Adjusted quantile residual for generalized linear models

Juliana Scudilio\textsuperscript{a,b,*}, Gustavo H. A. Pereira\textsuperscript{b}

\textsuperscript{a}Maths Science Institute and Computing, University of São Paulo, São Carlos, SP, Brazil.
\textsuperscript{b} Department of Statistics, Federal University of São Carlos, São Carlos, SP, Brazil.

Abstract

Generalized linear models are widely used in many areas of knowledge. As in other classes of regression models, it is desirable to perform diagnostic analysis in generalized linear models using residuals that are approximately standard normally distributed. Diagnostic analysis in this class of models are usually performed using the standardized Pearson residual or the standardized deviance residual. The former has skewed distribution and the latter has negative mean, specially when the variance of the response variable is high. In this work, we introduce the adjusted quantile residual for generalized linear models. Using Monte Carlo simulation techniques and two applications, we compare this residual with the standardized Pearson residual, the standardized deviance residual and two other residuals. Overall, the results suggest that the adjusted quantile residual is a better tool for diagnostic analysis in generalized linear models.

Keywords: Diagnostic analysis; generalized linear models; quantile residual.

1. Introduction

Generalized linear models (GLMs) were introduced by Nelder and Wedderburn (1972) and are widely used to study the relationship between a response variable and a set of predictor variables. It assumes that the distribution of the response variable given the set of predictors belongs to the exponential family and can be used to model many types of variables, such as binary variables, count variables and continuous skewed variables. The routines to fit GLMs are implemented in many statistical softwares and these models are described in details in McCullagh and Nelder (1989) and Hardin et al. (2007).

In regression models, it is important to check model adequacy and to identify outliers and influential observations. To reach these goals, it is interesting to use residuals whose distribution is well approximated by the standard normal distribution. When the residuals do not have this property, it is not unusual that they

\*e-mail: juliana-scudilio@uol.com.br
are also non-identically distributed (see, for example, Table 1 of Anholeto et al. (2014)). In these cases, it is hard to interpret residuals plots and well-fitted models can be discarded (see Application 3 of Pereira (2017)).

Many residuals are used to perform diagnostic analysis in GLMs, such as the deviance residual (Davison and Gigli, 1989), the Pearson residual (McCullagh and Nelder, 1989), the Williams residual (Williams, 1987) and the Anscombe residual (Pierce and Schafer, 1986). Usually they are standardized so that their variance become close to one. However, the distributions of these residuals are not approximated standard normal when the variance of the response variable is high (see Section 3). Some works introduced residuals for GLMs that are approximated standard normal distributed (Cordeiro, 2004; Urbano et al., 2012), but they are complex and are not implemented in the statistical softwares.

The quantile residual (Dunn and Smyth, 1996) is simple and general, but is not often used in GLMs. It is usually used to perform diagnostic analysis in complex regression models, such as the generalized additive models for location, scale and shape (Rigby and Stasinopoulos, 2005). The quantile residual is asymptotically standard normally distributed if the parameters are consistently estimated (Dunn and Smyth, 1996), but its distribution is unknown in small sample sizes. Feng et al. (2017) compared the quantile residual with the deviance residual and the Pearson residual in generalized linear models. Using simulation studies, the authors concluded that the distribution of the quantile residual is better approximated by the standard normal distribution than that of the others residuals. In addition, Feng et al. (2017) showed that the quantile residual is the best for detecting lack of fit. However, the authors considered sample size equal to 1000 and they did not note that the variance of the quantile residual is far from one in small sample sizes (see Section 2.2).

In this work, we introduce the adjusted quantile residual for GLMs. Using Monte Carlo simulation studies and two applications, we compare the behavior of the adjusted quantile residual and other residuals usually used in GLMs.

This paper is organized as follows. Section 2 presents the models used in this work, introduces the adjusted quantile residual and defines other residuals considered in this paper. In Section 3, Monte Carlo simulation studies are performed to compare the properties of the defined residuals in several scenarios. In the following section, we study the ability of the residuals to identify model misspecification and outliers. Finally, in Section 5, we present the conclusions.
2. Methodology

2.1. Generalized linear models

Let \( y_1, y_2, \ldots, y_n \) be independent random variables, where each \( y_i \) has a density functions in the linear exponential family defined as

\[
\pi(y; \theta_i, \phi) = \exp \left[ \phi \left( y \theta_i - b(\theta_i) + c(y) \right) + d(y; \phi) \right],
\]

where \( b(\cdot) \) and \( c(\cdot) \) are know functions and \( \phi \) is a precision parameter. It can be proved that \( E(y_i) = \mu_i = b'(\theta_i) \) and \( Var(y_i) = \phi^{-1} V(\mu_i) \), where \( V_i = V(\mu_i) = d\mu_i/d\theta_i \) is the variance function.

Generalized linear models (GLMs) are defined by the family of distributions specified in equation 1 and the systematic component

\[
g(\mu_i) = \eta_i = x_i^T \beta,
\]

where \( x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})^T \) is a vector of known covariates \( (k < n) \), \( \beta = (\beta_1, \beta_2, \ldots, \beta_k)^T \) is a vector of unknown parameters \( (\beta \in \mathbb{R}^k) \) and \( g(.) \) is a strictly monotonic and twice differentiable link function.

This paper focus on generalized linear models with continuous and asymmetric response variables. We considered only continuous variable as response variable, because in the discrete case the residual distribution cannot be continuous without randomization. We also use asymmetric distributions because in the normal case a residual with good properties is already known.

The gamma regression models and the inverse Gaussian regression models are the most important regression models in the class of GLMs with asymmetric and continuous response variable. They are considered in this paper and we define these models using the parameterization used by Stasinopoulos et al. (2014).

The gamma distribution with parameters \( \mu_i \) and \( \sigma \) has a density function defined as

\[
f_1(y; \mu_i, \sigma) = \frac{\exp\left(\frac{-1}{\sigma^2} \frac{y}{\mu_i} \right)}{\Gamma\left(\frac{1}{\sigma^2}\right) \left(\mu_i \sigma^2\right)^{\frac{1}{\sigma^2}}} y^{\frac{1}{\sigma^2}-1} \]

where \( y > 0, \mu_i > 0 \) and \( \sigma > 0 \). In this parameterization of the gamma distribution, \( E(y_i) = \mu_i \) and \( Var(y_i) = \sigma^2 \mu_i^2 \).

The inverse Gaussian distribution with parameters \( \mu_i \) and \( \sigma \) is defined as

\[
f_2(y; \mu_i, \sigma) = \frac{1}{\sqrt{2\pi\sigma^3}} \exp\left\{ -\frac{(y - \mu_i)^2}{2\mu_i^2\sigma^2 y} \right\},
\]

where \( y > 0, \mu_i > 0 \) and \( \sigma > 0 \). In this parameterization of the inverse Gaussian, \( E(y_i) = \mu_i \) and \( Var(y_i) = \sigma^2 \mu_i^3 \).
The parameters of the GLMs can be estimated by maximum likelihood using numerical optimization algorithms and hypothesis tests can be conducted using Wald, score or likelihood ratio statistic. The \textit{glm} function and the \textit{gamlss} package, which are available for the R statistical software, can be used for fitting gamma and inverse Gaussian regression models.

