6,00					
3	20,50	-0,06	0,84	0,00	-0,41	17,46	-0,06	0,97	0,03	-0,41	12,68					
4	23,86	-0,04	0,90	-0,06	-0,44	5,86	-0,04	1,00	-0,04	-0,44	6,19					
5	23,03	-0,04	0,86	-0,03	-0,32	8,15	-0,04	0,96	-0,01	-0,32	4,11					
6	20,82	-0,03	0,86	-0,04	-0,42	7,30	-0,03	0,99	-0,02	-0,42	4,30					
7	43,53	-0,04	0,42	-0,12	-0,31	186,83	-0,04	1,00	-0,01	-0,40	6,95					
8	26,83	-0,02	0,89	-0,04	-0,38	3,10	-0,01	1,04	-0,01	-0,38	4,01					
9	33,67	-0,05	0,56	-0,13	-0,41	89,70	-0,05	1,04	-0,04	-0,46	11,06					
10	32,31	-0,05	0,74	-0,10	-0,52	25,85	-0,05	1,03	-0,04	-0,54	13,17					
11	33,48	-0,01	0,87	-0,06	-0,44	4,26	0,00	1,01	-0,03	-0,44	3,49					
12	23,93	-0,04	0,71	-0,04	-0,44	30,81	-0,04	0,98	0,01	-0,44	6,11					
13	20,96	-0,04	0,86	-0,03	-0,42	9,12	-0,04	0,99	-0,01	-0,42	6,39					
14	24,63	-0,02	0,91	-0,05	-0,38	3,08	-0,02	0,99	-0,03	-0,38	2,79					
15	20,64	-0,05	0,85	-0,01	-0,44	12,74	-0,05	1,01	0,02	-0,44	11,35					
Média	26,78	-0,04	0,80	-0,06	-0,41	27,91	-0,04	1,00	-0,02	-0,42	7,03					
Desv.Pad	6,50	0,01	0,14	0,04	0,05	49,09	0,02	0,02	0,02	0,05	3,41					

Apêndice D

Tabelas dos estudos de simulação no modelo ZAGA

Tabela D.1: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 1.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,04	0,10	0,13	0,14	0,18	0,32
	<-2	2,28	2,04	2,44	2,65	2,63	2,80	3,08
	<-1	15,87	15,12	16,36	16,74	16,73	17,18	17,82
	>1	15,87	13,88	15,14	15,54	15,54	15,92	17,26
	>2	2,28	1,78	2,05	2,21	2,21	2,36	2,86
	>3	0,13	0,04	0,08	0,1	0,11	0,14	0,26
r_i^{**dev}	<-3	0,13	0,3	0,66	0,84	0,84	1,01	1,42
	<-2	2,28	3,08	4	4,22	4,20	4,435	4,98
	<-1	15,87	10,74	12,54	13,54	13,42	14,39	15,62
	>1	15,87	9,72	11,62	12,43	12,34	13,14	14,38
	>2	2,28	2,8	3,39	3,68	3,68	3,95	4,52
	>3	0,13	0,24	0,52	0,7	0,68	0,8	1,28
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,05	0,07	0,14
	<-2	2,28	1,54	1,92	2,09	2,08	2,23	2,64
	<-1	15,87	15,04	16,16	16,56	16,51	16,92	17,52
	>1	15,87	13,98	15,34	15,73	15,73	16,11	17,42
	>2	2,28	2,14	2,50	2,68	2,68	2,86	3,34
	>3	0,13	0,08	0,18	0,22	0,22	0,26	0,4
r_i^{**pea}	<-3	0,13	0,12	0,38	0,54	0,54	0,68	1
	<-2	2,28	2,6	3,42	3,67	3,63	3,91	4,5
	<-1	15,87	10,58	12,33	13,25	13,09	13,99	15,20
	>1	15,87	9,84	11,79	12,74	12,63	13,43	14,66
	>2	2,28	3,42	3,94	4,16	4,18	4,45	5,06
	>3	0,13	0,50	0,76	0,98	0,97	1,13	1,60
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,14	0,26
	<-2	2,28	1,68	2,10	2,32	2,30	2,51	2,78
	<-1	15,87	14,44	15,57	15,95	15,94	16,34	17,04
	>1	15,87	14,28	15,54	15,94	15,98	16,38	17,74
	>2	2,28	1,74	2,04	2,23	2,24	2,42	2,92
	>3	0,13	0,02	0,06	0,10	0,10	0,12	0,26
r_i^{**qu}	<-3	0,13	0,22	0,56	0,71	0,71	0,87	1,18
	<-2	2,28	2,60	3,56	3,82	3,78	4,04	4,54
	<-1	15,87	10,16	11,89	12,78	12,67	13,56	14,70
	>1	15,87	9,84	11,90	12,80	12,68	13,46	14,68
	>2	2,28	2,72	3,48	3,78	3,76	4,03	4,66
	>3	0,13	0,24	0,50	0,7	0,68	0,81	1,28
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,12	0,16	0,28
	<-2	2,28	1,86	2,28	2,50	2,47	2,69	3,00
	<-1	15,87	15,00	16,19	16,58	16,56	17,00	17,78
	>1	15,87	13,84	14,97	15,40	15,38	15,79	17,18
	>2	2,28	1,58	1,90	2,06	2,09	2,26	2,76
	>3	0,13	0,02	0,06	0,08	0,09	0,12	0,24
r_i^{**ans}	<-3	0,13	0,26	0,60	0,78	0,77	0,94	1,32
	<-2	2,28	2,82	3,80	4,08	4,02	4,29	4,80
	<-1	15,87	10,54	12,39	13,36	13,22	14,19	15,20
	>1	15,87	9,44	11,44	12,31	12,15	12,87	14,22
	>2	2,28	2,54	3,22	3,51	3,52	3,82	4,42
	>3	0,13	0,20	0,46	0,64	0,62	0,76	1,16
r_i^{*wil}	<-3	0,13	0,04	0,1	0,12	0,14	0,18	0,3
	<-2	2,28	2,04	2,44	2,63	2,61	2,79	3,04
	<-1	15,87	15,12	16,36	16,74	16,73	17,18	17,82
	>1	15,87	13,88	15,15	15,55	15,55	15,94	17,26
	>2	2,28	1,78	2,06	2,22	2,223	2,37	2,86
	>3	0,13	0,04	0,08	0,10	0,11	0,14	0,26
r_i^{**wil}	<-3	0,13	0,30	0,66	0,83	0,83	1	1,40
	<-2	2,28	3,06	3,94	4,21	4,19	4,43	4,98
	<-1	15,87	10,74	12,54	13,54	13,41	14,38	15,52
	>1	15,87	9,72	11,62	12,44	12,35	13,15	14,38
	>2	2,28	2,82	3,40	3,72	3,70	3,96	4,52
	>3	0,13	0,24	0,52	0,7	0,69	0,83	1,28

Tabela D.2: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 2a.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,00	0,12	0,14	0,14	0,16	0,28
	<-2	2,28	2,14	2,48	2,57	2,62	2,79	3,14
	<-1	15,87	15,26	16,66	17,03	17,02	17,39	18,20
	>1	15,87	14,42	15,24	15,57	15,59	15,89	16,70
	>2	2,28	1,70	2,06	2,23	2,21	2,36	2,86
	>3	0,13	0,00	0,08	0,10	0,10	0,13	0,26
r_i^{**dev}	<-3	0,13	0,16	0,30	0,40	0,39	0,45	0,70
	<-2	2,28	2,76	3,26	3,48	3,45	3,62	4,18
	<-1	15,87	13,74	15,23	15,74	15,70	16,29	17,18
	>1	15,87	12,20	13,91	14,36	14,35	14,75	15,84
	>2	2,28	2,18	2,78	2,95	2,97	3,18	4,06
	>3	0,13	0,10	0,24	0,30	0,31	0,38	0,58
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,04	0,06	0,16
	<-2	2,28	1,44	1,80	1,92	1,92	2,04	2,32
	<-1	15,87	14,78	16,18	16,50	16,51	16,82	17,78
	>1	15,87	14,62	15,72	16,01	16,04	16,43	17,10
	>2	2,28	2,34	2,64	2,78	2,80	2,94	3,40
	>3	0,13	0,10	0,20	0,24	0,25	0,30	0,58
r_i^{**pea}	<-3	0,13	0,04	0,12	0,17	0,18	0,22	0,40
	<-2	2,28	1,90	2,52	2,74	2,73	2,92	3,48
	<-1	15,87	13,26	14,70	15,23	15,16	15,69	16,48
	>1	15,87	12,52	14,45	14,87	14,84	15,28	16,32
	>2	2,28	2,84	3,38	3,60	3,60	3,81	4,64
	>3	0,13	0,24	0,44	0,55	0,56	0,66	0,84
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,11	0,14	0,26
	<-2	2,28	1,68	2,06	2,24	2,25	2,45	2,76
	<-1	15,87	14,26	15,62	16,02	16,01	16,41	17,34
	>1	15,87	14,70	15,62	16,00	16,03	16,40	17,12
	>2	2,28	1,70	2,06	2,20	2,23	2,38	2,86
	>3	0,13	0,00	0,06	0,10	0,10	0,12	0,26
r_i^{**qu}	<-3	0,13	0,10	0,24	0,32	0,31	0,37	0,58
	<-2	2,28	2,10	2,84	3,08	3,04	3,22	3,72
	<-1	15,87	12,80	14,33	14,72	14,72	15,29	16,20
	>1	15,87	12,42	14,40	14,76	14,75	15,14	16,12
	>2	2,28	2,00	2,84	3,04	3,01	3,24	3,98
	>3	0,13	0,10	0,22	0,30	0,31	0,38	0,58
r_i^{*ans}	<-3	0,13	0,00	0,10	0,12	0,12	0,14	0,28
	<-2	2,28	1,80	2,26	2,42	2,44	2,64	3,06
	<-1	15,87	15,00	16,32	16,73	16,74	17,12	18,14
	>1	15,87	14,04	14,94	15,38	15,32	15,63	16,44
	>2	2,28	1,52	1,90	2,05	2,06	2,22	2,64
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,20
r_i^{**ans}	<-3	0,13	0,10	0,26	0,34	0,34	0,40	0,62
	<-2	2,28	2,32	3,06	3,28	3,26	3,44	3,92
	<-1	15,87	13,52	15,00	15,47	15,42	15,98	16,98
	>1	15,87	11,78	13,68	14,12	14,07	14,46	15,40
	>2	2,28	1,88	2,64	2,81	2,80	3,01	3,78
	>3	0,13	0,08	0,20	0,26	0,28	0,34	0,58
r_i^{*wil}	<-3	0,13	0,00	0,12	0,14	0,14	0,16	0,28
	<-2	2,28	2,12	2,46	2,56	2,60	2,77	3,10
	<-1	15,87	15,26	16,66	17,00	17,01	17,38	18,20
	>1	15,87	14,42	15,26	15,60	15,61	15,93	16,70
	>2	2,28	1,70	2,08	2,24	2,23	2,38	2,86
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,26
r_i^{**wil}	<-3	0,13	0,16	0,30	0,38	0,38	0,44	0,70
	<-2	2,28	2,76	3,24	3,46	3,43	3,58	4,16
	<-1	15,87	13,72	15,21	15,74	15,68	16,28	17,14
	>1	15,87	12,22	13,92	14,37	14,36	14,77	15,90
	>2	2,28	2,28	2,80	2,98	2,99	3,19	4,08
	>3	0,13	0,10	0,24	0,32	0,32	0,38	0,58

Tabela D.3: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - Cenário 2b.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,10	0,14	0,14	0,18	0,28
	<-2	2,28	2,04	2,44	2,60	2,61	2,77	3,46
	<-1	15,87	9,86	13,32	14,78	14,45	15,75	17,02
	>1	15,87	9,12	12,85	14,08	13,72	14,94	15,90
	>2	2,28	1,66	2,06	2,29	2,28	2,47	2,92
	>3	0,13	0,00	0,06	0,10	0,09	0,12	0,22
r_i^{**dev}	<-3	0,13	0,84	1,42	1,63	1,61	1,86	2,14
	<-2	2,28	3,50	4,16	4,41	4,41	4,68	5,74
	<-1	15,87	6,34	8,63	9,51	9,55	10,55	12,90
	>1	15,87	5,90	8,04	8,97	8,90	9,85	11,56
	>2	2,28	2,92	3,72	4,01	3,98	4,26	4,86
	>3	0,13	0,64	1,22	1,37	1,39	1,59	2,08
r_i^{*pea}	<-3	0,13	0,00	0,04	0,06	0,06	0,08	0,18
	<-2	2,28	1,70	2,10	2,27	2,27	2,43	3,06
	<-1	15,87	9,86	13,32	14,77	14,43	15,74	16,90
	>1	15,87	9,12	12,86	14,11	13,74	14,95	16,00
	>2	2,28	1,84	2,34	2,59	2,56	2,75	3,22
	>3	0,13	0,02	0,12	0,16	0,18	0,23	0,42
r_i^{**pea}	<-3	0,13	0,58	1,08	1,36	1,32	1,56	1,92
	<-2	2,28	3,32	3,85	4,08	4,11	4,38	5,40
	<-1	15,87	6,30	8,50	9,34	9,43	10,39	12,64
	>1	15,87	5,94	8,11	9,06	9,02	10,00	11,84
	>2	2,28	3,02	3,95	4,24	4,24	4,54	5,08
	>3	0,13	0,94	1,48	1,64	1,65	1,85	2,38
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,14	0,26
	<-2	2,28	1,70	2,14	2,34	2,34	2,52	3,12
	<-1	15,87	9,54	12,94	14,39	14,02	15,21	16,48
	>1	15,87	9,42	13,21	14,44	14,06	15,24	16,42
	>2	2,28	1,64	2,12	2,33	2,32	2,56	3,06
	>3	0,13	0,00	0,06	0,10	0,09	0,12	0,24
r_i^{**qu}	<-3	0,13	0,70	1,24	1,45	1,43	1,67	2,02
	<-2	2,28	3,24	3,76	4,08	4,07	4,36	5,40
	<-1	15,87	5,94	8,25	9,02	9,11	10,05	12,28
	>1	15,87	6,06	8,23	9,20	9,14	10,13	11,64
	>2	2,28	2,96	3,80	4,12	4,07	4,39	5,00
	>3	0,13	0,52	1,22	1,40	1,42	1,67	2,12
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,12	0,16	0,26
	<-2	2,28	1,88	2,30	2,50	2,50	2,67	3,34
	<-1	15,87	9,84	13,32	14,77	14,42	15,71	16,90
	>1	15,87	9,12	12,79	14,07	13,68	14,87	15,80
	>2	2,28	1,48	1,94	2,21	2,19	2,39	2,86
	>3	0,13	0,00	0,04	0,08	0,08	0,12	0,18
r_i^{**ans}	<-3	0,13	0,72	1,34	1,55	1,53	1,79	2,10
	<-2	2,28	3,36	3,98	4,28	4,29	4,60	5,66
	<-1	15,87	6,28	8,51	9,43	9,45	10,41	12,56
	>1	15,87	5,82	7,96	8,91	8,80	9,77	11,16
	>2	2,28	2,82	3,60	3,90	3,86	4,16	4,80
	>3	0,13	0,42	1,14	1,32	1,32	1,53	2,06
r_i^{*wil}	<-3	0,13	0,02	0,10	0,14	0,13	0,18	0,28
	<-2	2,28	2,04	2,44	2,60	2,59	2,77	3,46
	<-1	15,87	9,86	13,32	14,78	14,45	15,74	17,02
	>1	15,87	9,12	12,85	14,08	13,72	14,94	15,90
	>2	2,28	1,66	2,06	2,29	2,29	2,48	2,92
	>3	0,13	0,00	0,06	0,10	0,10	0,12	0,24
r_i^{**wil}	<-3	0,13	0,80	1,42	1,63	1,61	1,86	2,14
	<-2	2,28	3,50	4,16	4,41	4,40	4,68	5,74
	<-1	15,87	6,34	8,61	9,50	9,55	10,53	12,90
	>1	15,87	5,90	8,04	8,98	8,90	9,85	11,56
	>2	2,28	2,94	3,72	4,02	3,99	4,26	4,86
	>3	0,13	0,66	1,22	1,38	1,40	1,61	2,08

