Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Estatística

Pós-graduação em Estatística

Modelos Não Lineares Generalizados com Superdispersão

Maria Lídia Coco Terra

Tese de Doutorado

Recife, fevereiro de 2013

Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza

Departamento de Estatística

Maria Lídia Coco Terra

Modelos Não Lineares Generalizados com Superdispersão

Trabalho apresentado ao Programa de Pós-graduação em Estatística do Departamento de Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Estatística.

Orientadora: Prof^{a}. Dr^{a}. Audrey Helen A. Cysneiros

Recife, fevereiro de 2013

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da Silva, CRB4-1217

Terra, Maria Lídia Coco
Modelos não lineares generalizados com superdispersão / Maria Lídia Coco Terra. - Recife: O Autor, 2013.
xii, 92 p.: il., fig., tab.

Orientadora: Audrey Helen A. Cysneiros.
Tese (doutorado) - Universidade Federal de Pernambuco. CCEN, Estatística, 2013.

Inclui bibliografia e apêndice.

1. Estatística matemática. 2. Teoria assintótica. I. Cysneiros, Audrey Helen A. (orientadora). II. Título.

519.5 CDD (23. ed.) MEI2013 – 020

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

Recife, 08 de fevereiro de 2013.

Nós recomendamos que a tese de doutorado de autoria de

Maria Lídia Coco Terra

Intitulada

"Modelos Não Lineares Generalizados com Superdispersão"

Seja aceita como cumprimento parcial dos requerimentos para o grau de Doutor(a) em Estatística.

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Audrey Helen Mariz de Aquino Cysneiros	Orientadora / UFPE
Klaus Leite Pinto Vasconcellos	UFPE
Francisco José de Azevedo Cysneiros	UFPE
Miguel Angel Uribe Opazo	UNIOESTE
Jalmar Manuel Farfan Carrasco	UFBA

Este documento será anexado à versão final da tese.

Aos meus pais, Paulo e Lia, e a minha irmã, Lígia

pelo apoio incondicional.

Agradecimentos

Aos meus pais e minha irmã, pelo carinho, dedicação, confiança, paciência e por terem contribuído para a minha formação moral e acadêmica.

À minha família e amigos pela enorme paciência, apoio e compreensão com a minha ausência.

À professora Audrey Helen Mariz de Aquino Cysneiros, pela incrível paciência, atenção e excelente orientação.

Aos professores da Pós-Graduação em Estatística da UFPE, os quais me ensinaram sobre Estatística e me receberam bem aqui no doutorado.

À Valéria Bittencourt, pelo enorme carinho, paciência e amizade com que sempre tratou a mim e aos demais alunos do doutorado.

Em especial a minha irmã Lígia, Juliana, Marcelo, Carlos e Larissa pela amizade, ajuda, apoio e pelos momentos de alegria.

À Capes, pelo apoio financeiro.

Aos participantes da banca examinadora pelas sugestões.

Resumo

Dey et al. (1997) propuseram uma classe de modelos que permite a introdução de um segundo parâmetro que controla a variância independentemente da média através de um modelo de regressão, chamada modelos lineares generalizados com superdispersão. Nesta tese, estendemos a classe de modelos proposta por Dey et al. (1997) permitindo que as funções de ligações da média e da dispersão possam ser funções não lineares obtendo expressões matriciais para os fatores de correção Bartlett e tipo-Bartlett para as estatísticas da razão da verossimilhanças e escore, respectivamente, na classe dos modelos não lineares generalizados com superdispersão (MNLGSs). Foi realizado um estudo de simulação para avaliar os desempenhos dos testes baseados nas estatísticas da razão de verossimilhanças e escore com suas respectivas versões corrigidas (Bartlett e tipo-Bartlett) com relação ao tamanho e poder em amostras de tamanhos finitos. Propomos também técnicas de diagnósticos para os MNLGSs, tais como: Alavancagem generalizada, Distância de Cook e Influência local. Finalmente, um conjunto de dados reais é utilizado para avaliar nossos resultados teóricos.

Palavras-chave: Alavancagem generalizada; Correção de Bartlett; Correção tipo-Bartlett; Influência Local; Métodos de diagnósticos; Modelos não lineares generalizados com superdispersão; Teste Escore; Teste da Razão de Verossimilhanças.

Abstract

Dey et al. (1997) proposed a class of models in which allows to insert a second parameter which controls the variance independently of the mean through a regression model, named overdispersed generalized linear models. In this work, we extend this class of models proposed by Dey et al. (1997), allowing that the mean and the dispersion link functions to be nonlinear functions, and we obtained matricial equations to the Bartlett correction and Bartlett-type correction for likelihood ratio and score tests, respectively, in this class of overdispersion generalized nonlinear models (OGNLM). Simulation study to evaluate the performance of the tests based on the likelihood ratio and score statistics and its Bartlett and Bartlett-type corrected versions, respectively on size and power in finite-samples. Also was calculated to this class of models some of the most common diagnostic methods like Generalized leverage, Cook distance and Local influence and application with real life data.

Keywords: Bartlett correction; Bartlett-type correction; Cook distance; Diagnostic methods; Generalized leverage; Local influence; Overdispersion generalized nonlinear models; Ratio likelihood test; Score test.

Índice

Lista de Figuras

Lista de Tabelas

CAPÍTULO 1

Introdução

O fenômeno da superdispersão tem sido amplamente discutido na literatura. O termo superdispersão se dá quando a média da variância da variável resposta excede da variância nominal predita do modelo proposto. O problema da superdispersão é facilmente detectado, mas é difícil de ser tratado.

Nesta tese utilizamos uma forma natural de representação da superdispersão através da classe de distribuições propostas por Dey et al. (1997). Nesta classe de modelos, considera-se um modelo de regressão para o parâmetro de locação e um outro modelo de regressão para o parâmetro de dispersão. Além disso, estes modelos de regressão para os parâmetros podem apresentar formas lineares ou não lineares.

Dey et al. (1997) generalizaram a família exponencial dupla proposta por Efron (1986) que permite a introdução de um segundo parâmetro que controla a variância independentemente da média através de um modelo de regressão em que os dois parâmetros dependem de covariadas observadas. Esta ligação entre os parâmetros da média e da dispersão com suas covariadas é feita através de uma função linear formando dois componentes sistemáticos, ou seja com uma estrutura similar aos modelos lineares generalizados (MLG). Esta classe de modelos é chamada modelos lineares generalizados com superdispersão. Mas, na prática, existem casos em que esta ligação se dá através de uma função não linear, desta forma propomos nesta tese uma classe de modelos não lineares generalizados com superdispersão (MNLGS). A família exponencial dupla tem as mesmas propriedades de uma família exponencial e pode ser generalizada para qualquer modelo de regressão da família exponencial.

Os testes baseados nas estatísticas da razão de verossimilhanças e escore são mais frequentemente utilizados na literatura. Em grandes amostras e sob a hipótese nula, estas estatísticas de teste têm uma distribuição qui-quadrado (χ^2_m) em que m é o número de restrições impostas pela hipótese nula, H_0, e o erro dessa aproximação é de ordem $O(n^{-1})$. No entanto, quando o tamanho de amostra é pequeno ou mesmo de tamanho moderado, esta aproximação pode não ser satisfatória.

A ideia de modificar a estatística da razão de verossimilhanças por um fator de correção, visando produzir uma estatística modificada com primeiro momento igual ao da distribuição χ^2 de referência é devida a Bartlett (1937). Um método geral para o cálculo do fator de correção foi proposto por Lawley (1956).

Cordeiro e Ferrari (1991) obtiveram um fator de correção para a estatística escore denominado fator de correção tipo-Bartlett, de forma que a estatística escore aperfeiçoada por este fator de correção tem distribuição χ^2_m até ordem $O(n^{-1})$ sob a hipótese nula. Sob certas condições de regularidade, a estatística escore modificada é obtida através da multiplicação da estatística escore original por um polinômio de segundo grau da própria estatística. A estatística proposta por Cordeiro e Ferrari (1991) nem sempre é uma transformação monótona da estatística escore original. Kakizawa (1996) sugere uma transformação monótona envolvendo a estatística escore corrigida em um polinômio de quinta ordem da estatística escore original. Também Cordeiro et al. (1998) propõem uma forma alternativa da estatística escore corrigida, através de uma transformação monótona da estatística escore baseada na função de distribuição acumulada da normal.

Alguns artigos surgiram envolvendo a classe de modelos proposta por Dey et al. (1997). Cordeiro e Botter (2001) derivaram uma fórmula geral para os vieses de segunda ordem via Cox e Snell (1968) para os estimadores de máxima verossimilhança dos parâmetros nos modelos lineares generalizados com superdispersão (MLGSs), generalizando os resultados de Cordeiro e

McCullagh (1991) e Botter e Cordeiro (1997). Cordeiro et al. (2006) derivaram uma fórmula geral para o fator de correção de Bartlett da estatística da razão de verossmilhanças na classe dos MLGSa, generalizando assim, os resultados de Botter e Cordeiro (1997) para os modelo lineares generalizados duplos e Cordeiro (1983) para os modelos lineares generalizados. Cordeiro et al. (2008) derivaram uma fórmula geral para o vieses de $O(n^{-1})$ via Cox e Snell (1968) para os estimadores de máxima verossimilhança dos parâmetros nos MNLGSs.

Dentro deste contexto, nosso interesse é a obtenção de um fator de correção de Bartlett em fórmula matricial para a estatística da razão de verossimilhanças na classe dos MNLGSs, generalizando assim os resultados de Cordeiro et al. (2006). Além disso, iremos obter também um fator de correção tipo-Bartlett para a estatística escore nesta classe de modelos.

É importante também verificar se o modelo ajustado se adequa bem aos dados e se existem observações discrepantes com alguma interferência desproporcional ou inferencial nos resultados do ajuste. Esta análise é feita assumindo o modelo como correto e investigando se as conclusões feitas são coerentes. A análise de diagnóstico começou com a análise de resíduos para detectar a presença de pontos aberrantes e avaliar a adequação da distribuição proposta para a variável resposta. Pregibon (1981) propõe o componente do desvio como resíduo na classe dos modelos lineares generalizados e sugere uma padronização. Paula (1995) apresenta uma forma padronizada para o componente do desvio em MLGs com parâmetros restritos e verifica, através de estudos de simulação, forte concordância na maioria dos modelos estudados entre a distribuição empírica do resíduo padronizado e a distribuição normal padrão.

Um outro tópico importante na análise de diagnóstico é a detecção de observações influentes, isto é, pontos que exercem um peso desproporcional nas estimativas dos parâmetros do modelo. Durante a década de 70 surgiram várias propostas relacionadas com a influência das observações nas estimativas dos coeficientes do modelo normal linear, como os pontos de alavanca, que têm um peso desproporcional no próprio valor ajustado. Tais pontos podem exercer forte influência nas estimativas dos coeficientes da regressão. Wei et al. (1998) propuseram uma forma bastante geral para obtenção da matrix $\partial\hat{\boldsymbol{y}}/\partial\boldsymbol{y}$ que pode ser aplicada em diversas situações de estimação.

A deleção de pontos talvez seja a técnica mais conhecida para avaliar o impacto da retirada de uma observação particular nas estimativas da regressão. A distância de Cook (1977), originalmente desenvolvida para modelos normais lineares, foi rapidamente assimilada e estendida para diversas classes de modelos. Uma das propostas mais inovadoras na área de diagnóstico em regressão foi apresentada por Cook (1986), que propõe avaliar a influência conjunta das observações sob pequenas mudanças (perturbações) no modelo ou nos dados, ao invés da avaliação pela retirada individual ou conjunta de pontos. Essa metodologia, denominada influência local, teve uma grande receptividade entre os usuários e pesquisadores de regressão, havendo inúmeras publicações no assunto em que a metodologia é aplicada em classes particulares de modelos ou estendida para situações mais gerais.

Vale ressaltar aqui que não há estudos sobre técnicas de diagnóstico dos MNLGSs na literatura. Sendo assim, nesta tese, serão adequadas técnicas de diagnósticos para os MNLGSs, a saber: Alavancagem generalizada, Distância de Cook e Influência local.

O objetivo principal desta tese é a obtenção de refinamentos para testes de hipóteses em modelos não lineares generalizados com superdispesão (MNLGSs) e estudo de diagnóstico de influência global e local. Especificamente:

(i) Obtivemos uma fórmula geral para o fator de correção de Bartllet, em notação matricial, para a estatística da razão de verossimilhanças original nos modelos não lineares generalizados com superdispesão (MNLGSs), considerando três tipos de hipótese.

(ii) Obtivemos uma fórmula geral para o fator correção tipo-Bartllet, em notação matricial, para a estatística escore original nos modelos não lineares generalizados com superdispesão, considerando três tipos de hipóteses.

(iii) Desenvolvemos métodos usuais de diagnósticos para a classe de modelos não lineares generalizados com superdispesão, como Alavancagem generalizada, Distância de Cook e Influência local.

Esta tese de doutorado está organizada da seguinte forma:

No Capítulo 2 apresenta-se a classe de modelos não lineares generalizados com superdispersão (MNLGS), aspectos inferenciais (função escore e matriz de informação de Fisher), assim como a estimação dos parâmetros em algumas distribuições pertencentes à classe dos MLGSs e como ficam os testes de hipóteses.

No Capítulo 3 apresenta-se a obtenção do fator de correção de Bartlett para a estatística da razão de verossimilhanças na classe dos MNLGSs. Os resultados deste capítulo generalizam o artigo de Cordeiro et al. (2006).

No Capítulo 4 apresenta-se o fator de correção tipo-Bartlett para a estatística escore na classe dos MNLGSs. Um estudo de simulação via Monte Carlo é apresentado no Capítulo 5 com o objetivo de comparar o desempenho dos testes baseados nas estatísticas da razão de verossimilhanças e escore com suas respectivas versões corrigidas em relação ao tamanho e poder em amostras finitas.

No Capítulo 6, são desenvolvidas técnicas de diagnósticos para os MNLGSs, tais como: alavancagem generalizada, Distância de Cook e influência local. Vale ressaltar aqui que os Capítulos 3, 4 e 6 são as principais contribuições teóricas desta tese.

O Capítulo 7 apresenta-se uma aplicação das técnicas de diagnóstico desenvolvidas no Capítulo 6 com dados reais. Finalmente, o último capítulo apresenta as conclusões e as conclusões.

Os desenvolvimentos teóricos dos capítulos citados acima e o conjunto de dados reais considerados neste trabalho encontram-se nos Apêndices A e B, respectivamente.

CAPÍTULO 2

Modelos Não Lineares Generalizados com Superdispersão

Encontra-se nos modelos de regressão a presença de superdispersão, ou seja, a ocorrência do fenômeno em que os dados apresentam uma variabilidade maior do que a prevista pelo modelo adotado. Existem várias propostas para se contornar a superdispersão como podem ser vistas em Efron (1986) e Smyth (1989), que consideram uma família exponencial dupla e Lindsay (1986), que propõe utilizar modelos lineares generalizados (MLG) mistos, isto é, adicionar estruturas de efeitos aleatórios no preditor linear.

A proposta que iremos utilizar nesta tese foi a de Dey et al (1997). Neste artigo eles definem uma classe de modelos lineares generalizados com superdispersão estendendo a metodologia de MLG numa classe maior, isto é, considerando um modelo de regressão adicional para o parâmetro de dispersão e incorporando-o na função de variância, podendo até obter os modelos não lineares generalizados com superdispersão em que os modelos de regressão para os parâmetros de locação e dispersão são funções não lineares.

2.1 Modelo

Dey et al (1997) estenderam a metologia dos modelos lineares generalizados proposta por McCullagh e Nelder (1989) aplicando na família exponencial dupla proposta por Efron (1986), em que consideram um modelo de regressão adicional para o parâmetro de dispersão, incorporando-o na função de variância.

Este modelo segue a mesma estrutura de um MLG, ou seja, é caracterizado por um componente aleatório e dois componentes sistemáticos com duas funções de ligação, uma para o parâmetro de locação e outra para o parâmetro de dispersão.

O componente aleatório é o que descreve a variável resposta $\boldsymbol{y}$, cuja função de densidade de probabilidade é apresentada por Dey et al (1997). Seja as variáveis aleatórias $y_1, y_2, \ldots, y_n$ são independentes e cada variável y_i tem uma função de densidade (ou função de probabilidade) pertencente à família de distribuição exponecial dupla dada por

$$\pi(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\phi}) = A(\boldsymbol{y}) \exp\{(\boldsymbol{y} - \boldsymbol{\mu})\Psi^{(0,1)}(\boldsymbol{\mu}, \boldsymbol{\phi}) + \boldsymbol{\phi} T(\boldsymbol{y}) + \Psi(\boldsymbol{\mu}, \boldsymbol{\phi})\}, \tag{2.1}$$

em que $A(\cdot)$, $T(\cdot)$ e $\Psi(\cdot, \cdot)$ são funções conhecidas e $\Psi^{(r,s)} = \partial\Psi^{r+s}(\boldsymbol{\mu}, \boldsymbol{\phi})/\partial\boldsymbol{\mu}^r\partial\boldsymbol{\phi}^s$, $r, s \geq 0$. A média e a variância de $\boldsymbol{y}$ são, respectivamente, $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{\mu}$ e $\mathrm{Var}(\boldsymbol{y}) = {\Psi^{(2,0)}}^{-1}$ e a média e a variância de $T(\boldsymbol{y})$ são $\mathrm{E}(T(\boldsymbol{y})) = -\Psi^{(0,1)}$ e $\mathrm{Var}(T(\boldsymbol{y})) = -\psi^{(0,2)}$, respectivamente. Pode ser mostrado facilmente que $\mathrm{E}(\partial^2 \log \pi(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\phi})/\partial\boldsymbol{\mu}\partial\boldsymbol{\phi}) = 0$, ou seja, que os parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\phi}$ são ortogonais.

Sejam os componentes sistemáticos funções não lineares para a média definido por

$$g(\boldsymbol{\mu}) = \boldsymbol{\eta} = f_1(X; \boldsymbol{\beta}), \tag{2.2}$$

em que $g(\cdot)$ é denominada função ligação da média, conhecida e biunívoca, sendo X a matriz conhecida de covariáveis, $\boldsymbol{\beta}$ é o vetor de p parâmetros desconhecidos a ser estimado, com $p < n$ observações e $f_1(\cdot, \cdot)$ uma função contínua e diferenciável com respeito aos componentes de $\boldsymbol{\beta}$ tal que a matriz de derivadas $\widetilde{X} = \partial f_1(X; \boldsymbol{\beta})/\partial\boldsymbol{\beta}$ tenha posto completo. E para o parâmetro da dispersão

$$h(\boldsymbol{\phi}) = \boldsymbol{\tau} = f_2(S; \gamma), \tag{2.3}$$

em que $h(\cdot)$ é denominada função ligação da dispersão, conhecida e biunívoca, sendo S a matriz conhecida de covariáveis, $\boldsymbol{\gamma}$ é o vetor de q parâmetros desconhecidos a ser estimado, com $q < n$ observações e $f_2(\cdot, \cdot)$ uma função contínua e diferenciável com respeito aos componentes de $\boldsymbol{\gamma}$ tal que a matriz de derivadas $\widetilde{S} = \partial f_2(S; \gamma)/\partial\boldsymbol{\gamma}$ tenha posto completo.

O modelo não linear generalizado com superdispersão (MNLGS) é definido pela componente aleatória com função de densidade (ou de probabilidade) dada em 2.1 e pelos componentes sistemáticos dados em 2.2 e 2.3.

Algumas distribuições pertencentes aos MNLGS serão apresentadas a seguir.

2.2 Distribuições pertencentes à classe de modelos com superdispersão

Na sequência, algumas distribuições de probabilidades pertencentes aos MNLGS são apresentadas.

Passeio Aleatório

O modelo do passeio aleatório tem sido associado ao desenvolvimento de inúmeros modelos financeiros, como por exemplo os preços no mercado futuro brasileiro, e é muito comum também adotá-lo para valores imobiliários.

Esta distribuição é uma recíproca de uma variável aleatória tendo uma distribuição inversa gaussiana (Johnson et al. (1994)). Se Y tem uma distribuição de passeio aleatório RW(θ, δ), sua função de densidade de probabilidade é dada por

$$\pi(y;\theta,\delta) = \left(\frac{\delta}{2\pi y}\right)^{1/2} \exp\left(-\frac{\delta y}{2} + \frac{\delta}{\theta} - \frac{\delta}{2\theta^2 y}\right), \quad y>0, \theta>0, \delta>0. \tag{2.4}$$

A média e a variância de Y são dadas, respectivamente, por $\mu = 1/\theta + 1/\delta$ e $\mathrm{Var}(Y) = 1/(\theta\delta) + 2/\delta^2$. É fácil verificar que a distribuição passeio aleatório dada em (2.4) é um modelo exponencial com dois parâmetros, mas não sendo membro da classe dos MNLGS. Para escrever a densidade (2.4) na forma da equação (2.1), é necessário utilizar a parametrização dada por Cordeiro, Botter (2001) pela expressão $\phi = -\delta/(2\theta^2)$ e

$$\begin{aligned}\psi(\mu,\phi) &= 2(-\phi)^{1/2}\{(2\mu-\phi)^{1/2} - (-\phi)^{1/2}\}^{-1} + \frac{1}{2}\log 2 - \log\{(2\mu-\phi)^{1/2} - (-\phi)^{1/2}\} \\ &\quad - \mu\{(2\mu-\phi)^{1/2} - (-\phi)^{1/2}\}^{-2}.\end{aligned}$$

Essa distribuição é considerada especial, pois ela pertence à classe de modelos MNLGS, no entanto não pertence aos MLG e nem aos modelos lineares duplos.

Distribuição inversa Gaussiana generalizada (IGG)

Introduzida por Good (1953) e mais tarde estudada com interesse por Wise (1971) e Jϕrgensen (1983), a distribuição IGG abaixo é dada por Johnson et al. (1994; vol 1); que utilizam três parâmetros para definir a função densidade como

$$\pi(y;\alpha,\delta,r)=\frac{\left(\frac{\alpha}{\delta}\right)^{\frac{r}{2}}}{2K_r\left(\sqrt{\alpha\delta}\right)}y^{r-1}\exp\left\{-\frac{1}{2}(\delta y^{-1}+\alpha y)\right\},\quad y,\alpha,\delta>0, \tag{2.5}$$

em que $K_r(\cdot)$ é uma função de Bessel modificada de terceira ordem.

É utilizada quando a variável resposta tem assimetria considerável e intervalo de variação positivo. Ela é uma alternativa para as distribuições Gama e Passeio aleatório.

A expressão acima pode ser escrita como

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

Fazendo r conhecido e considerando que

$$\begin{aligned}&\theta=\Psi^{(1,0)}(\mu,\phi)=-\frac{\alpha}{2}, \qquad T(y)=y^{-1},\\ &\phi=-\frac{\delta}{2}, \quad \text{e} \quad A(y)=y^{r-1},\end{aligned}$$

como os MNLGS tem origem na forma da família de Geland e Dalal (1990) dada por

$$f(y;\mu,\phi)=A(y)\exp\left\{\theta y+\tau T(y)-\rho(\theta,\phi)\right\},$$

e a partir de $\rho(\theta,\phi)$ pode-se determinar a $\Psi(\mu,\phi)$, então

$$\begin{aligned}\rho(\theta,\phi)&=-\log\left(\frac{\alpha}{\delta}\right)^{\frac{r}{2}}+\log 2+\log K_r\left(\sqrt{\alpha\delta}\right)\ \text{e se } \alpha\delta=4\theta\phi,\\ \rho(\theta,\phi)&=-\frac{r}{2}\log\left(\frac{\theta}{\phi}\right)+\log 2+\log K_r\left(2\sqrt{\theta\phi}\right).\end{aligned}$$

Derivando-se $\rho(\theta,\phi)$ com relação a θ tem-se $\Psi(\mu,\phi)$

$$\begin{aligned}\frac{\partial\rho(\theta,\phi)}{\partial\theta}&=-\frac{r}{2}+\frac{K_r'\left(2\sqrt{\theta\phi}\right)}{K_r\left(2\sqrt{\theta\phi}\right)}\frac{1}{\sqrt{\theta\phi}}\\ &=-\frac{r}{2}+M(\theta,\phi)\\ &=\mu.\end{aligned}$$

Geland e Dalal (1990) apresentam que $\theta = \Psi^{(1,0)}(\mu, \phi)$, ou seja,

$$\begin{aligned}\frac{\partial \Psi(\mu, \phi)}{\partial \mu} &= \theta \Rightarrow \Psi(\mu, \phi) = \int \theta du \\ &= \Psi(\mu, \phi) = \int \Psi(u, \phi) du,\end{aligned}$$

que só pode ser calculada numericamente, ou seja, não tem solução explicita.

Desse modo, fica demonstrado que a distribuição inversa gaussiana generalizada pode ser escrita sob a forma dos modelos não lineares generalizados com superdispersão. E ainda, quando os parâmetros α, δ e r assumem alguns valores na IGG, essa distribuição cai em alguns casos especiais, como:

i) Distribuição inversa gaussiana, se $r = -\frac{1}{2}$;

ii) Distribuição recíproca da inversa gaussiana, se $r = \frac{1}{2}$;

iii) Distribuição gama, $\delta = 0$ e $r > 0$;

iv) Distribuição recíproca gama, se $r = 0$.

Desse modo, pode-se dizer que a distribuição inversa gaussiana generalizada é uma subfamília dos modelos não lineares generalizados com superdispersão.

Família exponencial dupla

Segundo Efron (1986), a família exponencial dupla pode ser escrita pela seguinte função de densidade

$$\begin{aligned}\pi(y; \theta, \rho, a) = c(\theta, \rho, a)\rho^{\frac{1}{2}} &\left\{\exp\left\{a(y\theta - \chi^2(\theta))\right\} h(y; a)\right\}^{\rho} \\ &\left\{\exp\left\{a(y\theta - \chi^2(\theta))\right\} h(y; a)\right\}^{1-\rho} \left[dh(y; a)\right], \\ = c(\theta, \rho, a)\rho^{\frac{1}{2}} &\exp\left\{a\rho(y\theta - \chi^2(\theta)) \right. \\ &\left. + a(1-\rho)\left(y\theta(y) - \chi^2(\theta(y))\right)\right\} h(y; a),\end{aligned}$$

com os parâmetros ρ, θ, a conhecidos, $\Psi(\mu) = \chi^2(\theta)$, a função $h(y; a)$ também é conhecida e a constante $c(\theta, \rho, a)$ aproximadamente igual a 1. A média e a variância são aproximadamente θ

e $V(\theta)/a\theta$, respectivamente. Essa família de distribuição desfruta das propriedades da família exponencial tanto para o parâmetro da média quanto para o parâmetro de dispersão.

Fazendo, $T(y) = y\theta(y) - \chi^2(\theta(y))$, $a(1-\rho) = \phi$, ou seja, $\rho = 1 - \frac{\phi}{a}$, temos

$$\begin{aligned}
\alpha &= a\rho\theta, \\
\alpha &= a\left(1 - \frac{\phi}{a}\right)\theta \text{ e} \\
\theta &= \frac{\alpha}{a-\phi}.
\end{aligned}$$

Agora, basta substituir estas quantidades na expressão da densidade acima, e temos que

$$\begin{aligned}
\pi(y;\theta,\rho,a) = c(\theta,\rho,a)\rho^{\frac{1}{2}} \exp\left\{a\rho y\theta - a\rho\chi^2(\theta)\right. \\
\left. a(1-\rho)T(y)h(y;a)\right\}, \\
= c(\theta,\rho,a)\rho^{\frac{1}{2}} \exp\left\{\alpha y + T(y) - \rho(\alpha,\phi)h(y;a)\right\},
\end{aligned}$$

e assim, fica demonstrado que a família exponencial dupla apresentada por Efron (1986) pode ser escrita sob a forma da família exponencial de dois parâmetros apresentada por Gelfand e Dalal (1990), que derivaram os modelos lineares generalizados com superdispersão.