2.2. Standardized quantile residual

The quantile residual can be used in a wide class of regression models. When the response variable is discrete, it uses uniform distributed random variables and it is called randomized quantile residual. For the subclass of GLMs considered in this paper, it is defined as

$$r_{qi} = \Phi^{-1}\{F(y_i; \hat{\mu}_i, \hat{\sigma})\}, \quad (2)$$

where $\Phi(.)$ is the cumulative distribution function of the standard normal distribution, $F(.)$ is the cumulative distribution function of the gamma distribution or inverse Gaussian distribution and $\hat{\mu}_i$ and $\hat{\sigma}$ are the maximum likelihood estimates of the parameters $\mu_i$ and $\sigma$, respectively.

According to Dunn and Smyth (1996), $r_{qi}$ is asymptotically standard normally distributed, but it is important to study its properties when sample size is not large. We used Monte Carlo simulation studies to study the properties of the quantile residual in small sample sizes. Gamma and inverse Gaussian regression models were considered and we assumed the model $\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$. The covariate values were generated as independent draws from the standard uniform distribution and remained constant throughout the simulations. We considered $\beta_0 = 3$, $\beta_1 = 2$ and $\beta_2 = 1$, which resulted in $\mu \in (20.085; 403.429)$ and used $\sigma = 0.1$ for gamma regression model and $\sigma = 0.02$ for inverse Gaussian regression model.

The left side of Tables 1 and 2 present the sample mean, variance, skewness and kurtosis coefficients for $r_{qi}$ for gamma and inverse Gaussian regression models, respectively. The tables also contain the value of the Anderson-Darling (AD) statistic (Anderson and Darling, 1952), used to test whether each residual is standard normally distributed. The value of the Anderson-Darling statistic is used as a closeness measure between each residual distribution and standard normal distribution. Results are based on 5000 Monte Carlo replications. Note that for both models the variance of $r_{qi}$ is far from 1.

In GLMs, Pearson and deviance residuals are standardized so that their variance becomes close to one. Our proposal is to use the same standardization term used for these residual in the quantile residual, ie, divide the residual $r_{qi}^{\text{nu}}$ by $\sqrt{1 - \hat{h}_{ii}}$, where $\hat{h}_{ii}$ is the $i$-th diagonal element of matrix $\hat{H} = \hat{W}^{1/2}X(X^T\hat{W}X)^{-1}X^T\hat{W}^{1/2}$. Then, the
adjusted quantile residual is defined as

\[ r_{i}^{*} = \frac{r_{i}^{qu}}{\sqrt{(1 - \hat{h}_{ii})}}. \]

(3)

We used the term adjusted instead of standardized because Klar and Meintanis (2012) introduced a different residual called standardized quantile residual to derive a goodness-of-fit test for generalized linear models. We did not consider the standardized quantile residual in this work because it assumes that the variance of the quantile residual is constant across observations. Tables 1 and 2 suggest that this assumption is not reasonable.

The results of the adjusted quantile residual are shown on the right side of Tables 1 and 2. Note that dividing \( r_{i}^{qu} \) by \( \sqrt{(1 - \hat{h}_{ii})} \) results in a residual with mean, skewness and kurtosis similar to \( r_{i}^{qu} \), a variance close to 1 and smaller AD than the \( r_{i}^{qu} \). In other words, the distribution of \( r_{i}^{*qu} \) is closer to the standard normal distribution than that of the \( r_{i}^{qu} \). We did the same analysis in several scenarios, which were omitted in this paper because they resulted on similar conclusions.
Table 2: Comparison between the quantile residual and the adjusted quantile residual - inverse Gaussian distribution.

|   | Mean | Variance | Skewness | Kurtosis | AD  | Mean | Variance | Skewness | Kurtosis | AD  |
|---|------|----------|----------|----------|-----|------|----------|----------|----------|-----|
| 1 | 67.742 | 0.007 | 0.851 | -0.060 | -0.424 | 5.708 | 0.008 | 1.034 | -0.060 | -0.424 | 5.000 |
| 2 | 100.306 | 0.007 | 0.864 | -0.123 | -0.307 | 5.644 | 0.007 | 1.003 | -0.123 | -0.307 | 2.700 |
| 3 | 215.703 | 0.039 | 0.932 | -0.171 | -0.320 | 9.226 | 0.042 | 1.072 | -0.171 | -0.321 | 15.513 |
| 4 | 115.585 | 0.019 | 0.935 | -0.093 | -0.369 | 3.078 | 0.020 | 1.053 | -0.093 | -0.369 | 6.654 |
| 5 | 131.633 | 0.034 | 0.913 | -0.111 | -0.426 | 7.442 | 0.036 | 1.014 | -0.111 | -0.426 | 9.541 |
| 6 | 200.681 | 0.020 | 0.901 | -0.172 | -0.240 | 5.628 | 0.021 | 1.035 | -0.172 | -0.240 | 6.316 |
| 7 | 33.844 | 0.054 | 0.528 | -0.165 | -0.390 | 109.272 | 0.074 | 1.012 | -0.166 | -0.395 | 22.802 |
| 8 | 80.061 | 0.013 | 0.870 | -0.077 | -0.379 | 5.361 | 0.014 | 1.028 | -0.077 | -0.379 | 5.174 |
| 9 | 48.232 | 0.007 | 0.554 | -0.131 | -0.373 | 87.104 | 0.010 | 1.096 | -0.130 | -0.377 | 3.868 |
| 10 | 52.030 | -0.001 | 0.684 | -0.080 | -0.372 | 35.364 | -0.001 | 1.001 | -0.080 | -0.372 | 2.527 |
| 11 | 48.729 | 0.012 | 0.854 | -0.095 | -0.361 | 6.530 | 0.013 | 1.020 | -0.095 | -0.361 | 4.495 |
| 12 | 114.403 | 0.013 | 0.725 | -0.150 | -0.333 | 26.833 | 0.016 | 1.028 | -0.149 | -0.341 | 6.055 |
| 13 | 194.271 | -0.003 | 0.892 | -0.150 | -0.309 | 3.629 | -0.003 | 1.024 | -0.150 | -0.309 | 3.474 |
| 14 | 103.731 | 0.018 | 0.953 | -0.066 | -0.412 | 2.651 | 0.019 | 1.038 | -0.066 | -0.412 | 5.574 |
| 15 | 208.921 | 0.001 | 0.881 | -0.157 | -0.350 | 4.847 | 0.001 | 1.036 | -0.156 | -0.351 | 5.324 |

Mean 0.016   0.822 | -0.120 | -0.358 | 21.221 | 0.018 | 1.026 | -0.120 | -0.359 | 7.001 |

SD 0.016   0.136 | 0.040 | 0.050 | 32.846 | 0.020 | 0.020 | 0.040 | 0.050 | 5.407 |

2.3. Other residuals in GLMs

The other residuals considered in this paper are the deviance residual, the Pearson residual, the Anscombe residual and the William residual.