Tabela D.4: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - Cenário 3.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,10	0,14	0,14	0,18	0,30
	<-2	2,28	2,06	2,46	2,60	2,59	2,70	3,12
	<-1	15,87	15,40	16,40	16,73	16,74	17,07	18,26
	>1	15,87	14,34	15,12	15,59	15,55	15,94	16,74
	>2	2,28	1,62	2,08	2,22	2,23	2,37	2,84
	>3	0,13	0,02	0,08	0,10	0,10	0,12	0,26
r_i^{**dev}	<-3	0,13	0,34	0,66	0,80	0,82	0,96	1,44
	<-2	2,28	3,38	3,96	4,26	4,20	4,42	4,96
	<-1	15,87	11,08	12,84	13,51	13,47	14,19	16,22
	>1	15,87	9,64	11,69	12,43	12,35	13,16	14,08
	>2	2,28	2,72	3,44	3,69	3,68	3,93	4,48
	>3	0,13	0,28	0,54	0,70	0,69	0,84	1,26
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,05	0,06	0,12
	<-2	2,28	1,58	1,94	2,07	2,06	2,20	2,50
	<-1	15,87	15,34	16,18	16,52	16,52	16,85	17,92
	>1	15,87	14,52	15,34	15,73	15,75	16,14	16,86
	>2	2,28	1,90	2,54	2,70	2,70	2,88	3,44
	>3	0,13	0,08	0,18	0,22	0,23	0,27	0,46
r_i^{**pea}	<-3	0,13	0,14	0,38	0,52	0,53	0,64	1,06
	<-2	2,28	2,70	3,33	3,64	3,61	3,86	4,48
	<-1	15,87	10,90	12,57	13,20	13,16	13,79	15,88
	>1	15,87	9,86	11,90	12,64	12,65	13,47	14,54
	>2	2,28	3,18	3,96	4,17	4,17	4,40	5,02
	>3	0,13	0,52	0,82	1,00	0,98	1,13	1,52
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,14	0,22
	<-2	2,28	1,70	2,18	2,28	2,28	2,42	2,78
	<-1	15,87	14,64	15,66	15,92	15,97	16,28	17,38
	>1	15,87	14,72	15,55	16,02	15,99	16,41	17,22
	>2	2,28	1,58	2,06	2,24	2,26	2,42	3,00
	>3	0,13	0,00	0,06	0,10	0,10	0,12	0,26
r_i^{**qu}	<-3	0,13	0,22	0,56	0,68	0,69	0,82	1,20
	<-2	2,28	2,86	3,52	3,78	3,78	4,06	4,58
	<-1	15,87	10,42	12,09	12,81	12,71	13,39	15,32
	>1	15,87	9,88	12,00	12,74	12,70	13,54	14,64
	>2	2,28	2,74	3,49	3,80	3,75	4,03	4,48
	>3	0,13	0,28	0,56	0,70	0,69	0,84	1,18
r_i^{*ans}	<-3	0,13	0,02	0,10	0,12	0,12	0,16	0,26
	<-2	2,28	1,84	2,32	2,44	2,45	2,60	2,98
	<-1	15,87	15,30	16,21	16,50	16,57	16,96	17,96
	>1	15,87	14,18	14,95	15,41	15,39	15,77	16,58
	>2	2,28	1,44	1,90	2,10	2,09	2,26	2,68
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,22
r_i^{**ans}	<-3	0,13	0,26	0,60	0,76	0,75	0,90	1,28
	<-2	2,28	2,98	3,78	4,04	4,02	4,30	4,88
	<-1	15,87	10,88	12,60	13,37	13,28	14,02	15,96
	>1	15,87	9,40	11,46	12,28	12,16	12,96	13,88
	>2	2,28	2,58	3,26	3,53	3,52	3,80	4,22
	>3	0,13	0,24	0,48	0,64	0,63	0,78	1,12
r_i^{*wil}	<-3	0,13	0,02	0,10	0,14	0,14	0,18	0,28
	<-2	2,28	2,04	2,46	2,58	2,58	2,68	3,10
	<-1	15,87	15,40	16,40	16,72	16,73	17,07	18,26
	>1	15,87	14,36	15,12	15,60	15,56	15,94	16,74
	>2	2,28	1,66	2,08	2,23	2,24	2,38	2,88
	>3	0,13	0,02	0,08	0,10	0,11	0,12	0,26
r_i^{**wil}	<-3	0,13	0,34	0,66	0,80	0,81	0,96	1,42
	<-2	2,28	3,34	3,94	4,22	4,18	4,42	4,96
	<-1	15,87	11,08	12,84	13,50	13,46	14,16	16,20
	>1	15,87	9,64	11,69	12,43	12,36	13,18	14,12
	>2	2,28	2,74	3,48	3,70	3,69	3,94	4,50
	>3	0,13	0,28	0,56	0,70	0,70	0,84	1,26

Tabela D.5: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 4a.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,04	0,10	0,14	0,13	0,16	0,24
	<-2	2,28	2,04	2,36	2,50	2,51	2,64	3,24
	<-1	15,87	15,10	15,99	16,28	16,39	16,87	17,80
	>1	15,87	14,56	15,38	15,82	15,80	16,18	17,12
	>2	2,28	1,86	2,18	2,33	2,33	2,48	2,74
	>3	0,13	0,02	0,08	0,12	0,11	0,14	0,24
r_i^{**dev}	<-3	0,13	0,34	0,66	0,78	0,80	0,94	1,34
	<-2	2,28	3,00	3,78	4,07	4,04	4,25	4,90
	<-1	15,87	10,54	12,53	13,30	13,16	13,90	15,18
	>1	15,87	10,30	11,94	12,60	12,59	13,39	14,92
	>2	2,28	3,04	3,62	3,86	3,83	4,09	4,40
	>3	0,13	0,28	0,56	0,71	0,70	0,82	1,24
r_i^{*pea}	<-3	0,13	0,02	0,06	0,08	0,08	0,11	0,20
	<-2	2,28	1,76	2,12	2,23	2,26	2,40	2,88
	<-1	15,87	15,08	15,88	16,21	16,29	16,77	17,58
	>1	15,87	14,60	15,43	15,90	15,90	16,29	17,22
	>2	2,28	2,08	2,44	2,56	2,57	2,73	3,08
	>3	0,13	0,04	0,14	0,16	0,17	0,20	0,32
r_i^{**pea}	<-3	0,13	0,20	0,52	0,63	0,65	0,78	1,20
	<-2	2,28	2,72	3,54	3,80	3,77	3,99	4,68
	<-1	15,87	10,48	12,36	13,15	12,99	13,74	14,88
	>1	15,87	10,42	12,04	12,75	12,74	13,53	15,02
	>2	2,28	3,36	3,88	4,11	4,09	4,32	4,72
	>3	0,13	0,42	0,70	0,84	0,86	0,98	1,40
r_i^{*qu}	<-3	0,13	0,02	0,06	0,10	0,11	0,14	0,24
	<-2	2,28	1,80	2,14	2,28	2,30	2,45	3,08
	<-1	15,87	14,78	15,49	15,88	15,91	16,41	17,24
	>1	15,87	14,76	15,52	15,92	15,94	16,34	17,28
	>2	2,28	1,80	2,12	2,30	2,29	2,45	2,78
	>3	0,13	0,02	0,08	0,10	0,10	0,12	0,24
r_i^{**qu}	<-3	0,13	0,20	0,58	0,71	0,71	0,85	1,22
	<-2	2,28	2,64	3,52	3,76	3,75	4,01	4,68
	<-1	15,87	10,02	12,01	12,90	12,68	13,38	14,64
	>1	15,87	10,40	11,99	12,70	12,67	13,50	14,98
	>2	2,28	2,82	3,55	3,83	3,78	4,07	4,36
	>3	0,13	0,18	0,54	0,70	0,68	0,80	1,12
r_i^{*ans}	<-3	0,13	0,20	0,58	0,71	0,71	0,85	1,22
	<-2	2,28	2,64	3,52	3,76	3,75	4,01	4,68
	<-1	15,87	10,02	12,01	12,90	12,68	13,38	14,64
	>1	15,87	10,40	11,99	12,70	12,67	13,50	14,98
	>2	2,28	2,82	3,55	3,83	3,78	4,07	4,36
	>3	0,13	0,18	0,54	0,70	0,68	0,80	1,12
r_i^{**ans}	<-3	0,13	0,20	0,62	0,73	0,74	0,86	1,24
	<-2	2,28	2,70	3,65	3,88	3,87	4,14	4,70
	<-1	15,87	10,32	12,29	13,13	12,96	13,71	14,84
	>1	15,87	10,14	11,81	12,38	12,40	13,18	14,68
	>2	2,28	2,68	3,42	3,73	3,66	3,93	4,24
	>3	0,13	0,14	0,52	0,65	0,65	0,76	1,08
r_i^{*wil}	<-3	0,13	0,04	0,10	0,14	0,13	0,16	0,24
	<-2	2,28	2,00	2,34	2,49	2,50	2,64	3,24
	<-1	15,87	15,10	15,99	16,28	16,39	16,87	17,78
	>1	15,87	14,58	15,39	15,82	15,81	16,18	17,12
	>2	2,28	1,86	2,20	2,34	2,34	2,49	2,74
	>3	0,13	0,02	0,08	0,12	0,11	0,14	0,24
r_i^{**wil}	<-3	0,13	0,32	0,66	0,78	0,80	0,94	1,34
	<-2	2,28	3,00	3,78	4,06	4,03	4,25	4,90
	<-1	15,87	10,54	12,51	13,29	13,15	13,90	15,16
	>1	15,87	10,30	11,94	12,60	12,60	13,39	14,92
	>2	2,28	3,04	3,64	3,86	3,83	4,10	4,40
	>3	0,13	0,30	0,56	0,71	0,71	0,82	1,24

Tabela D.6: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 4b.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,10	0,20	0,24	0,25	0,28	0,42
	<-2	2,28	2,90	3,32	3,48	3,49	3,67	4,08
	<-1	15,87	17,40	18,79	19,30	19,27	19,78	20,96
	>1	15,87	11,86	13,07	13,32	13,33	13,66	14,52
	>2	2,28	1,24	1,52	1,67	1,66	1,81	2,10
	>3	0,13	0,00	0,04	0,06	0,06	0,08	0,18
r_i^{**dev}	<-3	0,13	0,66	1,02	1,20	1,23	1,41	1,86
	<-2	2,28	4,50	5,16	5,42	5,40	5,62	6,18
	<-1	15,87	12,66	14,88	16,00	15,80	16,68	18,86
	>1	15,87	8,24	9,80	10,42	10,39	11,11	12,20
	>2	2,28	1,88	2,70	2,89	2,87	3,05	3,54
	>3	0,13	0,14	0,36	0,48	0,48	0,59	0,86
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,00
	<-2	2,28	0,08	0,18	0,26	0,28	0,36	0,70
	<-1	15,87	16,66	17,49	17,88	17,96	18,42	19,44
	>1	15,87	13,02	13,76	14,16	14,12	14,40	15,18
	>2	2,28	2,40	3,22	3,47	3,42	3,61	4,14
	>3	0,13	0,38	0,60	0,70	0,69	0,80	1,02
r_i^{**pea}	<-3	0,13	0,00	0,00	0,00	0,02	0,02	0,16
	<-2	2,28	0,54	1,06	1,49	1,50	1,89	2,82
	<-1	15,87	11,46	13,15	13,87	13,72	14,28	15,48
	>1	15,87	8,96	10,71	11,52	11,52	12,43	13,66
	>2	2,28	3,56	4,48	4,72	4,69	4,91	5,34
	>3	0,13	0,96	1,56	1,68	1,69	1,86	2,22
r_i^{*qu}	<-3	0,13	0,04	0,08	0,12	0,12	0,16	0,26
	<-2	2,28	1,94	2,14	2,34	2,32	2,46	2,76
	<-1	15,87	14,04	15,52	15,89	15,91	16,34	17,36
	>1	15,87	14,44	15,76	16,01	16,02	16,37	16,96
	>2	2,28	1,52	2,02	2,23	2,22	2,40	2,74
	>3	0,13	0,00	0,04	0,08	0,09	0,12	0,30
r_i^{**qu}	<-3	0,13	0,32	0,58	0,75	0,73	0,86	1,18
	<-2	2,28	2,86	3,62	3,82	3,81	4,06	4,56
	<-1	15,87	9,90	11,95	12,85	12,65	13,24	14,62
	>1	15,87	10,14	12,08	12,69	12,71	13,48	14,52
	>2	2,28	2,30	3,54	3,71	3,73	3,99	4,48
	>3	0,13	0,16	0,50	0,68	0,66	0,80	1,12
r_i^{*ans}	<-3	0,13	0,06	0,14	0,16	0,17	0,20	0,30
	<-2	2,28	2,56	3,06	3,22	3,20	3,35	3,80
	<-1	15,87	17,26	18,66	19,10	19,10	19,57	20,56
	>1	15,87	11,08	12,88	13,23	13,17	13,50	14,42
	>2	2,28	0,92	1,39	1,53	1,52	1,70	1,96
	>3	0,13	0,00	0,02	0,04	0,04	0,06	0,14
r_i^{**ans}	<-3	0,13	0,42	0,86	1,06	1,04	1,23	1,62
	<-2	2,28	3,94	4,88	5,13	5,11	5,36	5,86
	<-1	15,87	12,50	14,73	15,77	15,59	16,55	18,28
	>1	15,87	8,04	9,66	10,27	10,21	10,86	11,88
	>2	2,28	1,52	2,50	2,72	2,69	2,91	3,40
	>3	0,13	0,08	0,28	0,42	0,41	0,50	0,78
r_i^{*wil}	<-3	0,13	0,08	0,18	0,22	0,23	0,26	0,36
	<-2	2,28	2,82	3,24	3,40	3,41	3,58	4,04
	<-1	15,87	17,40	18,76	19,26	19,24	19,73	20,88
	>1	15,87	12,12	13,12	13,32	13,36	13,68	14,56
	>2	2,28	1,32	1,60	1,74	1,73	1,86	2,20
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,18
r_i^{**wil}	<-3	0,13	0,58	0,99	1,17	1,18	1,36	1,82
	<-2	2,28	4,30	5,10	5,34	5,32	5,53	6,10
	<-1	15,87	12,62	14,82	15,95	15,77	16,68	18,70
	>1	15,87	8,38	9,84	10,46	10,44	11,12	12,38
	>2	2,28	2,08	2,78	2,99	2,94	3,11	3,62
	>3	0,13	0,16	0,40	0,52	0,51	0,62	0,94

Tabela D.7: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo no modelo ZAGA - Cenário 5.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,00	0,06	0,12	0,13	0,18	0,34
	<-2	2,28	1,84	2,76	3,04	3,02	3,32	3,82
	<-1	15,87	15,82	17,49	18,19	18,29	18,95	22,06
	>1	15,87	12,48	14,18	14,69	14,65	15,16	16,20
	>2	2,28	1,22	1,68	1,92	1,91	2,14	2,62
	>3	0,13	0,00	0,02	0,04	0,05	0,08	0,26
r_i^{**dev}	<-3	0,13	0,30	0,76	0,92	0,91	1,08	1,34
	<-2	2,28	3,70	4,56	4,90	4,84	5,16	5,78
	<-1	15,87	11,14	13,80	14,93	14,94	16,02	19,90
	>1	15,87	9,16	11,15	11,68	11,60	12,15	13,32
	>2	2,28	2,00	3,01	3,31	3,30	3,68	4,30
	>3	0,13	0,08	0,36	0,49	0,50	0,64	1,02
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,02
	<-2	2,28	0,02	0,71	1,22	1,12	1,58	2,12
	<-1	15,87	15,68	16,98	17,49	17,44	17,96	18,88
	>1	15,87	13,30	14,70	15,21	15,19	15,64	16,58
	>2	2,28	1,76	2,88	3,21	3,17	3,53	3,90
	>3	0,13	0,02	0,26	0,40	0,39	0,55	0,82
r_i^{**pea}	<-3	0,13	0,00	0,02	0,12	0,16	0,29	0,56
	<-2	2,28	0,28	2,07	2,91	2,67	3,37	4,42
	<-1	15,87	10,80	13,14	13,82	13,72	14,41	16,04
	>1	15,87	9,34	11,68	12,49	12,37	13,14	14,24
	>2	2,28	3,56	4,42	4,61	4,64	4,99	5,38
	>3	0,13	0,66	1,22	1,34	1,32	1,49	1,92
r_i^{*qu}	<-3	0,13	0,00	0,02	0,06	0,07	0,10	0,26
	<-2	2,28	1,24	2,00	2,29	2,26	2,55	3,00
	<-1	15,87	15,12	15,88	16,22	16,20	16,60	17,24
	>1	15,87	14,78	15,71	16,22	16,22	16,60	17,94
	>2	2,28	1,38	2,04	2,25	2,24	2,45	3,02
	>3	0,13	0,00	0,04	0,06	0,07	0,08	0,28
r_i^{**qu}	<-3	0,13	0,08	0,48	0,62	0,62	0,81	1,10
	<-2	2,28	2,46	3,56	3,85	3,80	4,16	5,06
	<-1	15,87	10,18	12,48	13,09	12,98	13,78	15,10
	>1	15,87	9,66	12,26	13,15	12,98	13,72	14,98
	>2	2,28	2,60	3,52	3,88	3,82	4,10	4,68
	>3	0,13	0,12	0,47	0,58	0,59	0,78	1,16
r_i^{*ans}	<-3	0,13	0,00	0,04	0,08	0,10	0,14	0,32
	<-2	2,28	1,64	2,54	2,80	2,81	3,13	3,58
	<-1	15,87	15,70	17,39	18,10	18,11	18,79	21,26
	>1	15,87	12,40	14,07	14,54	14,49	15,04	16,10
	>2	2,28	0,94	1,54	1,80	1,77	2,03	2,58
	>3	0,13	0,00	0,02	0,04	0,04	0,06	0,24
r_i^{**ans}	<-3	0,13	0,10	0,67	0,81	0,80	1,00	1,26
	<-2	2,28	3,26	4,26	4,68	4,59	4,90	5,60
	<-1	15,87	10,88	13,62	14,77	14,73	15,81	19,06
	>1	15,87	9,00	10,91	11,52	11,41	12,01	13,20
	>2	2,28	1,64	2,80	3,16	3,14	3,54	4,26
	>3	0,13	0,06	0,30	0,46	0,45	0,60	1,00
r_i^{*wil}	<-3	0,13	0,00	0,06	0,12	0,12	0,18	0,34
	<-2	2,28	1,82	2,70	3,00	2,97	3,29	3,72
	<-1	15,87	15,82	17,48	18,17	18,26	18,93	21,86
	>1	15,87	12,48	14,21	14,71	14,67	15,16	16,20
	>2	2,28	1,34	1,74	1,97	1,95	2,16	2,66
	>3	0,13	0,00	0,02	0,06	0,06	0,08	0,26
r_i^{**wil}	<-3	0,13	0,26	0,74	0,88	0,88	1,06	1,34
	<-2	2,28	3,66	4,49	4,86	4,79	5,10	5,74
	<-1	15,87	11,14	13,78	14,92	14,91	15,99	19,66
	>1	15,87	9,18	11,18	11,76	11,63	12,19	13,32
	>2	2,28	2,20	3,05	3,38	3,35	3,69	4,32
	>3	0,13	0,12	0,36	0,51	0,52	0,65	1,02

