Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Programa de Pós-Graduação em Estatística

JOAS SILVA DOS SANTOS

CORREÇÃO TIPO-BARTLETT À ESTATÍSTICA GRADIENTE NOS MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS

Recife

2019

JOAS SILVA DOS SANTOS

CORREÇÃO TIPO-BARTLETT À ESTATÍSTICA GRADIENTE NOS MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de mestre em Estatística.

Área de Concentração: Ciência Exatas e da Terra

Orientadora: Profª. Drª. Audrey Helen Mariz de Aquino Cysneiros

Recife

2019

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

S237c Santos, Joas Silva dos
Correção tipo-Bartlett à estatística gradiente nos modelos lineares generalizados superdispersados / Joas Silva dos Santos. – 2019.
53 f.: il., fig., tab.

Orientadora: Audrey Helen Mariz de Aquino Cysneiros.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2019.
Inclui referências e apêndices.

1. Estatística. 2. Teste gradiente. I. Cysneiros, Audrey Helen Mariz de Aquino (orientadora). II. Título.

310 CDD (23. ed.) UFPE- MEI 2019-132

JOAS SILVA DOS SANTOS

CORREÇÃO TIPO-BARTLETT À ESTATÍSTICA GRADIENTE NO MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS

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

Aprovada em: 31 de julho de 2019.

BANCA EXAMINADORA

Prof.(ª) AUDREY HELEN MARIZ DE AQUINO CYSNEIROS
UFPE

Prof.(º) Vinícius Quintas Souto Maior
UFPE

Prof.(ª) Mariana Correia de Araújo
UFRN

À minha mãe, Valmira Silva.

AGRADECIMENTOS

Primeiramente, agradeço a Deus, pela sua bondade e fidelidade.

À minha mãe, Valmira Silva, por seu amor e cuidado.

À Professora Audrey Helen Mariz de Aquino Cysneiros, pela orientação e disponibilidade.

Aos professores do Programa de Mestrado em Estatística da UFPE.

Aos participantes da banca examinadora, pelas sugestões.

Aos professores do Departamento de Estatística da UFS, em especial, aos Professores José Rodrigo Santos Silva, Sadraque Eneas de Figueiredo Lucena e Luiz Henrique Gama Dore de Araujo.

À CAPES, pelo apoio financeiro.

RESUMO

Os modelos lineares generalizados superdispersados (MLGS), propostos por Dey et al. (1997), permitem que tanto a média quanto a dispersão sejam modeladas simultaneamente no contexto dos modelos lineares generalizados. Os MLGS são muito úteis para modelar a dispersão quando a variância da variável resposta excede a variância nominal predita pelo modelo. Esta dissertação tem três objetivos. O primeiro, é reunir resultados importantes sobre correções de Bartlett e tipo-Bartlett para os testes da razão de verossimilhanças e escore nos MLGS, propostos na literatura. O segundo, é a obtenção de um fator de correção tipo-Bartlett, em notação matricial, à estatística gradiente para testar simultaneamente ou separadamente os efeitos da média e da dispersão. A estatística gradiente corrigida tem distribuição qui-quadrado até um erro de ordem n^{-1} sob a hipótese nula. O terceiro, é apresentar resultados de simulação para averiguar o efeito das correções nos MLGS, no que tange ao tamanho e poder, em amostras finitas.

Palavras-chave: Correção tipo-Bartlett. Distribuição Qui-Quadrado. Teste gradiente.

ABSTRACT

The overdispersed generalized linear models (OGLMs) allow, in general, that the mean and the dispersion to be modeled simultaneously in the context of generalized linear models. The OGLMs are very useful for modeling the dispersion when the variance of the response variable exceeds the nominal variance predicted by the model. This dissertation has three purposes. The first purpose is to describe important results on Bartlett and Bartlett-type corrections to the likelihood ratio and score statistics in OGLMs, proposed in the literature. The second purpose is to obtain a Bartlett-type correction factor to the gradient statistic, in matrix notation, for simultaneously and separately testing the effect of the mean and the dispersion. The corrected statistic gradient has a chi-square distribution up to an error of order n^{-1} under the null hypothesis. The third purpose is to present simulation results in order to evaluate the effect of the corrections in OGLMs, in terms of size and power, in finite-sample.

Keywords: Bartlett-type correction. Chi-squared distribution. Gradient test.

LISTA DE GRÁFICOS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

Os modelos lineares generalizados (MLG), propostos por Nelder e Wedderburn (1972), são amplamente utilizados em diversas situações práticas. Alguns exemplos podem ser vistos em Lindsey (1997). Em certas aplicações, no entanto, verifica-se que a variância da variável resposta excede a variância nominal predita pelo modelo. Este fenômeno é descrito como superdispersão e tem sido largamente considerado na literatura, onde algumas alternativas para o problema vêm sendo discutidas e propostas. Uma forma de lidar com a superdispersão é controlar a variabilidade modelando também a dispersão de forma independente da média. Este tipo de abordagem é visto nos estudos de Efron (1986), Smyth (1989), Ganio e Schafer (1992) e Dey et al. (1997).

Nesta dissertação, consideramos os modelos lineares generalizados superdispersados (MLGS) definidos por Dey et al. (1997). Esta classe de modelos permite que a média e a dispersão sejam modeladas simultaneamente no contexto dos MLG. Além disso, os MLGS são úteis para modelar a dispersão quando a variância da variável resposta excede a variância nominal predita pelo MLG. Vale ressaltar que os modelos lineares generalizados superdispersados são uma generalização da família exponencial dupla introduzida por Efron (1986).

Os testes de hipóteses mais utilizados em grandes amostras são: razão de verossimilhanças (Wilks, 1938), escore (Rao, 1948) e Wald (Wald, 1943). Recentemente, Terrell (2002) propôs o teste gradiente cuja estatística é calculada de forma simples, pois não envolve, por exemplo, operações matriciais como produto e inversão de matrizes. Além de não depender da matriz de informação esperada, nem da observada. As estatísticas desses testes são equivalentes em grandes amostras e convergem, sob a hipótese nula $(\mathcal{H}_0)$, para a distribuição χ^2_q, em que q é o número de restrições impostas por $\mathcal{H}_0$. Em pequenas amostras, no entanto, essa aproximação pela distribuição qui-quadrado pode não ser satisfatória, conduzindo a taxas de rejeição consideravelmente distorcidas.

Neste sentido, Bartlett (1937) apresentou a primeira proposta de melhoria de testes estatísticos. Ele considerou apenas a estatística da razão de verossimilhanças, calculando seu valor esperado segundo $\mathcal{H}_0$ até ordem n^{-1}, sendo n o tamanho amostral. Esse resultado foi então utilizado por Bartlett para produzir um fator de correção conhecido como *correção de Bartlett*, de modo que a estatística da razão de verossimilhanças corrigida por esse fator tem distribuição, segundo $\mathcal{H}_0$, mais próxima da distribuição qui-quadrado de referência. Posteriormente, a ideia

de Bartlett foi generalizada por Lawley (1956) que apresentou um método geral para obtenção de correções de Bartlett em modelos estatísticos regulares. Cordeiro (1983) derivou uma expressão geral para o fator de correção de Bartlett nos MLG com parâmetro de dispersão conhecido. Esse resultado é generalizado pelo próprio autor, considerando o parâmetro de dispersão desconhecido (Cordeiro, 1987).

Na mesma linha, Cordeiro e Ferrari (1991) apresentaram uma estatística escore aperfeiçoada com distribuição qui-quadrado até $O(n^{-1})$. Mais especificamente, a estatística escore usual é multiplicada por um fator de correção expresso como um polinômio de segundo grau na própria estatística, resultando numa estatística escore modificada cuja distribuição nula é aproximadamente χ^2 até ordem n^{-1}. Esse fator multiplicativo é denominado *correção tipo-Bartlett*. Os resultados obtidos pelos autores são bastante gerais, visto que permitem a presença de parâmetros de perturbação. O mesmo vale para o fator de correção de Bartlett. Cordeiro et al. (1993) e Cribari-Neto e Ferrari (1995) obtiveram correções tipo-Bartlett para o teste escore em MLG com parâmetro de dispersão conhecido e desconhecido, respectivamente. Para maiores detalhes ver Cordeiro e Cribari-Neto (2014).

A seguir, apresentamos alguns estudos referentes aos MLGS e às correções de Bartlett e tipo-Bartlett nesses modelos. Cordeiro e Botter (2001) obtiveram fórmulas gerais para os vieses de segunda ordem das estimativas de máxima verossimilhança (EMV) na classe dos MLGS. Cordeiro et al. (2008) deduziram fórmulas para os vieses de segunda ordem das EMV nos modelos não lineares generalizados superdispersados (MNLGS). Cordeiro et al. (2006) derivaram um fator de correção de Bartlett na classe dos MLGS, generalizando os resultados de Botter e Denise (1998) para modelos lineares generalizados duplos e de Cordeiro (1983) para modelos lineares generalizados. Andrade (2013) obteve um fator de correção de Bartlett (baseado na proposta de DiCiccio e Stern, 1994) para a estatística da razão de verossimilhanças perfiladas modificada (baseada na proposta de Cox e Reid, 1987) nos MLGS. Rodrigues (2013) propôs técnicas de diagnósticos, a saber: alavancagem generalizada, análise de resíduos, influência global e influência local sob três esquemas de perturbação, nos MLGS. Terra (2013) derivou um fator de correção de Bartlett para a estatística da razão de verossimilhanças nos MNLGS, generalizando assim os resultados de Cordeiro et al. (2006). Adicionalmente, obteve um fator de correção tipo-Bartlett para a estatística escore e propôs também técnicas de diagnósticos para os MNLGS, a saber: alavancagem generalizada, distância de Cook e influência local. Campos (2014) derivou um fator de correção tipo-Bartlett para a estatística gradiente para testar o efeito da dispersão nos MLGS.

Recentemente, Vargas et al. (2013) obtiveram uma expansão assintótica para a distribuição da estatística gradiente até um erro de ordem $O(n^{-1})$ sob a hipótese nula. Para isso, utilizaram uma abordagem Bayesiana com base no argumento de encolhimento descrito por Ghosh e Mukerjee (1991). Então, usando essa expansão e os resultados de Cordeiro e Ferrari (1991), propuseram uma correção tipo-Bartlett para a estatística gradiente. A estatística corrigida proposta pelos autores possui, sob $\mathcal{H}_0$, distribuição qui-quadrado até um erro de ordem n^{-1}, enquanto a estatística gradiente sem correção tem distribuição qui-quadrado até um erro de ordem $O(n^{-1/2})$, ou seja, a correção tipo-Bartlett reduz o erro de aproximação de $O(n^{-1/2})$ para $O(n^{-1})$. Para uma aplicação desse resultado em MLG, ver Vargas et al. (2014).

O objetivo desta dissertação é derivar um fator de correção tipo-Bartlett, em notação matricial, para a estatística gradiente (ST) para testar simultaneamente os efeitos da média e dispersão ou separadamente os efeitos da média e dispersão, respectivamente, na classe dos MLGS. A estatística gradiente aperfeiçoada (ST^*) tem distribuição qui-quadrado com um erro de aproximação de ordem $O(n^{-1})$ sob a hipótese nula.

Esta dissertação está organizada da seguinte forma. No Capítulo 2, definimos os modelos lineares generalizados superdispersados, discutimos a estimação dos parâmetros de regressão e mostramos como ficam os testes de hipóteses nesta classe de modelos. No Capítulo 3, encontra-se a principal contribuição deste trabalho que é derivar expressões gerais, em forma matricial, para o fator de correção tipo-Bartlett para a estatística gradiente em MLGS. Além disso, apresentamos alguns resultados existentes na literatura, a saber: os fatores de correção de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore, nesta classe, obtidos por Cordeiro et al. (2006) e Terra (2013), respectivamente. Estudos de simulações de Monte Carlo para avaliar e comparar os desempenhos dos testes da razão de verossimilhanças, gradiente, escore, suas respectivas versões corrigidas e de bootstrap em pequenas amostras são apresentados no Capítulo 4. Por fim, no Capítulo 5 estão as principais conclusões deste estudo.

2 MODELOS LINEARES GENERALIZADOS SUPERDISPERSADOS

O fenômeno da superdispersão tem sido amplamente considerado na literatura. Tal fenômeno é facilmente detectado, porém é difícil de ser tratado. Segundo Hinde e Demétrio (1998), o processo de coleta dos dados, a correlação entre respostas individuais e as variáveis omitidas são algumas das possíveis causas da superdispersão. Ainda segundo os autores, uma consequência deste problema é que os erros-padrão das estimativas do modelo estarão incorretos e, também, os desvios serão grandes, conduzindo à seleção de modelos complexos.

Existem muitos modelos específicos para superdispersão. Neste trabalho, lidamos com os modelos lineares generalizados superdispersados (Dey et al., 1997). Esta classe de modelos tem despertado o interesse de alguns pesquisadores. Cordeiro e Botter (2001), a partir de uma expressão geral para os vieses das EMV dada por Cox e Snell (1968), obtiveram fórmulas gerais para os vieses de segunda ordem das EMV em MLGS. Posteriormente, seus resultados foram generalizados por Cordeiro et al. (2008), considerando uma extensão dos MLGS, a saber, os modelos não lineares generalizados superdispersados. Rodrigues (2013) propôs técnicas de diagnósticos para os MLGS. Terra (2013) derivou correções de Bartlett para a estatística da razão de verossimilhanças e tipo-Bartlett para a estatística escore para testar conjuntamente os efeitos da média e da dispersão ou separadamente os efeitos, em MNLGS, generalizando os resultados de Cordeiro et al. (2006). Adicionalmente, Terra (2013) propôs métodos de diagnósticos para esses modelos. Andrade (2013) derivou um fator de correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificada (Cox e Reid, 1987) em MLGS. Campos (2014) obteve um fator de correção tipo-Bartlett à estatística gradiente para testar um subconjunto dos parâmetros de dispersão em MLGS.

2.1 MODELO

Dey et al. (1997) propuseram os modelos lineares superdispersados definidos a seguir. Sejam $Y_1, \ldots, Y_n$ variáveis aleatórias independentes em que cada Y_l tem função densidade (ou de probabilidade) dada por

$$\pi(y_l; \mu_l, \phi_l) = A(y_l) \exp\left\{(y_l - \mu_l)\Psi^{(1,0)}(\mu_l, \phi_l) + \phi_l T(y_l) + \Psi(\mu_l, \phi_l)\right\}, \qquad (2.1)$$

em que $A(\cdot)$, $T(\cdot)$ e $\Psi(\cdot, \cdot)$ são funções conhecidas e $\Psi^{(r,s)}(\mu, \phi) = \partial^{r+s}\Psi(\mu, \phi)/\partial\mu^r\partial\phi^s$, $r, s \geq 0$. Aqui, a média de Y é $\mathrm{E}(Y) = \mu$ e a variância é $\mathrm{Var}(Y) = {\Psi^{(2,0)}}^{-1}$. Já a média e a

variância de $T(Y)$ são, respectivamente, $\mathrm{E}\{T(Y)\} = -\Psi^{(0,1)}$ e $\mathrm{Var}\{T(Y)\} = -\Psi^{(0,2)}$. Gelfand e Dalal (1990) mostraram que, se $T(Y)$ é uma função convexa e (2.1) é integrável com relação a y, para uma média (μ) comum, então $\mathrm{Var}(Y)$ aumenta com ϕ.

A família exponencial uniparamétrica é obtida de (2.1) com $\phi = 0$, isto é, $\pi(y;\theta,0) = A(y)\exp\{y\theta - b(\theta)\}$, em que $\theta = \Psi^{(1,0)}(\mu,0)$ e $b(\theta) = -\Psi(\mu,0) + \mu\Psi^{(1,0)}(\mu,0)$. Então, se $T(Y)$ é convexa, a família de modelos (2.1) é superdispersada em relação à família exponencial uniparamétrica correspondente com $\phi = 0$.

Os modelos lineares generalizados superdispersados são definidos por (2.1) e pelos componentes sistemáticos da média e da dispersão dados, respectivamente, por

$$f(\boldsymbol{\mu}) = \boldsymbol{\eta} = \mathbf{X}\boldsymbol{\beta}, \tag{2.2}$$

$$g(\boldsymbol{\phi}) = \boldsymbol{\tau} = \mathbf{S}\boldsymbol{\gamma}. \tag{2.3}$$

Em (2.2), $\mathbf{X}$ é uma matriz $n \times p$ de posto completo $(p < n)$ cujas linhas denotamos por $\boldsymbol{x}_l = (x_{l1},\ldots,x_{lp})^\top$, $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^\top$ é um vetor de parâmetros desconhecidos a serem estimados e $f(\cdot)$ é uma função conhecida, monótona e diferenciável, denominada função de ligação da média. Já em (2.3), $\mathbf{S}$ é uma matriz $n \times q$ de posto completo $(q < n)$ cujas linhas representamos por $\boldsymbol{s}_l = (s_{l1},\ldots,s_{lq})^\top$, $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_q)^\top$ é um vetor de parâmetros desconhecidos a serem estimados e $g(\cdot)$ é a função de ligação da dispersão conhecida, monótona e diferenciável.

Destacamos que os MLGS permitem que ambos os parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ dependam de covariáveis conhecidas por meio de estruturas lineares, conforme (2.2) e (2.3). Assim, é possível modelar separadamente a média e a dispersão. Para ajustar um MLGS, devemos escolher funções de ligação adequadas tanto para média quanto para dispersão. Apresentamos, em seguida, algumas distribuições pertencentes à classe do MLGS.