The standardized deviance residual is defined as

\[
r_{i}^{\text{dev}} = \frac{d(y_i; \hat{\mu}_i)}{\hat{\sigma}\sqrt{1 - \hat{h}_{ii}}},
\]

where \(d(y_i; \hat{\mu}_i) = \left\{-\log\left(\frac{y_i}{\hat{\mu}_i}\right) + \frac{(y_i - \hat{\mu}_i)^2}{y_i \hat{\mu}_i^2}\right\}\) for the gamma distribution and \(d(y_i; \hat{\mu}_i) = \left\{\frac{(y_i - \hat{\mu}_i)}{y_i \hat{\mu}_i}\right\}\) for the inverse Gaussian distribution.

The standardized Pearson residual is given by

\[
r_{i}^{\text{pea}} = \frac{(y_i - \hat{\mu}_i)}{\hat{\sigma}\sqrt{\hat{\mu}_i^2(1 - \hat{h}_{ii})}}.
\]

The William residual is also used in GLMs and is defined as

\[
r_{i}^{\text{wil}} = \text{sign}(y_i - \hat{\mu}_i)(1 - \hat{h}_{ii})(r_{i}^{\text{dev}})^2 + \hat{h}_{ii}(r_{i}^{\text{pea}})^2)^{1/2}.
\]
The last residual used in this paper is the standardized Anscombe residual given by
\[
r_{i}^{ans} = \frac{\psi(y_i) - \psi(\hat{\mu}_i)}{\hat{\sigma} \sqrt{\hat{V}(\hat{\mu}_i)/(1 - \hat{h}_{ii})}},
\]
(7)
where \(\psi(\mu_i) = 3\mu_i^{1/3}\) and \(V(\mu_i) = \mu^2\) for the gamma distribution and \(\psi(\mu_i) = \log(\mu_i)\) and \(V(\mu_i) = \mu^3\) for the inverse Gaussian distribution.

3. Simulation studies

In this section, we use Monte Carlo simulation studies to compare the properties of the residuals in the GLMs. Here, we use the gamma and the inverse Gaussian regression models. We considered seven scenarios for each distribution to study the residuals’ properties. In the first six scenarios, we assumed the model as
\[
\log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}.
\]
In the first four scenarios and in the last scenario for gamma and inverse Gaussian regression models, the covariates values were generated as independent draws from the standard uniform distribution. For the others scenarios, \(x_{i1}\) values were generated from the normal distribution, \((x_{i1} \sim \text{NO}(0.5, 0.25^2))\), \(x_{i2}\) values were generated from the inverse Gaussian distribution in scenarios V-a and V-b \((x_{i2} \sim \text{IG}(0.4, 2))\) and from the gamma distribution in scenarios VI-a and VI-b \((x_{i2} \sim \text{GA}(0.4, 1))\). The covariates values remained constant throughout the simulations.

In order to check if the results change considerably when other link function is used, in the last scenario we used the canonical link function. In Scenario VII-a, we used the gamma regression model with the inverse link function, in which
\[
\frac{1}{\mu_i} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}.
\]
In the last scenario (Scenario VII-b), we used the inverse Gaussian regression model with the link function \(1/\mu^2\), in which
\[
\frac{1}{\mu_i} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}.
\]
Tables 3 and 4 present the description of the scenarios for gamma and inverse Gaussian regression models. In the first scenario (Scenario I-a e I-b), we considered \(\beta_0 = 3, \beta_1 = 2\) and \(\beta_2 = 1\), which resulted in \(\mu \in (20.085; 403.429)\) and used \(\sigma = 0.1\) for gamma regression model and \(\sigma = 0.02\) for inverse Gaussian regression model. In the second and third scenarios, we used the same \(\beta_j, j = 0, 1, 2,\) of the Scenario 1, but with different \(\sigma\) values. In scenarios II-a and II-b we decreased the variance and in scenarios III-a and
III-b we increased the variance. In the following scenarios (scenarios IV-a and IV-b), we changed the values of the coefficients to result in mean response values close to zero. In scenarios V and VI, we changed the generating distribution of \( x_1 \) and \( x_2 \). In the last scenarios (scenarios VII-a and VII-b), we changed the link function. For all scenarios, two different sample sizes were considered: \( n = 15 \) and \( n = 50 \). All results are based on 5000 Monte Carlo replications. All simulations were performed using the software R. Results for \( r_i^{2ru} \) were omitted from the tables, because the simulation studies showed that their distributions are worse approximated by standard normal distribution than that of \( r_i^{2pu} \) in all scenarios.

Table 3: Description of the scenarios for gamma regression model.

| Scenario | Link function | Coefficients | Covariates | \( \mu \) | \( \sigma \) |
|----------|---------------|--------------|------------|----------|----------|
| I-a      | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.10 |
| II-a     | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.05 |
| III-a    | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.50 |
| IV-a     | log(\( \mu_i \)) | \( \beta_0 = -3, \beta_1 = 1.5, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (0.049; 0.606) | 0.10 |
| V-a      | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim N(O(0.5,0.25^2)); x_2 \sim IG(0.4,2) \) | (20.085; 403.429) | 0.10 |
| VI-a     | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim N(O(0.5,0.25^2)); x_2 \sim GA(0.4,1) \) | (7.400; 1892.000) | 0.10 |
| VII-a    | \( \frac{1}{\mu_i} \) | \( \beta_0 = 0.0025, \beta_1 = 0.04, \beta_2 = 0.01 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (19.04; 4000) | 0.10 |

Table 4: Description of the scenarios for inverse Gaussian regression model.

| Scenario | Link function | Coefficients | Covariates | \( \mu \) | \( \sigma \) |
|----------|---------------|--------------|------------|----------|----------|
| I-b      | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.02 |
| II-b     | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.01 |
| III-b    | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (20.085; 403.429) | 0.03 |
| IV-b     | log(\( \mu_i \)) | \( \beta_0 = -3, \beta_1 = 1.5, \beta_2 = 1 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (0.049; 0.606) | 0.50 |
| V-b      | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim N(O(0.5,0.25^2)); x_2 \sim IG(0.4,2) \) | (20.085; 403.429) | 0.02 |
| VI-b     | log(\( \mu_i \)) | \( \beta_0 = 3, \beta_1 = 2, \beta_2 = 1 \) | \( x_1 \sim N(O(0.5,0.25^2)); x_2 \sim GA(0.4,1) \) | (7.400; 1892.000) | 0.02 |
| VII-b    | \( \frac{1}{\mu_i} \) | \( \beta_0 = 0.000006, \beta_1 = 0.002, \beta_2 = 0.001 \) | \( x_1 \sim U(0,1); x_2 \sim U(0,1) \) | (18.24; 408.25) | 0.02 |