Desse modo, conclui-se que os MNLGS englobam a família exponencial dupla apresentada por Efron (1986) e que, portanto, todas as distribuições denominadas família binomial, família gama e Poisson dupla podem ser escritas sob a forma dos MNLGS.

Na sequência, a família Poisson dupla foi escrita, com certo rigor nos detalhes, sob a forma dos MNLGS, uma vez que foi utilizada para ajustar um conjunto de dados no Capítulo 7.

Família Poisson dupla

Essa família é baseada na distribuição de Poisson, sendo que a variável resposta é inteira não negativa. Pode ser utilizada em controle de qualidade, visto que é comum haver interesse em determinar o número de falhas de equipamentos em função do tempo.

Segundo Efron (1986), assume-se que a variável resposta Y foi reescalonada, com função densidade de probabilidade dada por

$$\pi(y;\mu,\phi) = \left(\phi^{\frac{1}{2}} e^{-\phi\mu}\right)\left(\frac{e^{-y}y^y}{y!}\right)\left(\frac{e\mu}{y}\right)^{\phi y}, \tag{2.6}$$

com $y = 0, 1, 2, 3, \ldots$ e média μ e $\mathrm{Var}(y) = \frac{\mu}{\phi}$.

Considerando que $A(y) = \left(\frac{e^{-y}y^y}{y!}\right)$ e aplicando exponencial no logaritmo da função densidade de probabilidade tem-se que:

$$
\begin{aligned}
\pi(y;\mu,\phi) &= A(y)\exp\left\{\log\left[\phi^{\frac{1}{2}}e^{-\phi\mu}\left(\frac{e\mu}{y}\right)^{\phi y}\right]\right\} \\
&= A(y)\exp\left\{-\phi\mu + \frac{1}{2}\log\phi + \phi y\log\mu - \phi y\log y\right\} \\
&= A(y)\exp\left\{(y-\mu)\phi - \phi y\log y + \phi y\log\mu + \frac{1}{2}\log\phi\right\} \\
&= A(y)\exp\left\{(y-\mu)\phi + \phi\mu\log\mu - \phi\mu\log\mu - \phi y\log y + \phi y\log\mu + \frac{1}{2}\log\phi\right\} \\
&= A(y)\exp\left\{(y-\mu)\left[\phi + \phi\log\mu\right] - \phi y\log y + \phi\mu\log\mu + \frac{1}{2}\log\phi\right\}
\end{aligned}
$$

em que $\mathrm{T}(y) = -y\log y$, $\Psi^{(1,0)} = \phi + \phi\log\mu$ e

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

Dessa forma, determinaram-se as componentes da família Poisson dupla sob a forma MNLGS.

No contexto dos modelos não lineares generalizados, cabe ressaltar que, a distribuição binomial negativa é bastante utilizada para dados que apresentam superdispersão, no entanto, não pertence a classe apresentada por Dey et al. (1997), em outras palavras, esta distribuição não pode ser escrita sob a forma dos MNLGS.

2.3 Matriz Informação de Fisher

Seja as variáveis aleatórias $y_1, y_2, \ldots, y_n$, independentes. A função de verossimilhança é a densidade conjunta das n observações dada pela expressão (2.1). Aplicando-se o logaritmo, toma a seguinte forma:

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{i=1}^{n}\left\{(y_i-\mu_i)\Psi^{(0,1)}(\mu_i,\phi_i) + \phi_i T(y_i) + \Psi(\mu_i,\phi_i)\right\} + \sum_{i=1}^{n}\log A(y_i), \qquad (2.7)$$

com $T(y)$, $\Psi(\mu,\phi)$ e $A(y)$ funções conhecidas e $\Psi^{(r,s)} = \partial\Psi^{r+s}(\mu_i,\phi_i)/\partial\mu_i^r\partial\phi_i^s$.

A função escore total é definida por

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

em que as componentes são as primeiras derivadas da função $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$, em (2.7), com respeito aos parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$. Em notação matricial essas derivadas tomam as seguintes formas

$$U_{\boldsymbol{\beta}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \widetilde{X}^{\top}\Psi^{(2,0)}M_1(\boldsymbol{y} - \boldsymbol{\mu}), \tag{2.8}$$

e

$$U_{\boldsymbol{\gamma}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1(\boldsymbol{y} - \boldsymbol{\mu}) + \widetilde{S}^{\top}\Phi_1 T(\boldsymbol{y}) + \widetilde{S}^{\top}\Psi^{(0,1)}\Phi_1, \tag{2.9}$$

em que $M_1 = \operatorname{diag}\{\mu_1, \ldots, \mu_n\}$, com $\mu_i = \partial\mu_i/\partial\eta_i$ e $\Phi_1 = \operatorname{diag}\{\phi_1, \ldots, \phi_n\}$, com $\phi_i = \partial\phi_i/\partial\tau_i$, $i = 1, \ldots, n$.

A matriz de informação total de Fisher $K = K(\boldsymbol{\beta}, \boldsymbol{\gamma})$ para $(\boldsymbol{\beta}^{\top}, \boldsymbol{\gamma}^{\top})^{\top}$ é bloco diagonal, da forma,

$$K(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} K_{\boldsymbol{\beta},\boldsymbol{\beta}} & 0 \\ 0 & K_{\boldsymbol{\gamma},\boldsymbol{\gamma}} \end{pmatrix},$$

em que

$$K_{\boldsymbol{\beta},\boldsymbol{\beta}} = \mathrm{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\top}}\right] = \widetilde{X}^{\top}\Psi^{(2,0)}M_1^2\widetilde{X}, \tag{2.10}$$

$$K_{\boldsymbol{\gamma},\boldsymbol{\gamma}} = \mathrm{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^{\top}}\right] = -\widetilde{S}^{\top}\Psi^{(0,2)}\Phi_1^2\widetilde{S}, \tag{2.11}$$

sendo que $M_1^2 = \operatorname{diag}\{\mu_1^2, \ldots, \mu_n^2\}$ e $\Phi_1^2 = \operatorname{diag}\{\phi_1^2, \ldots, \phi_n^2\}$.

Como K é uma matriz bloco diagonal, os parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são globalmente ortogonais, e seus estimadores de máxima verossimilhança $\widehat{\boldsymbol{\beta}}$ e $\widehat{\boldsymbol{\gamma}}$, respectivamente, são assintoticamente independentes.

Para a obtenção dos estimadores de máxima verossimilhança é necessário utilizar técnicas iterativas de otimização; ver Nocedal e Wright (1999, Capítulo 18) e Luenberger (1973, Capítulo 12), pois, como estamos trabalhando com modelos não lineares, não é possível expressar os estimadores de máxima verossimilhança em uma forma analítica fechada. O processo iterativo utilizado foi o método de otimização não linear quasi-Newton BFGS, usando a linguagem de programação matricial `Ox` (Doornik, 2001).

2.4 Testes de Hipóteses

É comum que inferências para um determinado modelo envolvam apenas alguns, mas não todos os parâmetros do modelo. Nestes casos dizemos que os parâmetros envolvidos são parâmetros de interesse, enquanto que os demais são chamados de parâmetros de perturbação. Nesta seção, iremos mostrar como ficam os testes da razão de verossimilhanças e escore para testar um subconjunto do vetor de parâmetros de interesse dos modelos não lineares generalizados com superdispersão.

Considere que o vetor de parâmetros do modelo seja decomposto da seguinte forma: $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ e $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$, em que $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_{p_1})^\top$, $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, $\boldsymbol{\gamma}_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ e $\boldsymbol{\gamma}_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$ com $p_1 \leq p$ e $q_1 \leq q$. Essas decomposições induzem as correspondentes partições das matrizes

$$\widetilde{X} = (\widetilde{X}_1, \widetilde{X}_2), \quad \widetilde{S} = (\widetilde{S}_1, \widetilde{S}_2),$$

em que $\widetilde{X}_1$, $\widetilde{X}_2$, $\widetilde{S}_1$ e $\widetilde{S}_2$ são sub-matrizes conhecidas de posto completo e dimensões $n \times p_1$, $n \times (p - p_1)$, $n \times q_1$, $n \times (q - q_1)$, respectivamente.

Estamos interessados em testar as seguintes hipóteses:

(i) $\mathrm{H}_1 : \beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$, contra H_1' : pelo menos uma das igualdades é violada (testar conjuntamente uma parte do vetor de parâmetros de interesse conjuntamente, enquanto que, o vetor de parâmetros $(\beta_2^\top, \gamma_2^\top)^\top$ são considerados vetores de parâmetros de perturbação);

(ii) $\mathrm{H}_2 : \beta_1 = \beta_1^{(0)}$ contra $\mathrm{H}_2' : \beta_1 \neq \beta_1^{(0)}$ (testar isoladamente uma parte do vetor de parâmetros de interesse, enquanto que, o vetor de parâmetros $(\beta_2^\top, \gamma^\top)^\top$ í considerado vetor de parâmetros de perturbação);

(iii) $\mathrm{H}_3 : \gamma_1 = \gamma_1^{(0)}$ contra $\mathrm{H}_3' : \gamma_1 \neq \gamma_1^{(0)}$ (testar isoladamente uma parte do vetor de parâmetros de interesse, enquanto que, o vetor de parâmetros $(\beta^\top, \gamma_2^\top)^\top$ é considerado vetor de parâmetros de perturbação).

Em que $\beta_1^{(0)}$ e $\gamma_1^{(0)}$ são vetores de valores nominais. A seguir veremos dois testes muito utilizados em inferência estatística.

2.4.1 Teste da Razão de Verossimilhanças

Suponha inicialmente que estamos interessados no teste da hipótese $\mathrm{H}_1\colon \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$, $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ contra H_1': pelo menos uma das igualdades é violada. Denotamos por $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^\top, \hat{\boldsymbol{\beta}}_2^\top)^\top$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$, $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}_1^\top, \hat{\boldsymbol{\gamma}}_2^\top)^\top$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\gamma}$, sendo $\tilde{\boldsymbol{\beta}}_2$ e $\tilde{\boldsymbol{\gamma}}_2$ os estimadores de máxima verossimilhança restritos de $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}_2$, respectivamente. A estatística da razão de verossimilhanças é dada por

$$LR_1 = 2\{\ell(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2) - \ell(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2, \boldsymbol{\gamma}_1^{(0)}, \tilde{\boldsymbol{\gamma}}_2)\}. \tag{2.12}$$

A estatística da razão de verossimilhanças para o teste de $\mathrm{H}_2\colon \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ contra $\mathrm{H}_2'\colon \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_1^{(0)}$, é dada por

$$LR_2 = 2\{\ell(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}}) - \ell(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2, \tilde{\boldsymbol{\gamma}})\}, \tag{2.13}$$

em que $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^\top, \hat{\boldsymbol{\beta}}_2^\top)^\top$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$, $\hat{\boldsymbol{\gamma}}$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\gamma}$ e $\tilde{\boldsymbol{\beta}}_2$ o estimador de máxima verossimilhança restrito de $\boldsymbol{\beta}_2$.

A estatística da razão de verossimilhanças para o teste de $\mathrm{H}_3\colon \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ contra $\mathrm{H}_3'\colon \boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_1^{(0)}$ é dada por

$$LR_3 = 2\{\ell(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2) - \ell(\tilde{\boldsymbol{\beta}}, \boldsymbol{\gamma}_1^{(0)}, \tilde{\boldsymbol{\gamma}}_2)\}, \tag{2.14}$$

em que $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}_1^\top, \hat{\boldsymbol{\gamma}}_2^\top)^\top$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\gamma}$, $\hat{\boldsymbol{\beta}}$ o vetor de estimadores de máxima verossimilhança irrestrito de $\boldsymbol{\beta}$ e $\tilde{\boldsymbol{\beta}}$, $\tilde{\boldsymbol{\gamma}}_2$ os estimadorers de máxima verossimilhança restritos de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}_2$, respectivamente.

Ou seja, para grandes tamanhos de amostras, e sob a hipótese nula, as estatísticas LR_1, LR_2 e LR_3 têm uma distribuição aproximadamente $\chi^2_{(p_1+q_1)}$, $\chi^2_{p_1}$ e $\chi^2_{q_1}$, respectivamente, em que p_1 é o número de parâmetros a serem testados no vetor $\boldsymbol{\beta}$ e q_1 é o número de parâmetros a serem testados no vetor $\boldsymbol{\gamma}$. No entanto, quando o tamanho da amostra é pequeno estas aproximações para a distribuição χ^2 não é muito boa.

2.4.2 Teste Escore

O teste escore requer somente estimação sob a hipótese nula, o que o torna mais atraente quando a estimação sob a hipóteses alternativa exige um extenso trabalho computacional.

Assumimos inicialmente, testar a seguinte hipótese $\mathrm{H}_1: \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$, $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ contra H_1': pelo menos uma das igualdades é violada, em que $\boldsymbol{\beta}_1^{(0)}$ e $\boldsymbol{\gamma}_1^{(0)}$ são os vetores conhecidos de dimensões p_1 e q_1, respectivamente.

Seja $\boldsymbol{\rho}_1 = (\boldsymbol{\beta}_1^\top, \boldsymbol{\gamma}_1^\top)^\top$ o vetor de parâmetros de interesse e $\boldsymbol{\rho}_2 = (\boldsymbol{\beta}_2^\top, \boldsymbol{\gamma}_2^\top)^\top$ o vetor de parâmetros de perturbação. A função escore correspondente a $\boldsymbol{\rho}_1$ e $\boldsymbol{\rho}_2$ é dada por

$$U = (U_1^\top, U_2^\top)^\top,$$

em que

$$U_1^\top = \partial\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\boldsymbol{\rho}_1 = \Big(U_{\boldsymbol{\beta}_1}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2), U_{\boldsymbol{\gamma}_1}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)\Big)^\top$$

e

$$U_2^\top = \partial\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\boldsymbol{\rho}_2 = \Big(U_{\boldsymbol{\beta}_2}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2), U_{\boldsymbol{\gamma}_2}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)\Big)^\top.$$

Além disso, a matriz de informação total de Fisher correspondente é dada por

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} K_{\boldsymbol{\beta}_{11}} & 0 & K_{\boldsymbol{\beta}_{12}} & 0 \\ 0 & K_{\boldsymbol{\gamma}_{11}} & 0 & K_{\boldsymbol{\gamma}_{12}} \\ K_{\boldsymbol{\beta}_{21}} & 0 & K_{\boldsymbol{\beta}_{22}} & 0 \\ 0 & K_{\boldsymbol{\gamma}_{21}} & 0 & K_{\boldsymbol{\gamma}_{22}} \end{pmatrix},$$

em que as matrizes são definidas por $K_{\boldsymbol{\beta}_{11}} = \widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_1$, $K_{\boldsymbol{\beta}_{12}} = K_{\boldsymbol{\beta}_{21}}^\top = \widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2$, $K_{\boldsymbol{\beta}_{22}} = \widetilde{X}_2^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2$, $K_{\boldsymbol{\gamma}_{11}} = -\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_1$, $K_{\boldsymbol{\gamma}_{12}} = K_{\boldsymbol{\gamma}_{21}}^\top = -\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2$ e $K_{\boldsymbol{\gamma}_{22}} = -\widetilde{S}_2^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2$.

Temos também que,

$$K^{-1} = \begin{pmatrix} K^{11} & K^{12} \\ K^{21} & K^{22} \end{pmatrix},$$

sendo, K^{11} dado por

$$K^{11} = \begin{pmatrix} K_{\boldsymbol{\beta}_{11}} & \mathbf{0} \\ \mathbf{0} & K_{\boldsymbol{\gamma}_{11}} \end{pmatrix}^{-1} = \begin{pmatrix} K_{\boldsymbol{\beta}_{11}}^{-1} & \mathbf{0} \\ \mathbf{0} & K_{\boldsymbol{\gamma}_{11}}^{-1} \end{pmatrix},$$

em que

$$K_{\boldsymbol{\beta}_{11}}^{-1} = \left(\widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_1\right)^{-1} = (R_1^\top R_1)^{-1}$$

com $R_1 = M_1(\Psi^{(2,0)})^{1/2}\widetilde{X}_1$ e

$$K_{\boldsymbol{\gamma}_{11}}^{-1} = \left(-\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_1\right)^{-1} = (R_2^\top R_2)^{-1}$$

com $R_2 = \Phi_1(-\Psi^{(0,2)})^{1/2}\widetilde{S}_1$.

De acordo com Rao (1947), a estatística escore para o teste de H_1 é dada por

$$SR_1 = \widetilde{U}_1^{0^\top} \widetilde{K}^{11} \widetilde{U}_1^0,$$

em que,

$$\widetilde{U}_1^0 = \begin{pmatrix} U_{\boldsymbol{\beta}_1} \\ U_{\boldsymbol{\gamma}_1} \end{pmatrix} = \begin{pmatrix} \widetilde{X}_1^\top \Psi^{(2,0)} M_1(\boldsymbol{y} - \boldsymbol{\mu}) \\ \widetilde{S}_1^\top \Psi^{(1,1)} \Phi_1(\boldsymbol{y} - \boldsymbol{\mu}) + \widetilde{S}_1^\top \Phi_1 T(\boldsymbol{y}) + \widetilde{S}_1^\top \Psi^{(0,1)} \Phi_1 \mathbf{1} \end{pmatrix},$$

em que, $\widetilde{U}_1^0$ e $\widetilde{K}^{11}$ são avaliadas em $(\widehat{\boldsymbol{\beta}}^0, \widehat{\boldsymbol{\gamma}}^0) = \left(\boldsymbol{\beta}_1^{(0)^\top}, \tilde{\boldsymbol{\beta}}_2, \boldsymbol{\gamma}_1^{(0)^\top}, \tilde{\boldsymbol{\gamma}}_2\right)^\top$ sendo $\tilde{\boldsymbol{\beta}}_2$ e $\tilde{\boldsymbol{\gamma}}_2$ os estimadores de máxima verossimilhança restritos de $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}_2$, respectivamente. Para grandes amostras e sob a hipótese nula, SR_1 tem uma distribuição aproximadamente $\chi^2_{(p_1+q_1)}$. Assim, a estatística escore para o teste de H_1 é dada por

$$SR_1 = U_{\boldsymbol{\beta}_1}^\top (R_1^\top R_1)^{-1} U_{\boldsymbol{\beta}_1} + U_{\boldsymbol{\gamma}_1}^\top (R_2^\top R_2)^{-1} U_{\boldsymbol{\gamma}_1}. \tag{2.15}$$

Nosso interesse agora é testar as seguintes hipóteses $\mathrm{H}_2\colon \boldsymbol{\beta}_1 = \boldsymbol{\beta}_1^{(0)}$ contra $\mathrm{H}_2'\colon \boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_1^{(0)}$. Nesse caso, $\boldsymbol{\rho}_1 = \boldsymbol{\beta}_1$ é o vetor de parâmetros de interesse e $\boldsymbol{\rho}_2 = (\boldsymbol{\beta}_2^\top, \boldsymbol{\gamma}^\top)^\top$ é o vetor de parâmetros de perturbação. A função escore correspondente a $\boldsymbol{\rho}_1$ e $\boldsymbol{\rho}_2$ é dada por

$$U = (U_1^\top, U_2^\top)^\top,$$

em que

$$U_1^\top = \partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial \boldsymbol{\rho}_1 = U_{\boldsymbol{\beta}_1}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma})$$

e

$$U_2^\top = \partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial \boldsymbol{\rho}_2 = \left(U_{\boldsymbol{\beta}_2}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}), U_{\boldsymbol{\gamma}}^\top(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma})\right)^\top.$$

A matriz de informação de Fisher correspondente é dada por

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} K_{\boldsymbol{\beta}_{11}} & K_{\boldsymbol{\beta}_{12}} & 0 \\ K_{\boldsymbol{\beta}_{21}} & K_{\boldsymbol{\beta}_{22}} & 0 \\ 0 & 0 & K_{\boldsymbol{\gamma},\boldsymbol{\gamma}} \end{pmatrix},$$

em que $K_{\boldsymbol{\beta}_{11}} = \widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_1$, $K_{\boldsymbol{\beta}_{12}} = K_{\boldsymbol{\beta}_{21}}^\top = \widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2$, $K_{\boldsymbol{\beta}_{22}} = \widetilde{X}_2^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2$, $K_{\boldsymbol{\gamma},\boldsymbol{\gamma}} = -\widetilde{S}^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}$. Temos

$$K^{-1} = \begin{pmatrix} K_{\boldsymbol{\beta}}^{11} & K_{\boldsymbol{\beta}}^{12} & 0 \\ K_{\boldsymbol{\beta}}^{21} & K_{\boldsymbol{\beta}}^{22} & 0 \\ 0 & 0 & K_{\boldsymbol{\gamma},\boldsymbol{\gamma}}^{-1} \end{pmatrix},$$

sendo $\widetilde{K}_{\boldsymbol{\beta}}^{11}$ é dado por

$$\widetilde{K}_{\boldsymbol{\beta}}^{11} = \left(\widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_1\right) + \widetilde{X}_1^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2 \left(\widetilde{X}_2^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2\right)^{-1} \widetilde{X}_2 M_1^2 \Psi^{(2,0)} \widetilde{X}_1,$$

e $K_{\boldsymbol{\beta}}^{12} = -K_{\boldsymbol{\beta}}^{11} K_{\boldsymbol{\beta}_{12}} K_{\boldsymbol{\beta}_{22}}^{-1}$, $K_{\boldsymbol{\beta}}^{21} = -K_{\boldsymbol{\beta}_{22}}^{-1} K_{\boldsymbol{\beta}_{21}} K_{\boldsymbol{\beta}}^{11}$ e $K_{\boldsymbol{\beta}}^{22} = K_{\boldsymbol{\beta}_{22}}^{-1} + K_{\boldsymbol{\beta}_{22}}^{-1} K_{\boldsymbol{\beta}_{21}} K_{\boldsymbol{\beta}}^{11} K_{\boldsymbol{\beta}_{12}} K_{\boldsymbol{\beta}_{22}}^{-1}$.

A estatística escore para o teste de H_2 é dada por

$$SR_2 = \widetilde{U}_1^{0\top} \widetilde{K}_{\boldsymbol{\beta}}^{11} \widetilde{U}_1^0,$$

ou seja, $\widetilde{U}_1^0$ e $\widetilde{K}_{\boldsymbol{\beta}}^{11}$ são avaliadas em $\left(\boldsymbol{\beta}_1^{(0)\top}, \widetilde{\boldsymbol{\beta}}_2, \widetilde{\boldsymbol{\gamma}}^\top\right)^\top$ sendo $\widetilde{\boldsymbol{\beta}}_2$ e $\widetilde{\boldsymbol{\gamma}}$ os estimadores de máxima verossimilhança restritos de $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}$, respectivamente. Para grandes amostras e sob a hipótese nula, SR_2 tem uma distribuição aproximadamente $\chi^2_{p_1}$.

Mas, $\widetilde{U}_1^0$ é dada por

$$\widetilde{U}_1^0 = U_{\boldsymbol{\beta}_1} = \widetilde{X}_1^\top \Psi^{(2,0)} M_1 (\boldsymbol{y} - \boldsymbol{\mu}).$$

Assim, a estatística escore para o teste de H_2 é dada por

$$SR_2 = U_{\boldsymbol{\beta}_1}^\top \widetilde{K}_{\boldsymbol{\beta}}^{11} U_{\boldsymbol{\beta}_1}. \tag{2.16}$$

Finalmente, para testar $\mathrm{H}_3\colon \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1^{(0)}$ contra $\mathrm{H}_3'\colon \boldsymbol{\gamma}_1 \neq \boldsymbol{\gamma}_1^{(0)}$, seja $\boldsymbol{\rho}_1 = \boldsymbol{\gamma}_1$ o vetor de parâmetros de interesse e $\boldsymbol{\rho}_2 = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}_2^\top)^\top$ o vetor de parâmetros de perturbação. A função escore correspondente a $\boldsymbol{\rho}_1$ e $\boldsymbol{\rho}_2$ é dada por

$$U = (U_1^\top, U_2^\top)^\top,$$

em que

$$U_1^\top = \partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\boldsymbol{\rho}_1 = U_{\boldsymbol{\gamma}_1}^\top(\boldsymbol{\beta},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2)$$

e

$$U_2^\top = \partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\boldsymbol{\rho}_2 = \left(U_{\boldsymbol{\gamma}_2}^\top(\boldsymbol{\beta},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2), U_{\boldsymbol{\beta}}^\top(\boldsymbol{\beta},\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2)\right)^\top.$$

A matriz de informação total de Fisher correspondente é dada por

$$K = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} = \begin{pmatrix} K_{\boldsymbol{\gamma}_{11}} & K_{\boldsymbol{\gamma}_{12}} & 0 \\ K_{\boldsymbol{\gamma}_{21}} & K_{\boldsymbol{\gamma}_{22}} & 0 \\ 0 & 0 & K_{\boldsymbol{\beta},\boldsymbol{\beta}} \end{pmatrix},$$

em que as matrizes são definidas por $K_{\boldsymbol{\beta},\boldsymbol{\beta}} = \widetilde{X}^\top \Psi^{(2,0)} M_1^2 \widetilde{X}$, $K_{\boldsymbol{\gamma}_{11}} = -\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_1$, $K_{\boldsymbol{\gamma}_{12}} = K_{\boldsymbol{\gamma}_{21}}^\top = -\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2$ e $K_{\boldsymbol{\gamma}_{22}} = -\widetilde{S}_2^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2$. Temos também que

$$K^{-1} = \begin{pmatrix} K_{\boldsymbol{\gamma}}^{11} & K_{\boldsymbol{\gamma}}^{12} & 0 \\ K_{\boldsymbol{\gamma}}^{21} & K_{\boldsymbol{\gamma}}^{22} & 0 \\ 0 & 0 & K_{\boldsymbol{\beta},\boldsymbol{\beta}}^{-1} \end{pmatrix},$$

sendo $K_{\boldsymbol{\gamma}}^{11}$ é dado por

$$K_{\boldsymbol{\gamma}}^{11} = \left(\widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_1\right) + \widetilde{S}_1^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2 \left(\widetilde{S}_2^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S}_2\right)^{-1} \widetilde{S}_2 \Phi_1^2 \Psi^{(0,2)} \widetilde{S}_1$$

e $K_{\boldsymbol{\gamma}}^{12} = -K_{\boldsymbol{\gamma}}^{11} K_{\boldsymbol{\gamma}_{12}} K_{\boldsymbol{\gamma}_{22}}^{-1}$, $K_{\boldsymbol{\gamma}}^{21} = -K_{\boldsymbol{\gamma}_{22}}^{-1} K_{\boldsymbol{\gamma}_{21}} K_{\boldsymbol{\gamma}}^{11}$ e $K_{\boldsymbol{\gamma}}^{22} = K_{\boldsymbol{\gamma}_{22}}^{-1} + K_{\boldsymbol{\gamma}_{22}}^{-1} K_{\boldsymbol{\gamma}_{21}} K_{\boldsymbol{\gamma}}^{11} K_{\boldsymbol{\gamma}_{12}} K_{\boldsymbol{\gamma}_{22}}^{-1}$.