Poisson Dupla

Efron (1986) propôs a distribuição Poisson dupla com parâmetros $\mu, \lambda > 0$, denotada por $\mathcal{PD}(\mu,\lambda)$, cuja função de probabilidade é dada por

$$\pi(y;\mu,\lambda) = c(\mu,\lambda)\,\frac{\lambda^{1/2}}{\mathrm{e}^{\lambda\mu}y!}\,\left(\frac{y}{\mathrm{e}}\right)^{y}\left(\frac{\mu\mathrm{e}}{y}\right)^{y\lambda}, \quad y = 0,1,2,..., \tag{2.4}$$

em que $c(\mu,\lambda)$ é uma constante de normalização aproximadamente igual a 1. O autor mostrou que, se $Y \backsim \mathcal{PD}(\mu,\lambda)$, sua média e variância são dadas, respectivamente, por

$$\mathrm{E}(Y) \backsimeq \mu \quad \text{e} \quad \mathrm{Var}(Y) \backsimeq \frac{\mu}{\lambda}.$$

Aqui, λ é um parâmetro de dispersão. O modelo Poisson duplo contempla tanto a superdispersão $(\lambda < 1)$ quanto a subdispersão $(\lambda > 1)$. Quando $\lambda = 1$, temos a distribuição Poisson. A função de probabilidade (2.4) pode ser reescrita da seguinte forma:

$$\begin{aligned}\pi(y;\mu,\phi) &= \frac{y^y}{\mathrm{e}^y\, y!}\, \exp\left\{\frac{1}{2}\, \log\, \lambda - \lambda\mu + y\,\lambda\, \log\, \mu + y\,\lambda - y\,\lambda\, \log\, y\right\} \\ &= \frac{y^y}{\mathrm{e}^y\, y!}\, \exp\left\{(\lambda + \lambda\, \log\, \mu)\, y - \lambda\, y\, \log\, y - \left(-\frac{1}{2}\, \log\, \lambda + \lambda\mu\right)\right\},\end{aligned} \tag{2.5}$$

em que

$$\phi = \lambda, \quad \Psi^{(1,0)}(\mu,\phi) = \lambda + \lambda \log \mu = \phi + \phi \log \mu \quad \text{e} \quad T(y) = -y \log y.$$

Como o componente aleatório de um MLGS é definido a partir da família exponencial biparamétrica (Gelfand e Dalal, 1990) da forma

$$\pi(y;\theta,\phi) = A(y) \exp\{\theta y + \phi T(y) - \rho(\theta,\phi)\},$$

com $\theta = \Psi^{(1,0)}(\mu,\phi)$ e $\rho(\theta,\phi) = \mu\Psi^{(1,0)}(\mu,\phi) - \Psi(\mu,\phi)$, pode-se determinar $\Psi(\mu,\phi)$ a partir $\rho(\theta,\phi)$. Assim, para o modelo (2.5), temos que

$$\begin{aligned}\theta = \Psi^{(1,0)}(\mu,\phi) &= \phi + \phi \log \mu \quad \text{e} \\ \rho(\theta,\phi) = -\frac{1}{2}\log\lambda + \lambda\mu &= -\frac{1}{2}\log\phi + \phi \exp\left(\frac{\theta}{\phi} - 1\right),\end{aligned}$$

então

$$\begin{aligned}\Psi(\mu,\phi) &= \mu\Psi^{(1,0)}(\mu,\phi) - \rho(\theta,\phi) \\ &= \mu(\phi + \phi\, \log\, \mu) - \left[-\frac{1}{2}\, \log\, \phi + \phi \exp\left(\frac{\theta}{\phi} - 1\right)\right] \\ &= \mu(\phi + \phi\, \log\, \mu) - \left(-\frac{1}{2}\, \log\, \phi + \mu\phi\right) \\ &= \mu\phi + \mu\phi \log\mu + \frac{1}{2}\, \log\phi - \mu\phi \\ &= \mu\phi \log\mu + \frac{1}{2}\log\phi.\end{aligned}$$

Gaussiana Inversa Generalizada

A distribuição Gaussiana inversa generalizada (GIG) introduzida por Good (1953) tem função densidade da forma

$$\pi(y;\alpha,\delta,r)=\frac{(\alpha/\delta)^{r/2}}{2K_r(\sqrt{\alpha\delta})}y^{r-1}\exp\left\{-\frac{1}{2}\left(\alpha y+\frac{\delta}{y}\right)\right\},\quad y>0. \tag{2.6}$$

Aqui, $-\infty<r<\infty$, $(\alpha,\delta)\in\Theta_r$, em que $\Theta_r=\{(\alpha,\delta):\alpha>0,\delta\geq 0\}$ se $r>0$, $\Theta_r=\{(\alpha,\delta):\alpha>0,\delta>0\}$ se $r=0$ e $\Theta_r=\{(\alpha,\delta):\alpha\geq 0,\delta>0\}$ se $r<0$. Além disso, $K_r(\cdot)$ é a função de Bessel modificada de terceira ordem. Um estudo detalhado sobre as propriedades da distribuição GIG pode ser encontrado em Jørgensen (1982), ver também Johnson et al. (1994).

Para r conhecido, a função densidade (2.6) pode ser expressa como

$$\pi(y;\alpha,\delta,r)=y^{r-1}\exp\left\{-\left(\frac{\alpha y}{2}+\frac{\delta}{2y}\right)+\log\left(\frac{(\alpha/\delta)^{r/2}}{2K_r(\sqrt{\alpha\delta})}\right)\right\},$$

com

$$\theta=-\frac{\alpha}{2},\quad \phi=-\frac{\delta}{2},\quad T(y)=\frac{1}{y}\quad \text{e}$$

$$\begin{aligned}\rho(\theta,\phi)&=-\frac{r}{2}\log\left(\frac{\alpha}{\delta}\right)+\log\ 2+\log\ K_r\left(2\sqrt{\alpha\delta}\right)\\&=-\frac{r}{2}\log\left(\frac{\theta}{\phi}\right)+\log\ 2+\log\ K_r\left(2\sqrt{\theta\phi}\right).\end{aligned}$$

A distribuição Gaussiana inversa, denotada por $\mathcal{IG}(\mu,\lambda)$, é um caso especial de (2.6) com $r=-1/2$. Assim, se a variável Y segue distribuição $\mathcal{IG}(\mu,\lambda)$, sua função densidade é dada por

$$\begin{aligned}f(y;\mu,\lambda)&=\left[\frac{\lambda}{2\pi y^3}\right]^{1/2}\exp\left\{-\frac{\lambda}{2\mu}\left(\frac{y}{\mu}-2+\frac{\mu}{y}\right)\right\}\\&=\left[\frac{\lambda}{2\pi y^3}\right]^{1/2}\exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 y}\right\},\quad y>0,\end{aligned} \tag{2.7}$$

em que $\mu,\lambda>0$. Note que (2.7) corresponde a (2.6) com $\delta=\lambda$ e $\alpha=\lambda/\mu^2$. Adicionalmente, $K_{-1/2}(\cdot)$ é a função de Bessel, isto é, $K_{-1/2}(\omega)=\sqrt{(\pi/2\omega)}\exp\{-\omega\}$.

Outros casos especiais da distribuição GIG incluem a distribuição recíproca gaussiana inversa ($r=1/2$), a distribuição gama ($\delta=0,r>0$), a distribuição recíproca gama ($\alpha=0,r<0$) e a distribuição hiperbólica ($r=0$).

Família Exponencial Dupla

Considere a família exponencial uniparamétrica com função densidade (ou de probabilidade) da forma

$$f(y;\theta,n) = A(y)\exp\{n[\theta y - b(\theta)]\}, \tag{2.8}$$

em que θ é o parâmetro canônico e y é a média de n variáveis aleatórias independentes e identicamente distribuídas, isto é,

$$y = \sum_{i=1}^{n} z_i/n, \quad \text{em que} \quad z_i \backsim f(z_i;\theta,1).$$

A família exponencial dupla de distribuições definida por Efron (1986) é dada por

$$\pi(y;\theta,\rho,n) = c(\theta,\rho,n)\rho^{1/2}\{f(y;\theta,n)\}^{\rho}\{f(y;y,n)\}^{1-\rho}, \tag{2.9}$$

em que θ é o parâmetro canônico, ρ é o parâmetro de dispersão e $f(y;\theta,n)$ está definida em (2.8). Uma vez que a densidade $f(\cdot)$ em (2.9) é um modelo arbitrário da família exponencial uniparamétrica, a família exponencial dupla pode assumir a seguinte forma:

$$\begin{aligned}\pi(y;\theta,\rho,n) &= c(\theta,\rho,n)\rho^{1/2}\left[\exp\left\{n(\theta y - b(\theta)\right\}\right]^{\rho}\left[\exp\left\{n[\theta(y)y - b(\theta(y))]\right\}\right]^{1-\rho} \\ &= c(\theta,\rho,n)\rho^{1/2}\exp\left\{n\rho(\theta y - b(\theta)) + n(1-\rho)\left[\theta(y)y - b(\theta(y))\right]\right\},\end{aligned}$$

em que $\theta(y) = (b^{'})^{-1}(y)$. Dessa forma, temos que

$$\theta = n\rho\theta, \quad \phi = n(1-\rho) \quad \text{e} \quad T(y) = \theta(y)y - b(\theta(y)),$$

sendo $T(y)$ uma função convexa. Para mais detalhes ver Gelfand e Dalal (1990).

2.2 ASPECTOS INFERENCIAIS

Seja y_l o valor observado da variável aleatória Y_l, $\boldsymbol{y} = (y_1,\ldots,y_n)^{\top}$ um vetor contendo n observações e $\boldsymbol{\theta} = (\boldsymbol{\beta}^{\top},\boldsymbol{\gamma}^{\top})^{\top}$ o vetor de parâmetros do modelo. Denotamos o logaritmo de função de verossimilhança (log-verossimilhança) por

$$\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{l=1}^{n}\left\{(y_l - \mu_l)\Psi^{(1,0)}(\mu_l,\phi_l) + \phi_l T(y_l) + \Psi(\mu_l,\phi_l)\right\} + \sum_{l=1}^{n}\log A(y_l). \tag{2.10}$$

A função $\ell(\boldsymbol{\theta})$ em (2.10) é regular com relação às derivadas em $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ até quarta ordem.

A seguir, introduzimos algumas notações. Denotamos por $m_{il} = \partial^i \mu_l / \partial \eta_l^i$ e $\phi_{il} = \partial^i \phi_l / \partial \tau_l^i$ as i-ésimas derivadas das funções de ligação inversas $\mu = f^{-1}(\eta)$ e $\phi = g^{-1}(\tau)$, $i = 1, 2$ e $l = 1, \ldots, n$. Adicionalmente, definimos as seguintes matrizes diagonais $n \times n$: $\boldsymbol{\Psi}^{(2,0)} = \text{diag}\{\Psi_1^{(2,0)}, \ldots, \Psi_n^{(2,0)}\}$, $\boldsymbol{\Psi}^{(0,2)} = \text{diag}\{\Psi_1^{(0,2)}, \ldots, \Psi_n^{(0,2)}\}$, $\boldsymbol{M}_r = \text{diag}\{m_{r1}, \ldots, m_{rn}\}$ e $\boldsymbol{\Phi}_r = \text{diag}\{\phi_{r1}, \ldots, \phi_{rn}\}$, com $r = 1, 2$. A função escore total particionada conforme $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ é dada por

$$\boldsymbol{U}(\boldsymbol{\theta}) = \boldsymbol{U}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} \\ \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\gamma} \end{pmatrix},$$

com

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \boldsymbol{X}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 (\boldsymbol{y} - \boldsymbol{\mu})$$

e

$$\frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}} = \boldsymbol{S}^\top \boldsymbol{\Phi}_1 \, \boldsymbol{v},$$

sendo $(\boldsymbol{y} - \boldsymbol{\mu}) = (y_1 - \mu_1, \ldots, y_n - \mu_n)^\top$, $\boldsymbol{v} = (v_1, \ldots, v_n)^\top$, com $v_l = \Psi_l^{(1,1)}(y_l - \mu_l) + T(y_l) + \Psi_l^{(0,1)}$.

Devido à ortogonalidade dos parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, a matriz de informação de Fisher é bloco-diagonal e pode ser escrita na forma

$$\boldsymbol{K}(\boldsymbol{\theta}) = \boldsymbol{K}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{bmatrix} \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{K}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} \end{bmatrix},$$

com

$$\boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \text{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right] = \boldsymbol{X}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1^2 \, \boldsymbol{X}$$

e

$$\boldsymbol{K}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} = \text{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^\top} \right] = -\boldsymbol{S}^\top \boldsymbol{\Psi}^{(0,2)} \, \boldsymbol{\Phi}_1^2 \, \boldsymbol{S}.$$

Uma vez que os EMV $\widehat{\boldsymbol{\beta}}$ e $\widehat{\boldsymbol{\gamma}}$ dos parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ não podem ser expressos em forma fechada, deve-se utilizar algum método de otimização não-linear como, por exemplo, o método de Newton-Raphson ou escore de Fisher para obtê-los, ver Nocedal e Wright (1999). Neste trabalho, usamos o método de otimização não-linear BFGS implementado na linguagem de programação `Ox` (Doornik, 2006).

2.3 TESTES DE HIPÓTESES

2.3.1 Testes de Hipóteses sobre $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$

Na maioria dos problemas, as restrições em teste envolvem um subconjunto do vetor de parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$. Particionando $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ como $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ e $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$, em que $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_{p_1})^\top$, $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, com $p_1 \leq p$, $\boldsymbol{\gamma}_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ e $\boldsymbol{\gamma}_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$, com $q_1 \leq q$. Queremos testar a hipótese

$$\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}, \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$$

contra $\mathcal{H}_1^1$: pelo menos uma igualdade é violada, em que $\boldsymbol{\beta}_1^{(0)}$ e $\boldsymbol{\gamma}_1^{(0)}$ são vetores de constantes conhecidas de dimensões p_1 e q_1, respectivamente. Neste caso, $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}_2$ são parâmetros de perturbação. As matrizes modelos particionadas conforme $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são dadas por $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$ e $\boldsymbol{S} = (\boldsymbol{S}_1, \boldsymbol{S}_2)$. A hipótese nula anterior induz à correspondente partição: $\boldsymbol{U}(\boldsymbol{\theta}) = (\boldsymbol{U}_1(\boldsymbol{\theta})^\top, \boldsymbol{U}_2(\boldsymbol{\theta})^\top)^\top$, em que

$$\boldsymbol{U}_1(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{U}_{\boldsymbol{\beta}_1}(\boldsymbol{\theta}) \\ \boldsymbol{U}_{\boldsymbol{\gamma}_1}(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\beta}_1 \\ \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\gamma}_1 \end{pmatrix}$$

e

$$\boldsymbol{U}_2(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{U}_{\boldsymbol{\beta}_2}(\boldsymbol{\theta}) \\ \boldsymbol{U}_{\boldsymbol{\gamma}_2}(\boldsymbol{\theta}) \end{pmatrix} = \begin{pmatrix} \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\beta}_2 \\ \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\gamma}_2 \end{pmatrix},$$

cujos elementos são

$$\begin{aligned} \frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}_1} &= \boldsymbol{X}_1^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1(\boldsymbol{y} - \boldsymbol{\mu}), \\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\gamma}_1} &= \boldsymbol{S}_1^\top \boldsymbol{\Phi}_1 \boldsymbol{v}, \\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}_2} &= \boldsymbol{X}_2^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1(\boldsymbol{y} - \boldsymbol{\mu}), \\ \frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\gamma}_2} &= \boldsymbol{S}_2^\top \boldsymbol{\Phi}_1 \boldsymbol{v}. \end{aligned} \tag{2.11}$$

A matriz de informação de Fisher particionada da mesma forma que $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ pode ser escrita como

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{K}_{11} & \boldsymbol{K}_{12} \\ \boldsymbol{K}_{21} & \boldsymbol{K}_{22} \end{bmatrix},$$

com

$$\begin{aligned}
\boldsymbol{K}_{11} &= \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}_{11}}, \boldsymbol{K}_{\boldsymbol{\gamma}_{11}}\big\},\\
\boldsymbol{K}_{12} &= \boldsymbol{K}_{21}^{\top} = \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}_{12}}, \boldsymbol{K}_{\boldsymbol{\gamma}_{12}}\big\},\\
\boldsymbol{K}_{22} &= \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}_{22}}, \boldsymbol{K}_{\boldsymbol{\gamma}_{22}}\big\},
\end{aligned}$$

sendo

$$\begin{aligned}
\boldsymbol{K}_{\boldsymbol{\beta}_{11}} &= \boldsymbol{X}_1^{\top}\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1^2\boldsymbol{X}_1,\\
\boldsymbol{K}_{\boldsymbol{\beta}_{12}} &= \boldsymbol{K}_{\boldsymbol{\beta}_{21}}^{\top} = \boldsymbol{X}_1^{\top}\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1^2\boldsymbol{X}_2,\\
\boldsymbol{K}_{\boldsymbol{\beta}_{22}} &= \boldsymbol{X}_2^{\top}\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1^2\boldsymbol{X}_2,\\
\boldsymbol{K}_{\boldsymbol{\gamma}_{11}} &= -\boldsymbol{S}_1^{\top}\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1^2\boldsymbol{S}_1,\\
\boldsymbol{K}_{\boldsymbol{\gamma}_{12}} &= \boldsymbol{K}_{\boldsymbol{\gamma}_{21}}^{\top} = -\boldsymbol{S}_1^{\top}\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1^2\boldsymbol{S}_2,\\
\boldsymbol{K}_{\boldsymbol{\gamma}_{22}} &= -\boldsymbol{S}_2^{\top}\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1^2\boldsymbol{S}_2.
\end{aligned} \tag{2.12}$$