Tables 5 and 6 present the simulation results considering \( n = 15 \) in the Scenario I-a for gamma regression model and in the Scenario I-b for inverse Gaussian regression model, respectively. The means of \( r_i^{2ru} \) and \( r_i^{2pu} \) are very close to zero, but the other residuals have means slightly lower than zero. For all residuals, the variance is close to one, but none of the residuals have a variance close to one for all 15 observations. The skewness coefficient is considerably greater than zero for \( r_i^{2pu} \) specially for inverse Gaussian regression model, very close to zero for \( r_i^{2ul} \) and slightly lower than zero for the other residuals. The excess kurtosis coefficients are similar and not too close to zero for the five residuals. The residual \( r_i^{2ru} \) has smaller AD statistic than the other residuals in 10 out of 15 observations for gamma regression model and 14 out of 15 observations for inverse Gaussian regression model. The AD statistic is smaller for the adjusted quantile residual, than that for the other residuals, because the mean of \( r_i^{2ru} \) is closer to zero than that of \( r_i^{2uv} \), \( r_i^{2ul} \) and \( r_i^{2pu} \) and the skewness is closer to zero than that of \( r_i^{2pu} \).
Table 5: Simulation results for Scenario I-a - gamma regression model.

| i   | Mean   | Variance | Skewness | Kurtosis | A-D     |
|-----|--------|----------|----------|----------|---------|
|     | \(\mu\) | \(r_{j}^{\text{dev}}\) | \(r_{j}^{\text{up}}\) | \(r_{j}^{\text{low}}\) | \(r_{j}^{\text{dev}}\) | \(r_{j}^{\text{up}}\) | \(r_{j}^{\text{low}}\) | \(r_{j}^{\text{dev}}\) | \(r_{j}^{\text{up}}\) | \(r_{j}^{\text{low}}\) | \(r_{j}^{\text{dev}}\) | \(r_{j}^{\text{up}}\) | \(r_{j}^{\text{low}}\) | \(r_{j}^{\text{dev}}\) | \(r_{j}^{\text{up}}\) | \(r_{j}^{\text{low}}\) |
| 1   | 67.742 | -0.045   | -0.015   | -0.009   | -0.045   | -0.040   | -0.041   | -0.398   | -0.401   | -0.404   | 7.600   | 5.563   | 3.408   | 5.899   | 6.934   |
| 2   | 100.306| -0.011   | 0.018    | 0.024    | -0.011   | -0.012   | -0.015   | -0.452   | -0.464   | -0.463   | 2.751   | 3.512   | 4.191   | 2.751   | 2.596   |
| 3   | 215.703| -0.047   | -0.017   | -0.011   | -0.047   | -0.041   | -0.011   | 0.126    | -0.089   | -0.088   | 0.984   | 0.988   | 0.984   | 0.983   | 0.984   |
| 4   | 115.585| -0.032   | -0.001   | 0.003    | -0.032   | -0.028   | -0.006   | 0.016    | -0.454   | -0.455   | -0.454   | 8.819   | 6.580   | 3.924   | 8.810   | 8.071   |
| 5   | 131.633| -0.031   | -0.001   | 0.004    | -0.031   | -0.028   | -0.006   | 0.016    | -0.454   | -0.455   | -0.454   | 8.819   | 6.580   | 3.924   | 8.810   | 8.071   |
| 6   | 200.681| -0.012   | 0.017    | 0.024    | -0.012   | -0.006   | -0.006   | -0.446   | -0.444   | -0.447   | 5.349   | 5.357   | 5.349   | 5.349   | 5.349   |
| 7   | 33.844 | -0.020   | 0.007    | 0.019    | -0.020   | -0.012   | -0.006   | -0.336   | -0.335   | -0.337   | 1.384   | 1.884   | 3.543   | 1.386   | 1.777   |
| 8   | 80.061 | -0.025   | 0.005    | 0.011    | -0.025   | -0.020   | -0.006   | -0.372   | -0.371   | -0.373   | 2.631   | 2.484   | 1.759   | 2.629   | 2.325   |
| 9   | 48.232 | -0.022   | 0.004    | 0.018    | -0.022   | -0.013   | -0.006   | -0.343   | -0.343   | -0.344   | 3.077   | 2.835   | 3.506   | 3.076   | 2.618   |
| 10  | 52.030 | -0.029   | -0.002   | 0.009    | -0.029   | -0.022   | -0.006   | -0.374   | -0.374   | -0.375   | 3.213   | 1.865   | 2.006   | 3.212   | 2.501   |
| 11  | 48.729 | -0.039   | -0.009   | 0.004    | -0.039   | -0.034   | -0.011   | -0.372   | -0.371   | -0.373   | 2.631   | 2.484   | 1.759   | 2.629   | 2.325   |
| 12  | 114.403| -0.028   | 0.002    | 0.016    | -0.028   | -0.021   | -0.006   | -0.343   | -0.343   | -0.344   | 3.077   | 2.835   | 3.506   | 3.076   | 2.618   |
| 13  | 194.271| -0.046   | -0.015   | -0.010   | -0.046   | -0.404   | -0.012   | -0.343   | -0.343   | -0.344   | 6.159   | 5.694   | 6.488   | 6.153   | 5.574   |
| 14  | 103.731| -0.033   | -0.001   | 0.001    | -0.033   | -0.030   | -0.011   | -0.343   | -0.343   | -0.344   | 6.159   | 5.694   | 6.488   | 6.153   | 5.574   |
| 15  | 208.921| -0.017   | 0.013    | 0.021    | -0.017   | -0.010   | -0.006   | -0.343   | -0.343   | -0.344   | 6.159   | 5.694   | 6.488   | 6.153   | 5.574   |

Mean -0.029 0.000 0.007 -0.029 -0.023 1.004 1.001 1.004 1.004 1.004 -0.045 0.093 -0.045 -0.044 -0.018 -0.417 -0.418 -0.416 -0.418 -0.420 4.890 4.099 3.572 4.887 4.445

SD 0.012 0.011 0.012 0.012 0.012 0.023 0.023 0.023 0.023 0.023 0.031 0.032 0.031 0.031 0.031 0.037 0.032 0.037 0.037 0.036 2.333 1.818 1.294 2.333 2.236
| i  | µ  | Mean  | Variance  | Skewness  | Kurtosis  | A-D  |
|----|----|-------|-----------|-----------|-----------|------|
|    |    |       | r_{ij}^\text{dev} | r_{ij}^\text{var} | r_{ij}^\text{quasi} | r_{ij}^\text{ans} | r_{ij}^\text{wili} | r_{ij}^\text{dev} | r_{ij}^\text{var} | r_{ij}^\text{quasi} | r_{ij}^\text{ans} | r_{ij}^\text{wili} | r_{ij}^\text{dev} | r_{ij}^\text{var} | r_{ij}^\text{quasi} | r_{ij}^\text{ans} | r_{ij}^\text{wili} | r_{ij}^\text{dev} | r_{ij}^\text{var} | r_{ij}^\text{quasi} | r_{ij}^\text{ans} | r_{ij}^\text{wili} |
| 1  |  67.742 |  -0.080 | -0.006 | 0.008 | -0.079 | -0.067 |
| 2  |  100.306 |  -0.096 | -0.008 | 0.007 | -0.096 | -0.084 |
| 3  |  215.703 |  -0.107 | 0.029 | 0.042 | -0.105 | -0.089 |
| 4  |  115.585 |  -0.089 | 0.013 | 0.020 | -0.088 | -0.077 |
| 5  |  131.633 |  -0.080 | 0.024 | 0.036 | -0.079 | -0.069 |
| 6  |  200.681 |  -0.122 | 0.005 | 0.021 | -0.121 | -0.106 |
| 7  |  33.844 |  -0.004 | 0.037 | 0.074 | -0.004 | 0.015 |
| 8  |  80.061 |  -0.079 | 0.001 | 0.014 | -0.079 | -0.067 |
| 9  |  48.232 |  -0.080 | 0.030 | 0.010 | -0.080 | -0.058 |
| 10 |  52.030 |  -0.086 | -0.028 | -0.001 | -0.085 | -0.067 |
| 11 |  48.729 |  -0.060 | 0.002 | 0.013 | -0.060 | -0.050 |
| 12 |  114.403 |  -0.106 | -0.017 | 0.016 | -0.105 | -0.080 |
| 13 |  194.271 |  -0.144 | -0.019 | 0.003 | -0.142 | -0.128 |
| 14 |  103.731 |  -0.083 | 0.014 | 0.019 | -0.083 | -0.075 |
| 15 |  208.921 |  -0.147 | -0.018 | 0.001 | -0.145 | -0.128 |