Tabela D.8: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 6a.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,00	0,10	0,14	0,14	0,18	0,28
	<-2	2,28	2,06	2,38	2,54	2,58	2,76	3,12
	<-1	15,87	14,06	16,34	16,78	16,70	17,12	18,56
	>1	15,87	14,32	15,22	15,52	15,52	15,87	16,60
	>2	2,28	1,70	2,10	2,25	2,25	2,41	2,80
	>3	0,13	0,02	0,08	0,12	0,11	0,14	0,26
r_i^{**dev}	<-3	0,13	0,26	0,68	0,88	0,86	1,00	1,38
	<-2	2,28	2,98	3,96	4,23	4,23	4,55	5,08
	<-1	15,87	9,94	12,86	13,35	13,33	13,95	15,92
	>1	15,87	9,86	11,82	12,34	12,26	12,72	13,94
	>2	2,28	2,28	3,56	3,79	3,75	4,00	4,44
	>3	0,13	0,22	0,60	0,70	0,72	0,84	1,34
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,05	0,06	0,12
	<-2	2,28	1,46	1,90	2,04	2,05	2,22	2,60
	<-1	15,87	13,88	16,11	16,54	16,50	16,91	18,32
	>1	15,87	14,44	15,40	15,68	15,69	16,00	16,74
	>2	2,28	2,06	2,56	2,69	2,71	2,86	3,28
	>3	0,13	0,06	0,18	0,22	0,22	0,28	0,40
r_i^{**pea}	<-3	0,13	0,06	0,40	0,53	0,54	0,64	1,02
	<-2	2,28	2,54	3,36	3,70	3,67	3,96	4,56
	<-1	15,87	9,82	12,53	13,09	13,02	13,58	15,46
	>1	15,87	10,10	12,05	12,60	12,53	13,04	14,24
	>2	2,28	2,70	4,04	4,27	4,24	4,45	4,92
	>3	0,13	0,38	0,88	1,02	1,01	1,15	1,64
r_i^{*qu}	<-3	0,13	0,00	0,08	0,12	0,11	0,14	0,24
	<-2	2,28	0,68	2,10	2,30	2,27	2,50	2,84
	<-1	15,87	12,40	15,48	15,98	15,92	16,46	17,98
	>1	15,87	12,88	15,58	15,96	15,91	16,34	17,16
	>2	2,28	0,72	2,18	2,32	2,28	2,48	2,88
	>3	0,13	0,00	0,06	0,12	0,11	0,14	0,26
r_i^{**qu}	<-3	0,13	0,08	0,58	0,76	0,73	0,90	1,22
	<-2	2,28	1,28	3,64	3,87	3,81	4,17	4,64
	<-1	15,87	9,14	12,19	12,68	12,59	13,20	14,96
	>1	15,87	10,04	12,20	12,70	12,58	13,11	14,08
	>2	2,28	1,34	3,64	3,92	3,83	4,13	4,58
	>3	0,13	0,06	0,60	0,72	0,72	0,86	1,26
r_i^{*ans}	<-3	0,13	0,00	0,08	0,14	0,13	0,16	0,26
	<-2	2,28	0,78	2,28	2,42	2,43	2,68	3,06
	<-1	15,87	12,92	16,10	16,56	16,51	17,00	18,34
	>1	15,87	11,88	15,04	15,40	15,35	15,74	16,50
	>2	2,28	0,68	2,00	2,16	2,12	2,31	2,70
	>3	0,13	0,00	0,06	0,10	0,09	0,12	0,26
r_i^{**ans}	<-3	0,13	0,08	0,64	0,84	0,79	0,95	1,32
	<-2	2,28	1,42	3,86	4,13	4,06	4,44	4,90
	<-1	15,87	9,64	12,62	13,21	13,12	13,75	15,66
	>1	15,87	9,58	11,62	12,17	12,05	12,54	13,52
	>2	2,28	1,06	3,42	3,70	3,60	3,89	4,34
	>3	0,13	0,04	0,56	0,67	0,66	0,81	1,14
r_i^{*wil}	<-3	0,13	0,00	0,10	0,14	0,14	0,17	0,28
	<-2	2,28	1,98	2,38	2,54	2,56	2,75	3,12
	<-1	15,87	14,02	16,32	16,78	16,69	17,12	18,54
	>1	15,87	14,32	15,25	15,52	15,52	15,87	16,60
	>2	2,28	1,74	2,12	2,26	2,26	2,41	2,80
	>3	0,13	0,02	0,08	0,12	0,11	0,14	0,26
r_i^{**wil}	<-3	0,13	0,24	0,68	0,88	0,85	1,00	1,36
	<-2	2,28	2,88	3,96	4,23	4,22	4,54	5,06
	<-1	15,87	9,94	12,85	13,34	13,32	13,95	15,92
	>1	15,87	9,88	11,84	12,34	12,27	12,72	13,94
	>2	2,28	2,42	3,56	3,81	3,77	4,02	4,46
	>3	0,13	0,24	0,60	0,71	0,72	0,85	1,34

Tabela D.9: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 6b.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,10	0,13	0,14	0,18	0,30
	<-2	2,28	2,02	2,40	2,55	2,56	2,73	3,12
	<-1	15,87	14,76	16,45	16,72	16,72	17,02	17,94
	>1	15,87	13,54	15,21	15,53	15,50	15,88	16,74
	>2	2,28	1,68	2,06	2,22	2,23	2,40	3,02
	>3	0,13	0,04	0,08	0,12	0,12	0,15	0,24
r_i^{**dev}	<-3	0,13	0,24	0,66	0,84	0,83	1,00	1,42
	<-2	2,28	3,24	3,92	4,16	4,16	4,40	5,02
	<-1	15,87	10,18	12,68	13,46	13,40	14,12	15,64
	>1	15,87	9,18	11,52	12,29	12,28	13,05	14,70
	>2	2,28	2,82	3,45	3,72	3,69	3,94	4,50
	>3	0,13	0,20	0,58	0,71	0,72	0,88	1,34
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,05	0,08	0,14
	<-2	2,28	1,58	1,92	2,02	2,05	2,18	2,52
	<-1	15,87	14,72	16,19	16,55	16,51	16,82	17,58
	>1	15,87	13,58	15,36	15,68	15,69	16,05	17,02
	>2	2,28	2,04	2,51	2,71	2,69	2,86	3,58
	>3	0,13	0,08	0,20	0,24	0,24	0,28	0,44
r_i^{**pea}	<-3	0,13	0,14	0,40	0,52	0,53	0,65	1,02
	<-2	2,28	2,52	3,38	3,61	3,59	3,91	4,52
	<-1	15,87	10,04	12,43	13,13	13,09	13,80	15,12
	>1	15,87	9,30	11,72	12,56	12,56	13,35	15,10
	>2	2,28	3,30	4,02	4,18	4,19	4,44	4,94
	>3	0,13	0,46	0,85	1,02	1,00	1,17	1,72
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,15	0,28
	<-2	2,28	1,64	2,12	2,26	2,27	2,44	2,84
	<-1	15,87	13,96	15,55	15,99	15,95	16,32	17,10
	>1	15,87	13,78	15,59	15,97	15,92	16,34	17,34
	>2	2,28	1,46	2,06	2,25	2,26	2,46	3,08
	>3	0,13	0,02	0,08	0,12	0,11	0,14	0,24
r_i^{**qu}	<-3	0,13	0,14	0,54	0,72	0,70	0,85	1,24
	<-2	2,28	2,12	3,54	3,81	3,74	4,02	4,74
	<-1	15,87	9,34	12,00	12,73	12,66	13,37	14,60
	>1	15,87	9,20	11,84	12,69	12,61	13,37	15,12
	>2	2,28	2,24	3,51	3,81	3,76	4,12	4,56
	>3	0,13	0,14	0,58	0,73	0,73	0,87	1,34
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,13	0,16	0,28
	<-2	2,28	1,76	2,22	2,44	2,43	2,62	2,98
	<-1	15,87	14,60	16,20	16,61	16,55	16,87	17,74
	>1	15,87	13,32	15,05	15,37	15,34	15,76	16,62
	>2	2,28	1,30	1,92	2,11	2,11	2,32	2,84
	>3	0,13	0,02	0,06	0,10	0,10	0,14	0,24
r_i^{**ans}	<-3	0,13	0,14	0,59	0,76	0,76	0,92	1,30
	<-2	2,28	2,42	3,76	4,05	3,98	4,24	5,00
	<-1	15,87	9,74	12,57	13,35	13,20	13,92	15,12
	>1	15,87	8,80	11,36	12,15	12,08	12,77	14,54
	>2	2,28	2,14	3,28	3,58	3,53	3,83	4,40
	>3	0,13	0,10	0,52	0,68	0,66	0,81	1,30
r_i^{*wil}	<-3	0,13	0,02	0,10	0,13	0,14	0,18	0,28
	<-2	2,28	2,02	2,38	2,53	2,55	2,72	3,12
	<-1	15,87	14,76	16,44	16,72	16,72	17,02	17,94
	>1	15,87	13,54	15,21	15,53	15,51	15,88	16,76
	>2	2,28	1,68	2,08	2,23	2,24	2,41	3,04
	>3	0,13	0,04	0,08	0,12	0,12	0,15	0,24
r_i^{**wil}	<-3	0,13	0,24	0,66	0,84	0,82	1,00	1,40
	<-2	2,28	3,24	3,90	4,14	4,15	4,40	5,02
	<-1	15,87	10,18	12,68	13,45	13,39	14,11	15,60
	>1	15,87	9,20	11,52	12,29	12,29	13,05	14,72
	>2	2,28	2,84	3,48	3,72	3,71	3,94	4,50
	>3	0,13	0,20	0,58	0,72	0,73	0,88	1,34

Tabela D.10: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário γ_a .

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,04	0,10	0,14	0,14	0,18	0,32
	<-2	2,28	1,84	2,46	2,61	2,60	2,74	3,18
	<-1	15,87	15,38	16,52	16,86	16,88	17,23	18,32
	>1	15,87	14,26	15,12	15,59	15,59	15,95	16,80
	>2	2,28	1,60	2,04	2,22	2,20	2,35	2,72
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
r_i^{**dev}	<-3	0,13	0,32	0,50	0,56	0,58	0,68	0,86
	<-2	2,28	3,22	3,72	3,89	3,89	4,06	4,60
	<-1	15,87	12,14	14,15	14,58	14,63	15,14	16,38
	>1	15,87	11,36	12,91	13,38	13,41	13,96	15,04
	>2	2,28	2,78	3,16	3,39	3,38	3,56	3,94
	>3	0,13	0,18	0,38	0,48	0,47	0,56	0,74
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,04	0,06	0,18
	<-2	2,28	1,50	1,92	2,00	2,01	2,12	2,48
	<-1	15,87	15,12	16,10	16,56	16,55	16,82	17,96
	>1	15,87	14,58	15,54	15,91	15,91	16,30	17,20
	>2	2,28	1,86	2,58	2,74	2,74	2,94	3,28
	>3	0,13	0,10	0,18	0,24	0,24	0,28	0,42
r_i^{**pea}	<-3	0,13	0,14	0,26	0,32	0,32	0,38	0,58
	<-2	2,28	2,54	3,06	3,21	3,23	3,39	3,88
	<-1	15,87	11,88	13,76	14,16	14,20	14,74	16,00
	>1	15,87	11,72	13,31	13,78	13,80	14,41	15,50
	>2	2,28	3,24	3,74	3,98	3,95	4,14	4,74
	>3	0,13	0,36	0,62	0,75	0,73	0,84	1,00
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,11	0,14	0,32
	<-2	2,28	1,04	2,14	2,28	2,27	2,42	2,84
	<-1	15,87	13,76	15,72	16,03	16,00	16,36	17,40
	>1	15,87	14,66	15,70	16,06	16,06	16,45	17,36
	>2	2,28	1,10	2,04	2,28	2,23	2,42	2,82
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
r_i^{**qu}	<-3	0,13	0,24	0,42	0,48	0,48	0,56	0,74
	<-2	2,28	1,98	3,30	3,46	3,46	3,67	4,12
	<-1	15,87	10,28	13,29	13,73	13,76	14,37	15,72
	>1	15,87	10,54	13,36	13,78	13,80	14,49	15,44
	>2	2,28	1,98	3,24	3,45	3,43	3,66	4,16
	>3	0,13	0,18	0,38	0,46	0,47	0,56	0,74
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,12	0,16	0,32
	<-2	2,28	1,14	2,32	2,46	2,45	2,63	3,02
	<-1	15,87	14,54	16,34	16,68	16,67	17,02	18,00
	>1	15,87	13,96	14,96	15,36	15,38	15,80	16,76
	>2	2,28	0,96	1,90	2,10	2,07	2,26	2,62
	>3	0,13	0,00	0,06	0,08	0,10	0,14	0,20
r_i^{**ans}	<-3	0,13	0,24	0,46	0,52	0,52	0,60	0,80
	<-2	2,28	2,22	3,54	3,74	3,71	3,92	4,30
	<-1	15,87	10,82	13,91	14,44	14,39	15,01	16,24
	>1	15,87	9,88	12,74	13,21	13,20	13,78	14,80
	>2	2,28	1,84	3,00	3,24	3,21	3,42	3,86
	>3	0,13	0,16	0,34	0,42	0,42	0,51	0,70
r_i^{*wil}	<-3	0,13	0,04	0,10	0,14	0,14	0,18	0,32
	<-2	2,28	1,82	2,46	2,60	2,59	2,74	3,18
	<-1	15,87	15,32	16,50	16,86	16,87	17,22	18,30
	>1	15,87	14,26	15,12	15,60	15,60	15,98	16,82
	>2	2,28	1,66	2,04	2,24	2,22	2,38	2,72
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
r_i^{**wil}	<-3	0,13	0,32	0,50	0,56	0,58	0,67	0,84
	<-2	2,28	3,18	3,70	3,87	3,87	4,06	4,56
	<-1	15,87	12,08	14,13	14,56	14,61	15,14	16,38
	>1	15,87	11,38	12,93	13,41	13,43	13,98	15,04
	>2	2,28	2,84	3,18	3,40	3,40	3,57	3,94
	>3	0,13	0,18	0,38	0,48	0,47	0,56	0,74

Tabela D.11: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 7b.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,12	0,14	0,15	0,20	0,26
	<-2	2,28	1,94	2,42	2,60	2,60	2,76	3,22
	<-1	15,87	13,78	16,38	16,79	16,72	17,16	18,56
	>1	15,87	14,36	15,26	15,58	15,55	15,83	16,86
	>2	2,28	1,68	2,10	2,20	2,21	2,35	2,66
	>3	0,13	0,00	0,06	0,10	0,10	0,14	0,22
r_i^{**dev}	<-3	0,13	0,26	0,50	0,76	0,77	1,01	1,52
	<-2	2,28	2,94	3,72	4,06	4,06	4,35	5,08
	<-1	15,87	9,18	12,80	13,94	13,80	15,06	16,34
	>1	15,87	9,48	11,80	12,78	12,68	13,53	15,56
	>2	2,28	2,34	3,34	3,53	3,52	3,79	4,30
	>3	0,13	0,16	0,44	0,56	0,61	0,74	1,28
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,05	0,08	0,14
	<-2	2,28	1,50	1,90	2,06	2,06	2,18	2,72
	<-1	15,87	13,72	16,07	16,46	16,46	16,88	18,40
	>1	15,87	14,40	15,47	15,78	15,79	16,12	17,38
	>2	2,28	2,16	2,50	2,67	2,69	2,86	3,10
	>3	0,13	0,08	0,16	0,22	0,22	0,26	0,40
r_i^{**pea}	<-3	0,13	0,10	0,28	0,42	0,49	0,69	1,28
	<-2	2,28	2,48	3,08	3,48	3,47	3,79	4,44
	<-1	15,87	9,00	12,56	13,57	13,44	14,61	15,80
	>1	15,87	9,66	12,09	13,10	12,99	13,93	15,98
	>2	2,28	3,08	3,86	4,05	4,05	4,28	4,70
	>3	0,13	0,30	0,72	0,86	0,89	1,06	1,54
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,14	0,24
	<-2	2,28	1,72	2,14	2,26	2,29	2,42	2,92
	<-1	15,87	13,38	15,56	15,94	15,92	16,36	17,62
	>1	15,87	14,64	15,70	16,02	15,98	16,29	17,34
	>2	2,28	1,68	2,10	2,22	2,24	2,37	2,82
	>3	0,13	0,00	0,06	0,10	0,10	0,14	0,22
r_i^{**qu}	<-3	0,13	0,18	0,40	0,64	0,65	0,84	1,38
	<-2	2,28	2,62	3,38	3,67	3,65	3,97	4,62
	<-1	15,87	8,54	12,20	13,18	13,01	14,16	15,20
	>1	15,87	9,68	12,20	13,16	13,03	13,95	15,70
	>2	2,28	2,24	3,40	3,61	3,58	3,88	4,42
	>3	0,13	0,14	0,44	0,58	0,61	0,77	1,24
r_i^{*ans}	<-3	0,13	0,02	0,10	0,12	0,12	0,16	0,24
	<-2	2,28	1,86	2,30	2,44	2,46	2,58	3,08
	<-1	15,87	13,66	16,14	16,58	16,54	16,96	18,38
	>1	15,87	14,30	15,02	15,41	15,38	15,72	16,46
	>2	2,28	1,54	1,96	2,08	2,09	2,19	2,60
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,18
r_i^{**ans}	<-3	0,13	0,18	0,44	0,72	0,71	0,96	1,44
	<-2	2,28	2,74	3,54	3,92	3,88	4,17	4,86
	<-1	15,87	8,90	12,61	13,76	13,59	14,80	15,84
	>1	15,87	9,28	11,59	12,66	12,47	13,33	14,94
	>2	2,28	2,04	3,20	3,36	3,36	3,65	4,20
	>3	0,13	0,14	0,40	0,52	0,56	0,71	1,16
r_i^{*wil}	<-3	0,13	0,02	0,12	0,14	0,14	0,18	0,26
	<-2	2,28	1,94	2,42	2,60	2,59	2,73	3,20
	<-1	15,87	13,78	16,38	16,78	16,72	17,16	18,54
	>1	15,87	14,36	15,26	15,58	15,56	15,85	16,88
	>2	2,28	1,68	2,10	2,22	2,23	2,38	2,66
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
r_i^{**wil}	<-3	0,13	0,26	0,48	0,76	0,76	0,99	1,52
	<-2	2,28	2,90	3,68	4,05	4,04	4,33	5,00
	<-1	15,87	9,16	12,80	13,94	13,79	15,06	16,34
	>1	15,87	9,52	11,80	12,79	12,69	13,56	15,60
	>2	2,28	2,40	3,34	3,54	3,54	3,80	4,32
	>3	0,13	0,16	0,46	0,57	0,62	0,75	1,30