A matriz inversa da informação de Fisher é dada por

$$\boldsymbol{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} \boldsymbol{K}^{11} & \boldsymbol{K}^{12} \\ \boldsymbol{K}^{21} & \boldsymbol{K}^{22} \end{bmatrix},$$

com

$$\begin{aligned}
\boldsymbol{K}^{11} &= \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}}^{11}, \boldsymbol{K}_{\boldsymbol{\gamma}}^{11}\big\},\\
\boldsymbol{K}^{12} &= \boldsymbol{K}^{21\top} = \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}}^{12}, \boldsymbol{K}_{\boldsymbol{\gamma}}^{12}\big\},\\
\boldsymbol{K}^{22} &= \text{diag}\big\{\boldsymbol{K}_{\boldsymbol{\beta}}^{22}, \boldsymbol{K}_{\boldsymbol{\gamma}}^{22}\big\},
\end{aligned}$$

em que

$$\begin{aligned}
\boldsymbol{K}_{\boldsymbol{\beta}}^{11} &= \big[\boldsymbol{K}_{\boldsymbol{\beta}_{11}} - \boldsymbol{K}_{\boldsymbol{\beta}_{12}}\boldsymbol{K}_{\boldsymbol{\beta}_{22}}^{-1}\boldsymbol{K}_{\boldsymbol{\beta}_{21}}\big]^{-1},\\
\boldsymbol{K}_{\boldsymbol{\gamma}}^{11} &= \big[\boldsymbol{K}_{\boldsymbol{\gamma}_{11}} - \boldsymbol{K}_{\boldsymbol{\gamma}_{12}}\boldsymbol{K}_{\boldsymbol{\gamma}_{22}}^{-1}\boldsymbol{K}_{\boldsymbol{\gamma}_{21}}\big]^{-1},\\
\boldsymbol{K}_{\boldsymbol{\beta}}^{12} &= \boldsymbol{K}_{\boldsymbol{\beta}}^{21\top} = -\boldsymbol{K}_{\boldsymbol{\beta}}^{11}\boldsymbol{K}_{\boldsymbol{\beta}_{12}}\boldsymbol{K}_{\boldsymbol{\beta}_{22}}^{-1},\\
\boldsymbol{K}_{\boldsymbol{\gamma}}^{12} &= \boldsymbol{K}_{\boldsymbol{\gamma}}^{21\top} = -\boldsymbol{K}_{\boldsymbol{\gamma}}^{11}\boldsymbol{K}_{\boldsymbol{\gamma}_{12}}\boldsymbol{K}_{\boldsymbol{\gamma}_{22}}^{-1},\\
\boldsymbol{K}_{\boldsymbol{\beta}}^{22} &= \boldsymbol{K}_{\boldsymbol{\beta}_{22}}^{-1} + \boldsymbol{K}_{\boldsymbol{\beta}_{22}}^{-1}\boldsymbol{K}_{\boldsymbol{\beta}_{21}}\boldsymbol{K}_{\boldsymbol{\beta}}^{11}\boldsymbol{K}_{\boldsymbol{\beta}_{12}}\boldsymbol{K}_{\boldsymbol{\beta}_{22}}^{-1},\\
\boldsymbol{K}_{\boldsymbol{\gamma}}^{22} &= \boldsymbol{K}_{\boldsymbol{\gamma}_{22}}^{-1} + \boldsymbol{K}_{\boldsymbol{\gamma}_{22}}^{-1}\boldsymbol{K}_{\boldsymbol{\gamma}_{21}}\boldsymbol{K}_{\boldsymbol{\gamma}}^{11}\boldsymbol{K}_{\boldsymbol{\gamma}_{12}}\boldsymbol{K}_{\boldsymbol{\gamma}_{22}}^{-1}.
\end{aligned} \tag{2.13}$$

Para maiores detalhes sobre inversas de matrizes particionadas, ver Rao (1973, p. 33).

As estatísticas da razão de verossimilhanças (LR), gradiente (ST) e escore (SR) para testar a hipótese $\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}, \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ podem ser expressas por

$$\begin{aligned}
LR &= 2\big\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\big\},\\
ST &= \boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}})^\top\big(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)}\big) + \boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}})^\top\big(\widehat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1^{(0)}\big),\\
SR &= \boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}})^\top\boldsymbol{K}_{\boldsymbol{\beta}}^{11}(\widetilde{\boldsymbol{\theta}})\boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}}) + \boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}})^\top\boldsymbol{K}_{\boldsymbol{\gamma}}^{11}(\widetilde{\boldsymbol{\theta}})\boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}}),
\end{aligned}$$

em que $\widehat{\boldsymbol{\theta}} = \big(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \widehat{\boldsymbol{\gamma}}_1, \widehat{\boldsymbol{\gamma}}_2\big)^\top$ e $\widetilde{\boldsymbol{\theta}} = \big(\boldsymbol{\beta}_1^{(0)}, \widetilde{\boldsymbol{\beta}}_2, \boldsymbol{\gamma}_1^{(0)}, \widetilde{\boldsymbol{\gamma}}_2\big)^\top$ são as EMV irrestrita e restrita (sob $\mathcal{H}_0^1$) de $\boldsymbol{\theta}$, respectivamente. Em grandes amostras e sob a hipótese nula, as três estatísticas acima têm distribuição $\chi^2_{p_1+q_1}$, em que $p_1 + q_1$ é número de restrições impostas por $\mathcal{H}_0^1$.

2.3.2 Testes de Hipóteses sobre $\boldsymbol{\beta}$

Agora, as restrições em teste envolvem um subconjunto do vetor de parâmetros $\boldsymbol{\beta}$. Particionando $\boldsymbol{\beta}$ como $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$, onde $\boldsymbol{\beta_1} = (\beta_1, \ldots, \beta_{p_1})^\top$ e $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, com $p_1 \leq p$, conduzindo a uma partição da matriz modelo como $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$. O interesse é testar a hipótese

$$\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)} \quad \text{contra} \quad \mathcal{H}_1^2 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_1^{(0)},$$

em que $\boldsymbol{\beta}_1^{(0)}$ é um vetor p_1-dimensional de constantes conhecidas. Aqui, $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}$ são parâmetros de perturbação. A função escore total pode ser particionada da mesma forma que $\boldsymbol{\beta}$ como $\boldsymbol{U}(\boldsymbol{\theta}) = (\boldsymbol{U}_1(\boldsymbol{\theta})^\top, \boldsymbol{U}_2(\boldsymbol{\theta})^\top)^\top$, em que $\boldsymbol{U}_1(\boldsymbol{\theta}) = \boldsymbol{U}_{\boldsymbol{\beta}_1}$ e $\boldsymbol{U}_2(\boldsymbol{\theta}) = (\boldsymbol{U}_{\boldsymbol{\beta}_2}^\top, \boldsymbol{U}_{\boldsymbol{\gamma}}^\top)^\top$. As funções $\boldsymbol{U}_{\boldsymbol{\beta}_1}$ e $\boldsymbol{U}_{\boldsymbol{\beta}_2}$ são definidas em (2.11) e a função $\boldsymbol{U}_{\boldsymbol{\gamma}}$ é apresentada na Seção 2.3. A matriz de informação de Fisher é bloco-diagonal e, considerando a partição de $\boldsymbol{\beta}$ mencionada acima, pode ser escrita na forma

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{bmatrix}
\boldsymbol{K}_{\boldsymbol{\beta}_{11}} & \boldsymbol{K}_{\boldsymbol{\beta}_{12}} & \boldsymbol{0}\\
\boldsymbol{K}_{\boldsymbol{\beta}_{21}} & \boldsymbol{K}_{\boldsymbol{\beta}_{22}} & \boldsymbol{0}\\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{K}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}
\end{bmatrix}.$$

As matrizes $\boldsymbol{K}_{\boldsymbol{\beta}_{11}}$, $\boldsymbol{K}_{\boldsymbol{\beta}_{12}}$, $\boldsymbol{K}_{\boldsymbol{\beta}_{21}}$ e $\boldsymbol{K}_{\boldsymbol{\beta}_{22}}$ estão definidas em (2.12) e a matriz $\boldsymbol{K}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}$ é dada na Seção 2.3.

A inversa da matriz de informação de Fisher também é bloco-diagonal sendo dada por

$$\boldsymbol{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix}
\boldsymbol{K}_{\boldsymbol{\beta}}^{11} & \boldsymbol{K}_{\boldsymbol{\beta}}^{12} & \boldsymbol{0}\\
\boldsymbol{K}_{\boldsymbol{\beta}}^{21} & \boldsymbol{K}_{\boldsymbol{\beta}}^{22} & \boldsymbol{0}\\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{K}_{\boldsymbol{\gamma}\boldsymbol{\gamma}}^{-1}
\end{bmatrix},$$

em que $\boldsymbol{K}_{\boldsymbol{\beta}}^{11}$, $\boldsymbol{K}_{\boldsymbol{\beta}}^{12}$, $\boldsymbol{K}_{\boldsymbol{\beta}}^{21}$ e $\boldsymbol{K}_{\boldsymbol{\beta}}^{22}$ estão definidas em (2.13).

As estatísticas da razão de verossimilhanças, gradiente e escore para testar a hipótese $\mathcal{H}_0^2$ podem ser escritas, respectivamente, como

$$
\begin{aligned}
LR &= 2\big\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\big\},\\
ST &= \boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}})^\top \big(\widehat{\boldsymbol{\beta}}_1 - \boldsymbol{\beta}_1^{(0)}\big),\\
SR &= \boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}})^\top \boldsymbol{K}_{\boldsymbol{\beta}}^{11}(\widetilde{\boldsymbol{\theta}})\boldsymbol{U}_{\boldsymbol{\beta}_1}(\widetilde{\boldsymbol{\theta}}),
\end{aligned}
$$

em que $\widehat{\boldsymbol{\theta}} = \big(\widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\beta}}_2, \widehat{\boldsymbol{\gamma}}\big)^\top$ e $\widetilde{\boldsymbol{\theta}} = \big(\boldsymbol{\beta}_1^{(0)}, \widetilde{\boldsymbol{\beta}}_2, \widetilde{\boldsymbol{\gamma}}\big)^\top$ são as EMV irrestrita e restrita de $\boldsymbol{\theta}$, respectivamente. Assintoticamente, as estatísticas LR, ST e SR possuem distribuição $\chi^2_{p_1}$ sob a hipótese nula.

2.3.3 Testes de Hipóteses sobre $\boldsymbol{\gamma}$

Considere a partição de $\boldsymbol{\gamma}$ como $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$, em que $\boldsymbol{\gamma}_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ e $\boldsymbol{\gamma}_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$, com $q_1 \leq q$. Estamos interessados em testar a hipótese

$$
\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)} \quad \text{contra} \quad \mathcal{H}_1^3 : \boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_1^{(0)},
$$

em que $\boldsymbol{\gamma}_1^{(0)}$ é um vetor q_1-dimensional de constantes conhecidas. Neste caso, $\boldsymbol{\gamma}_2$ e $\boldsymbol{\beta}$ são parâmetros de perturbação. A matriz modelo particionada como $\boldsymbol{\gamma}$ é dada por $\boldsymbol{S} = (\boldsymbol{S}_1, \boldsymbol{S}_2)$. A hipótese anterior conduz às seguintes partições: $\boldsymbol{U} = (\boldsymbol{U}_1^\top(\boldsymbol{\theta}), \boldsymbol{U}_2^\top(\boldsymbol{\theta}))^\top$, com $\boldsymbol{U}_1(\boldsymbol{\theta}) = \boldsymbol{U}_{\boldsymbol{\gamma}_1}$ e $\boldsymbol{U}_2(\boldsymbol{\theta}) = \big(\boldsymbol{U}_{\boldsymbol{\gamma}_2}^\top, \boldsymbol{U}_{\boldsymbol{\beta}}^\top\big)^\top$,

$$
\boldsymbol{K}(\boldsymbol{\theta}) = \begin{bmatrix}
\boldsymbol{K}_{\boldsymbol{\gamma}_{11}} & \boldsymbol{K}_{\boldsymbol{\gamma}_{12}} & \boldsymbol{0}\\
\boldsymbol{K}_{\boldsymbol{\gamma}_{21}} & \boldsymbol{K}_{\boldsymbol{\gamma}_{22}} & \boldsymbol{0}\\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}
\end{bmatrix}.
$$

A matriz inversa da informação de Fisher é definida da seguinte forma:

$$
\boldsymbol{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix}
\boldsymbol{K}_{\boldsymbol{\gamma}}^{11} & \boldsymbol{K}_{\boldsymbol{\gamma}}^{12} & \boldsymbol{0}\\
\boldsymbol{K}_{\boldsymbol{\gamma}}^{21} & \boldsymbol{K}_{\boldsymbol{\gamma}}^{22} & \boldsymbol{0}\\
\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{K}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}
\end{bmatrix}.
$$

As estatísticas da razão de verossimilhanças, gradiente e escore para testar a hipótese $\mathcal{H}_0^3$ são dadas, respectivamente, por

$$
\begin{aligned}
LR &= 2\big\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\big\},\\
ST &= \boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}})^\top \big(\widehat{\boldsymbol{\gamma}}_1 - \boldsymbol{\gamma}_1^{(0)}\big),\\
SR &= \boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}})^\top \boldsymbol{K}_{\boldsymbol{\gamma}}^{11}(\widetilde{\boldsymbol{\theta}})\boldsymbol{U}_{\boldsymbol{\gamma}_1}(\widetilde{\boldsymbol{\theta}}),
\end{aligned}
$$

em que $\widehat{\boldsymbol{\theta}} = \left(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}_1, \widehat{\boldsymbol{\gamma}}_2\right)^\top$ e $\widetilde{\boldsymbol{\theta}} = \left(\widetilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}_1^{(0)}, \widetilde{\boldsymbol{\gamma}}_2\right)^\top$ são as EMV irrestrita e restrita (sob a hipótese nula) de $\boldsymbol{\theta}$, respectivamente. Em grandes amostras e sob $\mathcal{H}_0^3$, as estatísticas LR, ST e SR têm distribuição $\chi^2_{q_1}$.

2.4 TESTES DE HIPÓTESES BASEADOS EM BOOTSTRAP

Uma alternativa às aproximações assintóticas para as distribuições nulas das estatísticas da razão de verossimilhanças, gradiente e escore discutidas anteriormente é utilizar um esquema de reamostragem bootstrap (Efron, 1979) em sua versão paramétrica, para estimar a distribuição nula da estatística de teste. Efron e Tibshirani (1993) apresentam um estudo detalhado sobre a utilização da técnica bootstrap na realização de testes de hipóteses. Essa metodologia consiste em obter, da amostra original $\boldsymbol{y}$, um grande número (digamos, B) de pseudo-amostras $\boldsymbol{y}^* = (y_1^*, \ldots, y_n^*)^\top$. A hipótese nula deve ser imposta ao gerar as amostras artificiais. Em seguida, calcula-se para cada amostra $\boldsymbol{y}^*$ a estatística de teste denotada por S^*, ou seja, obtêm-se $S_1^*, \ldots, S_B^*$. Utiliza-se as estatísticas bootstrap para estimar a distribuição nula da estatística S. O valor crítico bootstrap denotado por $\widehat{q}_{1-\alpha}$ $(0 < \alpha < 1)$ é obtido como o quantil $1-\alpha$ da distribuição empírica das B realizações da estatística S^*. Assim, a hipótese nula é rejeitada se $S > \widehat{q}_{1-\alpha}$.

3 CORREÇÕES DE BARTLETT E TIPO-BARTLETT

A distribuição qui-quadrado pode não ser uma boa aproximação para a distribuição nula das estatísticas apresentadas na Seção 2.4 quando o tamanho da amostra é pequeno ou moderado, podendo conduzir a taxas de rejeição bastante distorcidas. Neste sentido, foram obtidos alguns refinamentos com o objetivo de melhorar a aproximação pela distribuição qui-quadrado. Cordeiro et al. (2006) derivaram um fator de correção de Bartlett em notação matricial para a estatística da razão de verossimilhanças em MLGS. Terra (2013) obteve fatores de correção de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore na classe dos MNLGS, uma extensão dos MLGS.

Neste capítulo, com base nos resultados de Vargas et al. (2013), derivamos um fator de correção tipo-Bartlett em notação matricial para a estatística gradiente na classe dos MLGS, de modo que a estatística gradiente corrigida tem distribuição, segundo a hipótese nula, mais próxima da distribuição χ^2 de referência, isto é, o erro desta aproximação é reduzido de $O(n^{-1/2})$ para $O(n^{-1})$.

3.1 CORREÇÃO TIPO-BARTLETT PARA A ESTATÍSTICA GRADIENTE

Inicialmente, apresentamos algumas notações. Sejam $y_1, \ldots, y_n$ n observações independentes onde cada y_l possui função densidade (ou de probabilidade) $f(y; \boldsymbol{\theta})$ e $\ell(\boldsymbol{\theta})$ é a correspondente função de log-verossimilhança que depende de um vetor de parâmetros desconhecidos $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)^\top$ p-dimensional. Utilizamos a seguinte notação para as derivadas da função de log-verossimilhança, em que todos índices variam de 1 a p:

$$U_r = \partial\ell(\boldsymbol{\theta})/\partial\theta_r, \quad U_{rs} = \partial^2\ell(\boldsymbol{\theta})/\partial\theta_r\partial\theta_s, \quad U_{rst} = \partial^3\ell(\boldsymbol{\theta})/\partial\theta_r\partial\theta_s\partial\theta_t,$$

etc. Os cumulantes de derivadas da função de log-verossimilhança são denotados por

$$\kappa_r = \mathrm{E}(U_r), \quad \kappa_{rs} = \mathrm{E}(U_{rs}), \quad \kappa_{rst} = \mathrm{E}(U_{rst}),$$
$$\kappa_{r,s} = \mathrm{E}(U_r U_s), \quad \kappa_{r,st} = \mathrm{E}(U_r U_{st})$$

e assim por diante. As derivadas desses cumulantes são representadas por

$$\kappa_{rs}^{(t)} = \partial\kappa_{rs}/\partial\theta_t, \quad \kappa_{rs}^{(tu)} = \partial^2\kappa_{rs}/\partial\theta_t\partial\theta_u,$$

etc. Esses cumulantes satisfazem certas equações que facilitam seus cálculos que são denominadas *identidades de Bartlett*. As mais importantes são $\kappa_r = 0$ e $\kappa_{r,s} + \kappa_{rs} = 0$.