Mean: -0.091 | 0.000 | 0.018 | -0.090 | -0.075 | 1.016 | 0.999 | 1.026 | 1.007 | 1.013 | -0.119 | 0.302 | -0.120 | -0.113 | -0.047 | -0.367 | -0.338 | -0.359 | -0.389 | -0.388 | 22.735 | 15.892 | 7.001 | 22.515 | 18.834

SD: 0.034 | 0.021 | 0.020 | 0.034 | 0.034 | 0.017 | 0.022 | 0.020 | 0.015 | 0.018 | 0.039 | 0.125 | 0.040 | 0.037 | 0.031 | 0.047 | 0.072 | 0.050 | 0.040 | 0.044 | 11.898 | 7.313 | 5.407 | 11.722 | 10.108
Table 7 presents the average values for the mean, variance, skewness and excess kurtosis for all residuals and both sample sizes for gamma regression model. Changes in the range of $\mu$, in the distribution of the covariates and in the link function does not seem to affect considerably the distribution of the residuals, since the results of Scenarios IV-a to VII-a are very similar to that of Scenario I-a. However, the distributions of the five residuals are worse approximated by the standard normal distribution when variance increases. The skewness coefficients are far from zero in Scenario III-a than in Scenario I-a and, except for $r_i^{pea}$, the mean and variance are also far from zero and one, respectively. As a consequence, the AD statistics increase substantially from Scenario I-a to Scenario III-a, but the rise is smaller for $r_i^{*qu}$.

In general, when sample size increases from $n = 15$ to $n = 50$, all residuals’s distributions are better approximated by the standard normal distribution. In all scenarios, the absolute value of the average skewness coefficient reduces considerably, except for the residual $r_i^{pea}$, in which the average skewness coefficient rises slightly. Additionally, the absolute average excess kurtosis coefficient also decreases in all scenarios for all residuals, except for $r_i^{pea}$ in Scenario III-a. Finally, in six out of seven scenarios, the AD statistic reduces for all residuals, but the reduction is greater for $r_i^{*qu}$.

Table 8 summarizes the results of the AD statistic in all scenarios for $n = 15$ and $n = 50$ considering the gamma regression model. When $n = 50$ the mean, the standard deviation and the maximum of the AD statistic are lower for $r_i^{*qu}$ than for other residuals. When sample size is 15, the sample mean of the AD statistic is also smaller for $r_i^{*qu}$ than for other residuals, except in Scenario IV-a in which $r_i^{pea}$ and $r_i^{*qu}$ have similar sample mean of the AD statistic. On the other hand, the standard deviation and the maximum are the lowest for $r_i^{*qu}$ in some scenarios and for $r_i^{pea}$ in others.

Tables 9 and 10 present the average of the sample moments for the residuals and a descriptive summary of the results of the AD statistic for the inverse Gaussian regression model. For all residuals, in the most of the scenarios, the average values of the AD statistic are higher for the inverse Gaussian regression model than for the gamma regression model. A plausible explanation for this result is that, for the same mean and variance, the inverse Gaussian distribution presents a greater skewness than the gamma distribution. In general, the standard deviation of AD statistic are also higher for the inverse Gaussian regression model. Residual $r_i^{*qu}$ presents lower mean of the AD statistic than the other residuals for all scenarios and both sample sizes. Additionally, the differences between the mean AD statistic for $r_i^{*qu}$ and for the other residuals are greater for the inverse Gaussian regression model. Finally, except for Scenario V-a when $n = 15$, the standard deviation of the AD statistic is lower for $r_i^{*qu}$ than for the other residuals.
Table 7: Distributional measures for the residual in Gamma regression model.