Tabela D.12: *Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 8a.*

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,12	0,14	0,15	0,18	0,30
	<-2	2,28	2,02	2,38	2,50	2,53	2,68	3,44
	<-1	15,87	15,30	16,24	16,65	16,62	16,94	18,30
	>1	15,87	14,00	15,04	15,40	15,45	15,84	16,94
	>2	2,28	1,62	2,00	2,15	2,16	2,30	2,72
	>3	0,13	0,02	0,08	0,10	0,12	0,14	0,28
r_i^{**dev}	<-3	0,13	0,28	0,60	0,80	0,80	0,96	1,34
	<-2	2,28	3,16	3,82	4,08	4,08	4,34	5,18
	<-1	15,87	10,40	12,49	13,36	13,36	14,29	16,14
	>1	15,87	9,30	11,63	12,18	12,29	13,09	15,06
	>2	2,28	2,54	3,30	3,56	3,55	3,77	4,52
	>3	0,13	0,26	0,48	0,63	0,65	0,81	1,20
r_i^{*pea}	<-3	0,13	0,00	0,04	0,06	0,06	0,08	0,16
	<-2	2,28	1,52	1,84	1,98	2,00	2,14	2,78
	<-1	15,87	15,02	16,04	16,43	16,39	16,70	18,12
	>1	15,87	14,08	15,23	15,62	15,65	16,02	17,20
	>2	2,28	2,00	2,44	2,62	2,62	2,78	3,28
	>3	0,13	0,08	0,18	0,24	0,24	0,28	0,50
r_i^{**pea}	<-3	0,13	0,06	0,34	0,47	0,50	0,64	1,08
	<-2	2,28	2,54	3,16	3,52	3,49	3,79	4,50
	<-1	15,87	10,34	12,22	13,04	13,03	13,90	15,76
	>1	15,87	9,38	11,87	12,48	12,59	13,44	15,38
	>2	2,28	3,04	3,84	4,02	4,05	4,27	4,96
	>3	0,13	0,46	0,78	0,92	0,94	1,09	1,62
r_i^{*qu}	<-3	0,13	0,02	0,08	0,12	0,13	0,16	0,26
	<-2	2,28	1,82	2,16	2,27	2,29	2,43	3,16
	<-1	15,87	14,52	15,56	15,93	15,90	16,20	17,52
	>1	15,87	14,56	15,60	15,92	15,98	16,38	17,32
	>2	2,28	1,68	2,08	2,26	2,25	2,40	2,82
	>3	0,13	0,02	0,08	0,12	0,12	0,16	0,28
r_i^{**qu}	<-3	0,13	0,18	0,54	0,70	0,70	0,84	1,26
	<-2	2,28	2,78	3,50	3,75	3,74	4,00	4,82
	<-1	15,87	9,92	11,94	12,70	12,70	13,60	15,36
	>1	15,87	9,54	12,05	12,67	12,75	13,59	15,46
	>2	2,28	2,72	3,48	3,72	3,71	3,95	4,68
	>3	0,13	0,26	0,50	0,66	0,68	0,84	1,30
r_i^{*ans}	<-3	0,13	0,02	0,10	0,14	0,14	0,18	0,26
	<-2	2,28	1,90	2,30	2,44	2,46	2,61	3,40
	<-1	15,87	15,12	16,14	16,56	16,53	16,86	18,18
	>1	15,87	13,92	14,96	15,32	15,36	15,79	16,86
	>2	2,28	1,58	1,94	2,10	2,09	2,22	2,72
	>3	0,13	0,02	0,08	0,10	0,11	0,14	0,26
r_i^{**ans}	<-3	0,13	0,18	0,58	0,76	0,76	0,92	1,30
	<-2	2,28	2,92	3,74	4,00	3,99	4,26	5,10
	<-1	15,87	10,38	12,42	13,28	13,26	14,19	16,02
	>1	15,87	9,16	11,52	12,13	12,19	13,00	14,82
	>2	2,28	2,48	3,20	3,46	3,47	3,70	4,46
	>3	0,13	0,22	0,46	0,60	0,62	0,78	1,14
r_i^{*wil}	<-3	0,13	0,02	0,10	0,14	0,15	0,18	0,28
	<-2	2,28	2,00	2,38	2,50	2,52	2,66	3,42
	<-1	15,87	15,30	16,24	16,64	16,61	16,94	18,26
	>1	15,87	14,00	15,04	15,40	15,45	15,84	16,94
	>2	2,28	1,64	2,00	2,16	2,16	2,30	2,72
	>3	0,13	0,02	0,08	0,11	0,12	0,14	0,28
r_i^{**wil}	<-3	0,13	0,28	0,60	0,80	0,79	0,96	1,34
	<-2	2,28	3,14	3,82	4,07	4,07	4,34	5,16
	<-1	15,87	10,40	12,49	13,36	13,35	14,28	16,14
	>1	15,87	9,30	11,63	12,18	12,29	13,11	15,06
	>2	2,28	2,56	3,30	3,56	3,56	3,78	4,52
	>3	0,13	0,26	0,48	0,64	0,65	0,82	1,22

Tabela D.13: *Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAGA - Cenário 8b.*

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,02	0,07	0,10	0,11	0,14	0,24
	<-2	2,28	2,20	2,57	2,70	2,74	2,90	3,40
	<-1	15,87	14,92	16,38	16,93	16,83	17,25	18,06
	>1	15,87	13,86	15,22	15,83	15,71	16,20	16,88
	>2	2,28	1,74	2,15	2,32	2,33	2,56	3,00
	>3	0,13	0,02	0,06	0,10	0,10	0,12	0,20
r_i^{**dev}	<-3	0,13	0,36	0,73	0,92	0,90	1,08	1,42
	<-2	2,28	3,28	4,20	4,36	0,40	4,67	5,36
	<-1	15,87	10,82	13,07	13,65	3,56	14,38	15,46
	>1	15,87	9,64	11,78	12,69	2,49	13,21	14,50
	>2	2,28	2,62	3,70	3,97	0,93	4,22	4,60
	>3	0,13	0,28	0,59	0,70	0,73	0,88	1,20
r_i^{*pea}	<-3	0,13	0,00	0,02	0,04	0,03	0,06	0,10
	<-2	2,28	1,60	2,06	2,20	2,21	2,38	2,68
	<-1	15,87	14,84	16,18	16,66	16,63	17,07	17,76
	>1	15,87	13,94	15,37	16,08	15,90	16,43	17,10
	>2	2,28	2,20	2,61	2,80	2,80	3,01	3,58
	>3	0,13	0,08	0,14	0,18	0,20	0,24	0,40
r_i^{**pea}	<-3	0,13	0,18	0,45	0,56	0,60	0,74	1,04
	<-2	2,28	2,58	3,66	3,87	3,86	4,10	4,68
	<-1	15,87	10,68	12,70	13,39	13,28	14,02	15,14
	>1	15,87	9,76	12,00	12,94	12,76	13,51	14,94
	>2	2,28	3,18	4,19	4,44	4,41	4,78	5,08
	>3	0,13	0,36	0,87	1,02	1,02	1,18	1,54
r_i^{*qu}	<-3	0,13	0,00	0,04	0,06	0,07	0,10	0,16
	<-2	2,28	1,24	2,08	2,30	2,27	2,53	3,08
	<-1	15,87	14,12	15,41	16,11	15,98	16,59	17,20
	>1	15,87	13,90	15,41	16,03	15,95	16,62	17,18
	>2	2,28	1,00	1,93	2,26	2,23	2,56	3,12
	>3	0,13	0,02	0,06	0,07	0,09	0,12	0,20
r_i^{**qu}	<-3	0,13	0,12	0,57	0,71	0,71	0,86	1,18
	<-2	2,28	1,96	3,52	3,85	3,82	4,24	4,88
	<-1	15,87	9,98	12,16	12,85	12,67	13,43	14,58
	>1	15,87	9,56	12,04	12,83	12,61	13,36	14,82
	>2	2,28	1,80	3,41	3,88	3,82	4,28	4,70
	>3	0,13	0,08	0,49	0,66	0,67	0,82	1,16
r_i^{*ans}	<-3	0,13	0,00	0,06	0,08	0,08	0,10	0,16
	<-2	2,28	1,36	2,23	2,45	2,43	2,70	3,22
	<-1	15,87	14,64	15,92	16,65	16,52	17,00	17,86
	>1	15,87	13,50	14,85	15,41	15,38	16,05	16,72
	>2	2,28	0,86	1,83	2,08	2,07	2,42	2,88
	>3	0,13	0,02	0,04	0,06	0,08	0,10	0,18
r_i^{**ans}	<-3	0,13	0,12	0,63	0,78	0,77	0,94	1,26
	<-2	2,28	2,06	3,82	4,07	4,06	4,51	5,16
	<-1	15,87	10,38	12,59	13,35	13,19	13,96	15,10
	>1	15,87	9,26	11,50	12,35	12,11	12,84	14,28
	>2	2,28	1,64	3,25	3,66	3,59	4,06	4,48
	>3	0,13	0,08	0,45	0,60	0,62	0,76	1,06
r_i^{*wil}	<-3	0,13	0,02	0,06	0,10	0,11	0,14	0,22
	<-2	2,28	2,16	2,53	2,68	2,71	2,88	3,40
	<-1	15,87	14,90	16,34	16,91	16,82	17,25	18,06
	>1	15,87	13,86	15,24	15,86	15,72	16,20	16,88
	>2	2,28	1,88	2,17	2,36	2,37	2,58	3,02
	>3	0,13	0,04	0,06	0,10	0,10	0,12	0,20
r_i^{**wil}	<-3	0,13	0,32	0,72	0,89	0,89	1,05	1,40
	<-2	2,28	3,20	4,17	4,34	4,36	4,66	5,32
	<-1	15,87	10,80	13,06	13,64	13,54	14,35	15,44
	>1	15,87	9,68	11,78	12,69	12,50	13,23	14,52
	>2	2,28	2,70	3,79	4,01	3,96	4,22	4,62
	>3	0,13	0,28	0,59	0,72	0,75	0,92	1,20

Apêndice E

Tabelas dos estudos de simulação no modelo ZAIG

Tabela E.1: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 1.

Resíduos	Intervalo do resíduo	Valor						
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,12	0,18	0,18	0,22	0,40
	<-2	2,28	2,12	2,70	2,90	2,93	3,14	3,76
	<-1	15,87	15,64	17,30	17,84	17,92	18,45	21,10
	>1	15,87	11,90	14,05	14,42	14,34	14,80	16,16
	>2	2,28	1,22	1,72	1,88	1,86	2,03	2,38
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,20
r_i^{**dev}	<-3	0,13	0,52	0,80	0,92	0,93	1,04	1,48
	<-2	2,28	3,78	4,42	4,65	4,63	4,81	5,46
	<-1	15,87	11,02	13,37	14,50	14,50	15,51	18,86
	>1	15,87	9,28	10,96	11,33	11,29	11,72	12,72
	>2	2,28	1,92	2,89	3,25	3,18	3,48	3,92
	>3	0,13	0,12	0,38	0,54	0,54	0,68	1,00
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,00	0,12
	<-2	2,28	0,12	0,84	1,28	1,21	1,56	2,28
	<-1	15,87	15,44	16,73	17,01	17,08	17,41	18,54
	>1	15,87	13,46	14,60	14,94	14,93	15,34	16,46
	>2	2,28	1,86	2,80	3,06	3,09	3,40	4,28
	>3	0,13	0,08	0,34	0,44	0,46	0,59	1,12
r_i^{**pea}	<-3	0,13	0,00	0,04	0,14	0,23	0,34	0,98
	<-2	2,28	0,54	2,12	2,90	2,68	3,29	4,00
	<-1	15,87	10,62	12,64	13,50	13,31	14,04	15,50
	>1	15,87	9,52	11,54	12,16	12,09	12,62	14,04
	>2	2,28	3,32	4,18	4,50	4,47	4,76	5,20
	>3	0,13	0,96	1,22	1,34	1,35	1,47	1,88
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,11	0,14	0,28
	<-2	2,28	1,52	2,14	2,28	2,31	2,48	2,90
	<-1	15,87	14,36	15,58	15,96	15,96	16,20	17,50
	>1	15,87	14,38	15,65	15,97	15,97	16,38	17,40
	>2	2,28	1,44	2,08	2,24	2,24	2,38	2,88
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,26
r_i^{**qu}	<-3	0,13	0,38	0,54	0,68	0,69	0,80	1,26
	<-2	2,28	2,90	3,62	3,80	3,79	3,99	4,70
	<-1	15,87	9,68	12,01	12,84	12,69	13,54	15,26
	>1	15,87	9,42	12,17	12,86	12,75	13,54	15,00
	>2	2,28	2,90	3,52	3,73	3,73	4,00	4,40
	>3	0,13	0,28	0,52	0,64	0,64	0,78	1,04
r_i^{*ans}	<-3	0,13	0,04	0,10	0,12	0,13	0,16	0,28
	<-2	2,28	1,66	2,50	2,71	2,72	2,95	3,56
	<-1	15,87	15,30	17,06	17,72	17,76	18,30	20,80
	>1	15,87	11,62	13,84	14,30	14,20	14,70	16,06
	>2	2,28	1,08	1,56	1,74	1,71	1,88	2,22
	>3	0,13	0,00	0,02	0,06	0,06	0,08	0,18
r_i^{**ans}	<-3	0,13	0,42	0,66	0,82	0,81	0,92	1,44
	<-2	2,28	3,28	4,18	4,40	4,40	4,63	5,22
	<-1	15,87	10,48	13,23	14,35	14,28	15,33	18,46
	>1	15,87	8,78	10,68	11,18	11,12	11,59	12,50
	>2	2,28	1,58	2,70	3,06	3,00	3,32	3,68
	>3	0,13	0,08	0,30	0,50	0,46	0,58	0,86
r_i^{*wil}	<-3	0,13	0,04	0,12	0,16	0,17	0,22	0,34
	<-2	2,28	2,08	2,68	2,85	2,89	3,11	3,66
	<-1	15,87	15,62	17,30	17,83	17,90	18,45	20,96
	>1	15,87	11,96	14,05	14,45	14,36	14,82	16,18
	>2	2,28	1,26	1,78	1,92	1,90	2,06	2,46
	>3	0,13	0,00	0,04	0,08	0,08	0,11	0,20
r_i^{**wil}	<-3	0,13	0,50	0,76	0,90	0,90	1,01	1,48
	<-2	2,28	3,72	4,36	4,60	4,59	4,80	5,42
	<-1	15,87	11,02	13,36	14,48	14,48	15,48	18,80
	>1	15,87	9,28	10,96	11,35	11,32	11,74	12,74
	>2	2,28	1,94	2,92	3,29	3,22	3,51	4,00
	>3	0,13	0,18	0,42	0,54	0,56	0,70	1,00

Tabela E.2: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 2a.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,14	0,18	0,18	0,22	0,36
	<-2	2,28	2,36	2,74	2,94	2,96	3,16	3,74
	<-1	15,87	16,54	17,73	18,22	18,37	19,03	20,62
	>1	15,87	12,30	13,49	14,34	14,19	14,84	15,56
	>2	2,28	1,26	1,70	1,87	1,85	2,04	2,44
	>3	0,13	0,00	0,04	0,08	0,08	0,10	0,20
r_i^{**dev}	<-3	0,13	0,24	0,38	0,44	0,45	0,52	0,70
	<-2	2,28	3,30	3,68	3,91	3,92	4,13	4,56
	<-1	15,87	14,06	15,96	16,93	17,01	18,02	19,66
	>1	15,87	11,34	12,56	13,03	12,96	13,36	14,12
	>2	2,28	1,64	2,22	2,63	2,55	2,83	3,52
	>3	0,13	0,04	0,16	0,21	0,23	0,29	0,68
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,04
	<-2	2,28	0,02	0,62	1,07	0,99	1,28	2,00
	<-1	15,87	15,20	16,18	16,55	16,59	17,02	17,78
	>1	15,87	13,88	15,13	15,40	15,41	15,80	16,60
	>2	2,28	2,50	3,14	3,33	3,41	3,65	4,64
	>3	0,13	0,22	0,42	0,50	0,56	0,69	1,12
r_i^{**pea}	<-3	0,13	0,00	0,00	0,01	0,04	0,06	0,42
	<-2	2,28	0,10	1,04	1,75	1,66	2,20	3,24
	<-1	15,87	13,24	14,69	15,10	15,05	15,40	16,54
	>1	15,87	12,96	13,93	14,33	14,31	14,61	16,00
	>2	2,28	3,42	3,92	4,16	4,18	4,40	4,96
	>3	0,13	0,54	0,80	0,92	0,95	1,06	1,56
r_i^{*qu}	<-3	0,13	0,00	0,08	0,12	0,11	0,14	0,30
	<-2	2,28	1,76	2,11	2,24	2,29	2,46	2,98
	<-1	15,87	14,46	15,62	16,05	16,01	16,41	17,40
	>1	15,87	14,44	15,72	16,05	16,05	16,45	17,34
	>2	2,28	1,52	2,02	2,25	2,22	2,40	2,78
	>3	0,13	0,00	0,06	0,09	0,10	0,14	0,24
r_i^{**qu}	<-3	0,13	0,14	0,26	0,32	0,31	0,36	0,56
	<-2	2,28	2,50	2,84	3,04	3,06	3,28	3,86
	<-1	15,87	12,46	14,20	14,82	14,73	15,29	16,38
	>1	15,87	12,42	14,26	14,82	14,76	15,20	16,82
	>2	2,28	2,40	2,80	2,98	3,00	3,22	3,60
	>3	0,13	0,10	0,24	0,26	0,29	0,35	0,58
r_i^{*ans}	<-3	0,13	0,00	0,10	0,14	0,13	0,16	0,30
	<-2	2,28	1,96	2,52	2,66	2,70	2,94	3,42
	<-1	15,87	15,68	17,47	17,97	18,06	18,83	20,22
	>1	15,87	11,94	13,29	14,02	13,92	14,51	15,16
	>2	2,28	0,96	1,53	1,70	1,69	1,88	2,28
	>3	0,13	0,00	0,04	0,04	0,06	0,08	0,16
r_i^{**ans}	<-3	0,13	0,16	0,30	0,36	0,36	0,42	0,56
	<-2	2,28	3,06	3,42	3,62	3,63	3,82	4,38
	<-1	15,87	13,28	15,70	16,66	16,68	17,62	19,32
	>1	15,87	10,92	12,31	12,65	12,68	13,11	13,80
	>2	2,28	1,38	2,03	2,41	2,34	2,64	3,10
	>3	0,13	0,02	0,12	0,18	0,19	0,25	0,54
r_i^{*wil}	<-3	0,13	0,04	0,14	0,18	0,17	0,22	0,34
	<-2	2,28	2,34	2,70	2,88	2,91	3,10	3,64
	<-1	15,87	16,48	17,68	18,20	18,34	18,98	20,54
	>1	15,87	12,44	13,53	14,39	14,23	14,87	15,60
	>2	2,28	1,38	1,75	1,92	1,91	2,10	2,50
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,20
r_i^{**wil}	<-3	0,13	0,22	0,36	0,43	0,43	0,50	0,68
	<-2	2,28	3,30	3,64	3,86	3,87	4,04	4,52
	<-1	15,87	13,98	15,92	16,91	16,97	17,98	19,60
	>1	15,87	11,44	12,62	13,05	13,00	13,38	14,18
	>2	2,28	1,78	2,30	2,69	2,61	2,90	3,54
	>3	0,13	0,06	0,18	0,24	0,25	0,32	0,72