A estatística escore para o teste de H_3 é dada por

$$SR_3 = \widetilde{U}_1^{0\top} \widetilde{K}_{\boldsymbol{\gamma}}^{11} \widetilde{U}_1^0,$$

ou seja, $\widetilde{U}_1^0$ e $\widetilde{K}_{\boldsymbol{\gamma}}^{11}$ são avaliadas em $\left(\widetilde{\boldsymbol{\beta}}^\top, \widetilde{\boldsymbol{\gamma}}^{(0)\top}, \widetilde{\boldsymbol{\gamma}}_2\right)^\top$ sendo $\widetilde{\boldsymbol{\beta}}$ e $\widetilde{\boldsymbol{\gamma}}_2$ os estimadores de máxima verossimilhança restritos de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}_2$, respectivamente. Em grandes amostras e sob a hipótese nula, SR_3 tem uma distribuição aproximadamente $\chi^2_{q_1}$.

Mas, $\widetilde{U}_1^0$ é dada por

$$\widetilde{U}_1^0 = U_{\boldsymbol{\gamma}_1} = \widetilde{S}_1^\top \Psi^{(1,1)} \Phi_1 (\boldsymbol{y} - \boldsymbol{\mu}) + \widetilde{S}_1^\top \Phi_1 T(\boldsymbol{y}) + \widetilde{S}_1^\top \Psi^{(0,1)} \Phi_1 \mathbf{1}.$$

Então, a estatística escore para o teste de H_3 é dada por

$$SR_3 = U_{\boldsymbol{\gamma}_1}^\top \widetilde{K}_{\boldsymbol{\gamma}}^{11} U_{\boldsymbol{\gamma}_1}. \tag{2.17}$$

O teste escore proposto por Rao (1947) tem uma grande vantagem em relação ao teste da razão de verossimilhanças, pois requer um custo computacional menos intenso, enquanto que o teste da razão de verossimilhanças requer estimação restrita e irrestrita dos parâmetros. No entanto, sabe-se que a presença de um número considerável de parâmetros de perturbação nos modelos, deteriora a qualidade das aproximações envolvidas nas inferências em resultados assintóticos. Desta forma, torna-se importante obter refinamento de testes que reduzam estes problemas.

CAPÍTULO 3

Correção de Bartlett para a estatística da razão de verossimilhanças

3.1 Introdução

Sabe-se que em problemas regulares a estatística da razão de verossimilhanças (LR) converge para a distribuição qui-quadrado com m graus de liberdade (χ^2_m), em que m é a diferença entre as dimensões dos espaços paramétricos sob as duas hipóteses testadas. Entretanto, em pequenas amostras, a aproximação de primeira ordem pode não ser satisfatória devido ao fato de que a aproximação se dá quando o tamanho da amostra tende a infinito. Uma alternativa para resolver esse problema é usar a teoria assintótica de alta ordem, que pode tornar métodos inferenciais tradicionais mais precisos em pequenas amostras.

Bartlett (1937), com o objetivo de melhorar esta aproximação, propôs multiplicar a estatística LR por um fator de correção, dado por $(1+c/m)^{-1}$, obtendo assim uma nova estatística conhecida como estatística da razão de verossimilhanças corrigida:

$$LR^* = \frac{LR}{1+c/m}, \tag{3.1}$$

em que c é uma constante conhecida de $O(n^{-1})$ escolhida apropriadamente tal que, sob H_0, $\mathrm{E}(LR^*) = m+O(n^{-3/2})$, sendo n o tamanho da amostra. É possivel mostrar que a estatística LR^* tem distribuição (χ^2_m) a menos de um erro de $O(n^{-2})$ sob H_0, isto é, $\mathrm{P}(LR^* > x_\alpha) = \alpha+O(n^{-2})$, enquanto que $\mathrm{P}(LR > x_\alpha) = \alpha + O(n^{-1})$, sendo x_α é o quantil de ordem $(1-\alpha)$ da distribuição

χ^2_m. Para maiores detalhes, ver Barndorff-Nielsen e Hall (1988).

Lawley (1956) desenvolveu uma fórmula geral para o fator de correção c que é uma função de vários cumulantes de derivadas do logaritmo da função de verossimilhança até a quarta ordem.

Cordeiro et al. (2006) obtiveram uma expressão matricial para o fator de correção de Bartlett à estatística da razão de verossimilhanças nos modelos lineares generalizados com superdispersão. Neste capítulo, nosso objetivo é generalizar os resultados obtidos por Cordeiro et al. (2006), isto é, derivar uma fórmula geral para o fator de correção à estatística da razão de verossimilhanças nos modelos não lineares generalizados com superdispersão.

3.2 O método de Lawley

Consideremos n variáveis aleatórias independentes $y_1, \ldots, y_n$, ou seja, o vetor de observações $y = (y_1, \ldots, y_n)^\top$. Seja $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$ o logaritmo da função de verossimilhança total dado em 2.1 que depende dos parâmetros desconhecidos $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ e $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$, em que $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_{p_1})^\top$, $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, $\boldsymbol{\gamma}_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ e $\boldsymbol{\gamma}_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$ com $p_1 \leq p$ e $q_1 \leq q$.

Segundo Cox e Hinkley, (1974), $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$ é regular com respeito às derivadas em relação aos componentes de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ até a quarta ordem. Nosso interesse aqui é testar a hipótese $\mathrm{H}_1 : \beta_1 = \beta_1^{(0)},\ \gamma_1 = \gamma_1^{(0)}$ contra H_1' : pelo menos uma das igualdades é violada (testar conjuntamente uma parte do vetor de parâmetros de interesse conjuntamente, enquanto que, o vetor de parâmetros $(\beta_2^\top, \gamma_2^\top)^\top$ é considerado vetor de parâmetros de perturbação). De acordo com Lawley (1956), podemos escrever o valor esperado da estatística da razão de verossimilhanças até a ordem $\mathrm{O}(n^{-1})$ como

$$\mathrm{E}\left[2\{\ell(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\gamma}}) - \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})\}\right] = p + q + \epsilon_{p+q}, \tag{3.2}$$

o termo ϵ_{p+q} é de $\mathrm{O}(n^{-1})$ sendo expresso como:

$$\epsilon_{p+q} = \sum_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \left(l_{rstu} - l_{rstuvw}\right), \tag{3.3}$$

em que $\sum_{\boldsymbol{\beta}, \boldsymbol{\gamma}}$ denota o somatório sobre todas as combinações de $p + q$ parâmetros de $(\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ e os índices r, s, t, u, v e w variam em ambos os vetores $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$. A notação do lado direito de (3.3) segue de Cordeiro (1987), e os l's são obtidos por:

$$l_{rstu} = \kappa^{rs}\kappa^{tu}(\kappa_{rstu}/4 - \kappa_{rst}^{(u)} - \kappa_{rt}^{(su)}) \tag{3.4}$$

e

$$l_{rstuvw} = \kappa^{rs}\kappa^{tu}\kappa^{vw}\left\{\kappa_{rtv}(\kappa_{suw}/6 - \kappa_{sw}^{(u)}) + \kappa_{rtu}(\kappa_{svw}/4 - \kappa_{sw}^{(v)}) + \kappa_{rt}^{(v)}\kappa_{sw}^{(u)} + \kappa_{rt}^{(u)}\kappa_{sw}^{(v)}\right\}. \tag{3.5}$$

Definimos a seguinte notação para as derivadas do logaritmo da função de verossimilhança total com respeito aos componentes de $\boldsymbol{\beta}$: $U_r = \partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\beta_r$, $U_{rs} = \partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\beta_r\partial\beta_s$ e assim por diante. Já a notação padrão para os cumulantes conjuntos são: $\kappa_{rs} = \mathrm{E}(U_{rs})$, $\kappa_{rst} = \mathrm{E}(U_{rst})$, $\kappa_{rstu} = \mathrm{E}(U_{rstu})$, $\kappa_{r,s} = \mathrm{E}(U_rU_s)$, $\kappa_{r,st} = \mathrm{E}(U_rU_{st})$, $\kappa_{r,stu} = \mathrm{E}(U_rU_{stu})$, $\kappa_{rs,tu} = \mathrm{E}(U_{rs}U_{tu}) - \kappa_{rs}\kappa_{tu}$, $\kappa_{r,s,tu} = \mathrm{E}(U_rU_sU_{tu}) - \kappa_{r,s}\kappa_{tu}$ e $\kappa_{r,s,t,u} = \mathrm{E}(U_rU_sU_tU_u) - \kappa_{r,s}\kappa_{t,u} - \kappa_{r,t}\kappa_{s,u} - \kappa_{r,u}\kappa_{s,u}$.

Denotamos as derivadas dos cumulantes em relação aos componentes do vetor $\boldsymbol{\beta}$ por $\kappa_{rs}^{(t)} = \partial\kappa_{rs}/\partial\beta_t$, $\kappa_{rs}^{(tu)} = \partial\kappa_{rs}/\partial\beta_t\partial\beta_u$, $\kappa_{rst}^{(u)} = \partial\kappa_{rst}/\partial\beta_u$, etc. A matriz de informação de Fisher para $\boldsymbol{\beta}$ tem elementos $\kappa_{r,s} = -\kappa_{rs}$, sendo $\kappa^{r,s} = -\kappa^{rs}$ os correspondentes de sua inversa. O mesmo pode ser feito para as derivadas do logaritmo da função de verossimilhança total com respeito aos componentes de $\boldsymbol{\gamma}$, mas para distinguir das derivadas de $\boldsymbol{\beta}$, as derivadas de $\boldsymbol{\gamma}$ foram utilizadas letra maiúsculas, por exemplo $U_R = \partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\gamma_R$. E o mesmo ocorre para as derivadas mista, por exemplo $U_{rS} = \partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})/\partial\beta_r\partial\gamma_S$.

Os cumulantes satisfazem relações denominadas identidades de Bartlett, que frequentemente facilitam seus cálculos. Alguns exemplos destas identidades são: $\kappa_{r,s} = -\kappa_{rs}$, $\kappa_{rst}^{(u)} = \kappa_{rstu} + \kappa_{u,rst}$, $\kappa_{r,s,t} = -\kappa_{rst} - \sum_{(3)}\kappa_{r,st}$, $\kappa_{r,s,t} = 2\kappa_{rst} - \sum_{(3)}\kappa_{rs}^{(t)}$, $\kappa_{r,s,t,u} = -3\kappa_{rstu} + 2\sum_{(4)}\kappa_{rst}^{(u)} - \sum_{(6)}\kappa_{rs}^{(tu)} + \sum_{(3)}\kappa_{rs,tu}$ e $\kappa_{r,s,tu} = \kappa_{rstu} - 2\kappa_{rtu}^{(s)} + \kappa_{tu}^{(rs)} - \kappa_{rs,tu}$, em que $\sum_{(k)}$ denota o somatório que varre todas as k permutações de índices. Assim, por exemplo, $\sum_{(3)}\kappa_{rs}^{(t)} = \kappa_{rs}^{(t)} + \kappa_{rt}^{(s)} + \kappa_{st}^{(r)}$. Nos MNLGSs, o cálculo do termo ϵ_{p+q} dado em (3.3) é simples devido à ortogonalidade dos parâmetros, de modo que vários cumulantes mistos são nulos e pois o inverso da matriz de informação de Fisher é bloco diagonal.

Segundo Botter e Cordeiro (1997), o termo ϵ_{p+q} pode ser decomposto da seguinte forma:

$$\epsilon_{p+q} = \epsilon_p(\boldsymbol{\beta}) + \epsilon_q(\boldsymbol{\gamma}) + \epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma}), \tag{3.6}$$

em que $\epsilon_p(\boldsymbol{\beta})$ depende apenas da componente sistemática para a média, $\epsilon_q(\boldsymbol{\gamma})$ depende apenas da componente sistemática para a dispersão e $\epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma})$ depende das duas componetes sistemáticas. Para o modelo não linear generalizado com superdispersão, os termos em (3.6) são apresentados a seguir

$$\begin{aligned}
\epsilon_p(\boldsymbol{\beta}) = {} & \frac{1}{4}\mathbf{1}^\top\Psi^{(4,0)}M_1^4 Z_{\boldsymbol{\beta}d}^2\mathbf{1} + \mathbf{1}^\top\Psi^{(3,0)}M_1^2M_2Z_{\boldsymbol{\beta}d}^2\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_2^2Z_{\boldsymbol{\beta}d}^2\mathbf{1} \\
& - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}d}D_{\boldsymbol{\beta}}\mathbf{1} - \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_1^2D_{\boldsymbol{\beta}}^2\mathbf{1} - \frac{1}{3}\mathbf{1}^\top\Psi^{(3,0)}M_1^3Z_{\boldsymbol{\beta}}^{(3)}M_1^3\Psi^{(3,0)}\mathbf{1} \\
& - \mathbf{1}^\top\Psi^{(3,0)}M_1^3Z_{\boldsymbol{\beta}}^{(3)}M_1M_2\Psi^{(2,0)}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}}^{(3)}M_1M_2\Psi^{(2,0)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\beta}}Z_{\boldsymbol{\beta}d}M_1M_2\Psi^{(2,0)}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\beta}}D_{\boldsymbol{\beta}}M_1^2\Psi^{(2,0)}\mathbf{1} \\
& - \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_1^2D_{\boldsymbol{\beta}}Z_{\boldsymbol{\beta}}D_{\boldsymbol{\beta}}M_1^2\Psi^{(2,0)}\mathbf{1},
\end{aligned} \tag{3.7}$$

$$\begin{aligned}
\epsilon_q(\boldsymbol{\gamma}) = {} & \frac{1}{4}\mathbf{1}^\top\Psi^{(0,4)}\Phi_1^4Z_{\boldsymbol{\gamma}d}^2\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^2\Phi_2Z_{\boldsymbol{\gamma}d}^2\mathbf{1} - \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_2^2Z_{\boldsymbol{\gamma}d}^2\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}d}D_{\boldsymbol{\gamma}}\mathbf{1} + \mathbf{1}^\top\frac{1}{2}\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}d}D_{\boldsymbol{\gamma}}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2D_{\boldsymbol{\gamma}}^2\mathbf{1} \\
& + \frac{1}{6}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}}^{(3)}\Phi_1^3\Psi^{(0,3)}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}}^{(3)}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}D_{\boldsymbol{\gamma}}\Phi_1^2\Psi^{(0,2)}\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}D_{\boldsymbol{\gamma}}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2D_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}}D_{\boldsymbol{\gamma}}\Phi_1^2\Psi^{(0,2)}\mathbf{1}
\end{aligned} \tag{3.8}$$

e

$$\begin{aligned}
\epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma}) = {} & \frac{1}{2}\mathbf{1}^\top\Psi^{(2,2)}M_1^2\Phi_1^2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}d}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}d}\mathbf{1} \\
& - \frac{1}{4}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_2Z_{\boldsymbol{\beta}d}D_{\boldsymbol{\gamma}}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}}^{(2)}\odot Z_{\boldsymbol{\gamma}}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\beta}d}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}D_{\boldsymbol{\gamma}}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1},
\end{aligned} \tag{3.9}$$

em que $Z_{\boldsymbol{\beta}} = \widetilde{X}(\widetilde{X}^\top\Psi^{(2,0)}\ M_1^2\widetilde{X})^{-1}\widetilde{X}^\top$, $Z_{\boldsymbol{\gamma}} = \widetilde{S}(-\widetilde{S}^\top\Psi^{(0,2)}\Phi_1^2\widetilde{S})^{-1}\widetilde{S}^\top$, $\mathbf{1}$ o vetor $n\times 1$ de uns, $Z_{\boldsymbol{\beta}}^{(3)} = Z_{\boldsymbol{\beta}}^{(2)}\odot Z_{\boldsymbol{\beta}}$, $Z_{\boldsymbol{\beta}}^{(2)} = Z_{\boldsymbol{\beta}}\odot Z_{\boldsymbol{\beta}}$ em que $\odot$ denota o produto de Hadamard (Rao, 1973, p. 30), ou seja, o elemento ij de $Z_{\boldsymbol{\beta}}^{(2)}$ é z_{ij}^2.

Utilizamos a notação $Z_{\boldsymbol{\beta}d}$ e $Z_{\boldsymbol{\gamma}d}$ para representar matrizes diagonais formadas pelos correspondentes elementos das diagonais das matrizes $Z_{\boldsymbol{\beta}}$ e $Z_{\boldsymbol{\gamma}}$, respectivamente. Definimos as matrizes $D_{\boldsymbol{\beta}} = \operatorname{diag}\{d_{\boldsymbol{\beta}1}, \ldots, d_{\boldsymbol{\beta}n}\}$ e $D_{\boldsymbol{\gamma}} = \operatorname{diag}\{d_{\boldsymbol{\gamma}1}, \ldots, d_{\boldsymbol{\gamma}n}\}$, cujos elementos são dados, respectivamente, por $d_{\boldsymbol{\beta}i} = \operatorname{tr}\left(\widetilde{\widetilde{X}}_i K^{-1}_{\boldsymbol{\beta},\boldsymbol{\beta}}\right)$ com $\widetilde{\widetilde{X}}_i = \partial^2\boldsymbol{\eta}/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top$ sendo a i-ésima linha da matriz $\widetilde{\widetilde{X}}$ e $d_{\boldsymbol{\gamma}i} = \operatorname{tr}\left(\widetilde{\widetilde{S}}_i K^{-1}_{\boldsymbol{\gamma},\boldsymbol{\gamma}}\right)$ com $\widetilde{\widetilde{S}}_i = \partial^2\boldsymbol{\tau}/\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^\top$ sendo a i-ésima linha da matriz $\widetilde{\widetilde{S}}$.

Nos MLGSs, temos que $\widetilde{X} = X$ e $\widetilde{S} = S$ e $\widetilde{\widetilde{X}} = \widetilde{\widetilde{S}} = D_{\boldsymbol{\beta}} = D_{\boldsymbol{\gamma}} = \mathbf{0}$. Portanto, os termos (3.7) a (3.9) são dados por:

$$
\begin{aligned}
\epsilon_p(\boldsymbol{\beta}) = {} & \frac{1}{4}\mathbf{1}^\top\Psi^{(4,0)}M_1^4Z_{\boldsymbol{\beta}d}^2\mathbf{1} + \mathbf{1}^\top\Psi^{(3,0)}M_1^2M_2Z_{\boldsymbol{\beta}d}^2\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_2^2Z_{\boldsymbol{\beta}d}^2\mathbf{1} \\
& - \frac{1}{3}\mathbf{1}^\top\Psi^{(3,0)}M_1^3Z_{\boldsymbol{\beta}}^{(3)}M_1^3\Psi^{(3,0)}\mathbf{1} - \mathbf{1}^\top\Psi^{(3,0)}M_1^3Z_{\boldsymbol{\beta}}^{(3)}M_1M_2\Psi^{(2,0)}\mathbf{1} \\
& - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}}^{(3)}M_1M_2\Psi^{(2,0)}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,0)}M_1M_2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\beta}}Z_{\boldsymbol{\beta}d}M_1M_2\Psi^{(2,0)}\mathbf{1},
\end{aligned}
\tag{3.10}
$$

$$
\begin{aligned}
\epsilon_q(\boldsymbol{\gamma}) = {} & \frac{1}{4}\mathbf{1}^\top\Psi^{(0,4)}\Phi_1^4Z_{\boldsymbol{\gamma}d}^2\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^2\Phi_2Z_{\boldsymbol{\gamma}d}^2\mathbf{1} - \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_2^2Z_{\boldsymbol{\gamma}d}^2\mathbf{1} \\
& + \frac{1}{6}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}}^{(3)}\Phi_1^3\Psi^{(0,3)}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}}^{(3)}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2Z_{\boldsymbol{\gamma}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1}
\end{aligned}
\tag{3.11}
$$

e

$$
\begin{aligned}
\epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma}) = {} & \frac{1}{2}\mathbf{1}^\top\Psi^{(2,2)}M_1^2\Phi_1^2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}d}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_2Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}d}\mathbf{1} \\
& - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}}^{(2)} \odot Z_{\boldsymbol{\gamma}}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\beta}d}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{\boldsymbol{\gamma}}Z_{\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1},
\end{aligned}
\tag{3.12}
$$

As equações (3.10) e (3.12), coincidem com as equações (3.2) a (3.4) de Cordeiro et. al. (2006), pp. 942–943.

Das equações (3.7) a (3.9) podemos facilmente obter nossas estatísticas de teste. O termo c_1 do fator de correção de Bartlett para o teste de H_1 é dado por

$$
c_1 = 1 + \frac{\epsilon_{p+q} - \epsilon_{p-p_1+q-q_1}}{p_1 + q_1}, \tag{3.13}
$$

em que ϵ_{p+q} pode ser calculado de (3.6) a (3.9) e $\epsilon_{p-p_1+q-q_1}$ também mas com $\widetilde{X}_2$, $\widetilde{\widetilde{X}}_2$, $\widetilde{S}_2$ e $\widetilde{\widetilde{S}}_2$ ao invés de $\widetilde{X}$, $\widetilde{\widetilde{X}}$, $\widetilde{S}$ e $\widetilde{\widetilde{S}}$, ou seja, $\epsilon_{p-p_1+q-q_1} = \epsilon_p(\boldsymbol{\beta}_2) + \epsilon_q(\boldsymbol{\gamma}_2) + \epsilon_{p,q}(\boldsymbol{\beta}_2, \boldsymbol{\gamma}_2)$.

Assim, para o teste de H_1, temos que a estatística da razão de verossimilhanças corrigida é

$$LR_1^* = \frac{LR_1}{c_1}, \tag{3.14}$$

com c_1 dado em (3.13) e LR_1 expressa como

$$LR_1 = 2\{\ell(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2) - \ell(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2, \boldsymbol{\gamma}_1^{(0)}, \tilde{\boldsymbol{\gamma}}_2)\}, \tag{3.15}$$

em que $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}}_1$ e $\hat{\boldsymbol{\gamma}}_2$ são os estimadores de máxima verossimilhança irrestritos de $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\gamma}_1$ e $\boldsymbol{\gamma}_2$, sendo $\widetilde{\beta}_2$ e $\widetilde{\boldsymbol{\gamma}}_2$ os estimadores de máxima verossimilhança restritos para $\boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}_2$, respectivamente. Assintoticamente e sob H_1, a estatística LR_1^* tem distribuição $\chi^2_{p_1+q_1}$.

O termo c_2 do fator de correção de Bartlett para o teste de H_2 é dado por

$$c_2 = 1 + \frac{\epsilon_p(\boldsymbol{\beta}) - \epsilon_{p_2}(\boldsymbol{\beta}_2) + \epsilon_{p,q}(\boldsymbol{\beta}, \boldsymbol{\gamma}) - \epsilon_{p_2,q}(\boldsymbol{\beta}_2, \boldsymbol{\gamma})}{p_1}, \tag{3.16}$$

em que ϵ_{p_2} e $\epsilon_{p_2,q}$ são

$$\begin{aligned}\epsilon_{p_2}(\boldsymbol{\beta}_2) = {} & \frac{1}{4}\mathbf{1}^\top \Psi^{(4,0)} M_1^4 Z_{2\boldsymbol{\beta}d}^2 \mathbf{1} + \mathbf{1}^\top \Psi^{(3,0)} M_1^2 M_2 Z_{2\boldsymbol{\beta}d}^2 \mathbf{1} + \frac{1}{4}\mathbf{1}^\top \Psi^{(2,0)} M_2^2 Z_{2\boldsymbol{\beta}d}^2 \mathbf{1} \\ & - \frac{1}{2}\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} D_{2\boldsymbol{\beta}} \mathbf{1} - \frac{1}{4}\mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}^2 \mathbf{1} - \frac{1}{3}\mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}}^{(3)} M_1^3 \Psi^{(3,0)} \mathbf{1} \\ & - \mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}}^{(3)} M_1 M_2 \Psi^{(2,0)} \mathbf{1} - \frac{1}{2}\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}}^{(3)} M_1 M_2 \Psi^{(2,0)} \mathbf{1} \\ & + \frac{1}{4}\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}} Z_{2\boldsymbol{\beta}d} M_1 M_2 \Psi^{(2,0)} \mathbf{1} - \frac{1}{2}\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}} D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1} \\ & - \frac{1}{4}\mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}} Z_{2\boldsymbol{\beta}} D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1},\end{aligned} \tag{3.17}$$

e

$$\begin{aligned}\epsilon_{p_2,q}(\boldsymbol{\beta}_2, \boldsymbol{\gamma}) = {} & \frac{1}{2}\mathbf{1}^\top \Psi^{(2,2)} M_1^2 \Phi_1^2 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}d} \mathbf{1} + \frac{1}{2}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_2 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}d} \mathbf{1} \\ & - \frac{1}{4}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_2 Z_{2\boldsymbol{\beta}d} D_{\boldsymbol{\gamma}} \mathbf{1} - \frac{1}{2}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}}^{(2)} \odot Z_{\boldsymbol{\gamma}} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\ & + \frac{1}{4}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}} Z_{2\boldsymbol{\beta}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \frac{1}{2}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}} Z_{\boldsymbol{\gamma}d} \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\ & + \frac{1}{2}\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}} Z_{\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \frac{1}{2}\mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 Z_{2\boldsymbol{\beta}d} Z_{\boldsymbol{\gamma}} D_{\boldsymbol{\gamma}} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1},\end{aligned} \tag{3.18}$$

em que $Z_{2\boldsymbol{\beta}} = \widetilde{X}_2(\widetilde{X}_2^\top \Psi^{(2,0)} M_1^2 \widetilde{X}_2)^{-1}\widetilde{X}_2^\top$, $Z_{2\boldsymbol{\beta}d}$ representa a matriz diagonal formada pelos correspondentes elementos da diagonal da matriz $Z_{2\boldsymbol{\beta}}$ e $D_{2\boldsymbol{\beta}} = \mathrm{diag}\{d_{\boldsymbol{\beta}1}, \ldots, d_{\boldsymbol{\beta}n}\}$, cujos elementos são dados respectivamente por $d_{\boldsymbol{\beta}i} = \mathrm{tr}\left(\widetilde{\widetilde{X}}_{2i} K^{-1}_{\boldsymbol{\beta}_2,\boldsymbol{\beta}_2}\right)$ com $\widetilde{\widetilde{X}}_{2i} = \partial^2\boldsymbol{\eta}/\partial\boldsymbol{\beta}_2\partial\boldsymbol{\beta}_2^\top$ sendo a i-ésima linha da matriz $\widetilde{\widetilde{X}}_2$, $i = 1, \ldots, n$.

Assim, para o teste de H_2, temos que a estatística da razão de verossimilhanças corrigida é

$$LR_2^* = \frac{LR_2}{c_2}, \tag{3.19}$$

com c_2 dado em (3.16) e LR_2 expressa como

$$LR_2 = 2\{\ell(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}}) - \ell(\boldsymbol{\beta}_1^{(0)}, \tilde{\boldsymbol{\beta}}_2, \hat{\boldsymbol{\gamma}})\}, \tag{3.20}$$

em que $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2$ e $\hat{\boldsymbol{\gamma}}$ são os estimadores de máxima verossimilhança irrestritos de $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ e $\boldsymbol{\gamma}$ sendo $\widetilde{\beta}_2$ o estimador de máxima verossimilhança restrito para $\boldsymbol{\beta}_2$. Assintoticamente e sob H_2, a estatística LR_2 tem distribuição $\chi^2_{p_1}$.