| Scenarios | $\hat{\mu}_{\hat{y}}$ | $\hat{\mu}_{\hat{y}}$ | $\hat{\mu}_{\hat{y}}$ | $\hat{\mu}_{\hat{y}}$ | $\hat{\mu}_{\hat{y}}$ |
|-----------|------------------------|------------------------|------------------------|------------------------|------------------------|
|           | n=15 n=50 | n=15 n=50 | n=15 n=50 | n=15 n=50 | n=15 n=50 |
| Mean      | -0.03 -0.03 | 0.00 0.00 | 0.01 0.00 | -0.03 -0.03 | -0.02 -0.03 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| I-a       | Skewness -0.04 -0.02 | 0.09 0.16 | -0.04 -0.02 | -0.04 -0.02 | -0.02 -0.01 |
|           | Kurtosis -0.42 -0.11 | -0.42 -0.08 | -0.42 -0.11 | -0.42 -0.11 | -0.42 -0.11 |
|           | AD Statistics 4.89 3.43 | 4.10 2.84 | 3.57 1.13 | 4.89 3.43 | 4.45 3.26 |
| Mean      | -0.01 -0.02 | 0.00 0.00 | 0.00 0.00 | -0.01 -0.02 | -0.01 -0.02 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| II-a      | Skewness -0.01 -0.01 | 0.06 0.09 | -0.01 -0.01 | -0.01 -0.01 | 0.00 0.00 |
|           | Kurtosis -0.42 -0.13 | -0.42 -0.12 | -0.42 -0.13 | -0.42 -0.13 | -0.43 -0.13 |
|           | AD Statistics 4.05 1.64 | 3.94 1.52 | 3.62 1.03 | 4.05 1.64 | 3.94 1.60 |
| Mean      | -0.15 -0.16 | 0.00 0.00 | 0.03 0.01 | -0.15 -0.16 | -0.13 -0.16 |
| Variance  | 1.08 1.04 | 1.00 1.00 | 1.07 1.03 | 1.07 1.03 | 1.06 1.04 |
| III-a     | Skewness -0.22 -0.09 | 0.46 0.79 | -0.26 -0.10 | -0.22 -0.07 | -0.09 -0.04 |
|           | Kurtosis -0.24 -0.10 | -0.30 0.62 | -0.02 -0.09 | -0.22 -0.16 | -0.32 -0.10 |
|           | AD Statistics 57.42 62.34 | 35.35 45.69 | 14.41 2.36 | 56.75 61.41 | 44.25 57.93 |
| Mean      | -0.03 -0.03 | 0.00 0.00 | 0.01 0.00 | -0.03 -0.03 | -0.02 -0.03 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| IV-a      | Skewness -0.06 -0.01 | 0.08 0.17 | -0.06 -0.01 | -0.06 -0.01 | -0.03 0.00 |
|           | Kurtosis -0.41 -0.12 | -0.42 -0.09 | -0.41 -0.12 | -0.41 -0.12 | -0.42 -0.12 |
|           | AD Statistics 4.97 3.57 | 4.00 3.08 | 4.01 1.17 | 4.97 3.57 | 4.53 3.41 |
| Mean      | -0.03 -0.03 | 0.00 0.00 | 0.01 0.00 | -0.03 -0.03 | -0.02 -0.03 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| V-a       | Skewness -0.03 -0.02 | 0.11 0.16 | -0.03 -0.02 | -0.03 -0.02 | 0.00 -0.01 |
|           | Kurtosis -0.43 -0.12 | -0.42 -0.09 | -0.43 -0.12 | -0.43 -0.13 | -0.43 -0.12 |
|           | AD Statistics 5.44 3.60 | 4.83 2.94 | 3.97 1.48 | 5.43 3.60 | 5.04 3.44 |
| Mean      | -0.03 -0.03 | 0.00 0.00 | 0.01 0.00 | -0.03 -0.03 | -0.02 -0.03 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| VI-a      | Skewness -0.04 -0.02 | 0.10 0.16 | -0.04 -0.02 | -0.04 -0.02 | -0.01 -0.01 |
|           | Kurtosis -0.41 -0.12 | -0.41 -0.09 | -0.41 -0.12 | -0.41 -0.12 | -0.41 -0.12 |
|           | AD Statistics 4.83 3.50 | 4.22 2.92 | 3.18 1.15 | 4.82 3.50 | 4.34 3.32 |
| Mean      | -0.03 -0.03 | 0.00 0.00 | 0.01 0.00 | -0.03 -0.03 | -0.02 -0.03 |
| Variance  | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 | 1.00 1.00 |
| VII-a     | Skewness -0.03 -0.02 | 0.10 0.16 | -0.03 -0.02 | -0.03 -0.02 | -0.01 -0.01 |
|           | Kurtosis -0.40 -0.11 | -0.40 -0.08 | -0.40 -0.11 | -0.41 -0.11 | -0.41 -0.11 |
|           | AD Statistics 5.32 3.35 | 4.64 2.79 | 3.82 1.01 | 5.31 3.35 | 4.87 3.18 |
Table 8: Comparison of the Anderson–Darling statistic for \( n = 15 \) and \( n = 50 \) - gamma regression model.

| Scenario | sigma | n  | Resíduos | Mean Deviation | Minimum | Q1 | Q2 | Q3 | Maximum |
|----------|-------|----|----------|----------------|---------|----|----|----|---------|
| I-a 0.1 15 | 1.38 | 6.52 | 8.82 |
| I-a 0.5 50 | 1.92 | 4.40 | 6.84 |
| II-a 0.1 15 | 1.40 | 4.64 | 9.14 |
| II-a 0.5 50 | 1.75 | 4.25 | 8.09 |
| III-a 0.05 15 | 1.14 | 3.76 | 4.31 |
| III-a 0.05 50 | 1.26 | 3.78 | 4.57 |
| IV-a 0.5 15 | 1.26 | 3.78 | 4.57 |
| IV-a 0.5 50 | 1.26 | 3.78 | 4.57 |

| Scenario | sigma | n  | Resíduos | Mean Deviation | Minimum | Q1 | Q2 | Q3 | Maximum |
|----------|-------|----|----------|----------------|---------|----|----|----|---------|
| V-a 0.1 15 | 1.84 | 3.82 | 4.29 |
| V-a 0.5 50 | 1.32 | 2.46 | 4.62 |
| VI-a 0.1 15 | 1.52 | 3.94 | 7.52 |
| VI-a 0.5 50 | 1.30 | 2.03 | 4.82 |
| VII-a 0.1 15 | 1.40 | 3.76 | 4.31 |
| VII-a 0.1 50 | 1.26 | 3.78 | 4.57 |
| VIII-a 0.1 50 | 1.26 | 3.78 | 4.57 |

| Scenario | sigma | n  | Resíduos | Mean Deviation | Minimum | Q1 | Q2 | Q3 | Maximum |
|----------|-------|----|----------|----------------|---------|----|----|----|---------|
| IX-a 0.1 15 | 1.38 | 6.52 | 8.82 |
| IX-a 0.5 50 | 1.92 | 4.40 | 6.84 |
| X-a 0.1 15 | 1.40 | 4.64 | 9.14 |
| X-a 0.5 50 | 1.75 | 4.25 | 8.09 |
| XI-a 0.05 15 | 1.14 | 3.76 | 4.31 |
| XI-a 0.05 50 | 1.26 | 3.78 | 4.57 |
| XII-a 0.5 15 | 1.26 | 3.78 | 4.57 |
| XII-a 0.5 50 | 1.26 | 3.78 | 4.57 |
Table 9: Distributional measures for the residuals in inverse Gaussian regression model.