Tabela E.3: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 2b.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,12	0,14	0,15	0,20	0,32
	<-2	2,28	1,86	2,66	2,91	2,89	3,14	4,06
	<-1	15,87	9,38	13,80	15,27	15,15	16,88	19,00
	>1	15,87	9,10	12,39	13,20	12,95	13,81	14,62
	>2	2,28	1,22	1,74	1,90	1,93	2,09	2,50
	>3	0,13	0,00	0,02	0,04	0,06	0,08	0,20
r_i^{**dev}	<-3	0,13	1,24	1,64	1,78	1,79	1,94	2,44
	<-2	2,28	2,70	4,46	4,87	4,82	5,30	6,56
	<-1	15,87	5,82	8,97	9,94	10,22	11,98	15,00
	>1	15,87	5,72	7,74	8,24	8,21	8,86	10,08
	>2	2,28	2,44	3,26	3,55	3,53	3,78	4,36
	>3	0,13	0,40	0,94	1,20	1,13	1,35	1,78
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,02	0,04
	<-2	2,28	0,40	1,54	1,79	1,71	2,02	2,40
	<-1	15,87	9,36	13,80	15,26	15,04	16,66	18,48
	>1	15,87	9,10	12,39	13,33	13,03	14,01	14,82
	>2	2,28	1,52	2,45	2,70	2,70	2,98	3,58
	>3	0,13	0,04	0,22	0,32	0,34	0,44	0,76
r_i^{**pea}	<-3	0,13	0,04	0,58	0,95	0,88	1,18	1,70
	<-2	2,28	1,80	3,52	3,78	3,72	4,03	4,82
	<-1	15,87	5,80	8,74	9,67	9,76	11,04	13,62
	>1	15,87	5,80	7,99	8,55	8,56	9,29	11,10
	>2	2,28	2,72	3,88	4,35	4,26	4,60	5,34
	>3	0,13	1,04	1,64	1,80	1,82	2,03	2,48
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,10	0,12	0,24
	<-2	2,28	1,36	2,16	2,36	2,35	2,56	3,14
	<-1	15,87	8,96	12,88	14,29	13,96	15,22	16,78
	>1	15,87	9,60	13,20	14,32	14,10	15,41	16,42
	>2	2,28	1,24	2,08	2,27	2,26	2,48	3,00
	>3	0,13	0,00	0,04	0,08	0,09	0,12	0,24
r_i^{**qu}	<-3	0,13	0,86	1,24	1,45	1,42	1,59	2,06
	<-2	2,28	2,30	3,86	4,13	4,10	4,44	5,18
	<-1	15,87	5,36	8,20	9,03	9,10	10,29	12,82
	>1	15,87	6,06	8,44	9,13	9,19	10,17	12,40
	>2	2,28	2,50	3,74	4,09	4,07	4,46	5,06
	>3	0,13	0,72	1,14	1,37	1,35	1,56	2,04
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,12	0,15	0,28
	<-2	2,28	1,56	2,52	2,79	2,74	3,02	3,74
	<-1	15,87	9,34	13,80	15,26	15,13	16,86	18,92
	>1	15,87	9,10	12,35	13,18	12,92	13,77	14,60
	>2	2,28	1,10	1,62	1,80	1,81	2,06	2,46
	>3	0,13	0,00	0,02	0,04	0,05	0,06	0,18
r_i^{**ans}	<-3	0,13	1,02	1,50	1,66	1,67	1,85	2,28
	<-2	2,28	2,44	4,30	4,80	4,68	5,14	6,20
	<-1	15,87	5,66	8,87	9,90	10,12	11,82	14,92
	>1	15,87	5,52	7,56	8,20	8,13	8,77	9,98
	>2	2,28	2,18	3,14	3,46	3,40	3,73	4,32
	>3	0,13	0,36	0,86	1,10	1,05	1,26	1,68
r_i^{*wil}	<-3	0,13	0,04	0,10	0,14	0,15	0,18	0,32
	<-2	2,28	1,84	2,62	2,90	2,87	3,12	4,04
	<-1	15,87	9,38	13,80	15,27	15,15	16,88	19,00
	>1	15,87	9,10	12,39	13,20	12,95	13,81	14,64
	>2	2,28	1,28	1,78	1,94	1,95	2,14	2,50
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,20
r_i^{**wil}	<-3	0,13	1,18	1,62	1,76	1,77	1,93	2,42
	<-2	2,28	2,70	4,44	4,86	4,80	5,28	6,44
	<-1	15,87	5,82	8,97	9,94	10,20	11,96	14,98
	>1	15,87	5,72	7,75	8,25	8,22	8,86	10,10
	>2	2,28	2,50	3,30	3,58	3,55	3,79	4,36
	>3	0,13	0,44	0,96	1,22	1,15	1,36	1,78

Tabela E.4: Estatística Descritiva da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 3.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,02	0,10	0,12	0,13	0,16	0,26
	<-2	2,28	1,78	2,30	2,45	2,47	2,64	3,06
	<-1	15,87	14,68	15,99	16,40	16,38	16,79	17,76
	>1	15,87	14,66	15,46	15,90	15,89	16,23	17,34
	>2	2,28	1,74	2,16	2,26	2,30	2,47	2,96
	>3	0,13	0,00	0,08	0,12	0,12	0,14	0,28
r_i^{**dev}	<-3	0,13	0,36	0,64	0,78	0,79	0,90	1,28
	<-2	2,28	3,06	3,79	4,03	4,00	4,22	4,88
	<-1	15,87	10,24	12,47	13,18	13,11	13,78	15,36
	>1	15,87	10,36	12,08	12,80	12,71	13,38	14,90
	>2	2,28	2,88	3,56	3,80	3,79	4,00	4,72
	>3	0,13	0,24	0,58	0,72	0,72	0,84	1,22
r_i^{*pea}	<-3	0,13	0,00	0,06	0,08	0,08	0,12	0,18
	<-2	2,28	1,74	2,10	2,22	2,27	2,45	2,92
	<-1	15,87	14,66	15,85	16,32	16,29	16,70	17,62
	>1	15,87	14,66	15,52	16,00	15,97	16,36	17,36
	>2	2,28	1,80	2,36	2,46	2,50	2,68	3,12
	>3	0,13	0,04	0,12	0,16	0,16	0,20	0,34
r_i^{**pea}	<-3	0,13	0,26	0,50	0,64	0,66	0,81	1,16
	<-2	2,28	2,82	3,56	3,80	3,77	4,05	4,80
	<-1	15,87	10,22	12,37	13,05	12,97	13,62	15,08
	>1	15,87	10,38	12,18	12,95	12,84	13,56	15,28
	>2	2,28	3,18	3,82	4,00	4,01	4,21	4,88
	>3	0,13	0,48	0,72	0,82	0,84	0,96	1,36
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,10	0,14	0,20
	<-2	2,28	1,40	2,08	2,26	2,28	2,46	3,00
	<-1	15,87	14,16	15,54	16,04	15,99	16,39	17,42
	>1	15,87	14,60	15,52	16,09	15,99	16,40	17,40
	>2	2,28	1,52	2,10	2,22	2,24	2,44	2,94
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,26
r_i^{**qu}	<-3	0,13	0,36	0,58	0,68	0,69	0,77	1,10
	<-2	2,28	2,92	3,48	3,76	3,73	3,96	4,66
	<-1	15,87	9,64	12,13	12,82	12,69	13,37	14,88
	>1	15,87	9,86	12,11	12,88	12,75	13,53	15,24
	>2	2,28	2,88	3,49	3,71	3,72	3,97	4,64
	>3	0,13	0,24	0,56	0,69	0,68	0,78	1,16
r_i^{*ans}	<-3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
	<-2	2,28	1,40	2,12	2,34	2,34	2,52	3,02
	<-1	15,87	14,34	15,85	16,28	16,24	16,66	17,64
	>1	15,87	14,48	15,32	15,77	15,76	16,16	17,30
	>2	2,28	1,50	2,04	2,16	2,18	2,35	2,90
	>3	0,13	0,00	0,06	0,10	0,10	0,13	0,26
r_i^{**ans}	<-3	0,13	0,36	0,60	0,70	0,72	0,79	1,14
	<-2	2,28	2,98	3,60	3,84	3,83	4,06	4,74
	<-1	15,87	9,76	12,36	13,03	12,92	13,59	15,14
	>1	15,87	9,74	11,96	12,69	12,53	13,25	14,68
	>2	2,28	2,70	3,34	3,62	3,62	3,88	4,60
	>3	0,13	0,24	0,54	0,66	0,66	0,76	1,14
r_i^{*wil}	<-3	0,13	0,02	0,08	0,12	0,12	0,16	0,24
	<-2	2,28	1,78	2,30	2,45	2,46	2,64	3,06
	<-1	15,87	14,68	15,98	16,39	16,38	16,79	17,76
	>1	15,87	14,66	15,46	15,91	15,89	16,23	17,34
	>2	2,28	1,74	2,16	2,26	2,31	2,47	2,96
	>3	0,13	0,00	0,08	0,12	0,12	0,14	0,28
r_i^{**wil}	<-3	0,13	0,36	0,64	0,78	0,78	0,90	1,28
	<-2	2,28	3,04	3,79	4,01	4,00	4,22	4,88
	<-1	15,87	10,24	12,47	13,17	13,11	13,77	15,34
	>1	15,87	10,36	12,08	12,80	12,72	13,38	14,92
	>2	2,28	2,88	3,56	3,80	3,80	4,01	4,72
	>3	0,13	0,24	0,58	0,72	0,72	0,84	1,22

Tabela E.5: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 4a.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,10	0,14	0,14	0,17	0,32
	<-2	2,28	1,78	2,46	2,64	2,62	2,84	3,42
	<-1	15,87	15,16	16,60	17,05	17,08	17,62	19,16
	>1	15,87	14,24	14,84	15,21	15,24	15,64	16,84
	>2	2,28	1,72	1,96	2,10	2,12	2,28	2,62
	>3	0,13	0,00	0,08	0,10	0,10	0,14	0,20
r_i^{**dev}	<-3	0,13	0,38	0,72	0,82	0,82	0,94	1,30
	<-2	2,28	3,34	4,08	4,26	4,26	4,44	5,16
	<-1	15,87	10,20	12,76	13,73	13,75	14,63	17,08
	>1	15,87	9,70	11,63	12,16	12,09	12,65	13,96
	>2	2,28	2,58	3,30	3,60	3,56	3,79	4,46
	>3	0,13	0,28	0,52	0,66	0,65	0,79	0,98
r_i^{*pea}	<-3	0,13	0,00	0,02	0,02	0,03	0,04	0,14
	<-2	2,28	1,00	1,66	1,84	1,82	2,00	2,92
	<-1	15,87	15,10	16,37	16,68	16,70	17,05	17,88
	>1	15,87	14,26	15,28	15,53	15,55	15,87	17,20
	>2	2,28	2,14	2,58	2,85	2,81	3,05	3,54
	>3	0,13	0,08	0,22	0,28	0,29	0,36	0,58
r_i^{**pea}	<-3	0,13	0,02	0,24	0,42	0,43	0,57	0,94
	<-2	2,28	2,02	3,02	3,43	3,35	3,67	4,58
	<-1	15,87	10,08	12,51	13,27	13,22	14,12	15,70
	>1	15,87	9,84	11,91	12,62	12,53	13,23	14,64
	>2	2,28	3,62	4,08	4,25	4,28	4,51	5,06
	>3	0,13	0,72	0,98	1,07	1,08	1,19	1,52
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,10	0,14	0,22
	<-2	2,28	1,40	2,12	2,28	2,25	2,42	3,12
	<-1	15,87	14,40	15,73	15,96	16,03	16,40	17,18
	>1	15,87	14,38	15,66	15,96	15,99	16,35	17,60
	>2	2,28	1,58	2,08	2,28	2,25	2,41	2,82
	>3	0,13	0,02	0,08	0,10	0,11	0,14	0,26
r_i^{**qu}	<-3	0,13	0,30	0,54	0,69	0,67	0,77	1,08
	<-2	2,28	2,72	3,54	3,72	3,73	3,95	4,88
	<-1	15,87	9,18	12,17	12,76	12,75	13,70	15,18
	>1	15,87	9,58	12,08	12,85	12,74	13,46	15,02
	>2	2,28	3,04	3,52	3,72	3,75	3,98	4,50
	>3	0,13	0,36	0,56	0,68	0,68	0,78	1,02
r_i^{*ans}	<-3	0,13	0,00	0,08	0,10	0,12	0,14	0,24
	<-2	2,28	1,44	2,32	2,50	2,47	2,68	3,36
	<-1	15,87	14,82	16,50	16,87	16,92	17,43	18,96
	>1	15,87	14,02	14,65	15,10	15,10	15,48	16,66
	>2	2,28	1,52	1,82	1,98	1,99	2,14	2,54
	>3	0,13	0,00	0,06	0,08	0,09	0,10	0,18
r_i^{**ans}	<-3	0,13	0,32	0,62	0,76	0,74	0,84	1,16
	<-2	2,28	2,84	3,88	4,10	4,07	4,27	5,08
	<-1	15,87	9,60	12,62	13,60	13,56	14,56	16,90
	>1	15,87	9,22	11,44	12,01	11,91	12,48	13,86
	>2	2,28	2,46	3,14	3,38	3,40	3,67	4,32
	>3	0,13	0,24	0,46	0,60	0,59	0,70	0,94
r_i^{*wil}	<-3	0,13	0,04	0,10	0,13	0,13	0,16	0,32
	<-2	2,28	1,74	2,44	2,62	2,61	2,82	3,42
	<-1	15,87	15,16	16,59	17,05	17,07	17,60	19,12
	>1	15,87	14,26	14,86	15,21	15,25	15,64	16,84
	>2	2,28	1,78	1,98	2,12	2,15	2,30	2,64
	>3	0,13	0,02	0,08	0,10	0,11	0,14	0,20
r_i^{**wil}	<-3	0,13	0,38	0,70	0,82	0,81	0,92	1,30
	<-2	2,28	3,32	4,06	4,23	4,24	4,41	5,16
	<-1	15,87	10,20	12,76	13,72	13,74	14,60	17,04
	>1	15,87	9,72	11,63	12,16	12,10	12,67	13,96
	>2	2,28	2,64	3,36	3,60	3,58	3,83	4,46
	>3	0,13	0,30	0,54	0,68	0,66	0,80	0,98

Tabela E.6: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 4b.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,14	0,18	0,19	0,24	0,36
	<-2	2,28	2,10	2,90	3,11	3,17	3,42	4,26
	<-1	15,87	15,72	17,89	18,54	18,71	19,45	22,18
	>1	15,87	10,82	13,00	13,66	13,56	14,22	15,70
	>2	2,28	1,04	1,42	1,62	1,62	1,83	2,22
	>3	0,13	0,00	0,02	0,04	0,05	0,08	0,18
r_i^{**dev}	<-3	0,13	0,60	0,90	1,06	1,04	1,16	1,52
	<-2	2,28	3,86	4,74	4,99	4,99	5,24	5,88
	<-1	15,87	11,22	13,96	15,04	15,22	16,70	20,12
	>1	15,87	8,64	10,24	10,60	10,52	10,93	12,04
	>2	2,28	1,54	2,48	2,83	2,81	3,17	3,74
	>3	0,13	0,10	0,30	0,46	0,47	0,62	0,88
r_i^{*pea}	<-3	0,13	0,00	0,00	0,04	0,11	0,15	0,66
	<-2	2,28	0,00	1,24	2,27	2,01	2,79	4,12
	<-1	15,87	10,84	12,76	13,49	13,30	13,94	15,16
	>1	15,87	9,02	11,31	11,69	11,63	12,08	12,90
	>2	2,28	3,20	4,36	4,62	4,59	4,82	5,36
	>3	0,13	0,76	1,46	1,60	1,57	1,72	2,04
r_i^{**pea}	<-3	0,13	0,00	0,00	0,04	0,11	0,15	0,66
	<-2	2,28	0,00	1,24	2,27	2,01	2,79	4,12
	<-1	15,87	10,84	12,76	13,49	13,30	13,94	15,16
	>1	15,87	9,02	11,31	11,69	11,63	12,08	12,90
	>2	2,28	3,20	4,36	4,62	4,59	4,82	5,36
	>3	0,13	0,76	1,46	1,60	1,57	1,72	2,04
r_i^{*qu}	<-3	0,13	0,02	0,08	0,10	0,11	0,14	0,26
	<-2	2,28	1,42	2,16	2,32	2,32	2,50	2,86
	<-1	15,87	14,26	15,53	15,87	15,91	16,30	17,22
	>1	15,87	14,36	15,73	16,10	16,05	16,45	17,22
	>2	2,28	1,22	2,04	2,19	2,20	2,36	2,78
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,20
r_i^{**qu}	<-3	0,13	0,26	0,58	0,70	0,71	0,83	1,22
	<-2	2,28	2,82	3,58	3,81	3,81	4,04	4,70
	<-1	15,87	9,50	11,99	12,80	12,70	13,49	15,22
	>1	15,87	9,18	12,18	12,85	12,76	13,52	14,92
	>2	2,28	2,72	3,42	3,69	3,69	3,95	4,66
	>3	0,13	0,22	0,52	0,64	0,65	0,80	0,98
r_i^{*ans}	<-3	0,13	0,02	0,10	0,12	0,12	0,16	0,28
	<-2	2,28	1,68	2,64	2,86	2,87	3,13	3,62
	<-1	15,87	15,38	17,75	18,37	18,53	19,31	21,68
	>1	15,87	10,58	12,81	13,56	13,40	14,09	15,60
	>2	2,28	0,68	1,23	1,46	1,46	1,68	2,04
	>3	0,13	0,00	0,02	0,02	0,04	0,06	0,16
r_i^{**ans}	<-3	0,13	0,24	0,70	0,90	0,85	0,99	1,34
	<-2	2,28	3,24	4,42	4,65	4,65	4,92	5,46
	<-1	15,87	10,62	13,76	14,86	14,99	16,45	19,58
	>1	15,87	8,12	10,04	10,41	10,32	10,76	11,94
	>2	2,28	1,34	2,16	2,67	2,60	3,05	3,48
	>3	0,13	0,02	0,20	0,39	0,38	0,52	0,84
r_i^{*wil}	<-3	0,13	0,04	0,12	0,18	0,18	0,22	0,36
	<-2	2,28	2,06	2,82	3,08	3,11	3,36	4,18
	<-1	15,87	15,70	17,88	18,51	18,68	19,42	22,12
	>1	15,87	10,84	13,04	13,72	13,58	14,24	15,72
	>2	2,28	1,10	1,46	1,68	1,68	1,88	2,26
	>3	0,13	0,00	0,04	0,06	0,06	0,09	0,18
r_i^{**wil}	<-3	0,13	0,58	0,86	1,02	1,00	1,12	1,50
	<-2	2,28	3,74	4,68	4,92	4,92	5,18	5,76
	<-1	15,87	11,16	13,94	15,01	15,19	16,66	19,96
	>1	15,87	8,66	10,31	10,63	10,56	10,95	12,06
	>2	2,28	1,60	2,57	2,89	2,87	3,23	3,78
	>3	0,13	0,10	0,32	0,50	0,50	0,66	0,90