Agora, considere a partição do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, em que $\boldsymbol{\theta}_1 = (\theta_1, \ldots, \theta_q)^\top$ é o vetor de parâmetros de interesse e $\boldsymbol{\theta}_2 = (\theta_{q+1}, \ldots, \theta_p)^\top$ é o vetor de parâmetros de perturbação. A função escore total também pode ser particionada conforme $\boldsymbol{\theta}$, ou seja, $U(\boldsymbol{\theta}) = (U_1(\boldsymbol{\theta})^\top, U_2(\boldsymbol{\theta})^\top)^\top$. Além disso, seja $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\theta}}_1, \widehat{\boldsymbol{\theta}}_2)^\top$ o estimador de máxima verossimilhança (EMV) irrestrito de $\boldsymbol{\theta}$ e $\widetilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_1^{(0)}, \widetilde{\boldsymbol{\theta}}_2)^\top$ o EMV restrito de $\boldsymbol{\theta}$.

Nosso objetivo é testar a hipótese $\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ contra $\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$, sendo $\boldsymbol{\theta}_1^{(0)}$ um vetor de constantes conhecidas de dimensão q. A estatística gradiente para testar $\mathcal{H}_0$ contra $\mathcal{H}_1$ é definida por

$$ST = \boldsymbol{U}(\widetilde{\boldsymbol{\theta}})^\top(\widehat{\boldsymbol{\theta}} - \widetilde{\boldsymbol{\theta}}),$$

podendo ser escrita como $ST = \boldsymbol{U}_1(\widetilde{\boldsymbol{\theta}})^\top(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^{(0)})$, pois $\boldsymbol{U}_2(\widetilde{\boldsymbol{\theta}}) = \mathbf{0}$. Em grandes amostras e sob a hipótese nula, a distribuição de ST é aproximadamente χ^2_q sendo q o número de restrições impostas por $\mathcal{H}_0$. Adicionalmente, a partição $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ induz também às seguintes partições:

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{K}_{11} & \boldsymbol{K}_{12} \\ \boldsymbol{K}_{21} & \boldsymbol{K}_{22} \end{bmatrix} \quad \text{e} \quad \boldsymbol{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} \boldsymbol{K}^{11} & \boldsymbol{K}^{12} \\ \boldsymbol{K}^{21} & \boldsymbol{K}^{22} \end{bmatrix},$$

em que $\boldsymbol{K}(\boldsymbol{\theta})$ é a matriz de informação de Fisher e $\boldsymbol{K}(\boldsymbol{\theta})^{-1}$ representa sua inversa.

Vargas et al. (2013) derivaram uma expansão assintótica para a distribuição da estatística gradiente sob a hipótese nula utilizando uma abordagem bayesiana baseada em um argumento de encolhimento sugerido por Ghosh e Mukerjee (1991). Tal expansão é dada por

$$\Pr(ST \leq x) = G_q(x) + \frac{1}{24n}\sum_{i=0}^{3} R_i G_{q+2i}(x) + O(n^{-1}), \tag{3.1}$$

em que $G_z(\cdot)$ é a função de distribuição de uma variável aleatória qui-quadrado com z graus de liberdade,

$$R_1 = 3A_3 - 2A_2 + A_1, \quad R_2 = A_2 - 3A_3, \quad R_3 = A_3, \quad R_0 = -(R_1 + R_2 + R_3),$$

$$\begin{aligned}
A_1 = {} & 3{\sum}' \kappa_{jrs}\kappa_{klu}\left[m^{jr}a^{lu}(m^{sk} + 2a^{sk}) + a^{jr}m^{sk}a^{lu} + 2m^{jk}a^{rl}a^{su}\right] \\
& - 12{\sum}' \kappa_{jr}^{(s)}\kappa_{kl}^{(u)}(\kappa^{sj}\kappa^{rk}\kappa^{lu} + a^{sj}a^{rk}a^{lu} + \kappa^{sk}\kappa^{lj}\kappa^{ru} + a^{sk}a^{lj}a^{ru}) \\
& - 6{\sum}' \kappa_{jrs}\kappa_{kl}^{(u)}[(a^{su} - \kappa^{su})(\kappa^{jk}\kappa^{lr} - a^{jk}a^{lr}) + m^{jr}(a^{sk}a^{lu} + \kappa^{sk}\kappa^{lu}) \\
& \qquad + 2a^{rs}(\kappa^{jk}\kappa^{lu} - a^{jk}a^{lu}) + 2a^{rk}a^{ls}m^{ju}] \\
& + 6{\sum}' \kappa_{jrsu}m^{jr}a^{su} - 6{\sum}' \kappa_{jrs}^{(u)}[m^{jr}(a^{su} - \kappa^{su}) + 2m^{ju}a^{rs}] \\
& + 12{\sum}' \kappa_{rs}^{(ju)}(\kappa^{jr}\kappa^{su} - a^{jr}a^{su}),
\end{aligned}$$

$$
\begin{aligned}
A_2 = -3\sum{}' \kappa_{jrs}\kappa_{klu}[m^{jr}m^{sk}a^{lu} + m^{jr}a^{sk}m^{lu} + 2m^{jk}m^{rl}a^{su} \\
+ \frac{1}{4}(3m^{jr}m^{sk}m^{lu} + 2m^{jk}m^{rl}m^{su})] \\
+ 6\sum{}' \kappa_{jrs}\kappa_{kl}^{(u)}[m^{su}(\kappa^{jk}\kappa^{lr} - a^{jk}a^{lr}) + m^{jr}(\kappa^{sk}\kappa^{lu} - a^{sk}a^{lu})] \\
+ 6\sum{}' \kappa_{jrs}^{(u)}m^{jr}m^{su} - 3\sum{}' \kappa_{jrsu}m^{jr}m^{su}
\end{aligned}
$$

e

$$
A_3 = \frac{1}{4}\sum{}' \kappa_{jrs}\kappa_{klu}(3m^{jr}m^{sk}m^{lu} + 2m^{jk}m^{rl}m^{su}),
$$

sendo a^{ij} e m^{ij} os elementos (i, j) das matrizes

$$
\boldsymbol{A} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{K}_{22}^{-1} \end{bmatrix} \quad \text{e} \quad \boldsymbol{M} = \boldsymbol{K}(\boldsymbol{\theta})^{-1} - \boldsymbol{A},
$$

respectivamente. Observe que $\boldsymbol{K}_{22}^{-1}$ representa a matriz de covariância assintótica de $\widetilde{\boldsymbol{\theta}}_2$. Os somatórios $\sum'$ são tomados sobre todos os parâmetros $\theta_1, \ldots, \theta_p$, ou seja, os índices j, r, s, k, l, u variam de 1 a p. As expressões A_1, A_2 e A_3 são funções de cumulantes e de derivadas desses cumulantes e são avaliadas sob a hipótese nula.

A estatística ST modificada é definida por

$$
ST^* = ST\,[1 - (c + bST + aST^2)], \tag{3.2}
$$

em que

$$
a = \frac{A_3}{12nq(q+2)(q+4)}, \quad b = \frac{A_2 - 2A_3}{12nq(q+2)} \quad \text{e} \quad c = \frac{A_1 - A_2 + A_3}{12nq}.
$$

O fator multiplicativo $[1 - (c + bST + aST^2)]$ é um fator de correção tipo-Bartlett. O erro de aproximação pela χ_q^2 para a distribuição de ST é de $O(n^{-1/2})$, enquanto o erro desta aproximação para a distribuição de ST^* é reduzido para $O(n^{-1})$.

Em vez de modificar a estatística ST como em (3.2), é possível modificar os quantis da distribuição qui-quadrado de referência através da inversão de (3.1) (Hill e Davis, 1968). Dessa forma, temos que (Vargas et al., 2013)

$$
\begin{aligned}
z_\alpha = x_\alpha + \frac{1}{12}\Bigg[\frac{A_3 x_\alpha}{q(q+2)(q+4)}\{x_\alpha^2 + (q+4)x_\alpha + (q+2)(q+4)\} \\
+ \frac{x_\alpha(x_\alpha + q + 2)}{q(q+2)}(A_2 - 3A_3) + \frac{x_\alpha}{q}(3A_3 - 2A_2 + A_1)\Bigg],
\end{aligned} \tag{3.3}
$$

com $\Pr(ST > z_\alpha) = \alpha$ e $\Pr(\chi_q^2 > x_\alpha) = \alpha$. Assim, um teste gradiente aperfeiçoado também pode ser conduzido comparando o valor observado da estatística ST com quantil modificado apresentado em (3.3).

3.2 CORREÇÃO TIPO-BARTLETT PARA A ESTATÍSTICA GRADIENTE EM MLGS

Apresentamos nesta seção a principal contribuição desta dissertação, a saber, derivamos expressões matriciais para os A's que definem a estatística gradiente corrigida em MLGS. Sejam as matrizes

$$\boldsymbol{Z}_{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1^2 \boldsymbol{X})^{-1} \boldsymbol{X}^\top, \quad \boldsymbol{Z}_{2\boldsymbol{\beta}} = \boldsymbol{X}_2(\boldsymbol{X}_2^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1^2 \boldsymbol{X}_2)^{-1} \boldsymbol{X}_2^\top,$$
$$\boldsymbol{Z}_{\boldsymbol{\gamma}} = \boldsymbol{S}(-\boldsymbol{S}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \boldsymbol{S})^{-1} \boldsymbol{S}^\top, \quad \boldsymbol{Z}_{2\boldsymbol{\gamma}} = \boldsymbol{S}_2(-\boldsymbol{S}_2^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \boldsymbol{S}_2)^{-1} \boldsymbol{S}_2^\top,$$

denotamos por $\boldsymbol{Z}_{\boldsymbol{\beta}d}$, $\boldsymbol{Z}_{2\boldsymbol{\beta}d}$, $\boldsymbol{Z}_{\boldsymbol{\gamma}d}$ e $\boldsymbol{Z}_{2\boldsymbol{\gamma}d}$ as matrizes diagonais compostas pelos correspondentes elementos de $\boldsymbol{Z}_{\boldsymbol{\beta}}$, $\boldsymbol{Z}_{2\boldsymbol{\beta}}$, $\boldsymbol{Z}_{\boldsymbol{\gamma}}$ e $\boldsymbol{Z}_{2\boldsymbol{\gamma}}$, respectivamente, $\boldsymbol{Z}_{\boldsymbol{\beta}}^{(2)} = \boldsymbol{Z}_{\boldsymbol{\beta}} \odot \boldsymbol{Z}_{\boldsymbol{\beta}}$, $\boldsymbol{Z}_{\boldsymbol{\beta}}^{(3)} = \boldsymbol{Z}_{\boldsymbol{\beta}}^{(2)} \odot \boldsymbol{Z}_{\boldsymbol{\beta}}$, em que $\odot$ denota o produto de Hadamard. Por exemplo, z_{ij}^2 denota o (i,j)-ésimo elemento da matriz $\boldsymbol{Z}_{\boldsymbol{\beta}}^{(2)}$.

As fórmulas matriciais para os A's que definem a estatística gradiente aperfeiçoada para testar $\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}, \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ podem ser escritas da seguinte forma:

$$A_1 = A_{1,\boldsymbol{\beta}} + A_{1,\boldsymbol{\gamma}} + A_{1,\boldsymbol{\beta\gamma}},$$
$$A_2 = A_{2,\boldsymbol{\beta}} + A_{2,\boldsymbol{\gamma}} + A_{2,\boldsymbol{\beta\gamma}},$$
$$A_3 = A_{3,\boldsymbol{\beta}} + A_{3,\boldsymbol{\gamma}} + A_{3,\boldsymbol{\beta\gamma}},$$

em que

$$\begin{aligned}
A_{1,\beta} = {} & 3\,\mathbf{1}^\top(2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})\boldsymbol{Z}_{2\beta d} \\
& + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\beta}\boldsymbol{Z}_{2\beta d} + \boldsymbol{Z}_{2\beta d}(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})\boldsymbol{Z}_{2\beta d} + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot \boldsymbol{Z}_{2\beta}^{(2)}] \\
& \times (2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1} - 12\,\mathbf{1}^\top(\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2) \\
& \times [-\boldsymbol{Z}_{\beta d}\boldsymbol{Z}_\beta\boldsymbol{Z}_{\beta d} + \boldsymbol{Z}_{2\beta d}\boldsymbol{Z}_{2\beta}\boldsymbol{Z}_{2\beta d} - \boldsymbol{Z}_\beta^{(3)} + \boldsymbol{Z}_{2\beta}^{(3)}](\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1} \\
& - 6\,\mathbf{1}^\top(2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)[(\boldsymbol{Z}_\beta + \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta^{(2)} - \boldsymbol{Z}_{2\beta}^{(2)}) \\
& + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_{2\beta}\boldsymbol{Z}_{2\beta d} + \boldsymbol{Z}_\beta\boldsymbol{Z}_{\beta d}) + 2\boldsymbol{Z}_{2\beta d}(\boldsymbol{Z}_\beta\boldsymbol{Z}_{\beta d} - \boldsymbol{Z}_{2\beta}\boldsymbol{Z}_{2\beta d}) \\
& + 2\boldsymbol{Z}_{2\beta}^{(2)} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})](\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1} - 6\,\mathbf{1}^\top(2\boldsymbol{\Psi}^{(4,0)}\boldsymbol{M}_1^4 \\
& + 12\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^2\boldsymbol{M}_2 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_2^2 + 4\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_3)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\beta d}]\mathbf{1} \\
& + 6\,\mathbf{1}^\top(2\boldsymbol{\Psi}^{(4,0)}\boldsymbol{M}_1^4 + 9\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^2\boldsymbol{M}_2 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_2^2 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_3) \\
& \times [(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(3\boldsymbol{Z}_{2\beta d} + \boldsymbol{Z}_{\beta d})]\mathbf{1} - 12\,\mathbf{1}^\top(\boldsymbol{\Psi}^{(4,0)}\boldsymbol{M}_1^4 + 5\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^2\boldsymbol{M}_2 \\
& + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_2^2 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_3)[\boldsymbol{Z}_{\beta d}^{(2)} - \boldsymbol{Z}_{2\beta d}^{(2)}]\mathbf{1},
\end{aligned} \tag{3.4}$$

$$
\begin{aligned}
A_{1,\gamma} = {} & 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\gamma d} \\
& + 2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d} + \boldsymbol{Z}_{2\gamma d}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\gamma d} + 2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot \boldsymbol{Z}_{2\gamma}^{(2)}] \\
& \times (\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} - 12\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2) \\
& \times [-\boldsymbol{Z}_{\gamma d}\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} + \boldsymbol{Z}_{2\gamma d}\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d} - \boldsymbol{Z}_\gamma^{(3)} + \boldsymbol{Z}_{2\gamma}^{(3)}](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} \\
& - 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma + \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma^{(2)} - \boldsymbol{Z}_{2\gamma}^{(2)}) \\
& + (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d} + \boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d}) + 2\boldsymbol{Z}_{2\gamma d}(\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} - \boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d}) \\
& + 2\boldsymbol{Z}_{2\gamma}^{(2)} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} + 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,4)}\boldsymbol{\Phi}_1^4 \\
& + 6\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^2\boldsymbol{\Phi}_2 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_2^2 + 4\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_3)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d\boldsymbol{Z}_{2\gamma d}]\mathbf{1} \\
& - 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,4)}\boldsymbol{\Phi}_1^4 + 6\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^2\boldsymbol{\Phi}_2 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_2^2 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_3) \\
& \times [(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(3\boldsymbol{Z}_{2\gamma d} + \boldsymbol{Z}_{\gamma d})]\mathbf{1} + 12\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,4)}\boldsymbol{\Phi}_1^4 + 5\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^2\boldsymbol{\Phi}_2 \\
& + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_2^2 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_3)[\boldsymbol{Z}_{\gamma d}^{(2)} - \boldsymbol{Z}_{2\gamma d}^{(2)}]\mathbf{1},
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
A_{1,\beta\gamma} = {} & 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot \boldsymbol{Z}_{2\beta} \odot \boldsymbol{Z}_{2\gamma} + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\beta d} \\
& + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\beta d} + \boldsymbol{Z}_{2\beta d}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\beta d} + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot \boldsymbol{Z}_{2\gamma} \odot \boldsymbol{Z}_{2\beta} \\
& + 2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot \boldsymbol{Z}_{2\beta}^{(2)}](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} - 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1) \\
& \times [(\boldsymbol{Z}_{2\gamma} + \boldsymbol{Z}_\gamma) \odot (\boldsymbol{Z}_\beta^{(2)} - \boldsymbol{Z}_{2\beta}^{(2)}) + 2\boldsymbol{Z}_{2\beta}^{(2)} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} \\
& - 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\beta d} + 2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \\
& \times \boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\beta d} + \boldsymbol{Z}_{2\gamma d}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\beta d}](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} - 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1) \\
& \times [(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\gamma d} + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d} + \boldsymbol{Z}_{2\beta d}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
& \times \boldsymbol{Z}_{2\gamma d}](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} + 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \\
& \times (\boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d} + \boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d}) + 2\boldsymbol{Z}_{2\beta d}(\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} - \boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d})](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} \\
& + 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,2)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1^2 + \boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d\boldsymbol{Z}_{2\beta d} + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{\gamma d}]\mathbf{1},
\end{aligned}
\tag{3.6}
$$

$$
\begin{aligned}
A_{2,\beta} = &-3\ \mathbf{1}^\top(2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})\boldsymbol{Z}_{2\beta d} \\
&+ (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\beta}(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d + 2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(2)} \odot \boldsymbol{Z}_{2\beta} + 3/4 \\
&\times (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d + 1/2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(3)}] \\
&\times (2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1} + 6\ \mathbf{1}^\top(2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2) \\
&\times [(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta^{(2)} - \boldsymbol{Z}_{2\beta}^{(2)}) + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\beta\boldsymbol{Z}_{\beta d} - \boldsymbol{Z}_{2\beta}\boldsymbol{Z}_{2\beta d})] \\
&\times (\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1} - 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(4,0)}\boldsymbol{M}_1^4 + 6\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^2\boldsymbol{M}_2 \\
&+ 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_2^2 + 2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_3)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d^{(2)}]\mathbf{1},
\end{aligned} \tag{3.7}
$$