| Scenarios | $r_{n,15}^{inv}$ n=15 | $r_{n,50}^{inv}$ n=50 | $r_{n,15}^{mnm}$ n=15 | $r_{n,50}^{mnm}$ n=50 | $r_{n,15}^{int}$ n=15 | $r_{n,50}^{int}$ n=50 |
|-----------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| Mean      | -0.09 -0.10            | 0.00 0.00              | -0.09 -0.10            | -0.08 -0.09            | -0.09 -0.10            | -0.08 -0.09            |
| Variance  | 1.02 1.00              | 1.00 1.00              | 1.01 1.01              | 1.01 1.01              | 1.01 1.01              | 1.01 1.01              |
| I-b       | Skewness -0.12 -0.06   | 0.30 0.48              | -0.12 -0.06            | -0.11 -0.05            | -0.05 -0.03            | -0.05 -0.03            |
|           | Kurtosis -0.37 -0.12   | -0.34 0.24             | -0.36 -0.12            | -0.39 -0.16            | -0.39 -0.12            | -0.39 -0.12            |
| AD Statistics 22.73 25.25 | 15.89 18.94           | 7.00 1.68              | 22.51 25.06            | 18.83 23.80            |
| Mean      | -0.05 -0.05            | 0.00 0.00              | 0.01 0.00              | -0.05 -0.05            | -0.04 -0.05            | -0.04 -0.05            |
| Variance  | 1.00 1.00              | 1.00 1.00              | 1.01 1.01              | 1.01 1.01              | 1.01 1.01              | 1.01 1.01              |
| II-b      | Skewness -0.06 -0.02   | 0.15 0.25              | -0.06 -0.02            | -0.06 -0.02            | -0.02 -0.01            | -0.02 -0.01            |
|           | Kurtosis -0.41 -0.12   | -0.41 -0.02            | -0.41 -0.12            | -0.42 -0.13            | -0.42 -0.12            | -0.42 -0.12            |
| AD Statistics 8.39 7.28 | 6.87 6.03             | 4.33 1.22              | 8.37 7.27              | 7.41 6.93              |
| Mean      | -0.14 -0.14            | 0.00 0.00              | 0.02 0.01              | -0.13 -0.14            | -0.11 -0.13            | -0.11 -0.13            |
| Variance  | 1.04 1.00              | 1.00 1.00              | 1.06 1.02              | 1.02 0.98              | 1.04 1.00              | 1.04 1.00              |
| AD Statistics 46.04 52.09 | 28.56 40.32           | 11.18 1.79             | 45.09 51.41            | 37.10 49.06            |
| Mean      | -0.05 -0.05            | 0.00 0.00              | 0.01 0.00              | -0.09 -0.10            | -0.08 -0.10            | -0.08 -0.10            |
| Variance  | 1.00 1.00              | 1.00 1.00              | 1.03 1.01              | 1.01 0.99              | 1.02 1.00              | 1.02 1.00              |
| IV-b      | Skewness -0.06 -0.02   | 0.30 0.53              | -0.13 -0.06            | -0.12 -0.05            | -0.05 -0.02            | -0.05 -0.02            |
|           | Kurtosis -0.41 -0.12   | -0.33 0.28             | -0.36 -0.09            | -0.39 -0.14            | -0.39 -0.10            | -0.39 -0.10            |
| AD Statistics 8.39 7.28 | 14.86 21.07           | 6.81 1.46              | 22.13 27.13            | 18.50 25.66            |
| Mean      | -0.01 -0.09            | 0.00 0.00              | 0.00 0.00              | -0.01 -0.09            | -0.01 -0.08            | -0.01 -0.08            |
| Variance  | 1.00 1.00              | 1.00 1.00              | 1.00 1.01              | 1.00 0.99              | 1.00 1.00              | 1.00 1.00              |
| AD Statistics 3.96 21.36 | 3.95 16.09            | 3.87 1.52              | 3.96 21.20             | 3.95 19.84            |
| Mean      | -0.08 -0.09            | 0.00 0.00              | 0.02 0.00              | -0.08 -0.09            | -0.07 -0.09            | -0.07 -0.09            |
| Variance  | 1.02 1.00              | 1.00 1.00              | 1.03 1.01              | 1.01 0.99              | 1.01 1.00              | 1.01 1.00              |
| VI-b      | Skewness -0.12 -0.05   | 0.26 0.47              | -0.12 -0.05            | -0.11 -0.05            | -0.04 -0.02            | -0.04 -0.02            |
|           | Kurtosis -0.39 -0.11   | -0.38 0.21             | -0.38 -0.10            | -0.41 -0.14            | -0.41 -0.10            | -0.41 -0.10            |
| AD Statistics 19.25 22.54 | 11.62 17.13           | 9.78 1.55              | 19.04 22.38            | 14.66 21.04            |
| Mean      | -0.05 -0.05            | -0.01 0.00             | 0.00 0.00              | -0.05 -0.05            | -0.05 -0.05            | -0.05 -0.05            |
| Variance  | 1.00 1.00              | 0.99 1.00              | 1.00 1.00              | 1.00 1.00              | 1.00 1.00              | 1.00 1.00              |
| VII-b     | Skewness -0.04 -0.03   | 0.15 0.26              | -0.04 -0.03            | -0.04 -0.03            | -0.01 -0.01            | -0.01 -0.01            |
|           | Kurtosis -0.41 -0.11   | -0.41 -0.02            | -0.41 -0.10            | -0.42 -0.11            | -0.42 -0.11            | -0.42 -0.11            |
| AD Statistics 8.57 6.98 | 6.44 5.56             | 2.52 1.26              | 8.56 6.97              | 7.73 6.42              |
| Scenario | sigma | n | Red | Mean | Deviation | Minimum | Q1 | Q2 | Q3 | Maximum |
|----------|-------|---|-----|-----|-----------|---------|----|----|----|---------|
| I-b      | 0.02  | 15 |     |     |           | 8.38    | 22.51| 4.91| 17.20| 26.18   |
|          |       |    |     |     |           | 11.99   | 14.43| 3.84| 13.92| 16.54   |
|          |       |    |     |     |           | 1.71    | 1.72 | 1.72| 1.72 | 1.72    |
| II-b     | 0.02  | 15 |     |     |           | 10.69   | 15.52| 15.52| 15.52| 15.52   |
|          |       |    |     |     |           | 11.69   | 17.12| 17.12| 17.12| 17.12   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| III-b    | 0.01  | 15 |     |     |           | 10.69   | 15.52| 15.52| 15.52| 15.52   |
|          |       |    |     |     |           | 11.69   | 17.12| 17.12| 17.12| 17.12   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| IV-b     | 0.03  | 15 |     |     |           | 10.69   | 19.65| 19.65| 19.65| 19.65   |
|          |       |    |     |     |           | 11.69   | 21.36| 21.36| 21.36| 21.36   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| V-b      | 0.02  | 50 |     |     |           | 10.69   | 19.65| 19.65| 19.65| 19.65   |
|          |       |    |     |     |           | 11.69   | 21.36| 21.36| 21.36| 21.36   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| VI-b     | 0.02  | 50 |     |     |           | 10.69   | 19.65| 19.65| 19.65| 19.65   |
|          |       |    |     |     |           | 11.69   | 21.36| 21.36| 21.36| 21.36   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| III-b    | 0.01  | 50 |     |     |           | 10.69   | 19.65| 19.65| 19.65| 19.65   |
|          |       |    |     |     |           | 11.69   | 21.36| 21.36| 21.36| 21.36   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
| IV-b     | 0.03  | 50 |     |     |           | 10.69   | 19.65| 19.65| 19.65| 19.65   |
|          |       |    |     |     |           | 11.69   | 21.36| 21.36| 21.36| 21.36   |
|          |       |    |     |     |           | 1.72    | 1.72 | 1.72 | 1.72 | 1.72    |
4. Application

In Section 3, we studied whether the distribution of the residuals considered in this paper are well approximated by the standard normal distribution. Other essential properties for residuals are their ability to identify model misspecification and outliers. We used two real datasets to study the residuals considered in this paper according to these aspects. In this section, we compare the $r_{i}^{*qu}$ with the two most commonly used residual in the GLMs, the $r_{i}^{dev}$ and the $r_{i}^{pea}$. Applications were performed using the gamlss package and the glm function of the R software.