O termo c_3 do fator de correção de Bartlett para o teste de H_3 é dado por

$$c_3 = 1 + \frac{\epsilon_q(\boldsymbol{\gamma}) - \epsilon_{q_2}(\boldsymbol{\gamma}_2) + \epsilon_{p,q}(\boldsymbol{\beta}, \boldsymbol{\gamma}) - \epsilon_{p,q_2}(\boldsymbol{\beta}, \boldsymbol{\gamma}_2)}{q_1}, \tag{3.21}$$

em que ϵ_{q_2} e ϵ_{p,q_2} são

$$\begin{aligned}
\epsilon_{q_2}(\boldsymbol{\gamma}_2) = {} & \frac{1}{4}\mathbf{1}^\top\Psi^{(0,4)}\Phi_1^4 Z_{2\boldsymbol{\gamma}d}^2\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^2\Phi_2 Z_{2\boldsymbol{\gamma}d}^2\mathbf{1} - \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_2^2 Z_{2\boldsymbol{\gamma}d}^2\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3 Z_{2\boldsymbol{\gamma}d} D_{2\boldsymbol{\gamma}}\mathbf{1} + \mathbf{1}^\top\frac{1}{2}\Phi_1\Phi_2 Z_{2\boldsymbol{\gamma}d} D_{2\boldsymbol{\gamma}}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2 D_{2\boldsymbol{\gamma}}^2\mathbf{1} \\
& + \frac{1}{6}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3 Z_{2\boldsymbol{\gamma}}^{(3)}\Phi_1^3\Psi^{(0,3)}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2 Z_{2\boldsymbol{\gamma}}^{(3)}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1\Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,3)}\Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} D_{2\boldsymbol{\gamma}}\Phi_1^2\Psi^{(0,2)}\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} D_{2\boldsymbol{\gamma}}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} + \frac{1}{4}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2 D_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}} D_{2\boldsymbol{\gamma}}\Phi_1^2\Psi^{(0,2)}\mathbf{1}
\end{aligned} \tag{3.22}$$

e

$$
\begin{aligned}
\epsilon_{p,q_2}(\boldsymbol{\beta},\boldsymbol{\gamma}_2) = {} & \frac{1}{2}\mathbf{1}^\top\Psi^{(2,2)}M_1^2\Phi_1^2Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}d}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_2Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}d}\mathbf{1} \\
& - \frac{1}{4}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_2Z_{\boldsymbol{\beta}d}D_{2\boldsymbol{\gamma}}\mathbf{1} - \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}}^{(2)}\odot Z_{2\boldsymbol{\gamma}}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + \frac{1}{4}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}}Z_{\boldsymbol{\beta}d}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}}Z_{2\boldsymbol{\gamma}d}\Phi_1^3\Psi^{(0,3)}\mathbf{1} \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(2,1)}M_1^2\Phi_1Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}}Z_{2\boldsymbol{\gamma}d}\Phi_1\Phi_2\Psi^{(0,2)}\mathbf{1} \qquad (3.23) \\
& + \frac{1}{2}\mathbf{1}^\top\Psi^{(0,2)}\Phi_1^2Z_{\boldsymbol{\beta}d}Z_{2\boldsymbol{\gamma}}D_{2\boldsymbol{\gamma}}M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1}, \qquad (3.24)
\end{aligned}
$$

em que $Z_{2\boldsymbol{\gamma}} = \widetilde{S}_2(\widetilde{S}_2^\top\Psi^{(0,2)}\Phi_1^2\widetilde{S}_2)^{-1}\widetilde{S}_2^\top$, $Z_{2\boldsymbol{\gamma}d}$ representa a matriz diagonal formada pelos correspondentes elementos da diagonal da matriz $Z_{2\boldsymbol{\gamma}}$ e $D_{2\boldsymbol{\gamma}} = \mathrm{diag}\{d_{\boldsymbol{\gamma}1},\ldots,d_{\boldsymbol{\gamma}n}\}$, cujos elementos são dados respectivamente por $d_{\boldsymbol{\gamma}i} = \mathrm{tr}\left(\widetilde{\widetilde{S}}_{2i}K^{-1}_{\boldsymbol{\gamma}_2,\boldsymbol{\gamma}_2}\right)$ e seja $\widetilde{\widetilde{S}}_{2i} = \partial^2\boldsymbol{\tau}/\partial\boldsymbol{\gamma}_2\partial\boldsymbol{\gamma}_2^\top$ a i-ésima linha da matriz $\widetilde{\widetilde{S}}_2$.

Assim, para o teste de H_3, temos que a estatística da razão de verossimilhanças corrigida é

$$LR_3^* = \frac{LR_3}{c_3}, \qquad (3.25)$$

com c_3 dado em (3.21) e LR_3 expressa como

$$LR_3 = 2\{\ell(\hat{\boldsymbol{\beta}},\hat{\boldsymbol{\gamma}}_1,\hat{\boldsymbol{\gamma}}_2) - \ell(\hat{\boldsymbol{\beta}},\boldsymbol{\gamma}_1^{(0)},\tilde{\boldsymbol{\gamma}}_2)\}, \qquad (3.26)$$

em que $\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}_1$ e $\hat{\boldsymbol{\gamma}}_2$ são os estimadores de máxima verossimilhança irrestritos de $\boldsymbol{\beta}, \boldsymbol{\gamma}_1$ e $\boldsymbol{\gamma}_2$ sendo $\widetilde{\gamma}_2$ o estimador de máxima verossimilhança restrito para $\boldsymbol{\gamma}_2$. Assintoticamente e sob H_3, a estatística LR_3 tem distribuição $\chi^2_{q_1}$.

Outra forma assintoticamente equivalente para a estatística da razão de verossimilhanças corrigida, para o teste de H_1, é dada por:

$$\widetilde{LR}_1 = LR_1 \cdot d_1, \qquad (3.27)$$

em que o termo d_1 do fator de correção de Bartlett é dado por

$$d_1 = 1 - \frac{\epsilon_{p+q} - \epsilon_{p-p_1+q-q_1}}{p_1+q_1}, \qquad (3.28)$$

e temos que $\widetilde{LR}_1$ tem distribuição assintoticamente $\chi^2_{p_1+q_1}$ sob H_1. Procedendo analogamente, para as hipóteses H_2 e H_3 teremos as estatísticas da razão de verossimilhanças corrigidas $\widetilde{LR}_2$ e $\widetilde{LR}_3$, como:

$$\widetilde{LR}_2 = LR_2 \cdot d_2, \tag{3.29}$$

$$\widetilde{LR}_3 = LR_3 \cdot d_3, \tag{3.30}$$

em que

$$d_2 = 1 - \frac{\epsilon_p(\boldsymbol{\beta}) - \epsilon_{p_2}(\boldsymbol{\beta}_2) + \epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma}) - \epsilon_{p_2,q}(\boldsymbol{\beta}_2,\boldsymbol{\gamma})}{p_1}, \tag{3.31}$$

e

$$d_3 = 1 - \frac{\epsilon_q(\boldsymbol{\gamma}) - \epsilon_{q_2}(\boldsymbol{\gamma}_2) + \epsilon_{p,q}(\boldsymbol{\beta},\boldsymbol{\gamma}) - \epsilon_{p,q_2}(\boldsymbol{\beta},\boldsymbol{\gamma}_2)}{q_1}, \tag{3.32}$$

e, sob as hipóteses nulas H_2 e H_3, as estatísticas da razão de verossimilhanças corrigidas, $\widetilde{LR}_2$ e $\widetilde{LR}_3$, têm distribuição assintoticamente $\chi^2_{p_1}$ e $\chi^2_{q_1}$, respectivamente.

CAPÍTULO 4

Correção tipo-Bartlett para a estatística escore

4.1 Introdução

Testes estatísticos clássicos são usualmente baseados em aproximações assintóticas de primeira ordem, e podem ser poucos precisos quando o tamanho da amostra é pequeno ou mesmo moderado. Esse também é o caso do teste escore (SR).

Sabe-se que em problemas regulares a estatística escore (SR) converge para a distribuição qui-quadrado com m graus de liberdade (χ^2_m), em que m é o número de restrições impostas sob a hipótese nula (H_0).

Para contornar este problema, Cordeiro e Ferrari (1991) obtiveram um fator de correção para a estatística escore, denominado fator de correção tipo-Bartlett, de forma que a estatística escore aperfeiçoada por este fator de correção tem distribuição χ^2_m até ordem $\mathrm{O}(n^{-1})$ sob a hipótese nula. Sob certas condições de regularidade, a estatística escore modificada proposta por Cordeiro e Ferrari (1991) (SR^*) é obtida através da multiplicação da estatística escore original por um polinômio de segundo grau da própria estatística e com esta modificação, a distribuição da estatística escore corrigida tem uma melhor aproximação para a distribuição qui-quadrado de referência.

A partir do trabalho de Cordeiro e Ferrari (1991), muitos resultados têm sido publicados sobre fatores de correção tipo-Bartlett para a estatística escore em várias classes de modelos de regressão. Cordeiro et al. (2003), por exemplo, derivam um fator de correção tipo-Bartlett para a estatística escore para modelos lineares com covariadas na dispersão, generalizando os resultados de Cordeiro, Ferrari e Paula (1993) e Cribari-Neto e Ferrari (1995). Recentemente, Cysneiros et al. (2010) derivam fator de correção tipo-Bartlett para a estatística escore em modelos de regressão simétricos não lineares.

Neste capítulo, nosso objetivo é derivar uma fórmula geral para o fator de correção tipo-Bartlett, proposta por Cordeiro e Ferrari (1991), para a estatística escore nos modelos não lineares generalizados com superdispersão (MNLGSs).

4.2 Correções tipo-Bartlett

Cordeiro e Ferrari (1991) mostraram que, sob certas condições de regularidade, a estatística escore SR pode ser aperfeiçoada por uma correção tipo-Bartlett, que não é bem uma correção tradicional de Bartlett pois envolve um polinômio de segundo grau da estatística original. A forma corrigida para a estatística escore SR, proposta por Cordeiro e Ferrari, é dada por:

$$SR^* = SR\left[1 - aSR^2 - bSR - c\right], \tag{4.1}$$

com

$$a = \frac{A_3}{12m(m+2)(m+4)}, \quad b = \frac{A_2 - 2A_3}{12m(m+2)} \quad c = \frac{A_1 - A_2 + A_3}{12m}$$

em que as quantidades a, b e c são de ordem $\mathrm{O}(n^{-1})$, sendo estas funções de cumulantes que envolvem derivadas do logaritmo da função de verossimilhança até quarta ordem. Assintoticamente e sob a hipótese nula, SR^* tem distribuição χ^2_m. Os coeficientes A_1, A_2 e A_3 podem ser escritos como:

$$\begin{aligned}
A_1 &= 3\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}}(\kappa_{ijk} + 2\kappa_{i,jk})(\kappa_{rst} + 2\kappa_{rs,t})a_{ij}a_{st}m_{kr} - 6\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}}(\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}a_{kr}m_{st} \\
&+ 6\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}}(\kappa_{i,jk} - \kappa_{i,j,k})(\kappa_{rst} + 2\kappa_{rs,t})a_{js}a_{kt}m_{ir} - 6\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}}(\kappa_{i,j,k,r} + \kappa_{i,j,kr})a_{kr}m_{ij} \\
&= 3A_{11} - 6A_{12} + 6A_{13} - 6A_{14},
\end{aligned}$$

$$
\begin{aligned}
A_2 &= -3\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} \kappa_{i,j,k}\kappa_{r,s,t}a_{kr}m_{ij}m_{st} + 6\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} (\kappa_{ijk} + 2\kappa_{i,jk})\kappa_{r,s,t}a_{ij}m_{kr}m_{st} \\
&\quad - 6\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} \kappa_{i,j,k}\kappa_{r,s,t}a_{kt}m_{ir}m_{js} + 3\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} \kappa_{i,j,k,r}m_{ij}m_{kr} \\
&= -3A_{21} + 6A_{22} - 6A_{23} + 3A_{24}, \\
A_3 &= 3\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} \kappa_{i,j,k}m_{ij}m_{kr}m_{st} + 2\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}} \kappa_{i,j,k}\kappa_{r,s,t}m_{ir}m_{js}m_{kt} \\
&= 3A_{31} + 2A_{32},
\end{aligned}
$$

sendo que os índices i, j, k, r, s, t variam sobre os componentes dos vetores $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ e $\sum_{\boldsymbol{\beta},\boldsymbol{\gamma}}$ denota todas as possíveis combinações de $p + q$ parâmetros de $\beta_1, \ldots, \beta_p$ e $\gamma_1, \ldots, \gamma_q$. As matrizes A e M são dadas por

$$
A = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & K^{22} \end{pmatrix}
$$

e $M = K^{-1} - A$.

Uma dificuldade encontrada com o uso da correção SR^* é o fato de que, em alguns casos, os coeficientes a, b e c podem ser difíceis de calcular. A equação (4.1) nem sempre é uma função monótona. Para contornar este problema, Kakizawa (1996) sugere uma transformação monótona

$$
K(SR) = SR^* + P(SR), \tag{4.2}
$$

envolvendo a estatística SR^* e os coeficientes a, b e c, em que $P(SR)$ é um polinômio de quinto grau da estatística original SR e é de ordem $O(n^{-2})$. A expressão $P(SR)$ é dada por

$$
P(SR) = \frac{1}{4}\left\{c^2 SR + 2bc SR^2 + \left(2ac + \frac{4}{3}b^2\right) SR^3 + 3ab SR^4 + \frac{9}{5}a^2 SR^5\right\}.
$$

Também, Cordeiro et al. (1998) apresentaram uma fórmula alternativa para melhorar a estatísica escore que é uma transformação monótona de SR. A estatística alternativa $\widetilde{SR}$ é expressa em termos da função de distribuição normal padrão, $\Phi(\cdot)$, por

$$
\widetilde{SR} = \sqrt{\frac{\pi}{3a}}\exp\left(\frac{b^2}{3a} - c\right)\left\{\Phi\left(\sqrt{6a}SR + \sqrt{\frac{2}{3a}}b\right) - \Phi\left(\sqrt{\frac{2}{3a}}b\right)\right\}, \tag{4.3}
$$

se $a > 0$ (a é sempre não negativa) e

$$\widetilde{SR} = \frac{1}{2b} \exp(-c)1 - \exp(-2bSR)$$

se $a = 0$ e $b \neq 0$. Note que, se $a = b = 0$, SR^* é uma transformação monótona de SR e não é necessário definir uma estatística corrigida alternativa. As três estatísicas SR^*, $K(SR)$ e $\widetilde{SR}$ são equivalentes até a segunda ordem, ou seja, elas diferem tipicamente pela $O(n^{-3/2})$.

Consideremos n variáveis aleatórias independentes $y_1, \ldots, y_n$, ou seja, o vetor de observações $y = (y_1, \ldots, y_n)^\top$. Seja $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$ o logaritmo da função de verossimilhança total dado y que depende dos parâmetros desconhecidos $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ e $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \boldsymbol{\gamma}_2^\top)^\top$, em que $\boldsymbol{\beta}_1 = (\beta_1, \ldots, \beta_{p_1})^\top$, $\boldsymbol{\beta}_2 = (\beta_{p_1+1}, \ldots, \beta_p)^\top$, $\boldsymbol{\gamma}_1 = (\gamma_1, \ldots, \gamma_{q_1})^\top$ e $\boldsymbol{\gamma}_2 = (\gamma_{q_1+1}, \ldots, \gamma_q)^\top$ com $p_1 \leq p$ e $q_1 \leq q$.

Nosso objetivo aqui é testar a hipótese $\mathrm{H}_1 : \beta_1 = \beta_1^{(0)}, \ \gamma_1 = \gamma_1^{(0)}$, contra H_1' : pelo menos uma das igualdades é violada (testar conjuntamente uma parte do vetor de parâmetros de interesse conjuntamente, enquanto que o vetor de parâmetros $(\beta_2^\top, \gamma_2^\top)^\top$ é considerado vetor de parâmetros de perturbação.

Definimos a seguinte notação para as derivadas do logaritmo da função de verossimilhança total com respeito aos componentes de $\boldsymbol{\beta}$: $U_r = \partial\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r$, $U_{rs} = \partial^2\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r\partial\beta_s$ e assim por diante. Já a notação padrão para os cumulantes conjuntos é: $\kappa_{rs} = \mathrm{E}(U_{rs})$, $\kappa_{rst} = \mathrm{E}(U_{rst})$, $\kappa_{rstu} = \mathrm{E}(U_{rstu})$, $\kappa_{r,s} = \mathrm{E}(U_r U_s)$, $\kappa_{r,st} = \mathrm{E}(U_r U_{st})$, $\kappa_{r,stu} = \mathrm{E}(U_r U_{stu})$, $\kappa_{rs,tu} = \mathrm{E}(U_{rs} U_{tu}) - \kappa_{rs}\kappa_{tu}$, $\kappa_{r,s,tu} = \mathrm{E}(U_r U_s U_{tu}) - \kappa_{r,s}\kappa_{tu}$ e $\kappa_{r,s,t,u} = \mathrm{E}(U_r U_s U_t U_u) - \kappa_{r,s}\kappa_{t,u} - \kappa_{r,t}\kappa_{s,u} - \kappa_{r,u}\kappa_{s,u}$.

Denotamos as derivadas dos cumulantes em relação aos componentes do vetor $\boldsymbol{\beta}$ por $\kappa_{rs}^{(t)} = \partial\kappa_{rs}/\partial\beta_t$, $\kappa_{rs}^{(tu)} = \partial\kappa_{rs}/\partial\beta_t\partial\beta_u$, $\kappa_{rst}^{(u)} = \partial\kappa_{rst}/\partial\beta_u$, etc. A matriz de informação de Fisher para $\boldsymbol{\beta}$ tem elementos $\kappa_{r,s} = -\kappa_{rs}$, sendo $\kappa^{r,s} = -\kappa^{rs}$ os correspondentes de sua inversa. O mesmo pode ser feito para as derivadas do logaritmo da função de verossimilhança total com respeito aos componentes de $\boldsymbol{\gamma}$, mas para distinguir das derivadas de $\boldsymbol{\beta}$, para as derivadas de $\boldsymbol{\gamma}$ foram utilizadas letras maiúsculas, por exemplo $U_R = \partial\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\gamma_R$, $U_{rS} = \partial^2\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})/\partial\beta_r\partial\gamma_S$, e outros.

Os cumulantes satisfazem relações denominadas identidades de Bartlett, que frequentemente facilitam seus cálculos. Alguns exemplos destas identidades são: $\kappa_{r,s} = -\kappa_{rs}$, $\kappa_{rst}^{(u)} = \kappa_{rstu} + \kappa_{u,rst}$,

$\kappa_{r,s,t} = -\kappa_{rst} - \sum_{(3)} \kappa_{r,st}$, $\kappa_{r,s,t} = 2\kappa_{rst} - \sum_{(3)} \kappa_{rs}^{(t)}$, $\kappa_{r,s,t,u} = -3\kappa_{rstu} + 2\sum_{(4)} \kappa_{rst}^{(u)} - \sum_{(6)} \kappa_{rs}^{(tu)} + \sum_{(3)} \kappa_{rs,tu}$ e $\kappa_{r,s,tu} = \kappa_{rstu} - 2\kappa_{rtu}^{(s)} + \kappa_{tu}^{(rs)} - \kappa_{rs,tu}$, em que $\sum_{(k)}$ denota o somatório que varre todas as k permutações de índices. Assim, por exemplo, $\sum_{(3)} \kappa_{rs}^{(t)} = \kappa_{rs}^{(t)} + \kappa_{rt}^{(s)} + \kappa_{st}^{(r)}$.

Nos MNLGSs, os termos A_{11} a A_{14}, A_{21} a A_{24}, A_{31} e A_{32} foram obtidos substituindo os cumulantes (ver Apêndice A) calculados para o modelo. Esses termos de A_{11} a A_{14} assumem as seguintes formas matriciais:

$$
\begin{aligned}
A_{11} = {}& \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) Z_{2\boldsymbol{\beta}d} M_1 M_2 \Psi^{(2,0)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \\
& \times D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} \\
& \times (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\beta}d} \Phi_1 M_1^2 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
& + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} \\
& \times \Phi_1^2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}), \\
& \times Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} \\
& \times (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 \\
& \times Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,2)} \\
& \times \Phi_1^2 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} \Phi_1^2 \Psi^{(0,2)} \mathbf{1},
\end{aligned} \tag{4.4}
$$

$$
\begin{aligned}
A_{12} = {}& \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}}(Z - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}} Z_{\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
& + \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}} Z_{\boldsymbol{\beta}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \\
& \times \Phi_1^2 \Psi^{(1,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\gamma}} \\
& \times (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 \\
& \times Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
& - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \\
& \times \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},
\end{aligned} \tag{4.5}
$$

$$
\begin{aligned}
A_{13} = &-2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 M_2 \Psi^{(2,0)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \\
&\times M_2 \Psi^{(2,0)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{\boldsymbol{\beta}} \odot Z_{\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&- \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1}
\end{aligned} \tag{4.6}
$$

e

$$
\begin{aligned}
A_{14} = &\mathbf{1}^\top \left[2\frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}} M_1^4 - \Psi^{(4,0)} M_1^4 - \Psi^{(3,0)} M_1^2 M_2 \right] Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(3,0)} M_1^3 D_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} + \mathbf{1}^\top \left[2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 + \Psi^{(1,2)} M_2 \Phi_1^2 \right] \\
&\times Z_{2\boldsymbol{\gamma}d} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} + \mathbf{1}^\top \left[2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 \right. \\
&\left. - \Psi^{(2,2)} M_1^2 \Phi_1^2 - \Psi^{(2,1)} M_1^2 \Phi_2 \right] Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1} - \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 D_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1} \\
&+ \mathbf{1}^\top \left[2 \times \frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} \Phi_1^4 - \Psi^{(0,4)} \Phi_1^4 - \Psi^{(0,3)} \Phi_1^2 \Phi_1 \right] Z_{2\boldsymbol{\gamma}d} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1} \\
&- \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1}.
\end{aligned} \tag{4.7}
$$

Os termos de A_{21} a A_{24} assumem as seguintes formas matriciais:

$$
\begin{aligned}
A_{21} = &\mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},
\end{aligned} \tag{4.8}
$$

$$
\begin{aligned}
A_{22} &= \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta} d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \\
&\times (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta} d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
&\times (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&- \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
&\times (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&- \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
&\times (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},
\end{aligned} \tag{4.9}
$$

$$
\begin{aligned}
A_{23} &= \mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\beta}}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}.
\end{aligned} \tag{4.10}
$$

e

$$
\begin{aligned}
A_{24} &= \mathbf{1}^\top \left[3\frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}} M_1^4 - \Psi^{(4,0)} M_1^4 \right] (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) \mathbf{1} \\
&+ 2 \left[\frac{\left(\Psi^{(3,0)}\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 + 2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 \right. \\
&\left. - \Psi^{(2,0)} \Psi^{(0,2)} M_1^2 \Phi_1^2 \right] (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1} \\
&+ \mathbf{1}^\top \left[3\frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} \Phi_1^4 - \Psi^{(0,4)} \Phi_1^4 \right] Z_{2\boldsymbol{\gamma} d}(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1}
\end{aligned} \tag{4.11}
$$

Os termos de A_{31} e A_{32} assumem as seguintes formas matriciais:

$$
\begin{aligned}
A_{31} &= \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ 2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \qquad (4.12)
\end{aligned}
$$

e

$$
\begin{aligned}
A_{32} &= \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&+ 3\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&+ 3\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&+ \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}, \qquad (4.13)
\end{aligned}
$$

em que as matrizes envolvidas: $Z_{\boldsymbol{\beta}} = \widetilde{X}(\widetilde{X}^\top \Psi^{(2,0)} M_1^2 \widetilde{X})^{-1} \widetilde{X}^\top$, $Z_{\boldsymbol{\gamma}} = \widetilde{S}(-\widetilde{S}^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S})^{-1} \widetilde{S}^\top$, $\mathbf{1}$ o vetor $n \times 1$ de uns, $Z_{\boldsymbol{\beta} d}$ e $Z_{\boldsymbol{\gamma} d}$ as matrizes diagonais com os elementos correspondentes de $Z_{\boldsymbol{\beta}}$ e $Z_{\boldsymbol{\gamma}}$, respectivamente. Assim como, as matrizes $D_{\boldsymbol{\beta}} = \mathrm{diag}\{d_{\boldsymbol{\beta} 1}, \ldots, d_{\boldsymbol{\beta} n}\}$, em que $d_{\boldsymbol{\beta} i} = \mathrm{tr}\left(\widetilde{\widetilde{X}}_i K^{-1}_{\boldsymbol{\beta},\boldsymbol{\beta}}\right)$ e $\widetilde{\widetilde{X}}_i = \partial^2 \boldsymbol{\eta} / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top$ e $D_{\boldsymbol{\gamma}} = \mathrm{diag}\{d_{\boldsymbol{\gamma} 1}, \ldots, d_{\boldsymbol{\gamma} n}\}$, em que $d_{\boldsymbol{\gamma} i} = \mathrm{tr}\left(\widetilde{\widetilde{S}}_i K^{-1}_{\boldsymbol{\gamma},\boldsymbol{\gamma}}\right)$ e $\widetilde{\widetilde{S}}_i = \partial^2 \boldsymbol{\tau} / \partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^\top$.

Pode-se calcular o fator de correção tipo-Bartlett para a estatística escore considerando as seguintes hipóteses, a saber: $\mathrm{H}_2 : \beta_1 = \beta_1^{(0)}$ contra $\mathrm{H}_2' : \beta_1 \neq \beta_1^{(0)}$ (testar isoladamente uma parte do vetor de parâmetros de interesse, enquanto que, o vetor de parâmetros $(\beta_2^\top, \gamma^\top)^\top$ é considerado vetor de parâmetros de perturbação) e $\mathrm{H}_3 : \gamma_1 = \gamma_1^{(0)}$ contra $\mathrm{H}_3' : \gamma_1 \neq \gamma_1^{(0)}$ (testar isoladamente uma parte do vetor de parâmetros de interesse, enquanto que, o vetor de parâmetros $(\gamma_2^\top, \beta^\top)^\top$ é considerado vetor de parâmetros de perturbação.