$$
\begin{aligned}
A_{2,\gamma} = &-3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\gamma d} \\
&+ (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d\boldsymbol{Z}_{2\gamma}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d + 2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})^{(2)} \odot \boldsymbol{Z}_{2\gamma} + 3/4 \\
&\times (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d + 1/2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})^{(3)}] \\
&\times (\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} + 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2) \\
&\times [(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma^{(2)} - \boldsymbol{Z}_{2\gamma}^{(2)}) + (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} - \boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d})] \\
&\times (\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} + 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,4)}\boldsymbol{\Phi}_1^4 + 6\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^2\boldsymbol{\Phi}_2 \\
&+ 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_2^2 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_3)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d^{(2)}]\mathbf{1},
\end{aligned} \tag{3.8}
$$

$$
\begin{aligned}
A_{2,\beta\gamma} = &-3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(2)} \odot \boldsymbol{Z}_{2\gamma} + 3/2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(2)}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
&+ (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})\boldsymbol{Z}_{2\beta d} + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\gamma}(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d + 3/4 \\
&\times (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d + 4(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot \boldsymbol{Z}_{2\beta}] \\
&\times (\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} + 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\beta^{(2)} - \boldsymbol{Z}_{2\beta}^{(2)})] \\
&\times (\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} + 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
&\times \boldsymbol{Z}_{2\beta d} + (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d\boldsymbol{Z}_{2\gamma}(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d + 3/4(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
&\times (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} + 3\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
&\times \boldsymbol{Z}_{2\gamma d} + (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d\boldsymbol{Z}_{2\gamma}(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d + 3/4(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
&\times (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1} - 6\ \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \\
&\times (\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} - \boldsymbol{Z}_{2\gamma}\boldsymbol{Z}_{2\gamma d})](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 2\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1},
\end{aligned} \tag{3.9}
$$

$$
\begin{aligned}
A_{3,\beta} = &\ \mathbf{1}^\top(2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)[3/4(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \\
&+ 1/2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(3)}](2\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3 + 3\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2)\mathbf{1},
\end{aligned} \tag{3.10}
$$

$$
\begin{aligned}
A_{3,\gamma} = \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[3/4(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \\
+ 1/2(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})^{(3)}](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1}
\end{aligned} \tag{3.11}
$$

e

$$
\begin{aligned}
A_{3,\beta\gamma} = \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[3/2(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})^{(2)} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) + 3/4 \\
\times (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} \\
- \mathbf{1}^\top(\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)[3/4(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \\
\times (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d](\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)\mathbf{1} - \mathbf{1}^\top(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{M}_1^2\boldsymbol{\Phi}_1)[3/4(\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \\
\times (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})(\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d](\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3 + 3\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2)\mathbf{1},
\end{aligned}
$$

em que $\mathbf{1} = (1, \ldots, 1)^\top$ é um vetor de uns n-dimensional.

As quantidades A_1, A_2 e A_3 são funções de cumulantes de derivadas da função de log-verossimilhança e de derivadas desses cumulantes, ver Apêndice B. Adicionalmente, se tais quantidades envolverem parâmetros desconhecidos, estes devem ser substituídos por suas respectivas EMV sob a hipótese nula $(\boldsymbol{\beta}_1^{(0)\top}, \widetilde{\boldsymbol{\beta}}_2^\top, \boldsymbol{\gamma}_1^{(0)\top}, \widetilde{\boldsymbol{\gamma}}_2^\top)$. Assim, a estatística gradiente corrigida é obtida como em (3.2) utilizando a estatística ST definida na Seção 2.4.1 e os A's obtidos acima.

Agora, suponha que o interesse é testar os efeitos da média e da dispersão separadamente. Primeiro, consideramos o teste sobre um subconjunto do vetor de parâmetros $\boldsymbol{\beta}$. A hipótese nula de interesse é $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ a ser testada contra $\mathcal{H}_1^2 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_1^{(0)}$. A estatística gradiente para testar $\mathcal{H}_0^2$ é apresentada na Seção 2.4.2. Neste caso, os A's são dados por

$$
A_1 = A_{1,\boldsymbol{\beta}} + A_{1,\boldsymbol{\beta\gamma}}, \quad A_2 = A_{2,\boldsymbol{\beta}} + A_{2,\boldsymbol{\beta\gamma}} \quad \text{e} \quad A_3 = A_{3,\boldsymbol{\beta}},
$$

em que $A_{1,\boldsymbol{\beta}}$, $A_{2,\boldsymbol{\beta}}$ e $A_{3,\boldsymbol{\beta}}$ são como nas equações (3.4), (3.7) e (3.10), respectivamente. Já as quantidades $A_{1,\boldsymbol{\beta\gamma}}$ e $A_{2,\boldsymbol{\beta\gamma}}$ são obtidas das expressões (3.6) e (3.9), respectivamente, com $\boldsymbol{Z}_\gamma = \boldsymbol{Z}_{2\gamma}$.

Por fim, consideramos o teste sobre um subconjunto do vetor de parâmetros $\boldsymbol{\gamma}$. A hipótese nula de interesse é $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1^3 : \boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_1^{(0)}$. A estatística gradiente para o teste de $\mathcal{H}_0^3$ é definida na Seção 2.4.3. As quantidades A_1, A_2 e A_3 podem ser expressas como

$$
A_1 = A_{1,\boldsymbol{\gamma}} + A_{1,\boldsymbol{\beta\gamma}}, \quad A_2 = A_{2,\boldsymbol{\gamma}} + A_{2,\boldsymbol{\beta\gamma}} \quad \text{e} \quad A_3 = A_{3,\boldsymbol{\gamma}},
$$

em que $A_{1,\gamma}$, $A_{2,\gamma}$ e $A_{3,\gamma}$ são como em (3.5), (3.8) e (3.11), respectivamente, e as quantidades $A_{1,\boldsymbol{\beta\gamma}}$ e $A_{2,\boldsymbol{\beta\gamma}}$ são dadas pelas expressões (3.6) e (3.9), respectivamente, com $\boldsymbol{Z}_\beta = \boldsymbol{Z}_{2\beta}$.

Os resultados obtidos nesta seção generalizam aqueles apresentados em Campos (2014), que derivou um fator de correção tipo-Bartlett à estatística gradiente para testar $\mathcal{H}_0^3$ em MLGS. Vale salientar aqui que a expressão do fator de correção tipo-Bartlett para esse teste é diferente daquela publicada em Campos (2014, p. 35 e 36). A diferença é que a autora utilizou no cálculo em A_1, o cumulante $\kappa_{jr}^{(su)}$ ao invés do cumulante $\kappa_{rs}^{(ju)}$ (cumulante apresentado em Vargas et al. (2013, p. 46)).

3.3 CORREÇÃO DE BARTLETT PARA A ESTATÍSTICA DA RAZÃO DE VEROSSIMILHANÇAS EM MLGS

Cordeiro et al. (2006) derivaram um fator de correção de Bartlett em notação matricial para a estatística LR (veja a Seção 2.4.1) para testar $\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}, \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ em MLGS. A estatística LR modificada é da forma

$$LR^* = \frac{LR}{c_{\,1}},$$

sendo $c_{\,1} = 1 + \big(\epsilon_{p+q} - \epsilon_{p-p_1+q-q_1}\big)\big/\big(p_1 + q_1\big)$ o fator de correção de Bartlett. O termo ϵ_{p+q} pode ser decomposto como

$$\epsilon_{p+q} = \epsilon_p\big(\beta\big) + \epsilon_q\big(\gamma\big) + \epsilon_{p,q}\big(\beta, \gamma\big), \tag{3.12}$$

em que

$$\begin{aligned}
\epsilon_p\big(\beta\big) = {} & \frac{1}{4}\,\mathbf{1}^\top\big(\boldsymbol{\Psi}^{(4,0)}\boldsymbol{M}_1^4 + 4\boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^2\boldsymbol{M}_2 + \boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_2^2\big)\boldsymbol{Z}_{\beta d}^{(2)}\mathbf{1} \\
& - \frac{1}{3}\,\mathbf{1}^\top\big[\big(\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2 + \boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3\big)\boldsymbol{Z}_\beta^{(3)}\big(2\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2 + \boldsymbol{\Psi}^{(3,0)}\boldsymbol{M}_1^3\big)\big]\mathbf{1} \\
& + \frac{1}{12}\,\mathbf{1}^\top\boldsymbol{\Psi}^{(2,0)}\boldsymbol{M}_1\boldsymbol{M}_2\big(2\boldsymbol{Z}_\beta^{(3)} + 3\boldsymbol{Z}_{\beta d}\boldsymbol{Z}_\beta\boldsymbol{Z}_{\beta d}\big)\boldsymbol{M}_1\boldsymbol{M}_2\boldsymbol{\Psi}^{(2,0)}\mathbf{1},
\end{aligned} \tag{3.13}$$

$$\begin{aligned}
\epsilon_q\big(\gamma\big) = {} & \frac{1}{4}\,\mathbf{1}^\top\big(\boldsymbol{\Psi}^{(0,4)}\boldsymbol{\Phi}_1^4 + 2\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^2\boldsymbol{\Phi}_2 - \boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_2^2\big)\boldsymbol{Z}_{\gamma d}^{(2)}\mathbf{1} \\
& + \frac{1}{12}\,\mathbf{1}^\top\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3(2\boldsymbol{Z}_\gamma^{(3)} + 3\boldsymbol{Z}_{\gamma d}\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d})\boldsymbol{\Phi}_1^3\boldsymbol{\Psi}^{(0,3)}\mathbf{1} \\
& + \frac{1}{4}\,\mathbf{1}^\top\boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2(\boldsymbol{Z}_{\gamma d}\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d} + 2\boldsymbol{Z}_\gamma^{(3)})\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2\boldsymbol{\Psi}^{(0,2)}\mathbf{1} \\
& + \frac{1}{2}\,\mathbf{1}^\top\boldsymbol{\Psi}^{(0,3)}\boldsymbol{\Phi}_1^3\boldsymbol{Z}_{\gamma d}\boldsymbol{Z}_\gamma\boldsymbol{Z}_{\gamma d}\boldsymbol{\Phi}_1\boldsymbol{\Phi}_2\boldsymbol{\Psi}^{(0,2)}\mathbf{1}
\end{aligned} \tag{3.14}$$

e

$$\epsilon_{p,q}(\beta,\gamma) = \frac{1}{2}\,\mathbf{1}^\top\big[\big(\boldsymbol{\Psi}^{(2,1)}\boldsymbol{\Phi}_2 + \boldsymbol{\Psi}^{(2,2)}\boldsymbol{\Phi}_1^2\big)\boldsymbol{M}_1^2\boldsymbol{Z}_{\beta d}\boldsymbol{Z}_{\gamma d} - \boldsymbol{\Psi}^{(2,1)}\boldsymbol{\Phi}_1\boldsymbol{M}_1^2\boldsymbol{Z}_\beta^{(2)}\boldsymbol{Z}_\gamma\boldsymbol{\Phi}_1\boldsymbol{M}_1^2\boldsymbol{\Psi}^{(2,1)}\big]\mathbf{1} + \frac{1}{4}\,\mathbf{1}^\top\boldsymbol{\Psi}^{(2,1)}\boldsymbol{\Phi}_1\boldsymbol{M}_1^2\boldsymbol{Z}_{\beta d}\boldsymbol{Z}_\gamma\big[\boldsymbol{Z}_{\beta d}\boldsymbol{M}_1^2\boldsymbol{\Psi}^{(2,1)} + 2\boldsymbol{Z}_{\gamma d}\big(\boldsymbol{\Phi}_1^2\boldsymbol{\Psi}^{(0,3)} + \boldsymbol{\Phi}_2\boldsymbol{\Psi}^{(0,2)}\big)\big]\boldsymbol{\Phi}_1\mathbf{1}. \tag{3.15}$$

O termo $\epsilon_{p-p_1+q-q_1}$ é obtido de (3.12)-(3.15) com $\boldsymbol{Z}_\gamma = \boldsymbol{Z}_{2\gamma}$ e $\boldsymbol{Z}_\beta = \boldsymbol{Z}_{2\beta}$, ou seja, $\epsilon_{p-p_1+q-q_1} = \epsilon_p\big(\beta_2\big) + \epsilon_q\big(\gamma_2\big) + \epsilon_{p-p_1,q-q_1}\big(\beta_2,\gamma_2\big)$. O erro de aproximação χ^2 para a distribuição de LR é de ordem $O(n^{-1})$, enquanto o erro desta aproximação para a distribuição de LR^* é reduzido para $O(n^{-2})$.

Para testar $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ contra $\mathcal{H}_1^2 : \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_1^{(0)}$, o fator de correção de Bartlett é dado por $c_{\,2} = 1 + \big\{\epsilon_p\big(\beta\big) - \epsilon_{p_2}\big(\beta_2\big) + \epsilon_{p,q}\big(\beta,\gamma\big) - \epsilon_{p_2,q}\big(\beta_2,\gamma\big)\big\}/p_1$, em que os termos $\epsilon_{p_2}\,(\beta_2)$ e $\epsilon_{p_2,q}\big(\beta_2,\gamma\big)$ são como (3.13) e (3.15), respectivamente, com $\boldsymbol{Z}_\beta = \boldsymbol{Z}_{2\beta}$.

Já a correção de Bartlett para o teste de $\mathcal{H}_0^3$ é da forma $c_{\,3} = 1 + \big\{\epsilon_q\big(\gamma\big) - \epsilon_{q_2}\big(\gamma_2\big) + \epsilon_{p,q}\big(\beta,\gamma\big) - \epsilon_{p,q_2}\big(\beta,\gamma_2\big)\big\}/q_1$, em que $\epsilon_{q_2}\big(\gamma_2\big)$ e $\epsilon_{p,q_2}\big(\beta,\gamma_2\big)$ são obtidos de (3.14) e (3.15), respectivamente, com $\boldsymbol{Z}_\gamma = \boldsymbol{Z}_{2\gamma}$.

3.4 CORREÇÃO TIPO-BARTLETT PARA A ESTATÍSTICA ESCORE EM MLGS

Terra (2013) obteve correções para o teste escore em MNLGS. Nesta seção, apresentamos um caso particular desse resultado, a saber, o fator de correção tipo-Bartlett para a estatística SR (ver a Seção 2.4.1) para testar $\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}, \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ em MLGS. A estatística escore aperfeiçoada foi obtida utilizando os resultados de Cordeiro e Ferrari (1991), e é dada por

$$SR^* = SR\big[1 - \big(c_R + b_R SR + a_R SR^2\big)\big],$$

em que

$$a_R = \frac{A_{R3}}{12\nu(\nu+2)(\nu+4)},$$

$$b_R = \frac{A_{R2} - 2A_{R3}}{12\nu(\nu+2)},$$

$$c_R = \frac{A_{R1} - A_{R2} + A_{R3}}{12\nu},$$

com $\nu = p_1 + q_1$. O fator $\big[1 - \big(c_R + b_R SR + a_R SR^2\big)\big]$ é denominado correção tipo-Bartlett. A estatística SR^* tem distribuição, segundo a hipótese nula, mais próxima da distribuição qui-

quadrado de referência do que a estatística SR, o erro de aproximação é reduzido de $O(n^{-1})$ para $O(n^{-2})$.