Tabela E.7: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 5.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,00	0,06	0,08	0,10	0,16	0,38
	<-2	2,28	1,52	2,54	2,89	2,83	3,08	3,70
	<-1	15,87	14,00	17,21	17,87	18,02	18,54	22,72
	>1	15,87	11,38	14,34	14,85	14,68	15,25	16,30
	>2	2,28	0,78	1,74	1,98	1,92	2,19	2,66
	>3	0,13	0,00	0,02	0,04	0,05	0,08	0,18
r_i^{**dev}	<-3	0,13	0,14	0,66	0,88	0,83	1,03	1,42
	<-2	2,28	3,10	4,42	4,70	4,66	4,96	5,42
	<-1	15,87	9,22	13,19	14,65	14,64	15,85	20,50
	>1	15,87	8,92	11,28	11,63	11,58	12,01	12,84
	>2	2,28	1,54	2,86	3,43	3,31	3,80	4,62
	>3	0,13	0,02	0,34	0,54	0,54	0,72	1,04
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,06
	<-2	2,28	0,00	0,78	1,52	1,33	1,88	2,44
	<-1	15,87	13,92	16,84	17,21	17,17	17,64	18,44
	>1	15,87	13,70	14,97	15,31	15,24	15,64	16,62
	>2	2,28	1,42	2,80	3,09	3,09	3,36	4,50
	>3	0,13	0,02	0,24	0,34	0,38	0,48	1,20
r_i^{**pea}	<-3	0,13	0,00	0,02	0,23	0,26	0,44	0,76
	<-2	2,28	0,10	2,09	3,31	2,84	3,65	4,28
	<-1	15,87	9,18	12,76	13,67	13,48	14,30	15,32
	>1	15,87	9,02	11,90	12,40	12,34	12,82	14,08
	>2	2,28	2,78	4,38	4,58	4,57	4,81	5,64
	>3	0,13	0,52	1,16	1,30	1,29	1,44	1,74
r_i^{*qu}	<-3	0,13	0,00	0,02	0,05	0,06	0,11	0,26
	<-2	2,28	1,00	1,98	2,23	2,23	2,55	3,00
	<-1	15,87	13,10	15,96	16,28	16,24	16,73	17,44
	>1	15,87	14,34	15,88	16,19	16,19	16,57	18,14
	>2	2,28	0,90	2,10	2,30	2,24	2,44	2,86
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,26
r_i^{**qu}	<-3	0,13	0,04	0,42	0,66	0,62	0,82	1,32
	<-2	2,28	2,22	3,38	3,85	3,79	4,21	4,72
	<-1	15,87	8,50	12,21	13,12	12,97	13,76	15,02
	>1	15,87	8,76	12,26	12,97	12,91	13,66	16,10
	>2	2,28	2,26	3,58	3,86	3,83	4,15	4,94
	>3	0,13	0,10	0,46	0,62	0,63	0,83	1,08
r_i^{*ans}	<-3	0,13	0,00	0,02	0,06	0,08	0,12	0,36
	<-2	2,28	1,14	2,36	2,64	2,60	2,90	3,50
	<-1	15,87	13,56	17,09	17,69	17,84	18,38	22,44
	>1	15,87	10,92	14,14	14,73	14,53	15,19	16,22
	>2	2,28	0,58	1,56	1,86	1,77	2,07	2,60
	>3	0,13	0,00	0,02	0,04	0,05	0,08	0,18
r_i^{**ans}	<-3	0,13	0,04	0,50	0,76	0,72	0,96	1,38
	<-2	2,28	2,58	4,08	4,43	4,39	4,76	5,34
	<-1	15,87	8,90	13,05	14,43	14,44	15,64	20,14
	>1	15,87	8,44	11,02	11,50	11,39	11,87	12,82
	>2	2,28	1,12	2,62	3,23	3,12	3,67	4,52
	>3	0,13	0,00	0,26	0,49	0,48	0,69	1,00
r_i^{*wil}	<-3	0,13	0,00	0,04	0,08	0,09	0,15	0,38
	<-2	2,28	1,52	2,48	2,84	2,78	3,04	3,66
	<-1	15,87	14,00	17,21	17,84	18,00	18,53	22,66
	>1	15,87	11,44	14,38	14,88	14,71	15,25	16,30
	>2	2,28	0,96	1,78	1,99	1,96	2,21	2,66
	>3	0,13	0,00	0,02	0,05	0,06	0,08	0,20
r_i^{**wil}	<-3	0,13	0,12	0,62	0,85	0,81	1,01	1,42
	<-2	2,28	3,06	4,34	4,64	4,61	4,92	5,42
	<-1	15,87	9,22	13,19	14,65	14,62	15,80	20,40
	>1	15,87	8,92	11,32	11,64	11,61	12,01	12,84
	>2	2,28	1,76	2,92	3,48	3,36	3,82	4,66
	>3	0,13	0,02	0,38	0,55	0,56	0,72	1,04

Tabela E.8: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 6a.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,06	0,12	0,16	0,17	0,20	0,30
	<-2	2,28	1,98	2,66	2,93	2,89	3,11	3,76
	<-1	15,87	14,82	17,16	17,76	17,76	18,41	21,28
	>1	15,87	9,84	14,14	14,50	14,47	14,98	16,10
	>2	2,28	0,50	1,74	1,92	1,90	2,08	2,48
	>3	0,13	0,00	0,06	0,08	0,08	0,10	0,18
r_i^{**dev}	<-3	0,13	0,56	0,84	0,98	0,97	1,08	1,52
	<-2	2,28	3,62	4,40	4,67	4,63	4,88	5,64
	<-1	15,87	10,20	13,12	14,22	14,26	15,42	19,60
	>1	15,87	8,90	10,80	11,34	11,21	11,70	12,78
	>2	2,28	1,02	3,03	3,32	3,24	3,57	4,08
	>3	0,13	0,02	0,44	0,58	0,58	0,74	1,02
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,02	0,06
	<-2	2,28	0,10	1,12	1,34	1,32	1,64	2,22
	<-1	15,87	14,66	16,61	17,15	17,02	17,50	18,38
	>1	15,87	12,70	14,62	15,06	15,00	15,37	16,40
	>2	2,28	2,00	2,80	3,04	3,04	3,29	3,88
	>3	0,13	0,12	0,32	0,43	0,43	0,52	0,86
r_i^{**pea}	<-3	0,13	0,00	0,06	0,18	0,26	0,45	0,90
	<-2	2,28	0,30	2,41	3,03	2,86	3,41	4,18
	<-1	15,87	9,94	12,50	13,32	13,20	14,04	15,08
	>1	15,87	9,52	11,29	12,04	11,96	12,61	14,12
	>2	2,28	3,52	4,26	4,47	4,45	4,66	5,22
	>3	0,13	0,88	1,24	1,34	1,34	1,44	1,76
r_i^{*qu}	<-3	0,13	0,00	0,08	0,12	0,12	0,16	0,28
	<-2	2,28	1,40	2,14	2,34	2,30	2,53	2,86
	<-1	15,87	13,54	15,66	16,04	15,95	16,41	17,36
	>1	15,87	14,54	15,56	16,07	16,01	16,41	17,54
	>2	2,28	1,14	2,10	2,23	2,25	2,44	2,94
	>3	0,13	0,00	0,06	0,10	0,10	0,14	0,22
r_i^{**qu}	<-3	0,13	0,28	0,62	0,75	0,74	0,86	1,20
	<-2	2,28	2,06	3,56	3,86	3,81	4,13	4,64
	<-1	15,87	8,84	11,94	12,79	12,57	13,46	14,54
	>1	15,87	9,18	11,84	12,67	12,56	13,34	15,04
	>2	2,28	1,90	3,56	3,80	3,74	4,03	4,60
	>3	0,13	0,02	0,56	0,69	0,69	0,84	1,14
r_i^{*ans}	<-3	0,13	0,02	0,08	0,13	0,13	0,18	0,30
	<-2	2,28	1,58	2,50	2,76	2,69	2,91	3,42
	<-1	15,87	14,36	17,05	17,68	17,59	18,30	19,90
	>1	15,87	8,72	13,96	14,39	14,32	14,88	16,04
	>2	2,28	0,26	1,56	1,79	1,76	1,98	2,34
	>3	0,13	0,00	0,04	0,06	0,06	0,08	0,18
r_i^{**ans}	<-3	0,13	0,26	0,74	0,86	0,87	0,99	1,46
	<-2	2,28	2,62	4,14	4,48	4,40	4,70	5,36
	<-1	15,87	9,46	12,99	14,11	14,05	15,15	18,00
	>1	15,87	7,74	10,63	11,21	11,02	11,54	12,58
	>2	2,28	0,46	2,82	3,18	3,07	3,44	4,02
	>3	0,13	0,02	0,36	0,52	0,52	0,69	0,96
r_i^{*wil}	<-3	0,13	0,06	0,12	0,16	0,16	0,20	0,30
	<-2	2,28	1,94	2,64	2,91	2,85	3,08	3,62
	<-1	15,87	14,82	17,14	17,76	17,75	18,41	20,92
	>1	15,87	10,40	14,20	14,53	14,49	15,00	16,10
	>2	2,28	0,94	1,78	1,96	1,94	2,08	2,54
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,18
r_i^{**wil}	<-3	0,13	0,46	0,82	0,96	0,95	1,07	1,52
	<-2	2,28	3,40	4,36	4,63	4,59	4,83	5,62
	<-1	15,87	10,18	13,10	14,21	14,23	15,33	19,18
	>1	15,87	9,24	10,80	11,35	11,24	11,74	12,80
	>2	2,28	1,38	3,08	3,35	3,28	3,59	4,10
	>3	0,13	0,04	0,44	0,58	0,60	0,76	1,06

Tabela E.9: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 6b.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,04	0,14	0,16	0,17	0,20	0,32
	<-2	2,28	1,96	2,70	2,90	2,91	3,11	3,66
	<-1	15,87	14,82	17,16	17,83	17,81	18,30	21,10
	>1	15,87	10,94	14,02	14,56	14,43	15,02	16,08
	>2	2,28	0,96	1,80	1,97	1,93	2,11	2,50
	>3	0,13	0,00	0,04	0,08	0,08	0,10	0,18
r_i^{**dev}	<-3	0,13	0,52	0,86	0,96	0,97	1,09	1,54
	<-2	2,28	3,76	4,44	4,70	4,67	4,85	5,42
	<-1	15,87	9,96	13,12	14,26	14,30	15,30	19,54
	>1	15,87	9,10	10,78	11,19	11,20	11,66	12,82
	>2	2,28	1,48	2,94	3,30	3,26	3,70	4,22
	>3	0,13	0,04	0,44	0,62	0,59	0,76	10,20
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,00	0,06
	<-2	2,28	0,00	1,02	1,37	1,30	1,62	2,08
	<-1	15,87	14,76	16,58	17,05	17,04	17,43	18,52
	>1	15,87	13,32	14,64	14,97	14,96	15,33	16,30
	>2	2,28	1,54	2,88	3,10	3,10	3,32	3,90
	>3	0,13	0,08	0,36	0,44	0,46	0,53	0,98
r_i^{**pea}	<-3	0,13	0,00	0,06	0,18	0,25	0,40	0,86
	<-2	2,28	0,10	2,46	2,96	2,83	3,42	4,16
	<-1	15,87	9,78	12,51	13,32	13,17	13,77	14,98
	>1	15,87	9,36	11,41	12,05	11,98	12,62	13,74
	>2	2,28	3,22	4,30	4,51	4,50	4,75	5,38
	>3	0,13	0,88	1,30	1,40	1,39	1,51	1,72
r_i^{*qu}	<-3	0,13	0,00	0,08	0,10	0,11	0,14	0,22
	<-2	2,28	1,18	2,14	2,32	2,31	2,50	2,74
	<-1	15,87	13,62	15,54	15,97	15,92	16,42	17,26
	>1	15,87	14,52	15,72	16,02	16,02	16,32	17,56
	>2	2,28	1,00	2,12	2,30	2,27	2,43	2,98
	>3	0,13	0,00	0,06	0,10	0,10	0,14	0,22
r_i^{**qu}	<-3	0,13	0,30	0,60	0,74	0,72	0,84	1,26
	<-2	2,28	2,54	3,62	3,87	3,82	4,08	4,52
	<-1	15,87	8,64	11,92	12,65	12,55	13,22	14,80
	>1	15,87	8,74	11,91	12,66	12,61	13,43	14,56
	>2	2,28	2,26	3,56	3,88	3,79	4,14	4,48
	>3	0,13	0,10	0,52	0,74	0,70	0,86	11,40
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,13	0,16	0,24
	<-2	2,28	1,30	2,54	2,74	2,71	2,92	3,28
	<-1	15,87	14,44	17,02	17,65	17,64	18,14	20,66
	>1	15,87	10,48	13,90	14,40	14,28	14,92	16,00
	>2	2,28	0,62	1,62	1,85	1,79	2,01	2,40
	>3	0,13	0,00	0,04	0,06	0,06	0,08	0,18
r_i^{**ans}	<-3	0,13	0,34	0,74	0,86	0,85	10,00	14,60
	<-2	2,28	2,84	4,16	4,47	4,42	4,65	5,24
	<-1	15,87	9,18	12,95	14,12	14,09	15,11	19,10
	>1	15,87	8,14	10,66	11,10	11,01	11,46	12,64
	>2	2,28	1,20	2,76	3,17	3,09	3,56	4,04
	>3	0,13	0,02	0,36	0,54	0,53	0,69	0,98
r_i^{*wil}	<-3	0,13	0,02	0,12	0,16	0,15	0,18	0,30
	<-2	2,28	1,92	2,68	2,87	2,87	3,04	3,44
	<-1	15,87	14,80	17,16	17,83	17,79	18,22	21,04
	>1	15,87	11,08	14,04	14,58	14,45	15,04	16,08
	>2	2,28	1,14	1,84	1,98	1,97	2,14	2,52
	>3	0,13	0,00	0,06	0,08	0,08	0,12	0,20
r_i^{**wil}	<-3	0,13	0,46	0,82	0,93	0,94	1,06	1,54
	<-2	2,28	3,72	4,40	4,64	4,62	4,80	5,38
	<-1	15,87	9,90	13,12	14,25	14,28	15,27	19,42
	>1	15,87	9,14	10,79	11,22	11,23	11,69	12,82
	>2	2,28	1,58	2,96	3,34	3,31	3,70	4,22
	>3	0,13	0,06	0,46	0,63	0,61	0,78	1,04