No caso do teste de hipótese H_2 temos que $Z_{\boldsymbol{\gamma}} = Z_{2\boldsymbol{\gamma}}$, e os termos A_{11} a A_{14} são dados por:

$$A_{11} = \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) Z_{2\boldsymbol{\beta}d} M_1 M_2 \Psi^{(2,0)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})$$
$$\times D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) D_{2\boldsymbol{\beta}} M_1^2 \Psi^{(2,0)} \mathbf{1},$$

$$A_{12} = \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}}(Z - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}} Z_{\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d}) M_1^3$$
$$\times \Psi^{(3,0)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}$$
$$\times (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1},$$

$$A_{13} = -2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 M_2 \Psi^{(2,0)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1$$
$$\times M_2 \Psi^{(2,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1}$$

e

$$A_{14} = \mathbf{1}^\top \left[2\frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}} M_1^4 - \Psi^{(4,0)} M_1^4 - \Psi^{(3,0)} M_1^2 M_2 \right] Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1}$$
$$+ \mathbf{1}^\top \Psi^{(3,0)} M_1^3 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} + \mathbf{1}^\top \left[2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 + \Psi^{(1,2)} M_2 \Phi_1^2 \right]$$
$$\times Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 D_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1}.$$

Os termos de A_{21} a A_{24} assumem as seguintes formas matriciais:

$$A_{21} = \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1}$$
$$+ \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) Z_{2\boldsymbol{\gamma}}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1},$$

$$A_{22} = \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})$$
$$\times (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1},$$

$$A_{23} = \mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\beta}}) M_1^3 \Psi^{(3,0)} \mathbf{1}$$
$$+ \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1}$$

e

$$A_{24} = \mathbf{1}^\top \left[3 \frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}} M_1^4 - \Psi^{(4,0)} M_1^4 \right] (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1}.$$

Os termos de A_{31} e A_{32} assumem as seguintes formas matriciais:

$$A_{31} = \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1}$$

e

$$A_{32} = \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^3 \Psi^{(3,0)} \mathbf{1}.$$

No caso do teste de hipótese H_3 temos que $Z_{\boldsymbol{\beta}} = Z_{2\boldsymbol{\beta}}$, então os termos de A_{11} a A_{14} são:

$$A_{11} = \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\beta}d} \Phi_1 M_1^2 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d}$$
$$\times \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}$$
$$\times (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2$$
$$\times Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 Z_{2\boldsymbol{\beta}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)}$$
$$\times \Phi_1^3 Z_{2\boldsymbol{\gamma}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)}$$
$$\times \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} + 2\mathbf{1}^\top$$
$$\times \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) D_{2\boldsymbol{\gamma}} \Phi_1^2 \Psi^{(0,2)} \mathbf{1},$$

$$A_{12} = \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1^2 D_{2\boldsymbol{\beta}} Z_{\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})$$
$$\times M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}$$
$$- \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1}$$
$$- \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3$$
$$\times \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},$$

$$A_{13} = 2\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{\boldsymbol{\beta}} \odot Z_{\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}$$
$$- \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}} Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1}$$

e

$$\begin{aligned} A_{14} = \mathbf{1}^\top & \left[2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 - \Psi^{(2,1)} M_1^2 \Phi_2 \right] Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1} \\ & - \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 D_{2\boldsymbol{\beta}} (Z - Z_{2\boldsymbol{\gamma} d}) \mathbf{1} \\ & + \mathbf{1}^\top \left[2\frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} \Phi_1^4 - \Psi^{(0,4)} \Phi_1^4 - \Psi^{(0,3)} \Phi_1^2 \Phi_1 \right] Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1} \\ & - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1}. \end{aligned}$$

Os termos de A_{21} a A_{24} assumem as seguintes formas matriciais:

$$\begin{aligned} A_{21} = & +\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\ & + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{22} = & -\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\ & - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} \\ & - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1^2 D_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}, \end{aligned}$$

$$\begin{aligned} A_{23} = & +\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\ & + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \end{aligned}$$

e

$$A_{24} = +\mathbf{1}^\top \left[3\frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} \Phi_1^4 - \Psi^{(0,4)} \Phi_1^4 \right] Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \mathbf{1}.$$

Os termos de A_{31} e A_{32} assumem as seguintes formas matriciais:

$$A_{31} = \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}$$

e

$$A_{32} = \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}.$$

Nos MLGSs temos que envolvem as matrizes de segundas derivadas, $D_{\boldsymbol{\beta}}$ e $D_{\boldsymbol{\gamma}}$, são nulas. Os termos de A_{11} a A_{14} dados em (4.4) a (4.7), são

$$
\begin{aligned}
A_{11} = {} & \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) Z_{2\boldsymbol{\beta}d} M_1 M_2 \Psi^{(2,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\beta}d} \\
& \times \Phi_1 M_1^2 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} \\
& \times (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
& + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \\
& \times \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
& \times Z_{2\boldsymbol{\gamma}d} \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1},
\end{aligned}
$$

$$
\begin{aligned}
A_{12} = {} & \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}} (Z - Z_{2\boldsymbol{\beta}d}) M_1^3 \Psi^{(3,0)} \mathbf{1} + \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) M_1 \\
& \times \Phi_1^2 \Psi^{(1,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}d} \\
& \times Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \\
& \times \Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
& - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}d} Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},
\end{aligned}
$$

$$
\begin{aligned}
A_{13} = {} & -2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 M_2 \Psi^{(2,0)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta}}^{(2)} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \\
& \times M_2 \Psi^{(2,0)} \mathbf{1} + 2\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{\boldsymbol{\beta}} \odot Z_{\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \\
& \times M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma}} \\
& \odot Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1}
\end{aligned}
$$

e

$$
\begin{aligned}
A_{14} = {} & \mathbf{1}^\top \left[2 \frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}} M_1^4 - \Psi^{(4,0)} M_1^4 - \Psi^{(3,0)} M_1^2 M_2 \right] Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} \\
& + \mathbf{1}^\top \left[2 \frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 + \Psi^{(1,2)} M_2 \Phi_1^2 \right] Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d}) \mathbf{1} \\
& + \mathbf{1}^\top \left[2 \frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}} M_1^2 \Phi_1^2 - \Psi^{(2,2)} M_1^2 \Phi_1^2 - \Psi^{(2,1)} M_1^2 \Phi_2 \right] Z_{2\boldsymbol{\beta}d}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1} \\
& + \mathbf{1}^\top \left[2 \frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}} \Phi_1^4 - \Psi^{(0,4)} \Phi_1^4 - \Psi^{(0,3)} \Phi_1^2 \Phi_1 \right] Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d}) \mathbf{1}.
\end{aligned}
$$

Os termos de A_{21} a A_{24} dados em (4.8) a (4.11), assumem as seguintes formas matriciais:

$$
\begin{aligned}
A_{21} &= \mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&\quad + 2\mathbf{1}^\top \Psi^{(3,0)} M_1^3 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&\quad + 2\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) Z_{2\boldsymbol{\beta}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) Z_{2\boldsymbol{\gamma}} (Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1},
\end{aligned}
$$

$$
\begin{aligned}
A_{22} &= \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(2,0)} M_1 M_2 Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
&\quad \times (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&\quad - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \\
&\quad \times (Z_{\boldsymbol{\beta} d} - Z_{2\boldsymbol{\beta} d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
&\quad - \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma} d} (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma} d} - Z_{2\boldsymbol{\gamma} d}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1},
\end{aligned}
$$

$$
\begin{aligned}
A_{23} &= \mathbf{1}^\top \Psi^{(3,0)} M_1^3 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\beta}}) M_1^3 \Psi^{(3,0)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&\quad + 2\mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
&\quad + 2\mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(1,2)} M_1 \Phi_1^2 Z_{2\boldsymbol{\beta}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) M_1 \Phi_1^2 \Psi^{(1,2)} \mathbf{1} \\
&\quad + \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 Z_{2\boldsymbol{\gamma}} \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1}
\end{aligned}
$$

e

$$
\begin{aligned}
A_{24} = {} & \mathbf{1}^{\top}\left[3\frac{\left(\Psi^{(3,0)}\right)^2}{\Psi^{(2,0)}}M_1^4 - \Psi^{(4,0)}M_1^4\right](Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})\mathbf{1} \\
& + 2\left[\frac{\left(\Psi^{(3,0)}\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}}M_1^2\Phi_1^2 + 2\frac{\left(\Psi^{(2,1)}\right)^2}{\Psi^{(2,0)}}M_1^2\Phi_1^2 - \Psi^{(2,2)}M_1^2\Phi_1^2\right. \\
& \left. - \Psi^{(2,0)}\Psi^{(0,2)}M_1^2\Phi_1^2\right](Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}d})\mathbf{1} \\
& + \mathbf{1}^{\top}\left[3\frac{\left(\Psi^{(1,2)}\right)^2}{\Psi^{(2,0)}}\Phi_1^4 - \Psi^{(0,4)}\Phi_1^4\right]Z_{2\boldsymbol{\gamma}d}(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})\mathbf{1}.
\end{aligned}
$$

Os termos de A_{31} e A_{32} dados em (4.12) a (4.13), assumem as seguintes formas matriciais:

$$
\begin{aligned}
A_{31} = {} & \mathbf{1}^{\top}\Psi^{(3,0)}M_1^3(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d})M_1^3\Psi^{(3,0)}\mathbf{1} \\
& + 2\mathbf{1}^{\top}\Psi^{(3,0)}M_1^3(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z - Z_{2\boldsymbol{\gamma}d})M_1\Phi_1^2\Psi^{(1,2)}\mathbf{1} \\
& + 2\mathbf{1}^{\top}\Psi^{(3,0)}M_1^3(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + \mathbf{1}^{\top}\Psi^{(2,1)}M_1^2\Phi_1(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\beta}d} - Z_{2\boldsymbol{\beta}d})M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + \mathbf{1}^{\top}\Psi^{(1,2)}M_1\Phi_1^2(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})M_1\Phi_1^2\Psi^{(1,2)}\mathbf{1} \\
& + \mathbf{1}^{\top}\Psi^{(0,3)}\Phi_1^3(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})(Z_{\boldsymbol{\gamma}d} - Z_{2\boldsymbol{\gamma}d})\Phi_1^3\Psi^{(0,3)}\mathbf{1}
\end{aligned}
$$

e

$$
\begin{aligned}
A_{32} = {} & \mathbf{1}^{\top}\Psi^{(3,0)}M_1^3(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})M_1^3\Psi^{(3,0)}\mathbf{1} \\
& + 3\mathbf{1}^{\top}\Psi^{(2,1)}M_1^2\Phi_1(Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}})M_1^2\Phi_1\Psi^{(2,1)}\mathbf{1} \\
& + 3\mathbf{1}^{\top}\Psi^{(1,2)}M_1\Phi_1^2(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\beta}} - Z_{2\boldsymbol{\beta}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})M_1\Phi_1^2\Psi^{(1,2)}\mathbf{1} \\
& + \mathbf{1}^{\top}\Psi^{(0,3)}\Phi_1^3(Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}}) \odot (Z_{\boldsymbol{\gamma}} - Z_{2\boldsymbol{\gamma}})\Phi_1^3\Psi^{(0,3)}\mathbf{1}.
\end{aligned}
$$

CAPÍTULO 5

Resultados de simulação

Neste capítulo, nosso objetivo é comparar, através de simulações de Monte Carlo, os desempenhos de vários testes baseados nas seguintes estatísticas: razão de verossimilhanças (LR), razão de verossimilhanças corrigidas (LR^* e $\widetilde{LR}$), escore (SR) e escores corrigidas (SR^*, $\widetilde{LR}$ e $K(SR)$), respectivamente. Enfocaremos três situações em que as hipóteses nulas de interesse são as seguintes:

(i) $\mathrm{H}_1 : \beta_1 = \beta_1^{(0)}, \quad \gamma_1 = \gamma_1^{(0)}$ contra H'_1 : pelo menos uma das igualdades é violada;

(ii) $\mathrm{H}_2 : \beta_1 = \beta_1^{(0)}$ contra $\mathrm{H}'_2 : \beta_1 \neq \beta_1^{(0)}$;

(iii) $\mathrm{H}_3 : \gamma_1 = \gamma_1^{(0)}$ contra $\mathrm{H}'_3 : \gamma_1 \neq \gamma_1^{(0)}$.

As estatísticas LR e SR são dadas no Capítulo 2 em (2.12) e (2.15) para o teste de H_1, em (2.13) e (2.16) para o teste de H_2 e em (2.14) e (2.17) para o teste de H_3, respectivamente. Já as estatísticas LR^* e $\widetilde{LR}$ são dadas no Capítulo 3 em (3.14) e (3.27) para o teste de H_1, em (3.19) e (3.29) para o teste de H_2 e em (3.25) e (3.30) para o teste de H_3. Finalmente, as estatísticas SR^*, $\widetilde{SR}$ e $K(SR)$ são dadas no Capítulo 4 em (4.1), (4.3) e (4.2).

A comparação do desempenho dos testes é feita em função da proximidade das probabilidades de rejeição da hipótese nula, sendo esta verdadeira (erro do tipo I), aos seus respectivos níveis nominais dos testes.

As simulações realizadas são baseadas no modelo Passeio Aleatório. Para a obtenção da variável resposta com distribuição passeio aleatório, foi necessário primeiro gerar uma distribuição inversa gaussiana, conforme sugerem Cordeiro e Botter (2001), ou seja, $X \sim IG(\theta; \delta)$, sendo que os parâmetros θ e δ são considerados como média e dispersão. Fazendo $Y = 1/X$, então a variável aleatória Y tem distribuição passeio aleatório, isto é, $Y \sim PA(\mu; \phi)$ com parâmetros média e dispersão dados por

$$\mu = \frac{1}{\theta} + \frac{1}{\delta} \quad \text{ e } \phi = \frac{-\delta}{2\theta^2}.$$

Os componentes sistemáticos do passeio aleatório são dadas por:

$$\mu_\ell = \exp(\eta_\ell) = \exp(\beta_0 + \beta_1 x_{\ell 1} + \exp(\beta_2 x_{\ell 2})$$

e

$$\phi_\ell = \exp(\tau_\ell) = \exp(\gamma_0 + \gamma_1 s_{\ell 1} + \exp(\gamma_2 s_{\ell 2}),$$

em que $\ell = 1, \ldots, n$. A variável resposta foi gerada assumindo que $\beta = (1, 1, 1)^\top$ e $\gamma = (1, 1, 2)^\top$, $p = 3$ e $q = 3$. As covariadas foram geradas como amostras aleatórias da distribuição uniforme padrão. O número de réplicas de Monte Carlo foi fixado em 10000 e foram considerados os seguintes níveis nominais: $\alpha = 10\%$; $\alpha = 5\%$ e $\alpha = 1\%$. As simulações foram realizadas usando a linguagem de programação matricial `Ox` (Doornik, 2001).

Para cada tamanho da amostra e cada nível considerado, calculamos as taxas de rejeição dos testes citados acima, isto é, estimamos via simulação $\mathrm{P}(LR \geq \chi^2_{(\alpha;m)})$, $\mathrm{P}(LR^* \geq \chi^2_{(\alpha;m)})$, $\mathrm{P}(\widetilde{LR} \geq \chi^2_{(\alpha;m)})$, $\mathrm{P}(SR \geq \chi^2_{(\alpha;m)})$, $\mathrm{P}(SR^* \geq \chi^2_{(\alpha;m)})$, $\mathrm{P}(\widetilde{SR} \geq \chi^2_{(\alpha;m)})$ e $\mathrm{P}(K(SR) \geq \chi^2_{(\alpha;m)})$, em que $\chi^2_{(\alpha;m)}$ é o percentil $(1-\alpha)$ da distribuição χ^2 e m é o número de parâmetros a serem testados. Ao testarmos as hipóteses H_1, H_2 e H_3 temos que m assume os valores $p_1 + q_1$, p_1 e q_1, respectivamente.

Nas Tabelas 5.1 a 5.3, fixamos os valores de p_1 e q_1 em 2 e variamos o tamanho da amostra em $n = 20, 30, 40, 50, 60$ e 80. Vale salientar aqui que o teste da razão de verossimilhanças e suas versões corrigidas foram apresentados para testar todas as hipóteses definidas neste capítulo, enquanto que, o teste escore e suas versões corrigidas somente foram apresentados para o teste de H_3, considerando os tamanhos de amostra $n = 20, 30, 40$ e 50.

Analisando essas tabelas, notamos que o teste razão de verossimilhança é notavelmente liberal, até mesmo para $n = 80$ $(13,80\%)$, apresentando taxas de rejeição bem maiores do que os níveis nominais correspondentes. Por outro lado, para os testes baseados nas estatísticas de razão de verossimilhanças corrigidas, o fator de correção de Bartlett aplicado à estatística da razão de verossimilhanças faz com que a taxa de rejeição empírica fique mais próxima do nível nominal do teste, mas não corrige sua tendência liberal completamente.

Observa-se que para $n = 20$ e $\alpha = 10\%$ o teste baseado na estatística LR apresenta taxa de rejeição igual a $31,93\%$, ou seja mais do que o triplo do nível nominal selecionado, enquanto que as taxas de rejeição correspondentes fornecidas pelos testes baseados nas estatísticas da razão de verossimilhanças corrigidas (LR^* e $\widetilde{LR}$) são iguais a $19,49\%$ e $15,62\%$, respectivamente, para a hipótese H_1. Para a hipótese H_2, o teste LR apresentou taxa de rejeição igual a $32,62\%$, e as taxas de rejeição dos testes LR^* e $\widetilde{LR}$ foram iguais a $19,35\%$ e $11,55\%$. E para a hipótese H_3, o teste LR apresentou taxa de rejeição igual a $23,71\%$, e os testes LR^* e $\widetilde{LR}$ apresentaram taxas de rejeição iguais a $15,64\%$ e $13,14\%$.

Também pelas Tabelas 5.1 a 5.3 é evidente que entre os testes baseados nas estatísticas LR^* e $\widetilde{LR}$ têm melhores desempenhos do que aquele baseado em LR.

No entanto, conforme cresce o tamanho da amostra, as taxas de rejeição do teste usual vão se aproximando dos respectivos níveis nominais. Já para os testes corrigidos, quando o tamanho da amostra aumenta, as taxas de rejeição permanecem mais estáveis em relação aos respectivos níveis nominais se comparadas às taxas de rejeição do teste baseado em LR.

Na Tabela 5.3 podemos observar que os testes baseados nas estatísticas da razão de verossimilhanças e escore corrigidas (LR^*, $\widetilde{LR}$, SR^*, $\widetilde{SR}$ e $K(SR)$) apresentam melhores desempenhos do que os testes baseados nas estatísticas da razão de verossimilhanças e escore usuais (LR e SR). Notamos ainda que estes testes são notavelmente liberais, apresentando taxas de rejeição superiores aos níveis nominais correspondentes, principalmente quando o tamanho da amostra é pequeno. Para $n = 20$ e $\alpha = 10\%$, por exemplo, os testes baseados em LR e SR apresentam taxas de rejeição iguais a $23,71\%$ e $17,38\%$, respectivamente, enquanto que as taxas correspondentes fornecidas pelos testes baseados em LR^*, $\widetilde{LR}$,SR^*, $\widetilde{SR}$ e $K(SR)$ são, espectivamente, $15,64$;

13, 14; 12, 82; 11, 20 e 11, 50. Conforme o tamanho da amostra aumenta, as taxas de rejeição de todos os testes vão se aproximando dos níveis nominais considerados. Mesmo para $n = 50$, as taxas dos testes LR e SR estão acima dos níveis nominais. Entre os testes não corrigidos, o teste escore apesar de liberal, tem um melhor desempenho do que o teste da razão de verossimilhanças. Observamos que entre os testes baseados nas estatísticas da razão de verossimilhanças corrigidas, o teste que apresentou melhor desempenho foi o teste baseado na estatística $\widetilde{LR}$. Por outro lado, entre os testes baseados nas estatísticas escore corrigidas, os teste que apresentou melhor desempenho foi o teste baseado em $\widetilde{SR}$, seguido pelos testes baseados em $K(SR)$ e SR^*, respectivamente.

Tabela 5.1: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_1.

n	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	31, 93	19, 49	15, 62
20	5%	22, 67	11, 37	8, 61
	1%	8, 91	3, 45	2, 40
	10%	24, 01	14, 79	12, 43
30	5%	15, 16	8, 14	6, 46
	1%	5, 22	2, 21	1, 38
	10%	19, 79	13, 27	11, 94
40	5%	12, 18	7, 23	6, 38
	1%	3, 76	1, 68	1, 27
	10%	17, 29	11, 68	10, 78
50	5%	9, 87	6, 21	5, 59
	1%	2, 70	1, 20	1, 04
	10%	15, 12	10, 75	10, 23
60	5%	8, 52	5, 66	5, 30
	1%	2, 42	1, 23	1, 10
	10%	13, 80	10, 46	10, 17
80	5%	7, 61	5, 37	5, 17
	1%	1, 81	1, 11	1, 02

Tabela 5.2: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_2.

n	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	$32,62$	$19,35$	$11,55$
20	5%	$23,04$	$11,02$	$5,33$
	1%	$9,78$	$2,51$	$0,92$
	10%	$25,39$	$15,68$	$11,55$
30	5%	$16,87$	$8,67$	$5,76$
	1%	$6,32$	$2,05$	$0,96$
	10%	$21,49$	$13,58$	$11,24$
40	5%	$13,26$	$7,30$	$5,65$
	1%	$4,32$	$1,68$	$1,17$
	10%	$18,53$	$12,59$	$11,05$
50	5%	$11,25$	$6,47$	$5,49$
	1%	$3,40$	$1,42$	$1,00$
	10%	$16,88$	$11,68$	$10,79$
60	5%	$9,87$	$6,08$	$5,35$
	1%	$2,86$	$1,48$	$1,17$
	10%	$14,38$	$10,51$	$10,00$
80	5%	$7,96$	$5,69$	$5,36$
	1%	$2,27$	$1,21$	$1,05$

Tabela 5.3: Tamanho dos testes baseados nas estatísticas LR, LR^*, $\widetilde{LR}$, SR, SR^*, $\widetilde{SR}$ e $K(SR)$ para o teste H_3.

n	α (%)	LR	LR^*	$\widetilde{LR}$	SR	SR^*	$\widetilde{SR}$	$K(SR)$
	10%	$23,71$	$15,64$	$13,14$	$17,38$	$12,82$	$11,20$	$11,50$
20	5%	$14,91$	$8,92$	$7,28$	$10,90$	$6,50$	$5,40$	$5,80$
	1%	$5,56$	$2,46$	$1,69$	$2,70$	$1,90$	$1,32$	$1,35$
	10%	$16,90$	$11,68$	$10,33$	$14,30$	$11,50$	$10,33$	$10,38$
30	5%	$10,01$	$6,10$	$5,16$	$8,20$	$5,70$	$5,10$	$5,30$
	1%	$2,85$	$1,37$	$1,03$	$2,30$	$1,40$	$1,10$	$1,20$
	10%	$15,20$	$11,03$	$10,27$	$12,70$	$10,38$	$10,10$	$10,20$
40	5%	$8,70$	$5,75$	$5,24$	$6,90$	$5,20$	$5,10$	$5,20$
	1%	$2,25$	$1,18$	$1,00$	$1,64$	$1,20$	$1,00$	$1,10$
	10%	$13,62$	$10,09$	$9,67$	$11,41$	$10,10$	$10,01$	$10,05$
50	5%	$7,36$	$5,12$	$4,75$	$5,35$	$5,10$	$5,00$	$5,05$
	1%	$1,79$	$1,00$	$0,89$	$1,16$	$1,10$	$1,00$	$1,05$

A Tabela 5.4 apresenta os resultados de simulação obtidos levando em consideração a hipótese alternativa H_3 para $n = 30$ ao nível $\alpha = 10\%$ e diferentes valores de $\boldsymbol{\gamma}_0 = \boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_a$, com $\boldsymbol{\gamma}_a$ variando de $0,5$ a $1,5$ foram considerados. É importante observar estas simulações de poder correspondem à situação abordada na Tabela 5.3 para $n = 30$.

Pode-se observar que à medida que os valores de $\boldsymbol{\gamma}_0$ e $\boldsymbol{\gamma}_1$ vão se afastando da hipótese nula o poder dos testes aumentam, e o poder dos testes baseados em $\widetilde{SR}$, SR^*, $K(SR)$ e $\widetilde{LR}$ são maiores do que os poderes dos testes baseados nas estatísticas LR e SR.

Os resultados na Tabela 5.4, indicam que não há nenhuma perda de poder derivada do fato de usar os fatores de correção de Bartlett e tipo-Bartlett derivados nos Capítulos 3 e 4. Os poderes dos testes baseados nas estatísticas escore e razão de verossimilhanças corrigidas são semelhantes, com leve vantagem dos testes baseados nas estatísticas corrigidas $\widetilde{SR}$ e $\widetilde{LR}$, vantagem esta que se origina de sua menor distorção de tamanho.

Tabela 5.4: Poder dos testes estimados baseados nas estatísticas LR, LR^*, $\widetilde{LR}$, SR, $\widetilde{SR}$ e $K(SR)$ para o teste H_3.

γ_a	LR	LR^*	$\widetilde{LR}$	SR	SR^*	$\widetilde{SR}$	$K(SR)$
$0,5$	$29,36$	$32,59$	$34,20$	$27,73$	$33,12$	$34,20$	$34,80$
$0,7$	$52,84$	$56,83$	$58,56$	$54,90$	$56,82$	$58,70$	$58,99$
$0,9$	$74,42$	$77,41$	$78,98$	$76,57$	$77,65$	$78,99$	$78,57$
$1,1$	$90,21$	$91,97$	$92,55$	$91,35$	$92,36$	$93,99$	$93,71$
$1,3$	$96,71$	$97,40$	$97,60$	$96,48$	$97,50$	$97,70$	$97,62$
$1,4$	$98,61$	$98,93$	$99,10$	$98,79$	$98,97$	$99,30$	$99,23$
$1,5$	$99,28$	$99,44$	$99,57$	$99,79$	$99,63$	$99,90$	$99,80$

Outro estudo de simulação foi realizado com o objetivo de analisar a influência do número de parâmetros de perturbação no desempenho dos testes apresentados. As hipóteses nulas consideradas aqui foram as seguintes: $H_2 : \beta_1 = \beta_1^{(0)}$ contra $H_2' : \beta_1 \neq \beta_1^{(0)}$, $H_3 : \gamma_1 = \gamma_1^{(0)}$ contra $H_3' : \gamma_1 \neq \gamma_1^{(0)}$, fixamos o tamanho da amostra $n = 30$ e consideramos os seguintes cenários:

1. $\eta_\ell = \beta_0 + \sum_{j=2}^{5} \beta_j x_{j\ell} + \exp(\beta_p x_{p\ell})$ e $\tau = \gamma_0 + \gamma_1 s_{1\ell} + \exp(\gamma_2 s_{2\ell})$;

2. $\eta_\ell = \beta_0 + \beta_1 x_{1\ell} + \exp(\beta_2 x_{2\ell})$ e $\tau = \gamma_0 + \sum_{j=2}^{5} \gamma_j s_{j\ell} + \exp(\gamma_q s_{q\ell})$,

com $\ell = 1, \ldots, n$. Vale ressaltar aqui que no primeiro cenário, o número de parâmetros do vetor γ é fixo em 3 e no segundo cenário, o número de parâmetros do vetor β é fixo em 3.

Nas Tabelas 5.5 e 5.6, são referentes aos dois cenários quando a hipótese nula é $\mathrm{H}_2\colon \boldsymbol{\beta}_0 = \boldsymbol{\beta}_1 = 0$. E as Tabelas 5.7 e 5.8, são referentes aos dois cenários quando a hipótese nula é $\mathrm{H}_3\colon \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}_1 = 0$.

Podemos observar na Tabela 5.5 à medida que o número de parâmetros de perturbação referentes ao aumentando os parâmetros de perturbação do vetor β a taxa de rejeição do teste baseado na estatística LR vai aumentando, tornando-se consideravelmente liberal. Observa-se que para $\alpha = 5\%$ e $j = 5$, a taxa de rejeição deste teste é $21,88\%$, isto é, 4 vezes maior do que o nível nominal selecionado. A tendência do teste em rejeitar com frequência demasiada a hipótese nula é atenuada pela correção de Bartlett, de forma que, o teste baseado na estatística LR^* apresentou distorções de tamanhos menores, porém o teste ainda é liberal. Novamente, para $\alpha = 5\%$ e $j = 5$, taxa de rejeição do teste baseado em LR^* é $9,95\%$. O teste baseado na estatística $\widetilde{LR}$ foi eficaz em reduzir as taxas de rejeição do teste baseado em LR exibindo taxas de rejeição menores que o nível nominal considerado. Temos que a sua taxa de rejeição para $\alpha = 5\%$ e $j = 5$ é $3,85\%$. O impacto do número de parâmetros de perturbação no cenário 1 é bem menos marcante no caso do teste baseado em $\widetilde{LR}$, seguido pelo teste baseado em LR^*. Enquanto que as taxas do teste baseado na estatística LR aumentam, ficando bem maior do que os níveis nominais correspondentes, à medida que o número de parâmetros aumenta.

Na Tabela 5.6, à medida que o número de parâmetros de perturbação referentes ao vetor γ aumenta, as taxas de rejeição do teste da razão de verossimilhanças também vão aumentando. Já as taxas de rejeição dos testes baseados nas estatísticas corrigidas se mantêm mais estáveis, sendo que o teste baseado em LR^* é mais liberal com taxas de rejeição mais próximas dos níveis nominais correspondentes e o teste baseado em $\widetilde{LR}$ é mais conservador com taxas de rejeição mais baixas que os níveis nominais correspondentes.