As quantidades A_{R1}, A_{R2} e A_{R3} podem ser escritas como

$$A_{R1} = 3A_{R11} - 6A_{R12} + 6A_{R13} - 6A_{R14},$$
$$A_{R2} = -3A_{R21} + 6A_{R22} - 6A_{R23} + 3A_{R24},$$
$$A_{R3} = 3A_{R31} + 2A_{R32},$$

em que

$$\begin{aligned}
A_{R11} = {} & \mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{Z}_{2\beta d} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{\Psi}^{(2,0)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\beta d} \boldsymbol{\Phi}_1 \boldsymbol{M}_1^2 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
& + 2\, \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{Z}_{2\gamma d} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1},
\end{aligned} \tag{3.16}$$

$$\begin{aligned}
A_{R12} = {} & \mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta d} \boldsymbol{Z}_{2\beta} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta d} \boldsymbol{Z}_{2\beta} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\gamma d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
& - \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\gamma d} \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma d})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1},
\end{aligned} \tag{3.17}$$

$$
\begin{aligned}
A_{R13} = &-2\,\mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 \boldsymbol{Z}_{2\beta}^{(2)} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{\Psi}^{(2,0)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta}^{(2)} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{\Psi}^{(2,0)} \mathbf{1} \\
&+ 2\,\mathbf{1}^\top \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{Z}_\beta \odot \boldsymbol{Z}_\gamma \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{Z}_{2\gamma}^{(2)} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma}^{(2)} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\gamma}^{(2)} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1},
\end{aligned}
\tag{3.18}
$$

$$
\begin{aligned}
A_{R14} = &\,\mathbf{1}^\top \left[2 \boldsymbol{\Psi}^{(3,0)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{M}_1^4 - \boldsymbol{\Psi}^{(4,0)} \boldsymbol{M}_1^4 - \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^2 \boldsymbol{M}_2 \right] \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \mathbf{1} \\
&+ \mathbf{1}^\top \left[2 \boldsymbol{\Psi}^{(2,1)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 - \boldsymbol{\Psi}^{(2,2)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 + \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_2 \boldsymbol{\Phi}_1^2 \right] \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \mathbf{1} \\
&+ \mathbf{1}^\top \left[2 \boldsymbol{\Psi}^{(2,1)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 - \boldsymbol{\Psi}^{(2,2)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 - \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_2 \right] \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \mathbf{1} \\
&+ \mathbf{1}^\top \left[2 \boldsymbol{\Psi}^{(1,2)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{\Phi}_1^4 - \boldsymbol{\Psi}^{(0,4)} \boldsymbol{\Phi}_1^4 - \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \right] \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \mathbf{1},
\end{aligned}
\tag{3.19}
$$

$$
\begin{aligned}
A_{R21} = &\,\mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{Z}_{2\beta} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
&+ 2\,\mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{Z}_{2\beta} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
&+ 2\,\mathbf{1}^\top \boldsymbol{M}_1^2 \boldsymbol{\Phi}_2 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{Z}_{2\beta} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{Z}_{2\gamma} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1},
\end{aligned}
\tag{3.20}
$$

$$
\begin{aligned}
A_{R22} = &\,\mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
&+ \mathbf{1}^\top \boldsymbol{\Psi}^{(2,0)} \boldsymbol{M}_1 \boldsymbol{M}_2 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma d})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} \\
&- \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1},
\end{aligned}
\tag{3.21}
$$

$$
\begin{aligned}
A_{R23} = {} & \mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 \boldsymbol{Z}_{2\beta} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\gamma} \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + 2\, \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{Z}_{2\beta} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + 2\, \mathbf{1}^\top \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{Z}_{2\gamma} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{Z}_{2\beta} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \boldsymbol{Z}_{2\gamma} \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{0,3} \mathbf{1},
\end{aligned}
\tag{3.22}
$$

$$
\begin{aligned}
A_{R24} = {} & \mathbf{1}^\top \left[3 \boldsymbol{\Psi}^{(3,0)^2} \boldsymbol{M}_1^4 \boldsymbol{\Psi}^{(4,0)^{-1}} \boldsymbol{M}_1^4 \right] (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \mathbf{1} \\
& + 2 \left[\boldsymbol{\Psi}^{(3,0)^2} \boldsymbol{\Psi}^{(1,2)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 + 2 \boldsymbol{\Psi}^{(2,1)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 \right. \\
& \left. - \boldsymbol{\Psi}^{(2,2)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 - \boldsymbol{\Psi}^{(2,0)} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1^2 \right] (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{\gamma d})_d \mathbf{1} \\
& + \mathbf{1}^\top \left[3 \boldsymbol{\Psi}^{(1,2)^2} \boldsymbol{\Psi}^{(2,0)^{-1}} \boldsymbol{\Phi}_1^4 - \boldsymbol{\Psi}^{(0,4)} \boldsymbol{\Phi}_1^4 \right] \boldsymbol{Z}_{2\gamma d} (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \mathbf{1},
\end{aligned}
\tag{3.23}
$$

$$
\begin{aligned}
A_{R31} = {} & \mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
& + 2\, \mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& + 2\, \mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1^{(2)} \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta})_d \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma})_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1}
\end{aligned}
\tag{3.24}
$$

e

$$
\begin{aligned}
A_{R32} = {} & \mathbf{1}^\top \boldsymbol{\Psi}^{(3,0)} \boldsymbol{M}_1^3 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1^3 \boldsymbol{\Psi}^{(3,0)} \mathbf{1} \\
& + 3\, \mathbf{1}^\top \boldsymbol{\Psi}^{(2,1)} \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \boldsymbol{M}_1^2 \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(2,1)} \mathbf{1} \\
& + 3\, \boldsymbol{\Psi}^{(1,2)} \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\beta - \boldsymbol{Z}_{2\beta}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{M}_1 \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(1,2)} \mathbf{1} \\
& + \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \odot (\boldsymbol{Z}_\gamma - \boldsymbol{Z}_{2\gamma}) \boldsymbol{\Phi}_1^3 \mathbf{1}.
\end{aligned}
\tag{3.25}
$$

O fator de correção tipo-Bartlett para a estatística escore para testar $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ é obtido de (3.16)-(3.25) substituindo $\boldsymbol{Z}_\gamma$ por $\boldsymbol{Z}_{2\gamma}$. Para testar $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$, a correção tipo-Bartlett também é obtida das expressões (3.16)-(3.25), porém, com $\boldsymbol{Z}_\beta = \boldsymbol{Z}_{2\beta}$. Adicionalmente, Terra (2013) somente apresentou resultados de simulação para a estatística escore e sua versão corrigida para o teste de $\mathcal{H}_0^3$.

4 RESULTADOS NUMÉRICOS

Neste capítulo, apresentamos os resultados de simulações de Monte Carlo para avaliar o desempenho dos testes baseados na seguintes estatísticas: gradiente (ST), gradiente corrigida (ST^*), gradiente bootstrap (ST_b), razão de verossimilhanças (LR), razão de verossimilhanças corrigida (LR^*), razão de verossimilhanças bootstrap (LR_b), escore (SR), escore corrigida (SR^*) e escore bootstrap (SR_b). Avaliamos o desempenho dos testes em função da proximidade da probabilidade de rejeição da hipótese nula, sendo esta verdadeira (probabilidade do erro tipo I), aos respectivos níveis nominais. Nas simulações consideramos três hipóteses nulas distintas:

$$\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \mathbf{0}, \boldsymbol{\gamma}_1 = \mathbf{0}, \quad \mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0} \quad \text{e} \quad \mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}$$

contra hipóteses alternativas bilaterais. As dimensões dos vetores $\boldsymbol{\beta}_1$ e $\boldsymbol{\gamma}_1$ são p_1 e q_1, respectivamente, e $\mathbf{0}$ representa vetores de zeros de dimensões apropriadas. Vale salientar que os testes da razão de verossimilhanças e escore e suas versões corrigidas e bootstrap somente foram apresentados para o teste de $\mathcal{H}_0^2$.

Os resultados de simulação são baseados no modelo Poisson duplo, cuja função de probabilidade é definida em (2.4). A seguir, descrevemos a geração da uma variável aleatória Y com distribuição Poisson dupla, denotada por $\mathcal{PD}(\mu, \phi)$. Seja X uma variável aleatória que segue distribuição Poisson com parâmetro $\mu\phi$, então, temos que

$$Y = \frac{X}{\phi}$$

segue distribuição $\mathcal{PD}(\mu, \phi)$. Os componentes sistemáticos da média e da dispersão são dados, respectivamente, por

$$\log(\mu_l) = \beta_0 + \beta_1 x_{l1} + \beta_2 x_{l2} + \cdots + \beta_p x_{lp},$$

$$\log(\phi_l) = \gamma_0 + \gamma_1 s_{l1} + \gamma_2 s_{l2} + \cdots + \gamma_q s_{lq},$$

com $l = 1, \ldots, n$. A variável resposta foi gerada assumindo que $\beta_0 = 3.5$, $\beta_1 = \cdots = \beta_p = 0.1$, $\gamma_0 = -0.5$ e $\gamma_1 = \cdots = \gamma_q = -0.01$, e as covariáveis foram geradas como amostras aleatórias da distribuição $\mathcal{U}(0, 1)$. Os tamanhos amostrais considerados são $n = 20$, 30, 40 e 50, e os níveis nominais são $\alpha = 5\%$ e 10%. As simulações foram realizadas utilizando a linguagem de programação `Ox` (Doornik, 2006). Para a maximização da função de log-verossimilhança utilizamos o algoritmo de otimização não-linear BFGS com derivadas analíticas. O número

de réplicas de Monte Carlo foi fixado em 10000 e 1000 réplicas bootstrap. Para cada tamanho amostral e cada nível considerado, calculamos as taxas de rejeição, ou seja, estimamos via simulação Pr($ST \geqslant x_\alpha$), Pr($ST^* \geqslant x_\alpha$), Pr($LR \geqslant x_\alpha$), Pr($LR^* \geqslant x_\alpha$), Pr($SR \geqslant x_\alpha$) e Pr($SR^* \geqslant x_\alpha$), em que x_α é o quantil $1-\alpha$ da distribuição qui-quadrado de referência.

Inicialmente, nos detemos em testar simultaneamente os efeitos da média e da dispersão. A hipótese nula a ser considerada é $\mathcal{H}_0^1 : \boldsymbol{\beta}_1 = \mathbf{0}, \boldsymbol{\gamma}_1 = \mathbf{0}$, com ($p_1 = q_1 = 1, p = q = 2$) e ($p_1 = q_1 = 2, p = q = 3$), sendo p_1 e q_1 as dimensões dos vetores $\boldsymbol{\beta}_1$ e $\boldsymbol{\gamma}_1$, respectivamente (ver a Tabela 1). Notamos que o teste gradiente é liberal, rejeitando a hipótese nula com maior frequência do que o esperado com base nos níveis nominais. Por exemplo, para $n = 20$ e ($p_1 = q_1 = 1$) e ($p_1 = q_1 = 2$), as taxas de rejeição nula do teste ST correspondente a $\alpha = 10\%$ são $12,08\%$ e $17,15\%$, respectivamente. Vale ressaltar, contudo, que o teste gradiente tende a funcionar melhor com o aumento do tamanho da amostra, como esperado. O teste ST^*, por outro lado, apresentou um bom desempenho, mesmo em tamanhos de amostra pequenos. Para $n = 30$, ($p_1 = q_1 = 2$) e $\alpha = 10\%$ a taxa de rejeição é $10,33\%$ para o teste ST^*. Em geral, a tendência do teste gradiente em rejeitar demasiadamente a hipótese nula é atenuada pela correção tipo-Bartlett, de modo que o teste corrigido, ST^*, apresenta taxas de rejeição mais próximas dos respectivos níveis nominais.

Nosso interesse agora é testar somente os efeitos da média. Neste caso, a hipótese nula a ser considerada é $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}$, com ($p_1 = 2, p = q = 3$), ($p_1 = 3, p = q = 4$) e ($p_1 = 4, p = q = 5$). Observe que consideramos três situações distintas entre si (ver a Tabela 2). Mais uma vez, notamos que o teste gradiente apresenta taxas de rejeição acima dos respectivos níveis nominais. Notamos também que o número de covariáveis de dispersão, bem como o número de parâmetros de interesse, p_1, impactam consideravelmente o desempenho do teste gradiente, resultando em taxas de rejeição ainda mais distorcidas. Por exemplo, para $n = 20$ e ($p_1 = 2, p = q = 3$) e ($p_1 = 4, p = q = 5$), as taxas de rejeição do teste ST correspondente a $\alpha = 10\%$ são $16,17\%$ e $20,58\%$, respectivamente. O teste gradiente corrigido, por sua vez, teve bom desempenho em todos os cenários descritos na Tabela 2. Considerando a mesma situação anterior, ou seja, quando $n = 20$ e ($p_1 = 2, p = q = 3$) e ($p_1 = 4, p = q = 5$), com $\alpha = 10\%$ as taxas de rejeição são, respectivamente, $10,28\%$ e $9,77\%$ para o teste ST^*. É interessante observar que, neste caso, o teste gradiente corrigido não foi afetado significativamente pelo número de covariáveis no modelo de dispersão, visto que este apresentou taxas de rejeição bem próximas dos níveis nominais considerados. O mesmo não ocorre com o teste gradiente que, nas mesmas circunstâncias, apresentou taxas de rejeição bastante afastadas do respectivos níveis

nominais.

Agora, consideramos a hipótese nula $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}$, com $(q_1 = 2, p = q = 3)$, $(q_1 = 3, p = q = 4)$ e $(q_1 = 4, p = q = 5)$. Aqui estamos interessados em testar somente os efeitos da dispersão. Os resultados são apresentados na Tabela 3. Verificamos que o teste gradiente apresentou taxas de rejeição muito distorcidas, mesmo em tamanhos de amostra relativamente grandes. Esse fato se agrava ainda mais, com o aumento do número de covariáveis da média. Por exemplo, quando $n = 20$, $(q_1 = 2, p = q = 3)$ e $\alpha = 10\%$ a taxa de rejeição do teste ST é $24,94\%$, mais que o dobro do nível nominal considerado. Para este mesmo caso, a taxa de rejeição do teste ST^* é $15,52\%$. É evidente que a teste gradiente corrigido mostrou desempenho superior quando comparado com o teste gradiente usual. No entanto, verificamos que as taxas de rejeição apresentadas pelo teste gradiente corrigido foram próximas, mas não o suficiente, dos respectivos níveis nominais.

Adicionalmente, analisamos a influência do número de parâmetros de interesse no desempenho dos testes. No primeiro caso, a hipótese nula considerada é $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}$, com $p_1 = 3, 4, 5, 6$, $q = 3$ e $p = p_1 + 1$, sendo p_1 e q as dimensões dos vetores $\boldsymbol{\beta}_1$ e $\boldsymbol{\gamma}$, respectivamente. Os resultados são apresentados na Tabela 4. Observamos, neste caso, que as taxas de rejeição do teste gradiente e do teste gradiente corrigido mantiveram-se estáveis diante do aumento do número de parâmetros de interesse, p_1. É possível notar que o teste gradiente usual é liberal. Já o teste gradiente corrigido apresenta leve distorção de tamanho. Por exemplo, para $p_1 = 6$ e $n = 30$, as taxas de rejeição dos testes ST e ST^* correspondente a $\alpha = 10\%$ são, respectivamente, $15,11\%$ e $10,84\%$.

No segundo caso, a hipótese nula em consideração é $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}$, com $q_1 = 3, 4, 5, 6$, $p = 3$ e $q = q_1 + 1$, sendo q_1 e p as dimensões dos vetores $\boldsymbol{\gamma}_1$ e $\boldsymbol{\beta}$, respectivamente (ver a Tabela 5). É possível observar aqui que o teste gradiente usual é bastante liberal, de modo que suas taxas de rejeição se afastam dos níveis nominais, à medida que o número de parâmetros de interesse aumenta. O teste gradiente corrigido, por sua vez, também é afetado pelo número de parâmetros de interesse, q_1, porém de forma menos acentuada. Por exemplo, para $q_1 = 4$, $n = 40$ e $\alpha = 5\%$ a taxa de rejeição do teste ST é $9,30\%$, enquanto a taxa de rejeição do teste ST^* é $5,55\%$.

Verificamos também o efeito do número de parâmetros de perturbação no desempenho dos testes. No primeiro cenário, a hipótese nula em consideração é $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}$, com $p_1 = 2$, $q = 3$ e $p = 4, 5, 6, 7$, sendo p a dimensão do vetor $\boldsymbol{\beta}$ (Tabela 6). Observamos que o teste gradiente apresenta taxas de rejeição distorcidas, as mesmas estão muito afastadas dos respectivos níveis nominais principalmente quando o número de covariáveis no modelo da

média aumenta. O mesmo ocorre com o teste gradiente corrigido, no entanto seu desempenho é visivelmente melhor do que o teste gradiente usual. Por exemplo, quando $p = 7$, $n = 30$ e $\alpha = 10\%$, as taxas de rejeição são $23,16\%$ e $15,24\%$ para os testes ST e ST^*, respectivamente.

No segundo cenário, a hipótese nula em teste é $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}$, com $q_1 = 2$, $p = 2$ e $q = 4, 5, 6, 7$, sendo q a dimensão do vetor $\boldsymbol{\gamma}$ (Tabela 7). Neste caso, o teste gradiente também apresenta taxas de rejeição afastadas do níveis nominais considerados, isso é mais evidente com o aumento do número de covariáveis no modelo de dispersão. Em contra partida, o teste gradiente corrigido apresentou um bom desempenho, com suas taxas de rejeição bem mais próximas dos níveis nominais. Para $q = 6$, $n = 20$ e $\alpha = 5\%$, as taxas de rejeição dos testes ST e ST^* são $13,69\%$ e $5,59\%$, respectivamente.

Na Gráfico 1 apresentamos gráficos quantil-quantil (quantil exato versus quantil assintótico) das estatísticas ST e ST^*. Consideramos a seguinte situação: $\mathcal{H}_0^1 : \beta_1 = 0, \gamma_1 = 0$, com $p = q = 2$. Os quantis assintóticos são obtidos da distribuição qui-quadrado com dois graus de liberdade. A linha contínua representa a identidade, ou seja, os quantis assintóticos são iguais aos quantis exatos. Verificamos que a estatística ST (linha pontilhada) apresenta quantis bem superiores aos quantis da distribuição qui-quadrado. Já os quantis da estatística ST^* (linha tracejada) estão bem próximos dos quantis da distribuição qui-quadrado. Os resultados evidenciam que a distribuição nula da estatística ST^* está muito próxima da respectiva distribuição assintótica.

Na Gráfico 2 também apresentamos gráficos quantil-quantil das estatísticas gradiente e gradiente corrigida. Consideramos, porém, $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}$, com $p_1 = 2$, $p = 3$ e $q = 3$. Novamente, a estatística ST apresenta quantis bem superiores aos quantis da distribuição qui-quadrado. A estatística ST^*, por outro lado, apresenta quantis mais próximos dos quantis da distribuição de referência. Dessa forma, o teste baseado na estatística ST^* tem desempenho superior ao do teste baseado na estatística ST.

Por fim, na Gráfico 3, apresentamos gráficos quantil-quantil das estatísticas ST e ST^*, considerando $\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}$, com $q_1 = 2$, $q = 3$ e $p = 3$. Observamos que os resultados são semelhantes aos que foram apresentados acima. Novamente, a distribuição nula da estatística ST^* é bem aproximada pela distribuição de referência.

Tabela 1 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* variando o número de parâmetros de interesse da média e da dispersão simultaneamente e o tamanho da amostra.

		$p_1 = q_1 = 1$		$p_1 = q_1 = 2$	
n	α	ST	ST^*	ST	ST^*
20	5%	6,25	5,32	9,72	8,30
	10%	12,08	10,18	17,15	13,95
30	5%	6,03	5,33	6,06	5,23
	10%	11,24	9,86	12,62	10,33
40	5%	5,63	5,15	5,60	5,19
	10%	10,91	10,00	11,08	9,88
50	5%	5,45	4,99	5,39	5,09
	10%	11,16	10,13	10,96	9,95

Tabela 2 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* variando o número de parâmetros de interesse da média e o tamanho da amostra.