Tabela E.10: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 7a.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,04	0,12	0,16	0,17	0,20	0,36
	<-2	2,28	1,06	2,72	2,90	2,88	3,07	3,52
	<-1	15,87	14,12	17,53	17,93	17,92	18,39	19,80
	>1	15,87	13,34	14,23	14,62	14,56	14,88	15,70
	>2	2,28	0,80	1,78	1,91	1,89	2,06	2,32
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,26
r_i^{**dev}	<-3	0,13	0,32	0,56	0,66	0,65	0,74	0,96
	<-2	2,28	2,18	4,04	4,28	4,24	4,50	4,78
	<-1	15,87	10,54	15,17	15,66	15,61	16,23	18,08
	>1	15,87	9,96	12,11	12,46	12,44	12,82	13,38
	>2	2,28	1,58	2,76	2,94	2,93	3,10	3,64
	>3	0,13	0,14	0,30	0,38	0,38	0,44	0,72
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,04
	<-2	2,28	0,68	1,06	1,24	1,26	1,45	1,82
	<-1	15,87	13,92	16,59	16,95	16,91	17,34	18,54
	>1	15,87	14,30	14,92	15,39	15,33	15,64	16,42
	>2	2,28	1,20	2,96	3,18	3,15	3,43	3,74
	>3	0,13	0,02	0,34	0,46	0,46	0,56	0,86
r_i^{**pea}	<-3	0,13	0,00	0,04	0,07	0,11	0,15	0,38
	<-2	2,28	1,28	2,04	2,38	2,38	2,71	3,18
	<-1	15,87	10,10	14,10	14,39	14,36	14,82	16,02
	>1	15,87	10,44	13,03	13,40	13,38	13,86	14,50
	>2	2,28	2,20	4,17	4,32	4,31	4,54	4,90
	>3	0,13	0,22	0,96	1,08	1,06	1,17	1,44
r_i^{*qu}	<-3	0,13	0,02	0,08	0,12	0,11	0,15	0,26
	<-2	2,28	0,62	2,18	2,32	2,29	2,44	2,84
	<-1	15,87	11,70	15,66	16,03	16,00	16,46	17,70
	>1	15,87	13,86	15,79	16,12	16,09	16,50	17,16
	>2	2,28	0,48	2,09	2,26	2,20	2,42	2,72
	>3	0,13	0,00	0,06	0,11	0,11	0,14	0,28
r_i^{**qu}	<-3	0,13	0,12	0,40	0,48	0,48	0,56	0,72
	<-2	2,28	1,14	3,32	3,50	3,46	3,69	4,10
	<-1	15,87	8,26	13,53	13,87	13,79	14,32	15,36
	>1	15,87	9,48	13,44	13,94	13,86	14,45	15,30
	>2	2,28	1,22	3,22	3,44	3,38	3,58	4,00
	>3	0,13	0,02	0,36	0,45	0,44	0,52	0,82
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,13	0,16	0,30
	<-2	2,28	0,76	2,54	2,72	2,69	2,88	3,40
	<-1	15,87	12,64	17,39	17,78	17,69	18,23	19,52
	>1	15,87	12,70	14,06	14,40	14,36	14,65	15,58
	>2	2,28	0,32	1,62	1,77	1,74	1,92	2,22
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,20
r_i^{**ans}	<-3	0,13	0,16	0,48	0,57	0,57	0,66	0,88
	<-2	2,28	1,40	3,82	4,06	4,00	4,27	4,64
	<-1	15,87	8,98	14,91	15,46	15,34	16,05	17,84
	>1	15,87	8,44	11,91	12,24	12,21	12,64	13,24
	>2	2,28	1,02	2,60	2,76	2,76	2,97	3,42
	>3	0,13	0,02	0,26	0,32	0,32	0,38	0,68
r_i^{*wil}	<-3	0,13	0,02	0,12	0,14	0,16	0,20	0,34
	<-2	2,28	1,00	2,68	2,84	2,84	3,04	3,48
	<-1	15,87	14,06	17,51	17,91	17,90	18,38	19,76
	>1	15,87	13,36	14,25	14,64	14,58	14,90	15,72
	>2	2,28	0,88	1,84	1,92	1,93	2,09	2,36
	>3	0,13	0,00	0,06	0,09	0,10	0,12	0,26
r_i^{**wil}	<-3	0,13	0,32	0,54	0,64	0,64	0,73	0,96
	<-2	2,28	2,00	4,02	4,24	4,20	4,45	4,78
	<-1	15,87	10,46	15,13	15,63	15,59	16,20	18,06
	>1	15,87	10,08	12,15	12,51	12,48	12,85	13,42
	>2	2,28	1,76	2,80	2,98	2,98	3,13	3,70
	>3	0,13	0,18	0,32	0,40	0,39	0,48	0,74

Tabela E.11: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 7b.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,02	0,14	0,16	0,17	0,21	0,30
	<-2	2,28	2,06	2,68	2,98	2,92	3,15	3,90
	<-1	15,87	14,40	17,25	17,91	17,96	18,65	20,90
	>1	15,87	12,22	13,89	14,41	14,31	14,80	16,66
	>2	2,28	1,28	1,74	1,83	1,84	2,01	2,34
	>3	0,13	0,00	0,06	0,08	0,08	0,10	0,22
r_i^{**dev}	<-3	0,13	0,30	0,62	0,84	0,84	1,02	1,66
	<-2	2,28	3,68	4,33	4,52	4,51	4,68	5,16
	<-1	15,87	9,22	13,35	14,80	14,86	16,50	20,02
	>1	15,87	8,58	11,29	11,68	11,51	12,04	13,06
	>2	2,28	1,62	2,64	3,20	3,03	3,40	3,92
	>3	0,13	0,06	0,32	0,52	0,51	0,68	1,12
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,00	0,00	0,06
	<-2	2,28	0,04	0,76	1,28	1,16	1,57	2,10
	<-1	15,87	14,36	16,54	16,93	16,88	17,26	18,50
	>1	15,87	13,38	14,74	15,02	15,03	15,30	16,84
	>2	2,28	2,06	2,84	3,16	3,14	3,41	4,24
	>3	0,13	0,06	0,36	0,50	0,50	0,60	1,16
r_i^{**pea}	<-3	0,13	0,00	0,02	0,14	0,22	0,34	0,92
	<-2	2,28	0,10	1,73	2,64	2,48	3,27	4,20
	<-1	15,87	8,90	12,67	13,74	13,51	14,68	15,96
	>1	15,87	8,78	11,81	12,61	12,44	13,34	14,84
	>2	2,28	3,56	4,28	4,47	4,44	4,63	5,12
	>3	0,13	0,92	1,16	1,27	1,28	1,38	1,68
r_i^{*qu}	<-3	0,13	0,02	0,06	0,10	0,10	0,13	0,24
	<-2	2,28	1,54	2,12	2,33	2,29	2,45	2,84
	<-1	15,87	13,30	15,60	15,98	15,93	16,27	17,36
	>1	15,87	13,76	15,73	16,01	16,04	16,32	17,58
	>2	2,28	1,50	2,08	2,22	2,20	2,35	2,94
	>3	0,13	0,00	0,06	0,10	0,10	0,14	0,24
r_i^{**qu}	<-3	0,13	0,16	0,42	0,62	0,63	0,82	1,26
	<-2	2,28	2,80	3,40	3,63	3,65	3,90	4,40
	<-1	15,87	8,28	12,01	13,18	12,96	14,13	15,90
	>1	15,87	8,78	12,39	13,23	13,06	14,11	16,02
	>2	2,28	2,60	3,30	3,61	3,58	3,88	4,32
	>3	0,13	0,18	0,40	0,62	0,60	0,74	1,14
r_i^{*ans}	<-3	0,13	0,02	0,08	0,12	0,12	0,15	0,26
	<-2	2,28	1,64	2,50	2,76	2,70	2,92	3,44
	<-1	15,87	14,24	17,08	17,77	17,78	18,52	20,58
	>1	15,87	11,88	13,72	14,27	14,15	14,63	16,60
	>2	2,28	1,06	1,58	1,71	1,70	1,88	2,26
	>3	0,13	0,00	0,04	0,06	0,06	0,08	0,18
r_i^{**ans}	<-3	0,13	0,18	0,48	0,77	0,73	0,93	1,46
	<-2	2,28	3,26	4,06	4,26	4,26	4,45	4,88
	<-1	15,87	8,70	13,27	14,72	14,65	16,27	19,58
	>1	15,87	8,28	11,08	11,52	11,31	11,87	12,94
	>2	2,28	1,50	2,44	2,98	2,85	3,23	3,76
	>3	0,13	0,06	0,28	0,44	0,44	0,58	0,96
r_i^{*wil}	<-3	0,13	0,02	0,12	0,16	0,16	0,20	0,30
	<-2	2,28	2,02	2,62	2,90	2,88	3,12	3,80
	<-1	15,87	14,40	17,25	17,91	17,94	18,61	20,82
	>1	15,87	12,32	13,90	14,43	14,32	14,82	16,66
	>2	2,28	1,30	1,76	1,88	1,88	2,06	2,38
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,22
r_i^{**wil}	<-3	0,13	0,30	0,58	0,82	0,82	1,00	1,62
	<-2	2,28	3,64	4,29	4,48	4,47	4,64	5,14
	<-1	15,87	9,20	13,32	14,80	14,83	16,45	20,02
	>1	15,87	8,64	11,29	11,69	11,53	12,06	13,08
	>2	2,28	1,68	2,70	3,21	3,08	3,46	3,98
	>3	0,13	0,10	0,32	0,54	0,53	0,70	1,16

Tabela E.12: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 8a.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,06	0,12	0,16	0,17	0,21	0,34
	<-2	2,28	1,96	2,64	2,82	2,85	3,04	3,72
	<-1	15,87	15,54	17,15	17,80	17,85	18,52	20,32
	>1	15,87	11,78	13,60	14,18	14,13	14,69	15,90
	>2	2,28	1,26	1,66	1,84	1,84	2,02	2,54
	>3	0,13	0,00	0,06	0,08	0,08	0,12	0,22
r_i^{**dev}	<-3	0,13	0,40	0,74	0,88	0,91	1,08	1,44
	<-2	2,28	3,70	4,32	4,53	4,54	4,74	5,42
	<-1	15,87	10,22	13,06	14,52	14,44	15,67	18,34
	>1	15,87	9,30	10,73	11,16	11,10	11,50	12,36
	>2	2,28	1,90	2,80	3,14	3,11	3,48	4,32
	>3	0,13	0,10	0,38	0,53	0,54	0,69	1,06
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,00	0,14
	<-2	2,28	0,04	0,76	1,18	1,13	1,58	2,12
	<-1	15,87	15,36	16,52	16,96	16,95	17,36	18,32
	>1	15,87	13,46	14,40	14,72	14,75	15,06	16,06
	>2	2,28	2,12	2,80	3,10	3,08	3,32	3,96
	>3	0,13	0,12	0,36	0,48	0,49	0,62	0,96
r_i^{**pea}	<-3	0,13	0,00	0,02	0,12	0,21	0,36	0,98
	<-2	2,28	0,42	1,82	2,60	2,54	3,29	3,96
	<-1	15,87	9,96	12,41	13,35	13,20	13,96	15,38
	>1	15,87	9,46	11,44	12,00	11,92	12,56	13,80
	>2	2,28	3,68	4,22	4,40	4,43	4,68	5,38
	>3	0,13	0,98	1,26	1,34	1,36	1,46	1,82
r_i^{*qu}	<-3	0,13	0,02	0,08	0,12	0,12	0,16	0,28
	<-2	2,28	1,68	2,18	2,28	2,29	2,42	3,06
	<-1	15,87	14,46	15,52	15,93	15,91	16,26	17,18
	>1	15,87	14,64	15,57	15,94	15,94	16,29	17,54
	>2	2,28	1,64	2,10	2,24	2,26	2,40	2,96
	>3	0,13	0,00	0,08	0,10	0,11	0,14	0,24
r_i^{**qu}	<-3	0,13	0,20	0,56	0,68	0,70	0,86	1,20
	<-2	2,28	2,98	3,56	3,78	3,76	3,96	4,78
	<-1	15,87	9,44	11,92	12,82	12,68	13,46	14,94
	>1	15,87	9,58	11,93	12,74	12,69	13,46	15,64
	>2	2,28	2,88	3,50	3,70	3,73	3,96	4,72
	>3	0,13	0,24	0,52	0,66	0,68	0,84	1,18
r_i^{*ans}	<-3	0,13	0,02	0,10	0,14	0,14	0,18	0,30
	<-2	2,28	1,84	2,54	2,71	2,72	2,90	3,46
	<-1	15,87	15,46	17,07	17,75	17,77	18,39	20,10
	>1	15,87	11,60	13,50	14,10	14,05	14,62	15,80
	>2	2,28	1,14	1,58	1,74	1,74	1,92	2,46
	>3	0,13	0,00	0,04	0,06	0,07	0,10	0,22
r_i^{**ans}	<-3	0,13	0,30	0,66	0,82	0,83	1,00	1,38
	<-2	2,28	3,58	4,18	4,39	4,39	4,59	5,32
	<-1	15,87	9,94	12,92	14,42	14,34	15,56	18,18
	>1	15,87	9,04	10,64	11,07	11,00	11,38	12,34
	>2	2,28	1,78	2,70	3,04	3,00	3,38	4,10
	>3	0,13	0,02	0,32	0,48	0,49	0,66	0,98
r_i^{*wil}	<-3	0,13	0,06	0,12	0,16	0,17	0,20	0,34
	<-2	2,28	1,94	2,62	2,80	2,83	3,01	3,72
	<-1	15,87	15,54	17,14	17,80	17,84	18,50	20,32
	>1	15,87	11,80	13,62	14,19	14,14	14,69	15,90
	>2	2,28	1,32	1,68	1,86	1,85	2,04	2,54
	>3	0,13	0,00	0,06	0,08	0,09	0,12	0,22
r_i^{**wil}	<-3	0,13	0,38	0,74	0,88	0,90	1,06	1,44
	<-2	2,28	3,70	4,30	4,50	4,52	4,72	5,38
	<-1	15,87	10,20	13,05	14,50	14,43	15,65	18,32
	>1	15,87	9,30	10,74	11,17	11,11	11,52	12,36
	>2	2,28	1,92	2,82	3,14	3,13	3,49	4,32
	>3	0,13	0,12	0,38	0,54	0,55	0,71	1,06

Tabela E.13: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo ZAIG - Cenário 8b.

Resíduos	Intervalo do resíduo	Valor		Simulado				
		Teórico	Min	Q1	Mediana	Média	Q3	Max
r_i^{*dev}	<-3	0,13	0,02	0,11	0,16	0,16	0,22	0,32
	<-2	2,28	2,10	2,75	3,06	3,09	3,49	4,14
	<-1	15,87	14,94	16,59	17,93	17,86	18,64	21,54
	>1	15,87	12,78	13,95	14,75	14,58	15,12	16,18
	>2	2,28	1,06	1,76	1,93	1,92	2,18	2,60
	>3	0,13	0,00	0,02	0,04	0,05	0,08	0,18
r_i^{**dev}	<-3	0,13	0,68	0,83	0,96	0,99	1,10	1,42
	<-2	2,28	3,62	4,59	4,86	4,82	5,12	5,84
	<-1	15,87	9,98	12,82	14,52	14,56	16,15	19,54
	>1	15,87	8,96	10,93	11,62	11,42	11,99	12,80
	>2	2,28	1,98	2,92	3,36	3,28	3,65	4,34
	>3	0,13	0,12	0,37	0,62	0,56	0,74	0,98
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,01	0,02	0,04
	<-2	2,28	0,14	0,98	1,50	1,31	1,74	2,12
	<-1	15,87	14,86	16,25	17,20	16,98	17,63	18,68
	>1	15,87	13,74	14,72	15,19	15,15	15,60	16,42
	>2	2,28	1,90	2,69	3,28	3,19	3,66	4,54
	>3	0,13	0,04	0,24	0,41	0,44	0,59	1,22
r_i^{**pea}	<-3	0,13	0,00	0,06	0,23	0,29	0,48	0,82
	<-2	2,28	0,78	2,23	3,19	2,82	3,51	3,92
	<-1	15,87	9,72	12,35	13,60	13,34	14,55	15,78
	>1	15,87	9,18	11,61	12,49	12,24	13,25	14,14
	>2	2,28	3,38	4,23	4,62	4,58	4,92	5,56
	>3	0,13	0,68	1,15	1,40	1,36	1,52	2,10
r_i^{*qu}	<-3	0,13	0,00	0,06	0,08	0,09	0,12	0,22
	<-2	2,28	1,36	2,01	2,40	2,32	2,61	3,04
	<-1	15,87	13,68	15,26	16,10	15,88	16,50	17,54
	>1	15,87	13,90	15,47	16,29	16,02	16,59	17,34
	>2	2,28	1,08	1,96	2,26	2,17	2,48	3,16
	>3	0,13	0,00	0,02	0,06	0,06	0,10	0,18
r_i^{**qu}	<-3	0,13	0,34	0,54	0,65	0,68	0,84	1,16
	<-2	2,28	2,80	3,46	3,90	3,78	4,09	4,36
	<-1	15,87	8,80	11,23	13,03	12,63	13,84	15,76
	>1	15,87	8,76	11,56	12,97	12,67	14,02	15,08
	>2	2,28	2,26	3,30	3,67	3,65	4,06	4,56
	>3	0,13	0,18	0,44	0,62	0,61	0,76	1,12
r_i^{*ans}	<-3	0,13	0,00	0,06	0,10	0,10	0,14	0,24
	<-2	2,28	1,44	2,30	2,81	2,72	3,10	3,66
	<-1	15,87	14,34	16,20	17,76	17,55	18,29	21,14
	>1	15,87	12,50	13,70	14,31	14,29	14,95	16,06
	>2	2,28	0,80	1,41	1,64	1,67	1,90	2,50
	>3	0,13	0,00	0,00	0,02	0,03	0,04	0,14
r_i^{**ans}	<-3	0,13	0,48	0,64	0,77	0,80	0,95	1,28
	<-2	2,28	3,06	4,01	4,56	4,39	4,78	5,38
	<-1	15,87	9,54	12,35	14,34	14,16	15,81	19,08
	>1	15,87	8,34	10,38	11,26	11,07	11,79	12,60
	>2	2,28	1,72	2,48	2,99	2,95	3,39	4,14
	>3	0,13	0,08	0,27	0,46	0,43	0,56	0,90
r_i^{*wil}	<-3	0,13	0,02	0,10	0,14	0,15	0,20	0,30
	<-2	2,28	2,02	2,68	3,02	3,01	3,41	4,00
	<-1	15,87	14,94	16,48	17,92	17,83	18,60	21,44
	>1	15,87	12,86	14,02	14,76	14,62	15,13	16,20
	>2	2,28	1,28	1,84	2,00	2,00	2,20	2,68
	>3	0,13	0,00	0,03	0,06	0,07	0,10	0,18
r_i^{**wil}	<-3	0,13	0,60	0,79	0,91	0,94	1,08	1,36
	<-2	2,28	3,52	4,43	4,79	4,73	5,04	5,70
	<-1	15,87	9,92	12,80	14,51	14,51	16,11	19,50
	>1	15,87	8,98	10,98	11,68	11,46	12,06	12,84
	>2	2,28	2,06	3,01	3,42	3,35	3,70	4,34
	>3	0,13	0,14	0,42	0,64	0,60	0,78	0,98