Na Tabela 5.7 observamos que ao aumentarmos o número de parâmetros de perturbação β, as taxas de rejeição do teste da razão de verossimilhanças aumenta. As taxas dos testes baseado em LR^* e $\widetilde{LR}$ aumentam um pouco e permanecem mais próximas dos níveis nominais. O teste baseado em $\widetilde{LR}$ apresenta taxas de rejeição mais próximas dos níveis nominais correspondentes do que as do teste baseado em LR^*. O mesmo ocorre para a Tabela 5.8, só que as taxas dos testes baseados em LR^* e $\widetilde{LR}$ são mais liberais do que na situação anterior.

Nas Tabelas 5.7 e 5.8, foi feito o estudo da influência do aumento do número de parâmetros de perturbação para a hipótese nula $\mathrm{H}_3\colon \boldsymbol{\gamma}_0 = \boldsymbol{\gamma}_1 = 0$ com os cenários descritos anteriormente. Podemos observar que, para os dois casos, as taxas de rejeição dos testes vão aumentando com o acréscimo de parâmetros de perturbação, sendo que as taxas dos testes baseados em LR^* e $\widetilde{LR}$ apresentam um aumento menor do que as da razão de verossimilhança, e o teste baseado em $\widetilde{LR}$ apresentou taxas menores e mais próximas dos níveis nominais correspondentes.

Tabela 5.5: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_2, aumentando os parâmetros de perturbação do vetor β.

j	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	28,60	17,05	11,35
2	5%	19,61	9,55	4,97
	1%	7,86	2,33	0,77
	10%	30,17	16,85	9,35
3	5%	20,48	9,29	3,71
	1%	8,16	2,15	0,48
	10%	30,32	17,71	8,77
4	5%	21,37	9,43	3,75
	1%	8,87	2,23	0,53
	10%	30,99	17,49	8,87
5	5%	21,88	9,95	3,85
	1%	9,47	2,55	0,57

Tabela 5.6: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_2, aumentando os parâmetros de perturbação referentes ao vetor γ.

j	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	25,54	15,75	11,58
2	5%	16,98	8,70	5,80
	1%	6,34	2,03	0,97
	10%	27,61	16,06	10,59
3	5%	18,22	8,50	5,06
	1%	6,48	1,79	0,69
	10%	29,33	16,48	10,17
4	5%	19,72	8,69	4,45
	1%	6,70	1,72	0,64
	10%	28,56	15,14	9,05
5	5%	18,61	7,84	4,00
	1%	6,54	1,59	0,48

Tabela 5.7: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_3, aumentando os parâmetros de perturbação referentes ao vetor β.

j	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	20,10	14,54	13,45
2	5%	12,72	8,50	7,43
	1%	4,33	2,18	1,85
	10%	23,49	16,39	14,33
3	5%	15,11	9,55	8,03
	1%	5,68	2,61	2,12
	10%	26,66	18,46	16,00
4	5%	17,95	11,27	9,35
	1%	7,26	3,38	2,75
	10%	28,44	19,06	15,77
5	5%	19,36	11,43	9,13
	1%	7,66	3,46	2,92

Tabela 5.8: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_3, aumentando os parâmetros de perturbação referentes ao vetor γ.

j	α (%)	LR	LR^*	$\widetilde{LR}$
	10%	18,65	13,78	12,89
2	5%	11,13	7,65	7,18
	1%	3,45	1,83	1,67
	10%	19,70	14,77	13,89
3	5%	11,57	8,24	7,77
	1%	3,56	2,08	1,87
	10%	19,37	15,12	14,53
4	5%	11,57	8,18	7,60
	1%	3,34	2,05	1,80
	10%	20,55	16,47	15,91
5	5%	12,70	9,77	9,37
	1%	4,23	2,70	2,55

Também analisou-se os desempenhos dos testes quando o número de parâmetros de interesse aumenta, ou seja, fixamos o tamanho da amostra $n = 30$ e consideramos os mesmos preditores não lineares acima, mas as hipóteses foram:

$$\mathrm{H}_4 : \beta_0 = \beta_1 = \cdots = \beta_j = 0, \quad j = 2, \ldots, 7$$

$$\mathrm{H}_5 : \gamma_0 = \gamma_1 = \cdots = \gamma_j = 0, \quad j = 2, \ldots, 7.$$

Nas Tabelas 5.9 e 5.10 foi feito o estudo da influência do aumento do número de parâmetros de interesse (j) para H_4 e H_5. Podemos observar que, para os dois casos, as taxas de rejeição dos testes vão aumentando com o acréscimo de parâmetros de interesse, principalmente as taxas do teste baseado em LR, que por exemplo, para $j = 5$ apresenta taxa $(50,04\%)$ 5 vezes maior que o nível nominal 10%. No entanto, as taxas dos testes baseados em LR^* e $\widetilde{LR}$ apresentam um aumento bem menor, ainda que levemente liberal, do que as taxas do teste da razão de verossimilhanças, e o teste baseado em $\widetilde{LR}$ apresentou taxas menores e mais próximas dos níveis nominais correspondentes.

Tabela 5.9: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_4.

j	α (%)	LR	LR^*	$\widetilde{LR}$
2	10%	$31,74$	$18,14$	$11,45$
	5%	$22,49$	$10,18$	$5,63$
	1%	$9,29$	$2,31$	$0,83$
3	10%	$38,70$	$20,31$	$11,61$
	5%	$27,82$	$11,32$	$5,23$
	1%	$11,92$	$2,88$	$0,84$
4	10%	$45,08$	$23,69$	$12,87$
	5%	$33,75$	$14,41$	$6,37$
	1%	$16,75$	$4,10$	$1,10$
5	10%	$50,04$	$26,00$	$13,47$
	5%	$38,53$	$16,10$	$6,50$
	1%	$20,57$	$4,49$	$0,99$

Tabela 5.10: Tamanho dos testes baseados nas estatísticas LR, LR^* e $\widetilde{LR}$ para o teste H_5.

j	α (%)	LR	LR^*	$\widetilde{LR}$
2	10%	$19,78$	$13,89$	$12,81$
	5%	$12,09$	$7,61$	$6,72$
	1%	$3,77$	$1,96$	$1,69$
3	10%	$22,24$	$15,07$	$13,77$
	5%	$13,77$	$8,49$	$7,63$
	1%	$4,53$	$2,32$	$1,99$
4	10%	$24,65$	$16,43$	$14,98$
	5%	$15,66$	$9,27$	$8,43$
	1%	$5,26$	$2,58$	$2,05$
5	10%	$27,23$	$18,00$	$16,34$
	5%	$17,57$	$10,11$	$9,12$
	1%	$5,87$	$2,62$	$2,36$

CAPÍTULO 6

Técnicas de Diagnósticos

Uma etapa importante na análise de um ajuste de regressão é a verificação de possíveis afastamentos das suposições feitas para o modelo, especialmente para o componente aleatório e para a parte sistemática do modelo, bem como a existência de observações discrepantes com alguma interferência desproporcional ou inferencial nos resultados do ajuste. Tal etapa, conhecida como análise de diagnóstico, tem longa data, e começou com a análise de resíduos para detectar a presença de pontos aberrantes e avaliar a adequação da distribuição proposta para a variável resposta.

Uma observação é dita influente se ela alterar o valor das estimativas quando a retiramos do conjunto de dados ou quando ela é submetida a perturbações. E é dita ponto de alavancagem quando uma observação tem um peso desproporcional no próprio valor ajustado.

Em modelos normais lineares, uma das técnicas de diagnósticos mais utilizadas é a análise da matriz de projeção $\boldsymbol{H}$, ela é também conhecida como matriz chapéu ou hat. A diagonal dessa matriz representa a influência do valor observado sobre o respectivo valor predito, e este valor corresponde ao peso desproporcional que tal observação exerce sobre seu valor estimado. Neste capítulo iremos apresentar uma forma bastante geral para se calcular alavancagem, que pode ser utilizada em modelos não lineares generalizados com superdispersão.

Contudo, uma das propostas mais inovadoras na área de diagnóstico em regressão foi apresentada por Cook (1986) que propõe avaliar a influência conjunta das observações sob pequenas mudanças (perturbações) no modelo ou nos dados, ao invés da avaliação pela retirada individual ou conjunta de pontos. Essa metodologia, denominada influência local, teve uma grande receptividade entre os usuários e pesquisadores de regressão, havendo inúmeras publicações no assunto em que a metodologia é aplicada em classes particulares de modelos ou estendida para situações mais gerais.

6.1 Alavancagem Generalizada

Wei et al. (1998) generalizaram a definição de pontos de alavanca para modelos bastante gerais cuja variável resposta seja contínua. A ideia de pontos de alavanca é estudar o efeito de cada observação, $\boldsymbol{y}$, no próprio valor ajustado, $\hat{\boldsymbol{y}}$. Segundo Wei et al (1998), a matriz $\partial\hat{\boldsymbol{y}}/\partial\boldsymbol{y}$ pode ser obtida da forma geral

$$GL(\boldsymbol{\theta}) = \frac{\partial\hat{\boldsymbol{y}}}{\partial\boldsymbol{y}} = \left\{ D_{\boldsymbol{\theta}} \left(-\ddot{L}_{\boldsymbol{\theta\theta}} \right)^{-1} \ddot{L}_{\boldsymbol{y\theta}} \right\}\bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}, \tag{6.1}$$

em que $D_{\boldsymbol{\theta}} = \frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}^\top}$, $\ddot{L}_{\boldsymbol{\theta\theta}} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}$ e $\ddot{L}_{\boldsymbol{y\theta}} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{y}^\top}$.

A matriz de informação observada será utilizada nos métodos de detecção de pontos de alavanca e influência local. A matriz de informação observada é denotada por $\ddot{L}_{\boldsymbol{\theta\theta}}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, e pode ser expressa da seguinte forma

$$\ddot{L}_{\boldsymbol{\theta\theta}} = \begin{bmatrix} \ddot{L}_{\boldsymbol{\beta\beta}} & \ddot{L}_{\boldsymbol{\beta\gamma}} \\ \ddot{L}_{\boldsymbol{\gamma\beta}} & \ddot{L}_{\boldsymbol{\gamma\gamma}} \end{bmatrix},$$

em que

$$\begin{aligned}
\ddot{L}_{\boldsymbol{\beta\beta}} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top} &= \widetilde{X}^\top\boldsymbol{\Psi}^{(3,0)}M_1^2(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{X} + \widetilde{X}^\top\boldsymbol{\Psi}^{(2,0)}M_2(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{X} - \widetilde{X}^\top\boldsymbol{\Psi}^{(2,0)}M_1^2\widetilde{X} \\
&\quad + \boldsymbol{\Psi}^{(2,0)}M_1(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{\widetilde{X}}, \\
\ddot{L}_{\boldsymbol{\beta\gamma}} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\gamma}^\top} &= \widetilde{X}^\top\boldsymbol{\Psi}^{(2,1)}M_1\Phi_1(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{S}, \\
\ddot{L}_{\boldsymbol{\gamma\beta}} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\beta}^\top} &= \widetilde{S}^\top\boldsymbol{\Psi}^{(2,1)}M_1\Phi_1(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{X} \quad \text{ e} \\
\ddot{L}_{\boldsymbol{\gamma\gamma}} = \frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\gamma}\partial\boldsymbol{\gamma}^\top} &= \widetilde{S}^\top\boldsymbol{\Psi}^{(1,2)}\Phi_1^2(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{S} + \widetilde{S}^\top\boldsymbol{\Psi}^{(1,1)}\Phi_2(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{S} + \widetilde{S}^\top\Phi_2 T(\boldsymbol{y})\widetilde{S} + \widetilde{S}^\top\boldsymbol{\Psi}^{(0,2)}\Phi_1^2\widetilde{S} \\
&\quad + \widetilde{S}^\top\boldsymbol{\Psi}^{(0,1)}\Phi_2\widetilde{S} + \boldsymbol{\Psi}^{(1,1)}\Phi_1(\boldsymbol{y}-\boldsymbol{\mu})\widetilde{\widetilde{S}} + T(\boldsymbol{y})\Phi_1\widetilde{\widetilde{S}} + \boldsymbol{\Psi}^{(0,1)}\Phi_1\widetilde{\widetilde{S}}.
\end{aligned}$$

Também precisaremos das matrizes $D_{\boldsymbol{\theta}}$ e $\ddot{L}_{\boldsymbol{\theta y}}$. Como

$$D_{\boldsymbol{\theta}}^{\boldsymbol{\mu}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^{\top}} = (D_{\boldsymbol{\beta}}^{\top}, D_{\boldsymbol{\gamma}}^{\top})^{\top},$$

temos que

$$D_{\boldsymbol{\beta}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\beta}} = \frac{\partial g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\beta}} = \frac{\partial g^{-1}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\beta}} = M_1 \widetilde{X}$$

e

$$D_{\boldsymbol{\gamma}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\gamma}} = \mathbf{0}.$$

Então, $D_{\boldsymbol{\theta}}^{\boldsymbol{\mu}} = (M_1 \widetilde{X}, \mathbf{0})^{\top}$. Analogamente,

$$D_{\boldsymbol{\theta}}^{\boldsymbol{\phi}} = \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}^{\top}} = (D_{\boldsymbol{\beta}}^{\top}, D_{\boldsymbol{\gamma}}^{\top})^{\top},$$

temos que

$$D_{\boldsymbol{\beta}} = \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\beta}} = \mathbf{0},$$

e

$$D_{\boldsymbol{\gamma}} = \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\gamma}} = \frac{\partial h^{-1}(\boldsymbol{\tau})}{\partial \boldsymbol{\gamma}} = \frac{\partial h^{-1}(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}} \frac{\partial \boldsymbol{\tau}}{\partial \boldsymbol{\gamma}} = \Phi_1 \widetilde{S}.$$

Então, $D_{\boldsymbol{\theta}}^{\boldsymbol{\phi}} = (\mathbf{0}, \Phi_1 \widetilde{S})^{\top}$e

$$\ddot{L}_{\boldsymbol{\theta y}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{y}^{\top}},$$

sendo

$$\begin{aligned} \ddot{L}_{\boldsymbol{\beta y}} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{y}^{\top}} = \widetilde{X}^{\top} \Psi^{(2,0)} M_1 \quad \text{ e} \\ \ddot{L}_{\boldsymbol{\gamma y}} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{y}^{\top}} = \widetilde{S}^{\top} \Psi^{(1,1)} \Phi_1 + \widetilde{S}^{\top} \Phi_1 T'(\boldsymbol{y}). \end{aligned}$$

Substituindo essas quantidades em (6.1) temos que

$$\begin{aligned} GL^{\boldsymbol{\mu}}(\boldsymbol{\theta}) &= \begin{bmatrix} M_1 \widetilde{X} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{L}_{\boldsymbol{\beta\beta}} & \ddot{L}_{\boldsymbol{\beta\gamma}} \\ \ddot{L}_{\boldsymbol{\gamma\beta}} & \ddot{L}_{\boldsymbol{\gamma\gamma}} \end{bmatrix}^{-1} \begin{bmatrix} \widetilde{X}^{\top} \Psi^{(2,0)} M_1 \\ \widetilde{S}^{\top} \Psi^{(1,1)} \Phi_1 + \widetilde{S}^{\top} \Phi_1 T'(\boldsymbol{y}) \end{bmatrix} \\ &= \begin{bmatrix} M_1 \widetilde{X} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \ddot{L}_{\boldsymbol{\beta\beta}}^{-1} + F E^{-1} F^{\top} & -F E^{-1} \\ -E^{-1} F^{\top} & E^{-1} \end{bmatrix} \begin{bmatrix} \widetilde{X}^{\top} \Psi^{(2,0)} M_1 \\ \widetilde{S}^{\top} \Psi^{(1,1)} \Phi_1 + \widetilde{S}^{\top} \Phi_1 T'(\boldsymbol{y}) \end{bmatrix}, \end{aligned}$$

em que $F = \ddot{L}_{\boldsymbol{\beta\beta}}^{-1} \ddot{L}_{\boldsymbol{\beta\gamma}}$, $F^{\top} = \ddot{L}_{\boldsymbol{\gamma\beta}} \ddot{L}_{\boldsymbol{\beta\beta}}^{-1}$ e $E^{-1} = \ddot{L}_{\boldsymbol{\gamma\gamma}} - \ddot{L}_{\boldsymbol{\gamma\beta}} \ddot{L}_{\boldsymbol{\beta\beta}}^{-1} \ddot{L}_{\boldsymbol{\beta\gamma}}$. Logo,

$$GL^{\boldsymbol{\mu}}(\boldsymbol{\theta}) = \left[\begin{array}{cc} M_1\widetilde{X}\left(\ddot{L}^{-1}_{\boldsymbol{\beta\beta}} + FE^{-1}F^{\top}\right) & M_1\widetilde{X}\left(-FE^{-1}\right) \end{array} \right] \left[\begin{array}{l} \widetilde{X}^{\top}\Psi^{(2,0)}M_1 \\ \widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y}) \end{array} \right]$$
$$= M_1\widetilde{X}\left(\ddot{L}^{-1}_{\boldsymbol{\beta\beta}} + FE^{-1}F^{\top}\right)\widetilde{X}^{\top}\Psi^{(2,0)}M_1 - M_1\widetilde{X}\left(FE^{-1}\right)\left[\widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y})\right].$$

Podemos, rescrever $GL^{\boldsymbol{\mu}}(\boldsymbol{\theta})$, da seguinte forma,

$$GL^{\mu}(\boldsymbol{\theta}) = GL^{\mu}_{\beta}(\boldsymbol{\theta}) + GL^{\mu}_{\gamma}(\boldsymbol{\theta}),$$

em que

$$GL^{\boldsymbol{\mu}}_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = M_1\widetilde{X}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\widetilde{X}^{\top}\Psi^{(2,0)}M_1$$

e

$$\begin{aligned} GL^{\boldsymbol{\mu}}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) &= M_1\widetilde{X}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}\ddot{L}^{-1}_{\boldsymbol{\gamma\gamma}}\ddot{L}_{\boldsymbol{\gamma\beta}}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\widetilde{X}^{\top}\Psi^{(2,0)}M_1 \\ &- M_1\widetilde{X}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}C(R^{\top}R)^{-1}C^{\top}\ddot{L}_{\boldsymbol{\gamma\beta}}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\widetilde{X}^{\top}\Psi^{(2,0)}M_1 \\ &- M_1\widetilde{X}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}\ddot{L}^{-1}_{\boldsymbol{\gamma\gamma}}\left[\widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y})\right] \\ &+ M_1\widetilde{X}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}C(R^{\top}R)^{-1}C^{\top}\left[\widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y})\right], \end{aligned}$$

sendo $C = \ddot{L}_{\boldsymbol{\gamma\gamma}}\ddot{L}_{\boldsymbol{\gamma\beta}}$ e $(R^{\top}R)^{-1} = \left(\ddot{L}_{\boldsymbol{\beta\beta}} + \ddot{L}_{\boldsymbol{\beta\gamma}}\ddot{L}^{-1}_{\boldsymbol{\gamma\gamma}}\ddot{L}_{\boldsymbol{\gamma\beta}}\right)^{-1}$.

Analogamente, temos que

$$GL^{\boldsymbol{\phi}}(\boldsymbol{\theta}) = \left[\begin{array}{cc} \mathbf{0} & \Phi_1\widetilde{S} \end{array} \right] \left[\begin{array}{cc} \ddot{L}^{-1}_{\boldsymbol{\beta\beta}} + FE^{-1}F^{\top} & -FE^{-1} \\ -E^{-1}F^{\top} & E^{-1} \end{array} \right] \left[\begin{array}{l} \widetilde{X}^{\top}\Psi^{(2,0)}M_1 \\ \widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y}) \end{array} \right],$$

em que $F = \ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}$, $F^{\top} = \ddot{L}_{\boldsymbol{\gamma\beta}}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}$ e $E^{-1} = \ddot{L}_{\boldsymbol{\gamma\gamma}} - \ddot{L}_{\boldsymbol{\gamma\beta}}\ddot{L}^{-1}_{\boldsymbol{\beta\beta}}\ddot{L}_{\boldsymbol{\beta\gamma}}$. Logo,

$$GL^{\boldsymbol{\phi}}(\boldsymbol{\theta}) = \left[\begin{array}{cc} -\Phi_1\widetilde{S}EF^{\top} & \Phi_1\widetilde{S}E^{-1} \end{array} \right] \left[\begin{array}{l} \widetilde{X}^{\top}\Psi^{(2,0)}M_1 \\ \widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \widetilde{S}^{\top}\Phi_1 T'(\boldsymbol{y}) \end{array} \right]$$
$$= -\Phi_1\widetilde{S}EF^{\top}\widetilde{X}^{\top}\Psi^{(2,0)}M_1 + \Phi_1\widetilde{S}E^{-1}F^{\top}\widetilde{S}^{\top}\Psi^{(1,1)} + \Phi_1\widetilde{S}E^{-1}\Phi_1 T'(\widetilde{y}).$$

Podemos, reescrever $GL^{\boldsymbol{\phi}}(\boldsymbol{\theta})$, da seguinte forma,

$$GL^{\phi}(\boldsymbol{\theta}) = GL^{\phi}_{\beta}(\boldsymbol{\theta}) + GL^{\phi}_{\gamma}(\boldsymbol{\theta}),$$

em que

$$GL^{\boldsymbol{\phi}}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \Phi_1\widetilde{S}(\ddot{L}_{\boldsymbol{\gamma\gamma}})^{-1}\widetilde{S}^{\top}\Psi^{(1,1)}\Phi_1 + \Phi_1\widetilde{S}(\ddot{L}_{\boldsymbol{\gamma\gamma}})^{-1}\widetilde{S}^{\top}T'(\boldsymbol{y})\Phi_1,$$

e

$$GL^{\phi}_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = -\Phi_1\widetilde{S}\left[(\ddot{L}_{\boldsymbol{\gamma\gamma}})^{-1} + C(R^\top R)^{-1}C^\top\right]\ddot{L}_{\boldsymbol{\gamma\beta}}(\ddot{L}_{\boldsymbol{\beta\beta}})^{-1}\widetilde{X}^\top\Psi^{(2,0)}M_1$$
$$+ \Phi_1\widetilde{S}C(R^\top R)^{-1}C^\top\widetilde{S}^\top\Psi^{(1,1)}\Phi_1 + \Phi_1\widetilde{S}C(R^\top R)^{-1}C^\top\Phi_1\widetilde{S}^\top T'(\boldsymbol{y})\Phi_1.$$

Como em geral $\widehat{GL}(\boldsymbol{\theta})$ depende de $\boldsymbol{\theta}$ sugerimos para detectar pontos de alavanca o gráfico dos elementos da diagonal principal de $\widehat{GL}(\boldsymbol{\theta})$ contra os valores ajustados.

6.2 Influência

A deleção de pontos talvez seja a técnica mais conhecida para avaliar o impacto da retirada de uma observação particular nas estimativas dos parâmetros do modelo proposto. Por exemplo, a distância de Cook, ver Cook e Weiberg (1982) originalmente desenvolvida para modelos normais lineares, que mede o afastamento pela verossimilhança quando eliminamos a i-ésima observação denotado por

$$LD_i = 2\left\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{(i)})\right\}$$

sendo portanto uma medida que verifica a influência da retirada da i-ésima observação em $\hat{\boldsymbol{\theta}}$, que denotaremos por $\hat{\boldsymbol{\theta}}_{(i)}$. Não sendo possível obtermos uma forma analítica para LD_i, é usual utilizarmos a segunda aproximação por série de Taylor em torno de $\hat{\boldsymbol{\theta}}$. Essa expansão leva ao seguinte resultado:

$$LD_i \cong (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^\top\left\{-\ddot{L}_{\boldsymbol{\theta\theta}}(\hat{\boldsymbol{\theta}})\right\}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}).$$

Substituindo $\ddot{L}_{\boldsymbol{\theta\theta}}(\hat{\boldsymbol{\theta}})$ pelo correspondente valor esperado e $\boldsymbol{\theta}$ por $\boldsymbol{\theta}_{(i)}$, obtemos

$$LD_i^{\beta} = (\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)})^\top(\widetilde{X}^\top\Psi^{(2,0)}M_1^2\widetilde{X})(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_{(i)}) \tag{6.2}$$

e

$$LD_i^{\gamma} = (\hat{\boldsymbol{\gamma}} - \hat{\boldsymbol{\gamma}}_{(i)})^\top(\widetilde{S}^\top\Psi^{(0,2)}\Phi_1^2\widetilde{S})(\hat{\boldsymbol{\gamma}} - \hat{\boldsymbol{\gamma}}_{(i)}). \tag{6.3}$$

Como em geral não é possível obtermos uma forma fechada para $\hat{\boldsymbol{\theta}}_{(i)}$, a aproximação de um passo tem sido utilizada (ver Cook e Weisberg, 1982), que consiste em tomarmos a primeira iteração do processo iterativo pelo método escore de Fisher quando o mesmo é iniciado em $\hat{\boldsymbol{\theta}}$.

Essa aproximação, introduzida por Pregibon (1981), é dada por

$$\hat{\boldsymbol{\theta}}^1_{(i)} = \hat{\boldsymbol{\theta}} + \{-\ddot{L}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}})\}^{-1}\ell_{(i)}(\hat{\boldsymbol{\theta}}),$$

em que $\ell_{(i)}(\hat{\boldsymbol{\theta}})$ é o logaritmo da função de verossimilhança sem a i-ésima observação. Substituindo novamente $\ddot{L}_{\boldsymbol{\theta}\boldsymbol{\theta}}(\hat{\boldsymbol{\theta}})$ por $K(\hat{\boldsymbol{\theta}})$ para cada elemento de $\boldsymbol{\theta}$ obtemos

$$\hat{\boldsymbol{\beta}}_{(i)} = \hat{\boldsymbol{\beta}} - \frac{(y_i - \hat{\mu}_i)}{\hat{V}_i^{1/2}(1-\hat{h}_{ii})}(\widetilde{X}^\top \Psi^{(2,0)} M_1^2 \widetilde{X})^{-1}\widetilde{X}_i(\Psi_i^{(2,0)} M_{1i}^2)^{1/2}, \qquad (6.4)$$

em que $\hat{h}_{ii}$ são elementos da diagonal principal da matriz

$$\widehat{H} = (\Psi^{(2,0)} M_1^2)^{1/2}\widetilde{X}(\widetilde{X}^\top \Psi^{(2,0)} M_1^2 \widetilde{X})^{-1}\widetilde{X}^\top(\Psi^{(2,0)} M_1^2)^{1/2},$$

e

$$\hat{\boldsymbol{\gamma}}_{(i)} = \hat{\boldsymbol{\gamma}} - \frac{(y_i - \hat{\mu}_i)}{\hat{V}_i^{1/2}(1-\hat{p}_{ii})}(\widetilde{S}^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S})^{-1}\widetilde{S}_i(\Psi_i^{(0,2)} \Phi_{1i}^2)^{1/2}, \qquad (6.5)$$

sendo $\hat{p}_{ii}$ são elementos da diagonal principal da matriz

$$\widehat{P} = (\Psi^{(0,2)} \Phi_1^2)^{1/2}\widetilde{S}(\widetilde{S}^\top \Psi^{(0,2)} \Phi_1^2 \widetilde{S})^{-1}\widetilde{S}^\top(\Psi^{(0,2)} \Phi_1^2)^{1/2}.$$

Assim, substituindo as expressões (6.4) e (6.5) acima em (6.2) e (6.3), respectivamente, temos que

$$LD_i^\beta = \frac{\hat{h}_{ii}}{1-\hat{h}_{ii}} t_{Si}^2,$$

em que

$$t_{Si} = \frac{(y_i - \hat{\mu}_i)}{(\Psi_i^{(2,0)})^{-1/2}\sqrt{1-\hat{h}_{ii}}}$$

e

$$LD_i^\gamma = \frac{\hat{p}_{ii}}{1-\hat{p}_{ii}} t_{Ti}^2,$$

sendo

$$t_{Ti} = \frac{(y_i - \hat{\mu}_i)}{(\Psi_i^{(2,0)})^{-1/2}\sqrt{1-\hat{p}_{ii}}}.$$

Gráficos de índices de LD_i e de $\hat{p}_{ii}$ $(\hat{h}_{ii})$ contra os valores ajustados são recomendados.