4.1. Oil dataset

The first application uses a database with 1000 observations about the daily prices of WTI (West Texas Intermediate) oil price traded by NYMEX (New York Mercantile Exchange). We used this dataset to investigate whether the $r_{i}^{dev}$, $r_{i}^{pea}$ and $r_{i}^{*qu}$ are able to identify model misspecification. The response variable is the price of oil. Here, we use two explanatory variables: the lagged version of the response variable and the log price of front month heating oil contract traded by NYMEX, as proposed by Stasinopoulos et al. (2017). We fitted inverse Gaussian regression model with logit link function for the oil data and obtained the three sets of residuals. Stasinopoulos et al. (2017) proposed the Sinh-Arcsinh (SHASH) distribution and splines to obtain a reasonable fit for this response variable. Therefore, lack of fit is expected when we fit the inverse Gaussian regression model. Figure 1 presents a plot of residuals versus linear predictor and a normal residual plot with simulated envelope (Atkinson, 1985) for $r_{i}^{dev}$, $r_{i}^{pea}$ and $r_{i}^{*qu}$. Clearly, the plots for the three residuals suggest lack of fit for the inverse Gaussian regression model, indicating that these residuals are able to identify this kind of problem. A similar analysis was performed using the gamma regression model. The conclusions were the same and the plots were omitted for the sake of brevity.
Figure 1: Residuals against linear predictor (left) and normal residual plots with simulated envelope (right) for oil dataset.

4.2. Turbine dataset

The second real dataset consists of the time to evaluate the performance of five types of high-speed turbines for plane engines (Lawless, 2011). The response variable is the time (in million of cycles units) until the loss of velocity and the covariate is the type of turbine. Our focus with this application is to check if $r_{dev}^i$, $r_{pea}^i$, $r_{qua}^i$ are good residuals to perform outlier identification. We fitted gamma regression model with identity link function to this dataset and obtained the three set of residuals. Figure 2 presents a plot of residuals against linear predictors and a normal residual plot with simulated envelope for $r_{dev}^i$, $r_{pea}^i$ and $r_{qua}^i$. None of the residuals suggest model misspecification.
For the three residuals, observation 40 has the highest value and observation 1 has the smallest. However, the value of the residual for observation 40 is considerably greater for $r_{i \text{pea}}$ than for the other residuals and the absolute value of the residual for observation 1 is substantially smaller for $r_{i \text{pea}}$ than for $r_{i \text{dev}}$ and $r_{i \text{qu}}$. In the simulation studies, we noted that the distribution of $r_{i \text{pea}}$ is right skewed and this explains the mentioned result. As a consequence, $r_{i \text{pea}}$ may incorrectly identify as an outlier an observation with positive value for
the residual and may not identify an outlier with value of \( r_{pi}^{pea} \) negative. On the other hand, the values of \( r_{pi}^{dev} \) and \( r_{pi}^{qua} \) are close for all observations. However, the values of the positive residual are slightly greater for \( r_{pi}^{qua} \) than for \( r_{pi}^{dev} \) and the absolute values of the negative residuals are slightly smaller for \( r_{pi}^{qua} \) than for \( r_{pi}^{dev} \). As \( r_{pi}^{dev} \) has mean far from zero than \( r_{pi}^{qua} \) and similar variance, skewness and excess kurtosis, the latter seems to be a better measure of the discrepancy of an observation and, consequently, better to perform outlier identification.

5. Conclusions

In this work, we introduced the adjusted quantile residual to perform diagnostic analysis in generalized linear models. We compared the distribution of the adjusted quantile residual with four other residuals using Monte Carlo simulation studies. Additionally, we used two applications to investigate if the residuals are able to identify model misspecification and outliers.

It is very desirable to find a residual whose distribution is well approximated by the standard normal distribution. Our simulation studies suggest that the adjusted quantile residual’s distribution better approximates to the standard normal distribution in all scenarios than the other residual’s distribution, specially when sample size is not small. Simulation results also suggest that, when variance increases, all residuals worsen regarding normal approximation. However, when variance is high, the distribution of the quantile residual is much better approximated by the standard normal distribution than that of the other residuals, even in small samples.

The applications investigated other properties of the residuals. The first application suggested that the three residuals considered in the analysis can detect lack of fit in generalized linear models. Based on the second application, the adjusted quantile residual seems to be the best to perform outlier identification.

The standardized Pearson residual and the standardized deviance residual are calculated by many packages and statistical softwares in their generalized linear models routine, but commonly the quantile residual is not implemented. However, adjusted quantile residual is simple and easy to calculate using any statistical package. Considering the results of the simulations studies, the applications and its simplicity, the adjusted quantile residual is a better choice of to perform diagnostic analysis in generalized linear models than the competing residuals.

6. Acknowledgements

The authors thank "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior" (CAPES) for the financial support received for this project.
7. References

References

Anderson, T. W. and Darling, D. A. (1952), ‘Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes’, *The annals of mathematical statistics* 23, 193–212.

Anholeto, T., Sandoval, M. C. and Botter, D. A. (2014), ‘Adjusted pearson residuals in beta regression models’, *Journal of Statistical Computation and Simulation* 84, 999–1014.

Atkinson, A. C. (1985), *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis*, Oxford.

Cordeiro, G. M. (2004), ‘On pearson’s residuals in generalized linear models’, *Statistics & probability letters* 66(3), 213–219.

Davison, A. and Gigli, A. (1989), ‘Deviance residuals and normal scores plots’, *Biometrika* 76(2), 211–221.

Dunn, P. K. and Smyth, G. K. (1996), ‘Randomized quantile residuals’, *Journal of Computational and Graphical Statistics* 5(3), 236–244.

Feng, C., Sadeghpour, A. and Li, L. (2017), ‘Randomized quantile residuals: an omnibus model diagnostic tool with unified reference distribution’, arXiv preprint arXiv:1708.08527.

Hardin, J. W., Hilbe, J. M. and Hilbe, J. (2007), *Generalized linear models and extensions*, Stata press.

Klar, B. and Meintanis, S. G. (2012), ‘Specification tests for the response distribution in generalized linear models’, *Computational Statistics* 27(2), 251–267.

Lawless, J. F. (2011), *Statistical models and methods for lifetime data*, Vol. 362, John Wiley & Sons.

McCullagh, P. and Nelder, J. A. (1989), *Generalized linear models*, 2 edn, CRC press.

Nelder, J. A. and Wedderburn, R. W. M. (1972), ‘Generalized linear models’, *Journal of the Royal Statistical Society* 10(3), 370–384.

Pereira, G. H. A. (2017), ‘On quantile residuals in beta regression’, *Communications in Statistics-Simulation and Computation*, in press.

Pierce, D. A. and Schafer, D. W. (1986), ‘Residuals in generalized linear models’, *Journal of the American Statistical Association* 81(396), 977–986.

Rigby, R. A. and 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.

Stasinopoulos, D. M., Rigby, R. A., Voudouris, V., Heller, G. and De Bastiani, F. (2017), *Flexible Regression and Smoothing: Using GAMLSS in R*, Taylor & Francis.

Stasinopoulos, D., Rigby, B. and Akantziliotou, C. (2014), ‘Instructions on how to use the gamlss package in r, 2008’.

Urbano, M. R., Borges Demétrio, C. G. and Cordeiro, G. M. (2012), ‘On wald residuals in generalized linear models’, *Communications in Statistics-Theory and Methods* 41(4), 741–758.

Williams, D. (1987), ‘Generalized linear model diagnostics using the deviance and single case deletions’, *Applied Statistics* pp. 181–191.