Apêndice F

Tabelas do ajuste do modelo incorreto

Tabela F.1: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - cenário I : Gerando ZAIG e ajustando ZAGA.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,00	0,00	0,02	0,25	0,19	3,14
	<-2	2,28	0,04	0,52	1,20	2,46	3,43	12,46
	<-1	15,87	11,78	14,29	16,18	17,13	19,27	28,44
	>1	15,87	10,60	12,14	13,49	14,02	15,71	19,60
	>2	2,28	0,02	0,64	1,40	2,21	3,38	8,20
	>3	0,13	0,00	0,00	0,06	0,32	0,42	2,42
r_i^{**dev}	<-3	0,13	0,00	0,10	0,24	0,75	0,95	5,78
	<-2	2,28	0,36	1,44	2,52	3,77	5,23	14,40
	<-1	15,87	5,30	9,47	11,98	13,24	16,76	27,02
	>1	15,87	4,02	8,07	10,19	10,73	13,63	18,84
	>2	2,28	0,26	1,47	2,54	3,25	4,85	9,40
	>3	0,13	0,00	0,14	0,39	0,81	1,22	4,00
r_i^{*pea}	<-3	0,13	0,00	0,00	0,00	0,02	0,01	0,48
	<-2	2,28	0,02	0,22	0,59	1,43	1,91	9,44
	<-1	15,87	11,54	13,91	15,60	16,61	18,59	27,84
	>1	15,87	10,78	12,38	13,90	14,39	16,18	20,20
	>2	2,28	0,08	1,00	2,10	2,95	4,53	9,70
	>3	0,13	0,00	0,06	0,20	0,70	0,98	4,24
r_i^{**pea}	<-3	0,13	0,00	0,02	0,06	0,21	0,22	2,08
	<-2	2,28	0,18	0,84	1,56	2,57	3,54	11,60
	<-1	15,87	4,74	8,74	11,23	12,48	15,92	26,22
	>1	15,87	4,40	8,58	10,84	11,25	14,15	19,26
	>2	2,28	0,60	2,05	3,39	4,06	5,97	10,74
	>3	0,13	0,02	0,34	0,76	1,37	2,18	5,58
r_i^{*qu}	<-3	0,13	0,00	0,00	0,02	0,17	0,10	2,36
	<-2	2,28	0,02	0,38	0,98	2,02	2,85	11,10
	<-1	15,87	10,20	12,51	14,47	15,48	17,79	26,90
	>1	15,87	12,00	13,46	14,75	15,16	16,71	20,32
	>2	2,28	0,04	0,71	1,55	2,36	3,69	8,26
	>3	0,13	0,00	0,00	0,07	0,33	0,45	2,34
r_i^{**qu}	<-3	0,13	0,00	0,06	0,16	0,57	0,71	4,46
	<-2	2,28	0,18	1,11	2,03	3,14	4,33	12,76
	<-1	15,87	3,90	8,11	10,58	11,79	15,24	25,42
	>1	15,87	4,88	9,21	11,18	11,60	14,43	19,16
	>2	2,28	0,34	1,73	2,82	3,50	5,23	9,46
	>3	0,13	0,00	0,16	0,44	0,85	1,33	4,02
r_i^{*ans}	<-3	0,13	0,00	0,00	0,02	0,21	0,16	2,52
	<-2	2,28	0,04	0,48	1,15	2,31	3,32	11,82
	<-1	15,87	11,52	14,20	16,04	16,94	19,05	28,04
	>1	15,87	10,38	12,00	13,41	13,88	15,57	19,36
	>2	2,28	0,02	0,60	1,31	2,10	3,27	7,78
	>3	0,13	0,00	0,00	0,05	0,28	0,36	2,08
r_i^{**ans}	<-3	0,13	0,00	0,10	0,22	0,67	0,87	4,88
	<-2	2,28	0,30	1,37	2,37	3,59	4,94	13,66
	<-1	15,87	5,10	9,26	11,84	13,03	16,55	26,56
	>1	15,87	3,80	7,92	10,06	10,55	13,40	18,48
	>2	2,28	0,24	1,41	2,42	3,13	4,75	9,00
	>3	0,13	0,00	0,12	0,35	0,74	1,13	3,58
r_i^{*wil}	<-3	0,13	0,00	0,00	0,02	0,24	0,19	2,92
	<-2	2,28	0,04	0,51	1,20	2,43	3,35	12,30
	<-1	15,87	11,76	14,27	16,18	17,11	19,25	28,40
	>1	15,87	10,60	12,14	13,49	14,04	15,72	19,66
	>2	2,28	0,02	0,64	1,44	2,24	3,46	8,24
	>3	0,13	0,00	0,00	0,06	0,34	0,43	2,56
r_i^{**wil}	<-3	0,13	0,00	0,10	0,24	0,73	0,93	5,50
	<-2	2,28	0,36	1,42	2,46	3,74	5,16	14,26
	<-1	15,87	5,30	9,46	11,97	13,22	16,76	26,92
	>1	15,87	4,04	8,07	10,20	10,74	13,65	18,88
	>2	2,28	0,26	1,47	2,55	3,28	4,87	9,54
	>3	0,13	0,00	0,14	0,39	0,83	1,23	4,16

Tabela F.2: Estatísticas descritivas da porcentagem observada dos resíduos para cada intervalo - cenário II : Gerando ZAGA e ajustando ZAIG.

Resíduo	Intervalo do resíduo	Valor Teórico	Simulado					
			Mínimo	Q1	Mediana	Média	Q3	Máximo
r_i^{*dev}	<-3	0,13	0,00	0,04	0,19	0,44	0,71	2,20
	<-2	2,28	0,02	0,95	2,80	2,96	5,01	7,28
	<-1	15,87	6,06	13,51	16,51	15,48	17,79	19,64
	>1	15,87	4,62	12,04	15,28	14,15	16,75	18,84
	>2	2,28	0,02	0,59	1,99	2,27	3,70	6,30
	>3	0,13	0,00	0,00	0,06	0,22	0,30	1,18
r_i^{**dev}	<-3	0,13	0,00	0,18	1,00	1,45	2,68	5,02
	<-2	2,28	0,08	1,78	4,40	4,38	6,99	9,44
	<-1	15,87	4,48	10,58	13,34	12,44	14,53	16,26
	>1	15,87	3,24	9,10	11,92	11,14	13,38	15,36
	>2	2,28	0,02	1,20	3,28	3,53	5,61	8,78
	>3	0,13	0,00	0,04	0,51	1,00	1,69	4,60
r_i^{*pea}	<-3	0,13	0,00	0,00	0,04	0,23	0,29	1,62
	<-2	2,28	0,00	0,54	2,12	2,36	4,18	6,86
	<-1	15,87	4,94	12,88	16,24	15,06	17,61	19,46
	>1	15,87	5,42	12,60	15,56	14,50	16,94	19,08
	>2	2,28	0,10	0,99	2,52	2,79	4,37	6,72
	>3	0,13	0,00	0,04	0,18	0,41	0,67	1,70
r_i^{**pea}	<-3	0,13	0,00	0,04	0,51	1,05	1,99	4,52
	<-2	2,28	0,00	1,02	3,62	3,72	6,29	9,10
	<-1	15,87	3,50	9,81	12,77	11,88	14,29	16,02
	>1	15,87	3,96	9,73	12,29	11,61	13,71	15,64
	>2	2,28	0,18	1,72	3,96	4,11	6,31	9,08
	>3	0,13	0,00	0,16	0,84	1,36	2,19	5,10
r_i^{*qu}	<-3	0,13	0,00	0,02	0,15	0,37	0,57	1,76
	<-2	2,28	0,02	0,80	2,45	2,64	4,65	6,84
	<-1	15,87	4,28	12,01	15,51	14,38	17,15	18,76
	>1	15,87	6,38	13,39	15,98	15,06	17,30	19,46
	>2	2,28	0,02	0,70	2,13	2,34	3,83	6,00
	>3	0,13	0,00	0,00	0,06	0,21	0,30	1,10
r_i^{**qu}	<-3	0,13	0,00	0,14	0,86	1,28	2,34	4,50
	<-2	2,28	0,04	1,40	3,89	3,95	6,54	8,80
	<-1	15,87	3,26	9,28	12,33	11,43	13,88	15,64
	>1	15,87	4,54	10,11	12,57	11,86	13,88	15,86
	>2	2,28	0,08	1,32	3,59	3,68	5,81	8,66
	>3	0,13	0,00	0,04	0,56	1,01	1,72	4,06
r_i^{*ans}	<-3	0,13	0,00	0,02	0,18	0,39	0,65	1,82
	<-2	2,28	0,02	0,90	2,64	2,82	4,91	7,10
	<-1	15,87	5,82	13,31	16,45	15,33	17,67	19,48
	>1	15,87	4,52	11,86	15,07	14,00	16,64	18,82
	>2	2,28	0,00	0,54	1,88	2,14	3,50	5,78
	>3	0,13	0,00	0,00	0,05	0,19	0,27	1,02
r_i^{**ans}	<-3	0,13	0,00	0,18	0,96	1,36	2,52	4,68
	<-2	2,28	0,06	1,68	4,22	4,22	6,79	9,02
	<-1	15,87	4,38	10,46	13,08	12,26	14,38	16,02
	>1	15,87	3,14	8,92	11,84	10,96	13,25	15,24
	>2	2,28	0,02	1,19	3,21	3,39	5,51	8,28
	>3	0,13	0,00	0,04	0,47	0,92	1,58	3,86
r_i^{*wil}	<-3	0,13	0,00	0,04	0,19	0,43	0,69	2,18
	<-2	2,28	0,02	0,95	2,78	2,94	5,01	7,28
	<-1	15,87	6,04	13,48	16,51	15,47	17,79	19,64
	>1	15,87	4,62	12,08	15,29	14,16	16,75	18,84
	>2	2,28	0,02	0,61	2,01	2,29	3,71	6,32
	>3	0,13	0,00	0,00	0,06	0,23	0,31	1,24
r_i^{**wil}	<-3	0,13	0,00	0,18	1,00	1,44	2,66	4,96
	<-2	2,28	0,08	1,74	4,38	4,36	6,94	9,38
	<-1	15,87	4,46	10,58	13,34	12,43	14,51	16,26
	>1	15,87	3,24	9,12	11,93	11,15	13,38	15,36
	>2	2,28	0,02	1,20	3,30	3,55	5,63	8,78
	>3	0,13	0,00	0,04	0,51	1,02	1,69	4,70

Referências Bibliográficas

- Abdo, A.,Ribeiro, F.,Juliane,Cirano,Casat,Zaffalon, M.,Ribeiro, F. V.,Giampaoli, V.,Viana Casarin, R. C. e Pimentel, S. P. (2013), ‘Influence of dyslipidemia and diabetes mellitus on chronic periodontal disease’, *Journal of periodontology* **84**(10), 1401–1408. 3
- Akaike, H. (1974), ‘A new look at the statistical model identification’, *IEEE Transactions on Automatic Control* **19**(6), 716–723. 14
- Anderson, T. W. e Darling, D. A. (1952), ‘Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes’, *The annals of mathematical statistics* **23**, 193–212. 19
- Anscombe, F. (1953), ‘Contribution to the discussion of h. hotellingâs paper’, *Journal of the Royal Statistical Society* **15**(2), 229–230. 16, 18
- Atkinson, A. C. (1985), *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis*, Oxford. 16, 45
- Bortoluzzo, A. B.,Claro, D. P.,Caetano, M. A. L. e Artes, R. (2011), ‘Estimating total claim size in the auto insurance industry: a comparison between tweedie and zero-adjusted inverse gaussian distribution’, *BAR-Brazilian Administration Review* **8**(1), 37–47. 4, 8
- 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. 13
- Cook, D. O.,Kieschnick, R. e McCullough, B. D. (2008), ‘Regression analysis of proportions in finance with self selection’, *Journal of empirical finance* **15**(5), 860–867. 4
- Cook, R. D. (1977), ‘Detection of influential observation in linear regression’, *Technometrics* **19**, 15–18. 43
- Cook, R. D. e Weisberg, S. (1982), ‘Residuals and influence in regression’, *Chapman and Hall* **19**, 15–18. 43
- Cordeiro, G. M. e Demétrio, C. G. (2008), *Modelos lineares generalizados e Extensões*, Recife: UFRPE. 16, 17, 43
- De Jong, P.,Heller, G. Z. et al. (2008), *Generalized linear models for insurance data*, Vol. 136, Cambridge University Press Cambridge. 8
- Dias, M. F. (2014), Modelos assimétricos in aficionados de zeros, Dissertação de mestrado em estatística., Universidade de São Paulo. 1, 3, 8
- Dunn, P. K. e Smyth, G. K. (1996), ‘Randomized quantile residuals’, *Journal of Computational and Graphical Statistics* **5**(3), 236–244. 1, 16, 17, 28

- Dunn, P. K. e Smyth, G. K. (2005), ‘Series evaluation of tweedie exponential dispersion model densities’, *Statistics and Computing* **15**(4), 267–280. 3
- Ferreira, F. (2005), Uma Abordagem Clássica e Bayesiana para o Modelo de Tweedie, Dissertação do programa de pós-graduação em estatística, Departamento de Estatística da Universidade Federal de São Carlos. 4
- Ferreira, J. (2008), Modelos de previsão de perdas para crédito massificado, Dissertação de mestrado, Ibmec. São Paulo. 8
- Feuerverger, A. (1979), ‘On some methods of analysis for weather experiments’, *Biometrika* **66**(3), 655–658. 1, 2, 3, 4, 8, 9
- Fonseca, N., Bressan, A., Iquiapaza, R., Guerra, J. et al. (2007), Análise do desempenho recente de fundos de investimento no brasil], Technical report, University Library of Munich, Germany. 49
- Heller, G., Stasinopoulos, M., Rigby, B. et al. (2006), ‘The zero-adjusted inverse gaussian distribution as a model for insurance claims’, **226233**. 2, 4, 8, 11
- Hoff, A. (2007), ‘Second stage dea: Comparison of approaches for modelling the dea score’, *European Journal of Operational Research* **181**(1), 425–435. 4
- Hosmer, D. W., Hosmer, T., Le Cessie, S., Lemeshow, S. et al. (1997), ‘A comparison of goodness-of-fit tests for the logistic regression model’, *Statistics in medicine* **16**(9), 965–980. 47
- Hosmer, D. W., Taber, S. e Lemeshow, S. (1991), ‘The importance of assessing the fit of logistic regression models: a case study.’, *American journal of public health* **81**(12), 1630–1635. 47
- Hosmer Jr, D. W. e Lemeshow, S. (2004), *Applied logistic regression*, John Wiley & Sons. 29, 47
- Johnson, N. L., Kotz, S. e Balakrishnan, N. (1994), *Continuous univariate distributions*, Vol. 1-2, New York: John Wiley & Sons. 65
- Jørgensen, B. (1997), *The theory of dispersion models*, Vol. 1, Chapman and Hall/CRC. 3
- Jørgensen, B. e Paes De Souza, M. C. (1994), ‘Fitting tweedie’s compound poisson model to insurance claims data’, *Scandinavian Actuarial Journal* **1994**(1), 69–93. 3
- Liu, Y. (2007), *On goodness-of-fit of logistic regression model*, ProQuest. 47
- McCullagh, P. (1985), ‘On the asymptotic distribution of pearson’s statistic in linear exponential-family models’, *International Statistical Review/Revue Internationale de Statistique* pp. 61–67. 47
- McCullagh, P. e Nelder, J. A. (1989), *Generalized linear models*, Vol. 37, CRC press. 16
- Nelder, J. A. e Wedderburn, R. W. M. (1972), ‘Generalized linear models’, *Journal of the Royal Statistical Society* **10**(3), 370–384. 10, 15

- Ospina, R. (2008), Modelos de regressão beta inflacionados, Tese do programa de pós-graduação em estatística, Instituto de Matemática e Estatística da Universidade de São Paulo. 13, 64
- Ospina, R. e Ferrari, S. L. (2012), ‘A general class of zero-or-one inflated beta regression models’, *Computational Statistics & Data Analysis* **56**(6), 1609–1623. 2, 4, 29, 32, 33
- Pace, L. e Salvan, A. (1997), *Principles of statistical inference: from a Neo-Fisherian perspective*, Vol. 4, World Scientific. 10
- Paula, G. A. (2013), *Modelos de regressão: com apoio computacional*, IME-USP São Paulo. 10, 16, 19, 43
- Pereira, G. H., Botter, D. A. e Sandoval, M. C. (2012), ‘The truncated inflated beta distribution’, *Communications in Statistics-Theory and Methods* **41**(5), 907–919. 4, 47
- Pereira, G. H., Botter, D. A. e Sandoval, M. C. (2014), ‘A new residual in zero inflated regression models’, *Proceedings of the 21th Simpósio Nacional de Probabilidade e Estatística.* (), . 1, 2, 28, 30, 32, 33
- Pereira, L., Novaes, D. A., Dias, P. L. et al. (2014), ‘Fundos de investimentos,’. **URL:** <http://www.investidor.gov.br> 49
- Pousinho, A. P. (2013), Pagamentos de sinistros, Dissertação de mestrado do instituto superior de economia e gestão, Instituto Superior de Economia e Gestão. 4
- Resti, Y., Ismail, N. e Jamaan, S. H. (2013), ‘Estimation of claim cost data using zero adjusted gamma and inverse gaussian regression models’, *Journal of Mathematics and Statistics* **9**(3), 186–192. 1, 3, 8
- 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. 13
- 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. 13, 14
- Rocha, L. O., Nakai, V. K., Braghini, R., Reis, T. A., Kobashigawa, E. e Corrêa, B. (2009), ‘Mycoflora and co-occurrence of fumonisins and aflatoxins in freshly harvested corn in different regions of brazil’, *International journal of molecular sciences* **10**(11), 5090–5103. 3
- Royston, P. (1992), ‘The use of cusums and other techniques in modelling continuous covariates in logistic regression’, *Statistics in medicine* **11**(8), 1115–1129. 47
- Schwarz, G. et al. (1978), ‘Estimating the dimension of a model’, *The annals of statistics* **6**(2), 461–464. 14
- 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. 13
- Stasinopoulos, D., Rigby, B. e Akantziliotou, C. (2014), ‘Instructions on how to use the gamlss package in r, 2008’. 5, 6, 7, 8, 20, 33, 41

- Stukel, T. A. (1988), ‘Generalized logistic models’, *Journal of the American Statistical Association* **83**(402), 426–431. 47
- Tong, E. N., Mues, C. e Thomas, L. (2013), ‘A zero-adjusted gamma model for mortgage loan loss given default’, *International Journal of Forecasting* **29**(4), 548–562. 1, 8
- Troiani, R. R. (2009), Aplicação do modelo ZAIG (Zero Adjusted Inverse Gaussian) na análise de cartões de créditos, Dissertação de mestrado profissional em economia., Faculdade Ibmec São Paulo. 1, 8
- Venezuela, M. K., Artes, R. et al. (2014), ‘Estimating equations and diagnostic techniques applied to zero-inflated models for panel data’, *Electronic Journal of Statistics* **8**(1), 1641–1660. 1, 8
- Williams, D. (1987), ‘Generalized linear model diagnostics using the deviance and single case deletions’, *Applied Statistics* pp. 181–191. 16, 19, 26