		$p_1 = 2, p = q = 3$		$p_1 = 3, p = q = 4$		$p_1 = 4, p = q = 5$	
n	α	ST	ST^*	ST	ST^*	ST	ST^*
20	5%	8,34	5,05	9,41	5,17	10,00	5,22
	10%	16,17	10,28	19,81	10,57	20,58	9,77
30	5%	7,05	5,07	7,88	4,94	9,62	4,91
	10%	13,85	10,35	16,32	10,15	19,76	10,22
40	5%	6,91	5,31	7,57	5,11	8,91	4,98
	10%	13,65	10,53	15,14	9,80	17,24	10,48
50	5%	6,27	5,05	6,92	4,82	7,91	4,85
	10%	11,96	10,03	13,67	9,85	15,50	10,06

Tabela 3 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* variando o número de parâmetros de interesse da dispersão e o tamanho da amostra.

		$q_1 = 2, p = q = 3$		$q_1 = 3, p = q = 4$		$q_1 = 4, p = q = 5$	
n	α	ST	ST^*	ST	ST^*	ST	ST^*
20	5%	16,89	8,99	24,96	11,40	40,29	15,16
	10%	24,94	15,52	35,28	18,02	51,88	23,13
30	5%	8,74	4,83	13,97	6,77	24,19	11,11
	10%	15,35	10,07	22,29	12,29	34,31	18,33
40	5%	7,98	5,28	10,24	5,58	15,06	6,55
	10%	14,48	10,58	17,82	10,88	24,22	12,88
50	5%	7,20	5,09	8,78	5,43	11,30	5,42
	10%	13,02	10,11	15,99	10,69	19,47	11,12

Tabela 4 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* aumentando o número de parâmetros de interesse da média com $q = 3$ $\left(\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}\right)$.

		$\alpha = 10\%$		$\alpha = 5\%$	
n	p_1	ST	ST^*	ST	ST^*
30	3	$14,59$	$9,99$	$6,59$	$4,59$
	4	$14,84$	$10,57$	$7,04$	$5,15$
	5	$15,09$	$10,73$	$6,51$	$4,86$
	6	$15,11$	$10,84$	$6,53$	$5,12$
40	3	$13,10$	$10,26$	$6,40$	$4,72$
	4	$13,00$	$10,14$	$6,29$	$4,92$
	5	$13,47$	$10,21$	$6,35$	$5,09$
	6	$13,82$	$10,87$	$6,36$	$5,15$

Tabela 5 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* aumentando o número de parâmetros de interesse da dispersão com $p = 3$ $\left(\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}\right)$.

		$\alpha = 10\%$		$\alpha = 5\%$	
n	q_1	ST	ST^*	ST	ST^*
30	3	$18,30$	$11,37$	$10,91$	$6,18$
	4	$20,93$	$12,39$	$12,47$	$6,56$
	5	$23,32$	$13,59$	$14,39$	$7,06$
	6	$25,94$	$14,44$	$15,71$	$8,18$
40	3	$14,92$	$10,63$	$8,36$	$5,46$
	4	$16,28$	$10,91$	$9,30$	$5,55$
	5	$17,01$	$10,73$	$9,49$	$5,65$
	6	$18,20$	$11,40$	$10,30$	$6,26$

Tabela 6 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* aumentando o número de parâmetros de perturbação da média com $p_1 = 2$ e $q = 3$ $\left(\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}\right)$.

		$\alpha = 10\%$		$\alpha = 5\%$	
n	p	ST	ST^*	ST	ST^*
20	4	$20,91$	$11,68$	$10,80$	$6,32$
	5	$24,89$	$15,28$	$13,75$	$9,76$
	6	$28,10$	$18,84$	$16,67$	$14,07$
	7	$32,48$	$23,83$	$20,13$	$18,71$
30	4	$15,58$	$10,35$	$7,80$	$5,46$
	5	$18,82$	$11,99$	$10,36$	$7,25$
	6	$20,48$	$13,31$	$11,67$	$8,83$
	7	$23,16$	$15,24$	$13,55$	$10,97$

Tabela 7 – Taxas de rejeição nula $(\%)$ dos testes ST e ST^* aumentando o número de parâmetros de perturbação da dispersão com $q_1 = 2$ e $p = 2$ $\left(\mathcal{H}_0^3 : \boldsymbol{\gamma}_1 = \mathbf{0}\right)$.

		$\alpha = 10\%$		$\alpha = 5\%$	
n	q	ST	ST^*	ST	ST^*
20	4	$19,73$	$11,23$	$11,64$	$6,13$
	5	$22,31$	$10,78$	$13,26$	$6,00$
	6	$21,98$	$10,02$	$13,69$	$5,59$
	7	$24,05$	$10,04$	$15,54$	$6,06$
30	4	$14,32$	$9,93$	$7,55$	$4,91$
	5	$15,63$	$9,42$	$8,53$	$4,93$
	6	$15,94$	$8,31$	$8,99$	$4,21$
	7	$17,34$	$8,00$	$9,97$	$4,13$

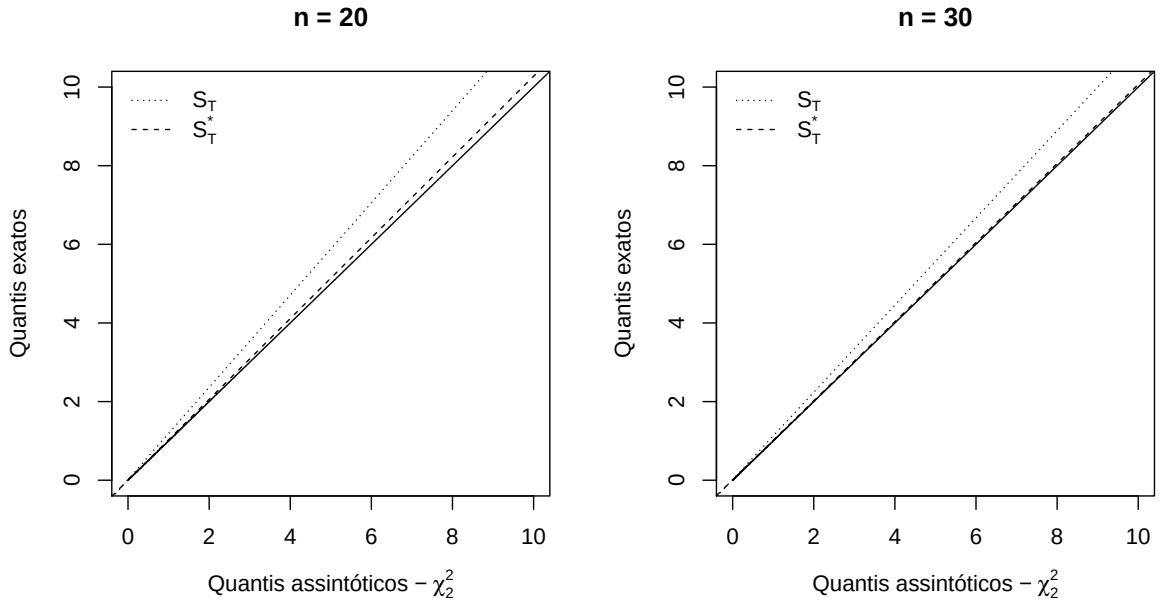

Gráfico 1 – Gráficos quantil-quantil das estatísticas S_T e S_T^*, considerando o teste dos efeitos da média e da dispersão simultaneamente.

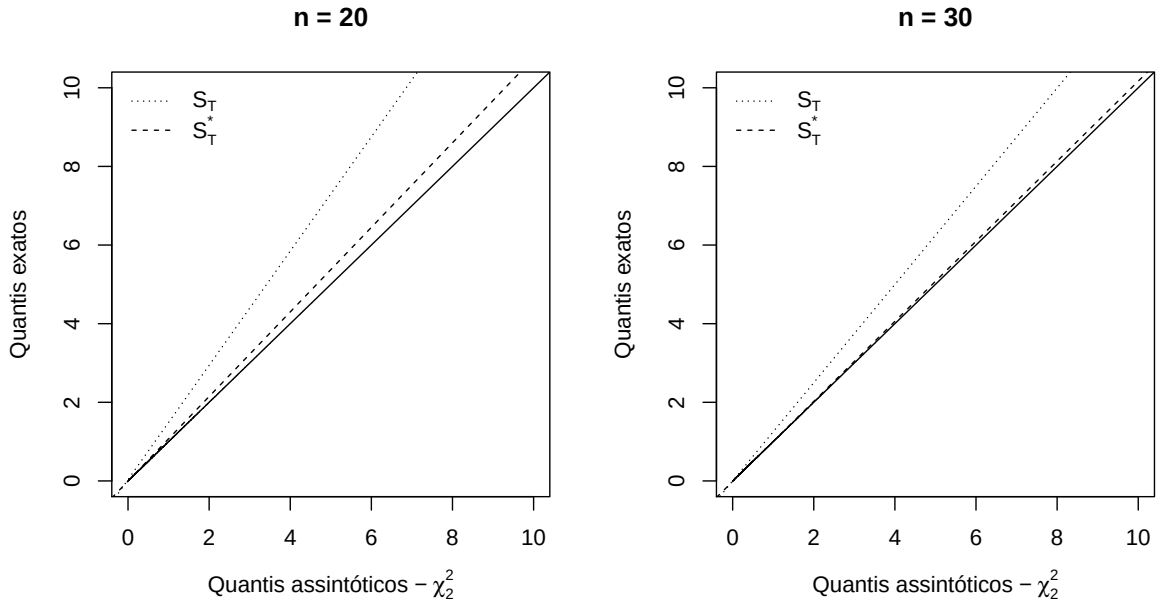

Gráfico 2 – Gráficos quantil-quantil das estatísticas S_T e S_T^*, considerando o teste dos efeitos da média.

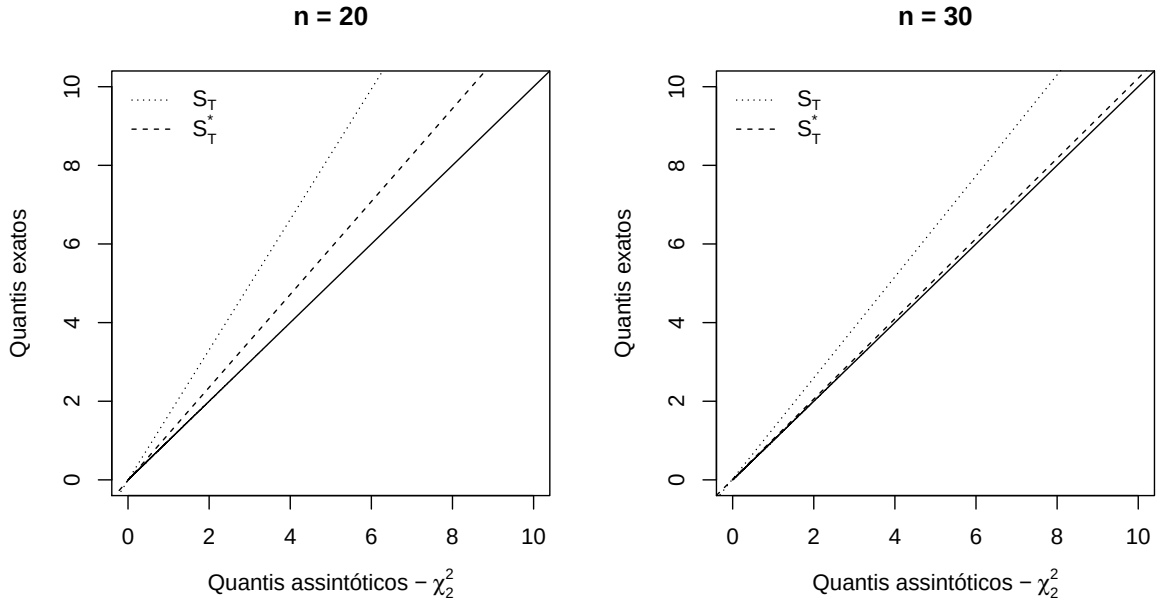

Gráfico 3 – Gráficos quantil-quantil das estatísticas S_T e S_T^*, considerando o teste dos efeitos da dispersão.

Na Tabela 8 apresentamos as taxas de rejeição dos testes LR, SR, ST, LR^*, SR^*, ST^*, LR_b, SR_b e ST_b, considerando a seguinte hipótese nula: $\mathcal{H}_0^2 : \boldsymbol{\beta}_1 = \mathbf{0}$, com $p_1 = 2$, $p = 3$ e

$q = 2$. É possível notar que os testes gradiente e razão de verossimilhanças são liberais, isto é, suas taxas de rejeição estão acima dos níveis nominais considerados. O teste escore, por sua vez, apresentou distorções de tamanho menores do que os testes mencionados anteriormente. Por exemplo, quando $n = 20$ e $\alpha = 5\%$, as taxas de rejeição dos testes ST, LR e SR são $7,70\%$, $9,78\%$ e $5,90\%$, respectivamente. Os testes corrigidos, por outro lado, apresentaram taxas de rejeição mais próximas dos níveis nominais. Considerando a mesma situação anterior, as taxas de rejeição são $5,26\%$, $6,54\%$ e $4,82\%$, para os testes ST^*, LR^* e SR^*, respectivamente. Vale salientar que os testes escore e gradiente corrigidos obtiveram desempenhos consideravelmente melhores. Observamos também que os testes bootstrap apresentaram leve distorção de tamanho. Por exemplo, para $n = 30$ e $\alpha = 5\%$, as taxas de rejeição dos testes ST_b, LR_b e SR_b correspondente a $\alpha = 5\%$ são $5,32\%$, $5,34\%$ e $5,04\%$, respectivamente. De modo geral, os testes corrigidos e os testes bootstrap apresentaram desempenhos similares. Note ainda que o teste LR_b apresentou desempenho superior ao do teste LR^*. Ressaltamos, no entanto, que os testes bootstrap podem ser custosos computacionalmente.

Já na Tabela 9 apresentamos os resultados de simulação para as taxas de rejeição não-nula (poder) dos testes LR^*, SR^*, ST^*, LR_b, SR_b e ST_b. É importante observar que estas simulações de poder correspondem à situação abordada na Tabela 8 para $n = 30$. Foram fixados $p_1 = 2$, $p = 3$, $q = 2$ e o nível nominal $\alpha = 10\%$. A hipótese alternativa em consideração é $\mathcal{H}_1^2 : \beta_1 = \beta_2 = \delta$, com diferentes valores para δ. Não consideramos os testes ST, LR e SR, visto que estes são bastante liberais, sob $\mathcal{H}_0^2$, e, portanto, não podem ser recomendados. A partir dos resultados de simulação é possível notar que, à medida que $|\delta|$ aumenta, o poderes dos testes também aumentam, como esperado. Entre os testes corrigidos, o teste LR^* foi o que apresentou desempenho levemente superior. Os resultados indicam que não há nenhuma perda de poder derivada do fato de usar o fator de correção de Bartlett e tipo-Bartlett derivado no Capítulo 3. Em geral, os testes corrigidos apresentaram desempenhos semelhantes aos dos testes bootstrap. Vale salientar aqui que as simulações das Tabelas 8 e 9 para o caso dos testes SR e SR^* são inéditas, pois em Terra (2013) as simulações para esses testes só foram apresentadas para a hipótese $\mathcal{H}_0^3$.

Tabela 8 – Taxas de rejeição nula $(\%)$ dos testes LR, SR, ST, LR^*, SR^*, ST^*, LR_b, SR_b e ST_b com $p_1 = 2$, $p = 3$, $q = 2$ e diferentes tamanhos amostrais.

n	α	ST	ST^*	ST_b	LR	LR^*	LR_b	SR	SR^*	SR_b
20	5%	7,70	5,26	5,12	9,78	6,54	4,52	5,90	4,82	4,78
	10%	14,30	10,12	9,68	16,56	12,04	9,20	11,84	9,56	9,54
30	5%	6,48	5,28	5,32	7,36	6,14	5,34	5,44	5,04	5,04
	10%	12,68	9,98	10,00	13,66	11,60	9,84	11,02	9,72	9,86
40	5%	6,60	5,30	5,26	7,36	6,02	5,16	5,86	5,16	5,36
	10%	11,82	10,02	9,94	12,46	10,88	9,86	10,88	9,96	9,98
50	5%	5,98	5,08	5,22	6,64	5,56	5,26	5,46	5,06	5,24
	10%	11,58	10,02	10,22	12,10	10,90	10,18	10,86	10,10	10,26

Tabela 9 – Taxas de rejeição não-nula $(\%)$ dos testes LR^*, SR^*, ST^*, LR_b, SR_b e ST_b com $p_1 = 2$, $p = 3$, $q = 2$, $n = 30$ e $\alpha = 10\%$.

δ	ST^*	ST_b	LR^*	LR_b	SR^*	SR_b
$-0,4$	94,08	94,20	94,62	93,86	93,72	93,76
$-0,3$	80,24	80,30	81,68	79,86	79,94	80,12
$-0,2$	52,28	52,00	54,64	51,60	52,40	52,20
$-0,1$	22,80	22,24	25,20	22,12	22,60	22,44
0,1	23,98	24,08	26,10	23,72	23,76	23,96
0,2	57,64	57,06	59,76	56,60	57,36	57,32
0,3	88,52	88,28	89,48	87,84	88,10	88,08
0,4	99,06	99,04	99,18	98,94	98,88	98,88

5 CONSIDERAÇÕES FINAIS

A principal contribuição teórica nesta dissertação foi a derivação de um fator de correção tipo-Bartlett à estatística gradiente para testar simultaneamente os efeitos da média e da dispersão em modelos lineares generalizados superdispersados. A partir dos resultados de simulações de Monte Carlo verificamos que o teste gradiente pode ser bastante liberal em amostras de tamanho pequeno, rejeitando a hipótese nula com maior frequência do que o esperado com base no nível nominal considerado. Nesse sentido, notamos que o fator de correção tipo-Bartlett que obtivemos neste trabalho melhorou a aproximação da distribuição da estatística gradiente pela distribuição qui-quadrado, uma vez que o teste corrigido apresentou taxas de rejeição empíricas mais próximas dos níveis nominais considerados. Além disso, tanto o teste gradiente usual quanto o teste gradiente corrigido são afetados pelo número de parâmetros de perturbação, principalmente, se estes são componentes do vetor de parâmetros da média. De modo geral, os testes baseados na estatística gradiente aperfeiçoada apresentaram melhores resultados do que aqueles baseados na estatística gradiente usual.