6.3 Influência Local

Um método muito conhecido na detecção de pontos influentes no modelo é o método de influência local proposto por Cook (1986), que avalia a sensibilidade dos estimadores de máxima verossimilhança com respeito a perturbações no modelo ou nos dados.

Perturbação é qualquer modificação nas suposições do modelo ou nos dados que causam diferenças substanciais nos resultados da análise. Os esquemas mais conhecidos são: perturbação de casos ponderados, perturbação aditiva na resposta, perturbação aditiva no preditor, entre outros. Neste trabalho vamos nos ater aos três primeiros casos.

Para um conjunto de dados observados seja $\ell(\boldsymbol{\theta})$ o logaritmo da função de verossimilhança do modelo onde $\boldsymbol{\theta}$ é um vetor $(p+q) \times 1$ de parâmetros desconhecidos. Podemos introduzir um vetor de perturbação $\boldsymbol{w}$ de dimensão $n \times 1$ cujo logaritmo da função de verossimilhança é $\ell(\boldsymbol{\theta}_{\boldsymbol{w}})$, e comparar os estimadores de máxima verossimilhança para se determinar a influência local. O método de Cook calcula LD, ou seja, desvio local, que nada mais é do que a diferença das verossimilhanças avaliadas em seus estimadores ($\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}_{\boldsymbol{w}}}$). Grandes valores de LD indicam que as estimações são altamente sensíveis à perturbação,

$$LD(\boldsymbol{w}) = 2\left\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{w}})\right\}. \tag{6.6}$$

Um gráfico de $LD(\boldsymbol{w})$ versus $\boldsymbol{w}$ contém informação sobre a influência do tipo de perturbação escolhido. Este gráfico forma uma superfície chamada de gráfico de influência (denotaremos por $\boldsymbol{\alpha}(\boldsymbol{w})$). Um estudo de influência local consiste em analisar como a superfície se desvia de seu plano tangente em $\boldsymbol{w}_0$, onde $\boldsymbol{w}_0$ é um vetor de não perturbação, tal que $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}_{\boldsymbol{w}_0})$. Esta análise pode ser feita estudando-se as curvaturas das seções normais à superfície $\boldsymbol{\alpha}(\boldsymbol{w})$ em $\boldsymbol{w}_0$, que são interseções de $\boldsymbol{\alpha}(\boldsymbol{w})$ com o plano que contém o vetor que é normal ao seu plano tangente em $\boldsymbol{w}_0$.

Para construir uma seção normal, considere uma linha aumentada (lifted line) em $\boldsymbol{w}$ que passa através de $\boldsymbol{w}_0$. Esta linha pode ser representada por:

$$\boldsymbol{w}(\boldsymbol{a}) = \boldsymbol{w}_0 + \boldsymbol{a}\boldsymbol{t},$$

onde $\boldsymbol{a} \in \mathbb{R}$ e $\boldsymbol{t}$ é um vetor unitário, ou seja, tem comprimento 1.

A curvatura $\boldsymbol{C_t}$, que é a curvatura da seção normal de $\boldsymbol{\alpha}(\boldsymbol{w})$ em $\boldsymbol{w}_0$ na direção de $\boldsymbol{t}$, é usada para caracterizar o comportamento do gráfico de influência em uma vizinhança de $\boldsymbol{w}_0$. A curvatura é dada por:

$$\boldsymbol{C_t}(\boldsymbol{\theta}) = 2\left|\boldsymbol{t}^\top \boldsymbol{\Delta}^\top \ddot{L}^{-1}_{\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}} \boldsymbol{\Delta t}\right|, \tag{6.7}$$

em que $\ddot{L}^{-1}_{\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}$, $\boldsymbol{\Delta}$ é uma matriz $(p+q) \times n$ que depende do esquema de pertubação cujos elementos são dados por $\Delta_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\theta}_i \partial w_j}$, com $i = 1, \ldots, p+q$ e $j = 1, \ldots, n$ avaliado em $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{w}_0$.

6.3.1 Perturbação aditiva na resposta

Considere uma perturbação aditiva sobre a i-ésima resposta, isto é, $y_{iw} = y_i + w_i$, onde w_i pode ser uma estimativa do desvio padronizado de y_i. Então, o logaritmo da função de verossimilhança é:

$$\ell(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n} \left\{ (y_{iw} - \mu_i)\Psi_i^{(1,0)}(\mu_i, \phi_i) + \phi_i T(y_{iw}) + \Psi_i(\mu_i, \phi_i) \right\} + \sum_{i=1}^{n} \log A(y_{iw}).$$

Vamos assumir que a matriz $\widehat{\boldsymbol{\Delta}}^\top = (\widehat{\boldsymbol{\Delta}}_{\boldsymbol{\beta}}^\top, \widehat{\boldsymbol{\Delta}}_{\boldsymbol{\gamma}}^\top)$ em que o (ji)-ésimo elemento de $\widehat{\boldsymbol{\Delta}}_{\boldsymbol{\beta}}^\top$ é dado por $\left.\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \beta_j \partial w_i}\right|_{(\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{w}=\boldsymbol{w}_0)}$, analogamente temos que o (ji)-ésimo elemento de $\widehat{\boldsymbol{\Delta}}_{\boldsymbol{\gamma}}^\top$ é dado por $\left.\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\delta})}{\partial \gamma_j \partial \delta_i}\right|_{(\boldsymbol{\gamma}=\hat{\boldsymbol{\gamma}}, \boldsymbol{w}=\boldsymbol{w}_0)}$.

Assim, temos que

$$\begin{aligned} \boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\beta} \partial \boldsymbol{w}^\top} \\ &= \sum_{i=1}^{n} \psi_i^{(2,0)} m_{1i}(s)_i, \end{aligned}$$

em que $(s)_i = \partial \eta_i / \partial \beta$. Portanto, em forma matricial, temos

$$\boldsymbol{\Delta}_{\boldsymbol{\beta}} = \widetilde{X}^\top \boldsymbol{\Psi}^{(2,0)} M_1. \tag{6.8}$$

Analogamente, temos que

$$\begin{aligned} \boldsymbol{\Delta}_{\gamma} &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{w}^\top} \\ &= \sum_{i=1}^{n} \left\{ \psi_i^{(1,1)} \phi_{1i}(S)_i + T'(y_{iw}) \phi_{1i}(S)_i \right\}. \end{aligned}$$

Portanto, em forma matricial, temos

$$\boldsymbol{\Delta_\gamma} = \widetilde{S}^\top \boldsymbol{\Psi}^{(1,1)}\boldsymbol{\Phi}_1 + \widetilde{S}^\top \boldsymbol{\Phi}_1 \boldsymbol{T'(y_w)}. \tag{6.9}$$

6.3.2 Perturbação aditiva nos preditores

Considere uma perturbação aditiva sobre a i-ésima obesrvação do preditor k, isto é, $x_{ikw} = x_{ik} + w_i$. Então, o logaritmo da função de verossimilhança é:

$$\ell(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n}\Big\{(y_i - \mu_{iw})\Psi_i^{(1,0)}(\mu_{iw}, \phi_{wi}) + \phi_i T(y_i) + \Psi_i(\mu_{iw}, \phi_{wi})\Big\} + \sum_{i=1}^{n}\log A(y_i).$$

Assim, temos que

$$\begin{aligned}
\boldsymbol{\Delta_\beta} &= \left.\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\beta}\partial \boldsymbol{w}^\top}\right|_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}, \boldsymbol{w}=\boldsymbol{w_0}} \\
&= \sum_{i=1}^{n}\Big\{\psi_i^{(3,0)} m_{1i}(s)_{iw}\frac{\partial \mu_{iw}}{\partial \eta_{iw}}\frac{\partial \eta_{lw}}{\partial w_i}(y_i - \mu_{iw}) + \psi_i^{(2,0)}\frac{\partial^2 \mu_{iw}}{\partial \eta_{iw}^2}\frac{\eta_{iw}}{\partial w_i}\frac{\partial \eta_{iw}}{\partial \beta_i}(y_i - \mu_{iw}) \\
&\quad - \psi_i^{(2,0)} m_{1i}\frac{\partial \mu_{iw}}{\partial w_i}\frac{\partial \eta_{iw}}{\partial \beta_i} + \psi_i^{(2,0)}(y_i - \mu_{iw}) m_{1i}\frac{\partial^2 \eta_{iw}}{\partial \beta_i \partial w_i}\Big\}.
\end{aligned}$$

Em forma matricial, temos

$$\boldsymbol{\Delta_\beta} = \widetilde{X}_w^\top \boldsymbol{\Psi}^{(3,0)} M_1^2(\boldsymbol{y} - \boldsymbol{\mu}_w)F_1 + \widetilde{X}_w^\top \boldsymbol{\Psi}^{(2,0)} M_2(\boldsymbol{y} - \boldsymbol{\mu}_w)F_1 - \widetilde{X}_w^\top \boldsymbol{\Psi}^{(2,0)} M_1^2 + \boldsymbol{\Psi}^{(2,0)} M_1(\boldsymbol{y} - \boldsymbol{\mu}_w)F_2, \tag{6.10}$$

em que $F_1 = \frac{\partial \eta_{iw}}{\partial w_i} = \frac{\partial \eta_{iw}}{\partial X_w}\frac{\partial X_w}{\partial w_i} = \frac{\partial \eta_{iw}}{\partial X_w} X$ e $F_2 = \frac{\partial^2 \eta_{iw}}{\partial \beta_i \partial w_i}$. Analogamente, temos

$$\begin{aligned}
\boldsymbol{\Delta_\gamma} &= \left.\frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\gamma}\partial \boldsymbol{w}^\top}\right|_{\boldsymbol{\gamma}=\hat{\boldsymbol{\gamma}}, \boldsymbol{w}=\boldsymbol{w_0}} \\
&= \sum_{i=1}^{n}\Big\{\psi_i^{(1,2)}\phi_{1iw}(y_i - \mu_i)(S)_{iw}\frac{\partial \phi_{iw}}{\partial \tau_{iw}}\frac{\partial \tau_{iw}}{\partial w_i} + \psi_i^{(1,1)}(y_i - \mu_i)(S)_{iw}\frac{\partial^2 \phi_{iw}}{\partial \tau_{iw}^2}\frac{\partial \tau_{iw}}{\partial w_i} \\
&\quad + \psi_i^{(1,1)}(y_i - \mu_i)\phi_{1iw}\frac{\partial^2 \tau_{iw}}{\partial \gamma_i \partial w_i} + T(y_i)\frac{\partial^2 \phi_{iw}}{\partial \tau_{iw}^2}\frac{\partial \tau_{iw}}{\partial w_i}(S)_{iw} + T(y_i)\phi_{1iw}\frac{\partial^2 \tau_{iw}}{\partial \gamma_i \partial w_i} \\
&\quad + \psi_i^{(0,2)}\phi_{1iw}\frac{\partial \phi_{iw}}{\partial \tau_{iw}}\frac{\partial \tau_{iw}}{\partial w_i}(S)_{iw} + \psi_i^{(0,1)}(S)_{iw}\frac{\partial^2 \phi_{iw}}{\partial \tau_{iw}^2}\frac{\partial \tau_{iw}}{\partial w_i} + \psi_i^{(0,1)}\phi_{1iw}\frac{\partial^2 \tau_{iw}}{\partial \gamma_i \partial w_i}\Big\}.
\end{aligned}$$

Em forma matricial, temos

$$\begin{aligned}
\boldsymbol{\Delta_\gamma} &= \widetilde{S}_w^\top \boldsymbol{\Psi}^{(1,1)}\boldsymbol{\Phi}_1^2(\boldsymbol{y} - \boldsymbol{\mu})G_1 + \widetilde{S}_w^\top \boldsymbol{\Psi}^{(1,1)}\boldsymbol{\Phi}_2(\boldsymbol{y} - \boldsymbol{\mu})G_1 + \boldsymbol{\Psi}^{(1,1)}\boldsymbol{\Phi}_1(\boldsymbol{y} - \boldsymbol{\mu})G_2 + \widetilde{S}_w^\top \boldsymbol{\Phi}_2 T(\boldsymbol{y})G_1 \\
&\quad + T(\boldsymbol{y})\boldsymbol{\Phi}_1 G_2 + \widetilde{S}_w^\top \boldsymbol{\Psi}^{(0,2)}\boldsymbol{\Phi}_1^2 G_1 + \widetilde{S}_w^\top \boldsymbol{\Psi}^{(0,1)}\boldsymbol{\Phi}_2 G_1 + \boldsymbol{\Psi}^{(0,1)}\boldsymbol{\Phi}_1 G_2,
\end{aligned} \tag{6.11}$$

em que $G_1 = \frac{\partial \tau_{iw}}{\partial w_i} = \frac{\partial \tau_{iw}}{\partial S_w}\frac{\partial S_w}{\partial w_i}$ e $G_2 = \frac{\partial^2 \tau_{iw}}{\partial \gamma_i \partial w_i}$.

6.3.3 Perturbação de casos ponderados

Com este tipo de perturbação se deseja avaliar se a contribuição das observações com ponderações afetam o estimador de máxima verossimilhança. O logaritmo da função de verossimilhança do modelo perturbado é dado por

$$\ell(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n} w_i \left\{ (y_i - \mu_i)\Psi_i^{(1,0)}(\mu_i, \phi_i) + \phi_i T(y_i) + \Psi_i(\mu_{iw}, \phi_i) \right\} + \sum_{i=1}^{n} w_i \log A(y_i).$$

Assim, temos que

$$\begin{aligned}\boldsymbol{\Delta}_{\boldsymbol{\beta}} &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\beta} \partial \boldsymbol{w}^\top} \\ &= \sum_{i=1}^{n} \psi_i^{(2,0)}(y_i - \mu_i) m_{1i}(s)_i.\end{aligned}$$

Em forma matricial, temos

$$\boldsymbol{\Delta}_{\boldsymbol{\beta}} = \widetilde{X}^\top \boldsymbol{\Psi}^{(2,0)} M_1^2 (\boldsymbol{y} - \boldsymbol{\mu}). \tag{6.12}$$

Analogamente, temos

$$\begin{aligned}\boldsymbol{\Delta}_{\gamma} &= \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{w}^\top} \\ &= \sum_{i=1}^{n} \left\{ \psi_i^{(1,1)}(y_i - \mu_i)\phi_{1i}(S)_i + T(y_i)\phi_{1i}(S)_i + \psi_i^{(0,1)}\phi_{1i}(S)_i \right\}.\end{aligned}$$

Em forma matricial, temos

$$\boldsymbol{\Delta}_{\gamma} = \widetilde{S}^\top \boldsymbol{\Psi}^{(1,1)} \boldsymbol{\Phi}_1 (\boldsymbol{y} - \boldsymbol{\mu}) + \widetilde{S}^\top \boldsymbol{\Phi}_1 \boldsymbol{T(y)} + \widetilde{S}^\top \boldsymbol{\Phi}_1 \boldsymbol{\Psi}^{(0,1)}. \tag{6.13}$$

Utilizando as derivadas da matriz $\boldsymbol{\Delta}$ para os esquemas de perturbação e calculando a curvatura da equação (6.7), podemos avaliar quais pontos são influentes no modelo proposto de acordo com cada caso. Um gráfico interessante de se observar é o gráfico $\boldsymbol{C_t}$ versus seus índice.

Escobar e Meeker (1992) sugerem que os elementos da diagonal principal da matriz $\boldsymbol{B}$, dada da seguinte forma $\boldsymbol{B} = -\boldsymbol{\Delta}^T \ddot{L}_{\hat{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}}^{-1} \boldsymbol{\Delta}\Big|_{w=w_0}$ devem ser tomados como medida de influência. Já Lesaffre e Verbeke (1998) sugerem que se deve avaliar a curvatura na direção da i-ésima observação, ou seja, avaliamos $\boldsymbol{C_t}$ no vetor $\boldsymbol{t}_i$ de zeros com um na i-ésima posição. Assim, obtemos que $C_i = 2\,|b_{ii}|$, em que b_{ii} é o elemento ii da diagonal da matriz $\boldsymbol{B}$. As observações tais que $C_i > 2\bar{C}$, onde $\bar{C}$ é a média das curvaturas, devem ser estudas com cautela.

CAPÍTULO 7

Aplicação das Técnicas de Diagnósticos

Para testar a aplicação das técnicas de diagnósticos dos MNLGS, empregou-se os dados experimentais obtidos durante um teste de aprendizagem e memória espaciais aplicado em ratos portadores de lesão cerebral isquêmica, ou seja, pela falta de sangue no cérebro, dados analisados por Cordeiro, Previdelli, Samohyl (2008). Esse teste utiliza o modelo do labirinto radial de 8 braços conforme Figura 7.1, usado pelo Laboratório da Isquemia Cerebral e Neuroproteção, do Departamento de Farmácia e Farmacologia da Universidade Estadual de Maringá.

7.1 Dados dos Ratos

A necessidade de desenvolvimento de estratégias farmacológicas para o tratamento e/ou prevenção de doenças neurodegenerativas agudas (lesão cerebral de origem isquêmica, infecciosa, traumática, etc) ou crônicas (doença de Parkinson, doença de Alzheimer, entre outras) é urgente. A população idosa do globo cresce em escala geométrica e essas doenças afetam fundamentalmente pessoas da terceira idade, ou seja, acima de 65 anos. reproduzem tanto as doenças quanto as sequelas das mesmas nos seres humanos representam o ponto de partida para a investigação dessas questões.

Na experimentação pré-clínica, o rato é o animal frequentemente utilizado, inclusive como modelo para as doenças neurodegenerativas e suas sintomatologias. Disfunções de aprendizagem e memória podem ser medidas nesses modelos animais mediante o uso de testes e/ou modelos comportamentais, tais como os chamados labirintos. Estes medem, de modo geral, a capacidade do animal para executar uma determinada tarefa que demanda “atenção”, “raciocínio”, “motivação”, ou seja, funções cognitivas (aprendizagem e memória). Além disso, o teste de labirinto é suficientemente sensível para detectar os efeitos das drogas no rato, isto é, na eficácia terapêutica em facilitar ou dificultar a capacidade de aprendizagem ou dos efeitos causados por lesões cerebrais, além dos efeitos do envelhecimento, entre outros.

Foi utilizado no experimento um labirinto radial de oito braços aversivo, o mesmo é considerado como um modelo de aprendizagem que pretende imitar situações que o animal possa se deparar no ambiente natural. Na Figura 7.1 temos uma representação esquemática do labirinto radial de oito braços aversivo. Esse tipo de experimento parte do pressuposto de que alguns comportamentos aprendidos pelo animal são úteis em sua sobrevivência no meio selvagem, como por exemplo, a procura por água e comida. No labirinto utilizado no experimento, os braços se originam num ponto central e a comunicação dos braços com a arena central tem trânsito livre. Nas extremidades dos braços, uma abertura permite o acesso do animal a uma pequena caixa escura localizada logo abaixo de cada orifício, a qual pode ser inserida e removida como uma gaveta abaixo, servindo como refúgio para o rato em relação às áreas iluminadas do labirinto. Dentre os oito braços, somente um contem o refúgio verdadeiro, sendo que nos demais braços os esconderijos são de fundo falso. As funções cognitivas de todos os ratos foram testadas através do teste do labirinto, no qual era avaliada a capacidade do rato em encontrar o esconderijo. Cerca de vinte dias após a indução da isquemia cerebral, os ratos foram colocados diariamente no labirinto. O experimento durou 15 dias, e a cada dia de teste foram dadas três tentativas ao animal para encontrar o esconderijo.

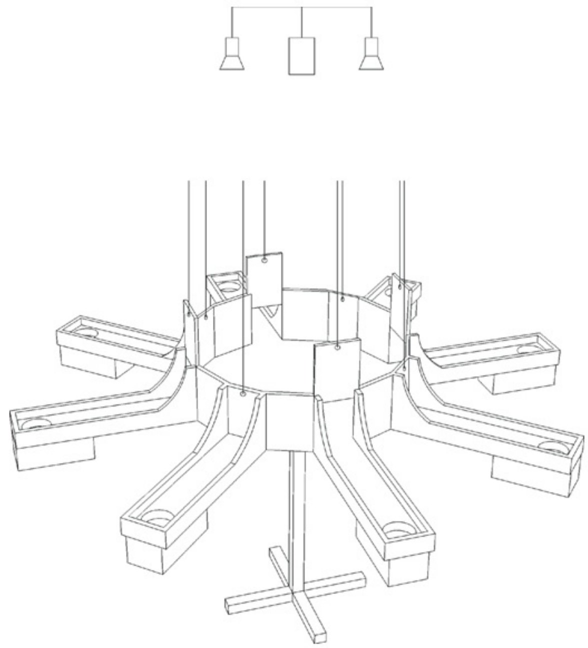

Figura 7.1: Representação esquemática do labirinto radial de oito braços aversivo.

Foram utilizados 51 ratos Wistar machos, adultos, pesando entre 280 e 300 gramas (aproximadamente 3 meses de idade). Eles permaneceram em condições padrão de alojamento, com ciclo de luz claro/escuro, temperatura controlada ($22 \pm 1°C$) e receberam água e ração à vontade. Vinte e cinco animais foram submetidos à isquemia cerebral global e transitória conforme o modelo de oclusão dos quatro vasos, e os outros 26 animais, designados como grupo falso isquêmico (não-lesionado), também foram submetidos ao mesmo procedimento cirúrgico, porém sem a oclusão das artérias vertebrais e carótidas.

No teste do labirinto radial avaliou-se a capacidade de aprendizagem e memória espacial no rato. A cada dia de teste foram dadas três tentativas ao animal para encontrar o esconderijo. Dentro de cada tentativa, um erro de referência é contado toda vez que o animal entrar num braço falso do esconderijo, pela primeira vez. Um erro operacional é registrado toda vez que o rato torna a entrar num braço previamente visitado. Por essa definição, o animal utiliza-se da memória de referência para se lembrar de que apenas um dos oito braços contém a recompensa (esconderijo), e que a sua relativa localização espacial se manteve inalterada ao longo das várias sessões (dias) de teste, indicando, portanto, uma memória de longo-prazo. Diferentemente, a memória operacional implica que o animal deve se recordar de qualquer dos outros sete braços (sem recompensa), os quais foram previamente visitados durante uma mesma tentativa, definida por memória de curto-prazo.

7.2 Modelo não Linear

O delineamento descrito é a contagem do número de erros observados no mesmo rato em tempos diferentes. Esses dados já foram analisados por Cordeiro, Previdelli, Samohyl (2008) pela metodologia de modelos lineares mistos, utilizando uma matriz de covariância não estruturada. No entanto, segundo Dobson (2001), uma análise alternativa é considerar os modelos da família exponencial, como, por exemplo, um modelo de regressão Poisson.

A variável resposta corresponde ao número de erros cometidos pelos ratos e as covariáveis foram

$$x_0 = \begin{cases} 1 & \text{, se o } i\text{-rato for não lesionado} \\ 0 & \text{, caso contrário,} \end{cases}$$

e

$$x_1 = \begin{cases} 1 & \text{, se o } i\text{-rato for lesionado} \\ 0 & \text{, caso contrário,} \end{cases}$$

e x_2 que corresponde ao tempo representado em cinco blocos de três dias cada, conforme as Tabelas B.1 e B.2.

Para comparar dois ou mais modelos para o mesmo conjunto de dados, utilizou-se o critério de AIC (critério de informação de Akaike), que penaliza o número de parâmetros no logaritmo da função de verossimilhança dado pela expressão

$$AIC = -2 \log \ell(\boldsymbol{\theta}) + 2n_{par},$$

em que n_{par} é o número de parâmetros do modelo ajustado. Sob essa definição, prefere-se o modelo com menor AIC.

No ajuste da Poisson obteve-se $AIC = 173,81$ e no ajuste da Poisson dupla um $AIC = 87,743$. Pode-se observar pelos valores dos critérios AIC que o modelo Poisson dupla se ajusta melhor aos dados, pois possui o menor valor do AIC.

O modelo utilizado por Cordeiro et al (2008) foi o modelo de regressão não linear generalizado superdispersado Poisson dupla, tendo com função de ligação para o parâmetro de locação a função identidade e no qual

$$\eta_i = \beta_0 x_{0i} + x_{1i} \exp(\beta_1 x_{2i}) + x_{0i} \exp(\beta_2 x_{2i}), \quad i = 1, \ldots, 255, \tag{7.1}$$

e o parâmetro de dispersão, com função de ligação exponencial, é estimado pela componente sistemática dada por

$$\phi_i = \exp(\gamma_0 + \gamma_1 x_{2i}), \quad i = 1, \ldots, 255. \tag{7.2}$$

Os parâmetros β's e γ's foram calculados, com os valores iniciais encontradas pelo modelo não linear normal na linguagem de progamação `Ox` (Doornik, 2001).

7.3 Ajuste do modelo de regressão Poisson Dupla

A Tabela 7.1 contém as estimativas dos parâmetros do submodelo da média com seus respectivos p-valores.

Tabela 7.1: Estimativas do submodelo da média da Poisson dupla.

estimadores	estimativas	p-valor
β_0	$0,0601$	$0,2879$
β_1	$-0,6488$	$< 0,0001$
β_2	$-0,3306$	$0,0043$

Pode-se observar que o estimador $\hat{\beta}_0$ foi não significativo, ou seja, há evidências de que a média dos erros dos ratos lesionados e a dos não lesionados inicialmente não diferem; no entanto, isso vai mudando ao longo do tempo. A estimativa $\hat{\beta}_1 = -0,6488$ corresponde ao déficit de aprendizagem para os ratos não lesionados e $\hat{\beta}_2 = -0,3306$ para os ratos lesionados. A Figura 7.2 evidencia que o decréscimo dos erros para os ratos lesionados é menor.

A Tabela 7.2 contém as estimativas dos parâmetros do submodelo da dispersão com seus respectivos p-valores.

Tabela 7.2: Estimativas do submodelo da dispersão da Poisson dupla.

estimadores	estimativas	p-valor
γ_0	-0.62679	$0,0025$
γ_1	0.44064	$< 0,0001$

A Figura 7.2 abaixo apresenta o submodelo da média ajustada pelo modelo Poisson dupla.

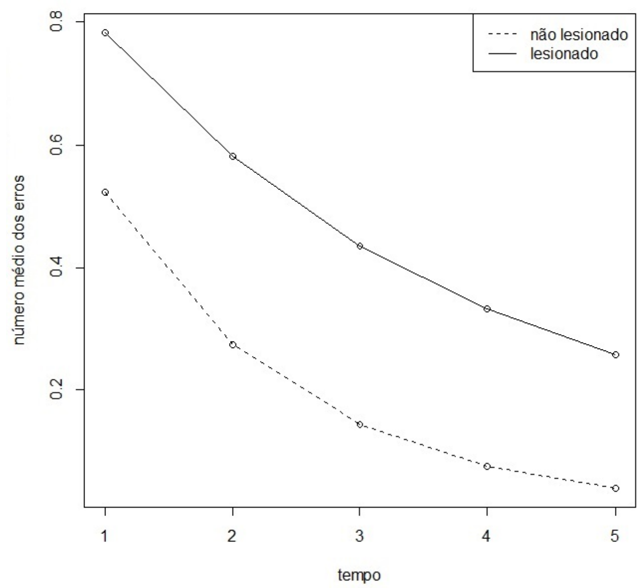

Figura 7.2: Modelo Poisson dupla ajustado.