Além disso, considerando o teste dos efeitos da média, notamos que os testes da razão de verossimilhanças, gradiente e escore corrigidos têm taxas de rejeição mais próximas dos respectivos níveis nominais do que suas versões sem correção. Verificamos também que os testes gradiente e escore corrigidos apresentaram desempenhos levemente melhores do que o teste da razão de verossimilhanças corrigido. Em geral, os desempenhos dos testes bootstrap são similares aos dos testes corrigidos, com uma pequena vantagem para o teste LR_b em relação ao teste LR^*. Os poderes dos testes bootstrap e dos testes corrigidos são semelhantes, com leve vantagem para a estatística da razão de verossimilhanças corrigida, LR^*.

REFERÊNCIAS

ANDRADE, T. A. N. *Verossimilhança Perfilada nos Modelos Lineares Generalizados com Superdispersão*. Dissertação (Mestrado) — Universidade Federal de Pernanbuco, 2013.

BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, v. 160, p. 268–282, 1937.

BOTTER, D. A.; DENISE, C. M. Improved estimators for generalized linear models with dispersion covariates. *Journal of Statistical Computation and Simulation*, v. 62, p. 91–104, 1998.

CAMPOS, R. P. S. *Aperfeiçoamento de Testes em Modelos Lineares Generalizados Superdispersados*. Dissertação (Mestrado) — Universidade Federal de Pernanbuco, 2014.

CORDEIRO, G. M. Improved likelihood ratio statistics for generalized linear models. *Journal of the Royal Statistical Society B*, v. 45, p. 404–413, 1983.

CORDEIRO, G. M. On the corrections to the likelihood ratio statistics. *Biometrika*, v. 74, p. 265–274, 1987.

CORDEIRO, G. M.; BOTTER, D. A. Second-order biases of maximum likelihood estimates in overdispersed generalized linear models. *Statistics & Probability Letters*, v. 55, p. 269–280, 2001.

CORDEIRO, G. M.; CRIBARI-NETO, F. *An Introduction to Bartlett Correction and Bias Reduction*. Springer, 2014.

CORDEIRO, G. M.; CYSNEIROS, A. H. M. A.; CYSNEIROS, F. J. A. Bartlett adjustments for overdispersed generalized linear models. *Communications in Statistics—Theory and Methods*, v. 35, n. 5, p. 937–952, 2006.

CORDEIRO, G. M.; FERRARI, S. L. P. A modified score test statistic having chi-squared distribution to order n^{-1}. *Biometrika*, v. 78, p. 573–582, 1991.

CORDEIRO, G. M.; FERRARI, S. L. P.; PAULA, G. A. Improved score tests for generalized linear models. *Journal of the Royal Statistical Society B*, v. 55, p. 661–674, 1993.

CORDEIRO, G. M.; PREVIDELLI, I.; SAMOHYL, R. W. Bias corrected maximum likelihood estimators in nonlinear overdispersed models. *Brazilian Journal of Probability and Statistics*, v. 22, p. 105–118, 2008.

COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, v. 49, p. 1–39, 1987.

COX, D. R.; SNELL, E. J. A general definition of residuals. *Journal of the Royal Statistical Society B*, v. 30, p. 248–275, 1968.

CRIBARI-NETO, F.; FERRARI, S. L. P. Second order asymptotics for score tests in generalised linear models. *Biometrika*, v. 82, p. 426–432, 1995.

DEY, D. K.; GELFAND, A. E.; PENG, F. Overdispersed generalized linear models. *Journal of Statistical Planning and Inference*, v. 64, p. 93–107, 1997.

DICICCIO, T. J.; STERN, S. E. Frequentist and bayesian bartlett correction of test statistics based on adjusted profile likelihoods. *Journal of the Royal Statistical Society B*, v. 56, p. 397–408, 1994.

DOORNIK, J. Ox: An object-oriented matrix programming language. Timberlake Consultants Press, London, v. 5th ed., 2006.

EFRON, B. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, v. 7, p. 1–26, 1979.

EFRON, B. Double exponential families and their use in generalized linear regression. *Journal of the American Statistical Association*, v. 81, p. 709–721, 1986.

EFRON, B.; TIBSHIRANI, R. J. *An Introduction to the Bootstrap*. Chapman and Hall, New York, 1993.

GANIO, L. M.; SCHAFER, D. W. Diagnostics for overdispersion. *Journal of the American Statistical Association*, v. 87, p. 795–804, 1992.

GELFAND, A. E.; DALAL, S. R. A note on overdispersed exponential families. *Biometrika*, v. 77, p. 55–64, 1990.

GHOSH, J. K.; MUKERJEE, R. Characterization of priors under which bayesian and frequentist bartlett corrections are equivalent in the multiparameter case. *Journal of Multivariate Analysis*, v. 38, p. 385–393, 1991.

GOOD, I. J. The population frequencies of species and the estimation of population parameters. *Biometrika*, v. 40, p. 237–264, 1953.

HILL, G. W.; DAVIS, A. W. Generalized asymptotic expansions of cornish-fisher type. *The Annals of Mathematical Statistics*, v. 39, p. 1264–1273, 1968.

HINDE, J.; DEMÉTRIO, C. G. B. Overdispersion: Models and estimation. *Computational Statistics & Data Analysis*, v. 27, p. 151–170, 1998.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous Univariate Distributions, Volume 1*. Wiley, New York, 1994.

JØRGENSEN, B. *Statistical Properties of the Generalized Inverse Gaussian Distribution*. Springer, New York, 1982.

LAWLEY, D. N. A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, v. 43, p. 295–303, 1956.

LINDSEY, J. K. *Applying Generalized Linear Models*. Springer, New York, 1997.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. *Journal of the Royal Statistical Society A*, v. 135, p. 370–384, 1972.

NOCEDAL, J.; WRIGHT, S. *Numerical Optimization*. Springer, New York, 1999.

RAO, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, v. 44, p. 50–57, 1948.

RAO, C. R. *Linear Statistical Inference and Its Applications*. Wiley, New York, 1973.

RODRIGUES, H. M. *Técnicas de Diagnóstico nos Modelos Lineares Generalizados com Superdispersão*. Dissertação (Mestrado) — Universidade Federal de Pernanbuco, 2013.

SMYTH, G. K. Generalized linear models with varying dispersion. *Journal of the Royal Statistical Society B*, v. 51, p. 47–60, 1989.

TERRA, M. L. C. *Modelos Não Lineares Generalizados com Superdispersão*. Tese (Doutorado) — Universidade Federal de Pernanbuco, 2013.

TERRELL, G. R. The gradient statistic. *Computing Science and Statistics*, v. 34, p. 206–215, 2002.

VARGAS, T. M.; FERRARI, S. L. P.; LEMONTE, A. J. Gradient statistic: Higher-order asymptotics and bartlett-type correction. *Electronic Journal of Statistics*, v. 7, p. 43–61, 2013.

VARGAS, T. M.; FERRARI, S. L. P.; LEMONTE, A. J. Improved likelihood inference in generalized linear models. *Computational Statistics & Data Analysis*, v. 74, p. 110–124, 2014.

WALD, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical society*, v. 54, p. 426–482, 1943.

WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, v. 9, p. 60–62, 1938.

APÊNDICE A – DERIVADAS DA FUNÇÃO $\Psi(\mu, \phi)$

Para o cálculo dos cumulantes e, consequentemente, da correção tipo-Bartlett para a estatística gradiente (ST) em modelos lineares generalizados superdispersados é necessário obter algumas derivadas da função $\Psi(\mu, \phi)$ definida a partir do modelo adotado. Neste trabalho, consideramos o modelo Poisson duplo, em que

$$\Psi(\mu, \phi) = \mu\phi \ln \mu + \frac{1}{2} \ln \phi.$$

Dada a função $\Psi(\mu, \phi)$, podemos facilmente calcular suas derivadas da seguinte forma:

$$\Psi^{(r,s)}(\mu, \phi) = \partial^{r+s}\Psi(\mu, \phi)/\partial\mu^r\partial\phi^s, \quad r, s \geq 0.$$

As derivadas da função $\Psi(\mu, \phi)$ utilizadas nesta dissertação são dadas a seguir.

$$\begin{aligned}
&\Psi^{(1,0)} = \phi \ln \mu + \phi, \quad \Psi^{(0,1)} = \phi \ln \mu + \frac{1}{2\phi}, \\
&\Psi^{(2,0)} = \frac{\phi}{\mu}, \quad \Psi^{(0,2)} = -\frac{1}{2\phi^2}, \quad \Psi^{(3,0)} = -\frac{\phi}{\mu^2}, \\
&\Psi^{(0,3)} = \frac{1}{\phi^3}, \quad \Psi^{(4,0)} = -\frac{2\phi}{\mu^3}, \quad \Psi^{(0,4)} = -\frac{3}{\phi^4}, \\
&\Psi^{(1,1)} = \ln \mu + 1, \quad \Psi^{(2,1)} = \frac{1}{\mu}, \quad \Psi^{(2,2)} = 0 \quad \text{e} \\
&\Psi^{(1,2)} = 0.
\end{aligned}$$

APÊNDICE B – CUMULANTES PARA OS MLGS

Neste apêndice, apresentamos os cumulantes e derivadas desses cumulantes que são necessários para a obtenção da correção tipo-Bartlett para a estatística gradiente no MLGS. Adota-se a seguinte notação para as derivadas da função de log-verossimilhança $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$:

$$U_r = \partial\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r, \quad U_{rS} = \partial^2\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r\partial\gamma_S, \quad U_{rsT} = \partial^3\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r\partial\beta_s\partial\gamma_T,$$

etc, em que as letras minúsculas e maiúsculas correspondem a componentes de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, respectivamente. Os cumulantes de derivadas da função de log-verossimilhança são denotados por

$$\kappa_r = \mathrm{E}(U_r), \quad \kappa_{rS} = \mathrm{E}(U_{rS}), \quad \kappa_{rsT} = \mathrm{E}(U_{rsT}),$$
$$\kappa_{r,S} = \mathrm{E}(U_r U_S), \quad \kappa_{r,sT} = \mathrm{E}(U_r U_{sT})$$

e assim por diante. As derivadas dos cumulantes são representadas por

$$\kappa_{rs}^{(t)} = \frac{\partial\kappa_{rs}}{\partial\beta_t}, \quad \kappa_{rs}^{(T)} = \frac{\partial\kappa_{rs}}{\partial\gamma_T},$$

etc. A seguir, apresentamos tais cumulantes na classe dos MLGS (Cordeiro et al., 2006).

- Cumulantes para o componente β

$$\kappa_{jr} = -\sum_l \Psi_l^{(2,0)} m_{1l}^2 x_{lj} x_{lr},$$
$$\kappa_{jrs} = -\sum_l (2\Psi_l^{(3,0)} m_{1l}^2 + 3\Psi_l^{(2,0)} m_{2l}) m_{1l} x_{lj} x_{lr} x_s,$$
$$\kappa_{jrsu} = -\sum_l (3\Psi_l^{(4,0)} m_{1l}^4 + 12\Psi_l^{(3,0)} m_{1l}^2 m_{m2l} + 3\Psi_l^{(2,0)} m_{2l}^2 + 4\Psi_l^{(2,0)} m_{1l} m_{3l}) \times x_{lj} x_{lr} x_{ls} x_{lu},$$
$$\kappa_{jr}^{(s)} = -\sum_l (\Psi_l^{(3,0)} m_{1l}^2 + 2\Psi_l^{(2,0)} m_{2l}) m_{1l} x_{lj} x_{lr} x_{ls}$$
$$\kappa_{rs}^{(ju)} = -\sum_l (\Psi_l^{(4,0)} m_{1l}^4 + 5\Psi_l^{(3,0)} m_{1l}^2 m_{2l} + 2\Psi_l^{(2,0)} m_{1l} m_{3l} + 2\Psi_l^{(2,0)} m_{2l}^2) x_{lj} x_{lr} x_{ls} x_{lu},$$
$$\kappa_{jrs}^{(u)} = -\sum_l (2\Psi_l^{(4,0)} m_{1l}^4 + 9\Psi_l^{(3,0)} m_{1l}^2 m_{2l} + 3\Psi_l^{(2,0)} m_{2l}^2 + 3\Psi_l^{(2,0)} m_{1l} m_{3l}) x_{lj} x_{lr} x_{ls} x_{lu}.$$

- Cumulantes para o componente γ

$$\kappa_{JR} = \sum_l \Psi_l^{(0,2)} \phi_{1l}^2 s_{lJ} s_{lR},$$

$$\kappa_{JRS} = \sum_l (\Psi_l^{(0,3)} \phi_{1l}^2 + 3\Psi_l^{(0,2)} \phi_{2l}) \phi_{1l} s_{lJ} s_{lR} s_{lS},$$

$$\kappa_{JRSU} = \sum_l (\Psi_l^{(0,4)} \phi_{1l}^4 + 6\Psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 3\Psi_l^{(0,2)} \phi_{2l}^2 + 4\Psi_l^{(0,2)} \phi_{1l} \phi_{3l}) s_{lJ} s_{lR} s_{lS} s_{lU},$$

$$\kappa_{JR}^{(S)} = \sum_l (\Psi_l^{(0,3)} \phi_{1l}^2 + 2\Psi_l^{(0,2)} \phi_{2l}) \phi_{1l} s_{lJ} s_{lR} s_{lS}$$

$$\kappa_{RS}^{(JU)} = \sum_l (\Psi_l^{(0,4)} \phi_{1l}^4 + 5\Psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 2\Psi_l^{(0,2)} \phi_{2l}^2 + 2\Psi_l^{(0,2)} \phi_{1l} \phi_{3l}) s_{lJ} s_{lR} s_{lS} s_{lU},$$

$$\kappa_{JRS}^{(U)} = \sum_l (\Psi_l^{(0,4)} \phi_{1l}^4 + 6\Psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 3\Psi_l^{(0,2)} \phi_{2l}^2 + 3\Psi_l^{(0,2)} \phi_{1l} \phi_{3l}) s_{lJ} s_{lR} s_{lS} s_{lU}.$$

- Cumulantes mistos

$$\kappa_{jRS} = 0, \quad \kappa_{jR}^{(s)} = 0, \quad \kappa_{jR}^{(S)} = 0, \quad \kappa_{JRSu} = 0$$

$$\kappa_{jRS}^{(u)} = 0, \quad \kappa_{jRS}^{(U)}, \quad \kappa_{rS}^{(ju)} = 0, \quad \kappa_{rS}^{(JU)} = 0,$$

$$\kappa_{jrS} = \kappa_{jr}^{(S)} = -\sum_l \Psi_l^{(2,1)} \phi_{1l} m_{1l}^2 x_{lj} x_{lr} s_{lS},$$

$$\kappa_{jrS}^{(u)} = -\sum_l (\Psi_l^{(3,1)} m_{1l}^3 + 2\Psi_l^{(2,1)} m_{1l} m_{2l}) \phi_{1l} x_{lj} x_{lr} x_{lu} s_{lS},$$

$$\kappa_{jrs}^{(U)} = -\sum_l (2\Psi_l^{(3,1)} m_{1l}^3 + 3\Psi_l^{(2,1)} m_{1l} m_{2l}) \phi_{1l} x_{lj} x_{lr} x_{ls} s_{lU},$$

$$\kappa_{JRS}^{(u)} = -\sum (\Psi_l^{(1,3)} m_{1l} \phi_{1l}^2 + 3\Psi_l^{(1,2)} m_{1l} \phi_{2l}) \phi_{1l} x_{lu} s_{lJ} s_{lR} s_{lS},$$

$$\kappa_{jrS}^{(U)} = -\sum_l (\Psi_l^{(2,2)} \phi_{1l}^2 m_{1l}^2 + \Psi_l^{(2,1)} \phi_{2l} m_{1l}^2) x_{lj} x_{lr} s_{lS} s_{lU},$$

$$\kappa_{JR}^{(s)} = \sum_l \Psi_l^{(1,2)} \phi_{1l}^2 m_{1l} x_{ls} s_{lJ} s_{lR},$$

$$\kappa_{rs}^{(jU)} = -\sum_l (\Psi_l^{(3,1)} m_{1l}^3 + 2\Psi_l^{(2,1)} m_{1l} m_{2l}) \phi_{1l} x_{lr} x_{ls} x_{lj} s_{lU},$$

$$\kappa_{rs}^{(JU)} = -\sum_l (\Psi_l^{(2,2)} \phi_{1l}^2 m_{1l}^2 + \Psi_l^{(2,1)} \phi_{2l} m_{1l}^2) x_{lr} x_{ls} s_{lJ} s_{lU},$$

$$\kappa_{RS}^{(ju)} = \sum_l (\Psi_l^{(2,2)} m_{1l}^2 + \Psi_l^{(1,2)} m_{2l}) \phi_{1l}^2 x_{lj} x_{lu} s_{lR} s_{lS},$$

$$\kappa_{RS}^{(jU)} = \sum_l (\Psi_l^{(1,3)} \phi_{1l}^3 + 2\Psi_l^{(1,2)} \phi_{1l} \phi_{2l}) m_{1l} x_{lj} s_{lR} s_{lS} s_{lU},$$

$$\kappa_{jrsU} = -\sum_l (2\Psi_l^{(3,1)} m_{1l}^3 + 3\Psi_l^{(2,1)} m_{1l} m_{2l}) \phi_{1l} x_{lj} x_{lr} x_{ls} s_{lU},$$

$$\kappa_{jrSU} = -\sum_l (\Psi_l^{(2,2)} \phi_{1l}^2 + \Psi_l^{(2,1)} \phi_{2l}) m_{1l}^2 x_{lj} x_{lr} s_{lS} s_{lU}.$$