Pela representação gráfica dos modelos ajustados, tanto para os ratos lesionados como para os não lesionados, pode-se afirmar que estes aprendem ao executar uma tarefa, pois os erros até achar a recompensa diminiu ao longo do tempo. No entanto, há uma declividade maior no número médio dos erros para os ratos não lesionados, isto é, estes aprendem numa velocidade maior ao longo do tempo.

As Figuras 7.3 a 7.10 apresentam alguns gráficos de diagnóstico para o ajuste do modelo. Na Figura 7.3 temos o gráfico de $\widehat{GL}(\boldsymbol{\mu})$ contra os valores ajustados e há um ponto em destaque, a observação 102. Já no gráfico de influência (Figura 7.4) foram destacados 2 pontos, 64 e 69, mas a partir do gráfico da Figura 7.5 identificamos que nenhum ponto se destaca. Na Figura 7.6 tem-se o gráfico de resíduos t_{Si} contra os valores ajustados, vários pontos estão fora do intervalo

$[-2, 2]$, as observações 64, 102, 192, 197 e 230 estão bem distantes das outras. As observações 64, 69, 102 (ratos lesionados) e 192, 197 e 230 (ratos não lesionados) são os ratos que apresentaram erros mais altos nos blocos 3 para as primeiras observações destacadas e bloco 5 para a última observação, respectivamente.

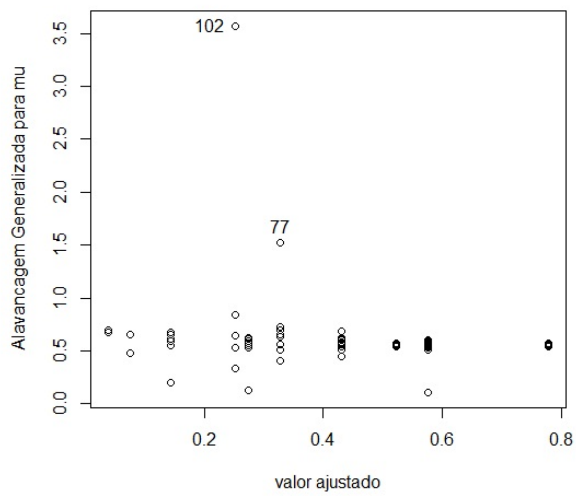

Figura 7.3: Gráfico de alavancagem generalizada referente ao submodelo da média da Poisson dupla.

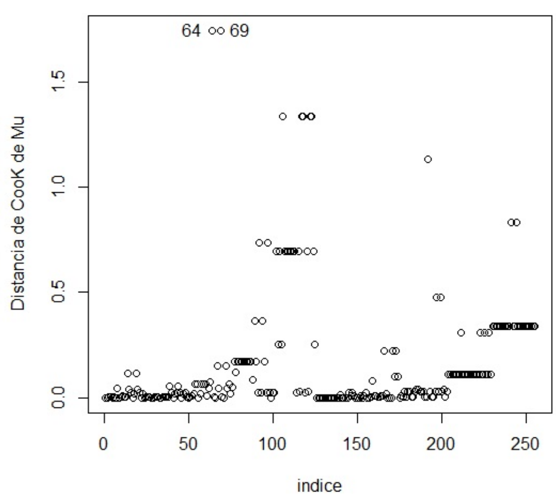

Figura 7.4: Gráfico de influência global referente ao submodelo da média da Poisson dupla.

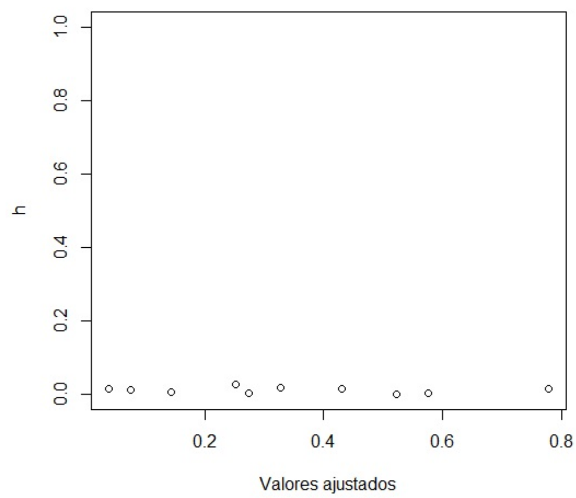

Figura 7.5: Gráfico de $\hat{h}_{ii}$ por valor ajustado referente ao submodelo da média da Poisson dupla.

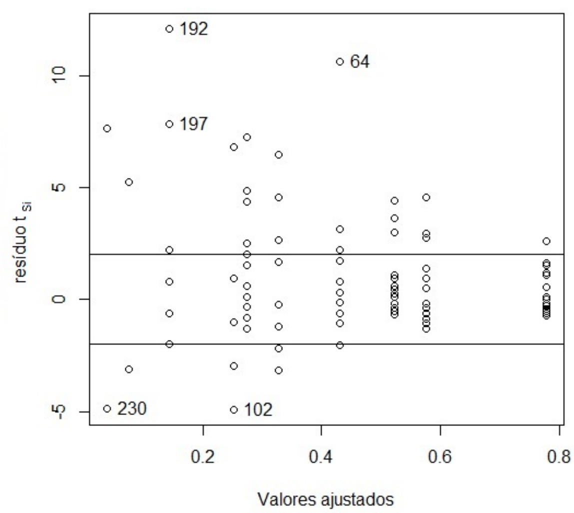

Figura 7.6: Gráfico de resíduo por valor ajustado referente ao submodelo da média da Poisson dupla.

Na Figura 7.7 temos o gráfico de $\widehat{GL}(\phi)$ contra os valores ajustados e há dois pontos em destaque, a observação 77 e a observação 204. Já no gráfico de influência (Figura 7.8) foram destacados 3 pontos, 192, 241 e 244, mas a partir do gráfico da Figura 7.9 identificamos que nenhum ponto se destaca. Na Figura 7.10 tem-se o gráfico de resíduos t_{Ti} contra os valores ajustados, vários pontos estão fora do intervalo $[-2, 2]$, as observações 64, 192 e 102, estão bem distantes das outras. As observações 64, 77, 102 (ratos lesionados) e 192, 204, 241 e 244 (ratos não lesionados) são os ratos que apresentaram erros mais altos no bloco 3 para as observações 64, 102 e 192, bloco 4 para as observações 77 e 204 destacadas e bloco 5 para as observações 241 e 244, respectivamente.

Retirando todas as observações destacadas pelos gráficos de diagnóstico, uma de cada vez, não houve mudança inferencial. Ou seja, apesar dos pontos se destacarem nos gráficos de diagnósticos, ao serem retirados as estimativas dos parâmetros não se alteraram. Assim, mesmo que alguns ratos tiveram erros maiores do que os outros em determinados blocos de tempo, a média dos erros ao longo do tempo continuaram a decrescer à medida que o tempo passava; e os ratos não lesionados erraram menos que os ratos lesionados ao passar do tempo como visto no gráfico 7.2. Levando a conclusão que a área afetada do cérebro dos ratos influencia na memória e aprendizagem dos ratos.

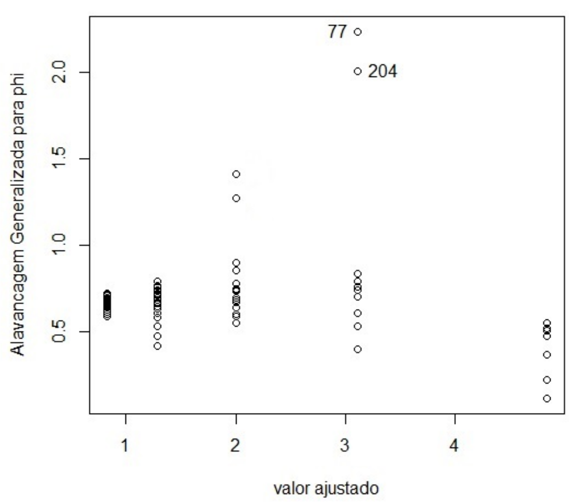

Figura 7.7: Gráfico de alavancagem generalizada referente ao submodelo da dispersão da Poisson dupla.

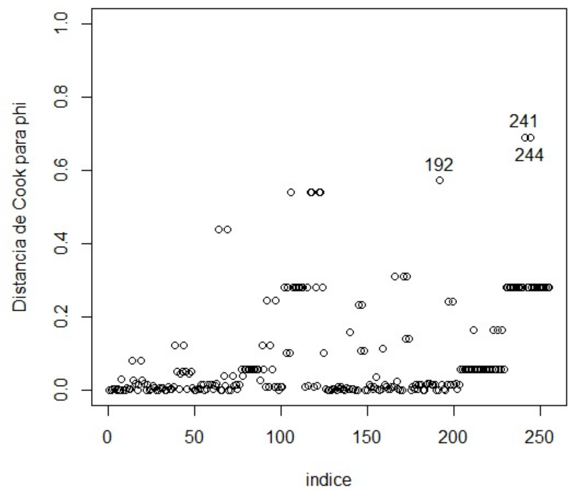

Figura 7.8: Gráfico de influência global referente ao submodelo da dispersão da Poisson dupla.

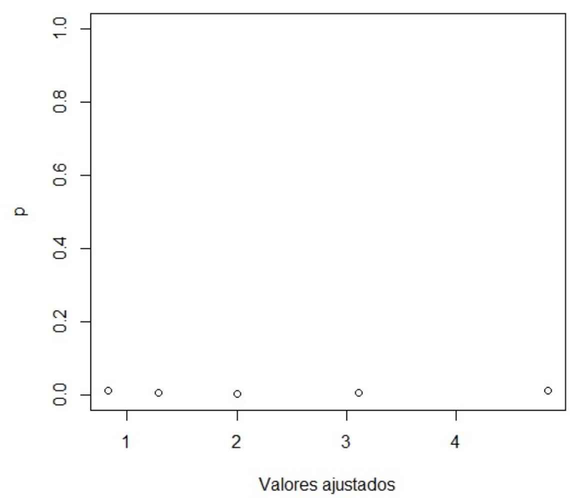

Figura 7.9: Gráfico de $\hat{h}_{ii}$ por valor ajustado referente ao submodelo da dispersão da Poisson dupla.

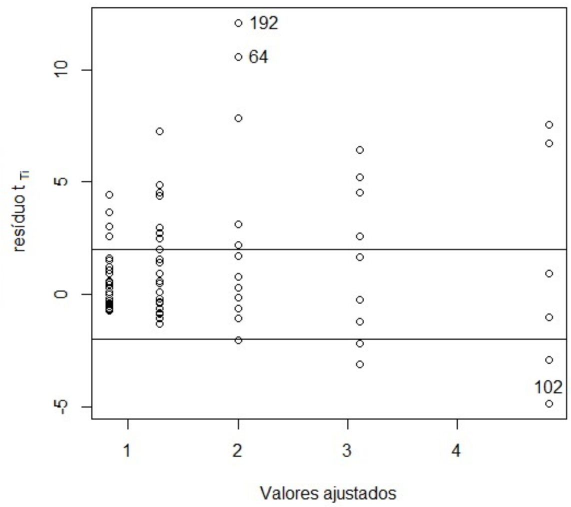

Figura 7.10: Gráfico de resíduo por valor ajustado referente ao submodelo da dispersão da Poisson dupla.

7.4 Influência Local

Como foi dito na seção 5.3, iremos analisar a influência local através dos seguintes esquemas de perturbação: perturbação de casos ponderados, perturbação aditiva na resposta, perturbação aditiva nos preditores, tanto para o submodelo da média quanto para o submodelo da dispersão.

As Figuras 7.11 a 7.13 apresentam os gráficos de curvaturas para o ajuste do modelo nos esquemas de perturbação citados acima. Pela Figura 7.11, podemos ver que, para o esquema de perturbação aditiva na resposta, os ratos não lesionados não apresentam curvaturas acima do ponto de corte ($2\bar{C}$), já os ratos não lesionados, observações 230 a 255 destacadas, apresentam um grupo com curvaturas muito acima dos demais. Na Figura 7.12, podemos notar que, para o esquema de perturbação aditiva nos preditores, não houve nenhuma observação que se destacou para ambos os ratos lesionados e não lesionados. E na Figura 7.13, podemos observar que, para o esquema de perturbação ponderada, os ratos lesionados, observações 101, 106, 114, 119 e 123, apresentaram curvaturas bem acima dos demais, enquanto que os ratos não lesionados apresentaram poucas curvaturas acima dos demais.

Retirando, uma de cada vez, todas as observações destacadas pelos gráficos de curvaturas não houve mudança inferencial nos parâmetros do modelo proposto. Com isso percebemos que o esquema de perturbação aditiva nos preditores não há nenhuma alteração no modelo proposto. E para os esquemas de perturbação aditiva na resposta e aditiva nos preditores, os estimadores são mais sensíveis a perturbações mas não ao ponto de interferir com as estimativas do modelo.

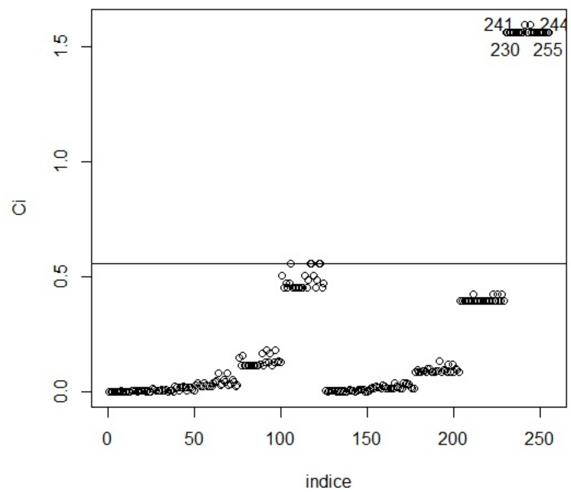

Figura 7.11: Gráfico de índice do C_i com perturbação aditiva na resposta referente ao modelo da Poisson dupla.

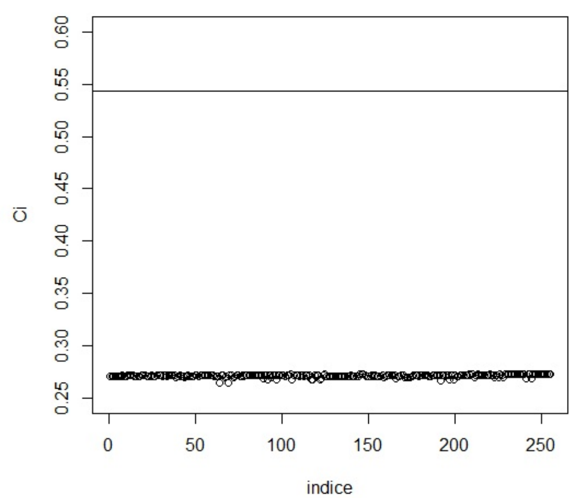

Figura 7.12: Gráfico de índice do C_i com perturbação aditiva nos preditores referente ao modelo da Poisson dupla.

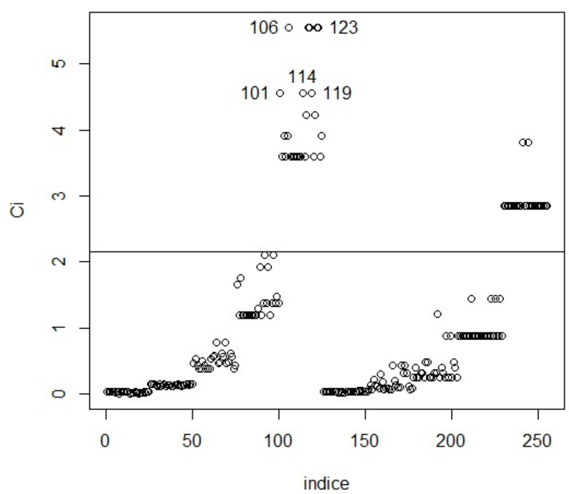

Figura 7.13: Gráfico de índice do C_i com perturbação ponderada referente ao modelo da Poisson dupla.

CAPÍTULO 8

Conclusão

Resumimos as principais contribuições teóricas desta tese nos seguintes itens:

(i) Obtivemos um fator de correção para o teste da razão de verossimilhanças nos modelos não lineares generalizados com superdispesão, estendendo assim a proposta de Cordeiro et al. (2006).

(ii) Obtivemos um fator de correção de tipo-Bartllet para o teste escore original nos modelos não lineares generalizados com superdispesão.

(iii) Aplicamos métodos usuais de diagnósticos para a classe de modelos não lineares generalizados com superdispesão, como alavancagem generalizada, distância de Cook e influência local.

Além dessas contribuições, podemos ter as seguintes conclusões:

Através de resultados de simulação, verificou-se que os testes baseados nas estatísticas escore e razão de verossimilhanças corrigidas, respectivamente, têm taxas de rejeição empíricas mais próximas dos níveis nominais selecionados do que os testes baseados nas estatísticas escore e razão de verossimilhanças não corrigidas, em amostras finitas. Também verificou-se através de estudos de simulação os desempenhos dos testes baseados nas estatísticas máxima verossimilhança e escore e suas versões corrigidas quando aumentava-se o número de parâmetros de perturbação e de interesse, onde os testes baseados nas estatísticas corrigidas apresentaram taxas menores do que os testes baseados nas estatísticas originais. A importância das correções fica ainda mais evidente no caso em que o tamanho da amostra não é grande ou o número de parâmetros de perturbação é considerável. E quanto ao poder do teste, à medida que nos afastamos da hipótese nula, os poderes dos testes baseados nas estatísticas corrigidas são maiores que os poderes dos testes baseados nas estatísticas originais.

As técnicas de diagnósticos apresentadas no Capítulo 6 para os modelos não lineares generalizados com superdispesão foram aplicadas em banco de dados reais. Assim, a partir do Capítulo 7, nota-se que as técnicas desenvolvidas nesta tese detectaram pontos discrepantes, mas uma análise mais profunda, verificou que apesar dos pontos se destacarem na análise, eles não afetam as estimativas dos parâmetros. Toda a análise de verificação de observações influentes foi realizada por métodos gráficos, construídos com os software `Ox` (Doornik, 2001) e `R`, este disponível gratuitamente em `www.r-project.org`.

Com isso, conclui-se que as técnicas de diagnósticos para os modelos não lineares generalizados com superdispersão desenvolvidas neste trabalho podem contribuir na área de modelagem em situações reais em análise de regressão.

Referências Bibliográficas

Bartlett, M. S. (1937). *Bartlett corrections for generalized linear models with dispersion covariates.* **Communications in Statistics - Theory and Methods**. **26**, 279–307.

Barndorff-Nielsen, O. E., Hall, P. (1988). *On the level-error after Bartlett adjustment of the likelihood ratio statistics.* **Biometrika** 75, 374–378.

Botter, D. A., Cordeiro, G. M. (1997). *Properties of sufficiency and statistical tests.* **Proceedings of the Royal Society of London A**. **160**, 268–282.

Cook, R. D. (1977). *Detection of influential observations in linear regressions.* **Technometrics** **19**, 15–18.

Cook, R. D. (1986). *Assessment of Local Influence (with discussion).* **Journal Royal Statistics Society B 48**, 133–169.

Cook, R. D., Weisberg, S. (1982). **Residuals and Influence in Regression**. Chapman and Hall, London.

Cordeiro, G. M. (1987). *On the Correction to the likelihood ratio statistics.* **Biometrika**, **74**, 265–274.

Cordeiro, G. M. (1983). *Improved likelihood ratio statistics for generalized linear models.* **Jour-**

nal of the Royal Statistical Society , **B** (Methodological) 45, 404–413.

Cordeiro, G. M., Botter, D. A. (2001). *Second-order biases of maximum likelihood estimates in overdispersed generalized linear models.* **Statistics and Probability Letters**, 55, 269–280.

Cordeiro, G. M., Botter, D. A., Ferrari, S. L. P. (2003). *Three corrected score tests for generalized linear models with dispersion covariates.* **Statistica Neerlandica**. **57**, nr. 4, 391–409.

Cordeiro, G. M., Cysneiros, A. H. M. A., Cysneiros, F. J. A. (2006). *Bartlett adjustments for Overdispersed Generalized linear Models.* **Communications in Statistics - Theory and Methods**. **35**, 937–952.

Cordeiro, G. M., Ferrari, S. L. P. (1991). *A general method for approximating to the distribuition of some statistics.* **Biometrika**. **78**, 573–582.

Cordeiro, G. M., Ferrari, S. L. P., Cysneiros, A. H. M. A. (1998). *A formula to improve score test statistics.* **Journal of Statistical Computation and Simulation**. **62**, 123–136.

Cordeiro, G. M., Ferrari, S. L. P., Paula, G. A. (1993). *Improved score tests for generalized linear models.* **Journal of the Royal Statistical Society**, **B**, **55**, 661–674.

Cordeiro, G. M., McCullagh, P. (1991). *Bias correction in generalized linear models.* **Journal of the Royal Statistical Society**, **B**, 53, 629–643.

Cordeiro, G. M., Previdelli, I. T. S., Samohyl, R. W. (2008). *Bias corrected maximum likelihood estimators in nonlinear overdispersed models.* **Brazilian Journal of Probability and Statistics**, v. 22, p. 105–118.

Cox, G. M., Hinkley, D. V. (1974). **Theoretical Statistics**. New York, John Wiley.

Cox, D. R., Snell, E. J. (1968). *A general definition of residuals.* **Journal of the Royal Statistical Society B 30**, 284–275.

Cribari-Neto, F., Ferrari, S. L. P. (1995). *Second order asymptotics for score tests in generalized linear models.* **Biometrika**. **82**, 426–432.

Cysneiros, A. H. M. A., Rodrigues, K. S. P., Cordeiro, G. M., Ferrari, S. L. P. (2010). *Three Bartlett-type corrections for score statistics in symmetric nonlinear regression models.* **Statistical Papers**, 51, 273–284.

Dey, D. K., Gelfand, A. E., Peng, F. (1997). *Overdispersed generalized linear models.* **Journal of Statistical Planning and Inference**. **64**, 93–107.

Dobson, A. J. (2001). **An introduction to generalized linear models**. 2nd ed. Chapman & Hall/CRC. Boca Raton, Florida.

Doornik, J. A. (2001). **Ox: an object-oriented matrix programming language.** 4th ed. London, Timberlake consultants and `Oxford.http://www.nuff.ox.ax.uk/Users/Doornik/`.

Efron, B. (1986). *Double exponencial families and their use in generalized linear regression.* **Journal of the American Statistical Association**. **81**, 709–721.

Escobar, L. A., Meeker, W. Q. (1992). *Assessing influence in regression analysis with censored data.* **Biometrics 48**, 507–528.

Gelfand, A.E., Dalal, S.R. (1990). *A note on overdispersed exponential families*, **Biometrika** 77, 55–64.

Good, I. J. (1953). *The population frequencies of species and the estimation of population parameters.* **Biometrika**, **40**, 237–264.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1994). **Continuos univariate distribuitions**. 2nd ed. New York: Jonh Wiley and Sons, p. 282.

Jϕrgensen, B. (1983). *Maximum Likelihood Estimation and Large-Simple Inference for Generalized Linear and Nonlinear Regression Models.* **Biometrika**, **70**, 19–28.

Kakizawa, Y. (1996). *Higher order monotone Bartlett-type adjustment for some multivariate test statistics.* **Biometrika**. **83**, 923–927.

Lawley, D. N. (1956). *A general method for approximating to the distribuition of the likelihood ratio criteria.* **Biometrika**. **43**, 295–303.

Lesaffre, F., Verbeke, G. (1998). *Local influence in linear mixed models.* **Biometrics 38**, 963–974.

Lindsay, B. (1986). *Exponential family mixture models (with least sqares estimators).* **Annals of Statistics**. **14**, 124–137.

Luenberger, D. G. (1973). **Introduction to linear and nonlinear programming**. Addison-Wesley Publishing Company Inc.

McCullagh, P., Nelder, J. A. (1989). **Generalized Linear Models**. 2^{nd} ed. London, Chapman and Hall.

Nocedal, J., Wright, S. J. (1999) **Numerical Optimization**, New York, NY: Springer.

Paula, G. A. (1995). *Influence and residuals in restricted generalized linear models.* **Journal of Statistical Computation and Simulation 51**, 315–352.

Pregibon, D. (1981). *Logistic regression diagnostics.* **Annals of Statistics 9**, 705–724.

Rao, C. R. (1947). *Large sample tests on statistical hypotheses concerning several parameters with applications to problems of estimation.* **Proceedings of the Cambridge Philosophical Society**. **44**, 50–57.

Rao, C. R. (1973). **Linear Statistical Inference and its Applications**. 2^{nd} ed. New York: Wiley. 625p.

Smyth, G. K. (1989). *Generalized linear models with varyng dispersion.* **Journal of the Royal**

Statistical Society. **B 51**, 47–60.

Wasan, M. T. (1968). **First passage time distribution of Brownian motion**. Monograph, Department of Mathematics, Kingston, Ontario: Queens's University.

Wei, B.C., Hu, Y.Q., Fung, W.K. (1998). *Generalized leverage and its applications.* **Scandinavian Journal of Statistics 25**, 25–37.

Wise, M. E. (1966). *Tracer-dilution curves in cardiology and randon walk and lognormal distributions.* **Acta Physiologica Pharmacologica Neerlandica**. **14**, 175–204.

Wise, M. E., (1971). *Skew probability curves with negative powers of time and related to random walks in series.* **Statistica Neerlandica**. **25**, 159–180.

APÊNDICE A

Cumulantes para os MNLGS

É importante lembrar algumas notações para o cálculo dos cumulantes, tais como:

$$\kappa_{rst} = \mathrm{E}\left[\frac{\partial^3 \ell(\theta)}{\partial\beta_r\partial\beta_s\partial\beta_t}\right] \quad \text{e} \quad \kappa_{rs,t} = \mathrm{E}\left[\left\{\frac{\partial^2 \ell(\theta)}{\partial\beta_r\partial\beta_s}\right\}\frac{\partial^2 \ell(\theta)}{\partial\beta_t}\right],$$

$$(r)_\ell = \frac{\partial\eta_\ell}{\partial\beta_r}, \qquad (rs)_\ell = \frac{\partial^2\eta_\ell}{\partial\beta_r\partial\beta_s} \qquad \kappa_{rs}^{(T)} = \frac{\partial\kappa_{rs}}{\partial\gamma_T}$$

Os κ's são todos de $O(n^{-1})$.

A.1 Cumulantes para o componente β

$$\kappa_{rs} = -\sum_\ell \psi_\ell^{(2,0)} m_{1\ell}^2 (r)_\ell (s)_\ell$$

$$\kappa_{rst} = -\sum_\ell \left[\left\{2\psi_\ell^{(3,0)} m_{1\ell}^3 + 3\psi_\ell^{(2,0)} m_{1\ell} m_{2\ell}\right\}(r)_\ell(s)_\ell(t)_\ell + \psi_\ell^{(2,0)} m_{1\ell}^2 \left\{(sr)_\ell(t)_\ell + (st)_\ell(r)_\ell\right.\right.$$
$$\left.\left. +\ (rt)_\ell(s)_\ell\right\}\right]$$

$$\kappa_{rstu} = -\sum_\ell \left[\left\{3\psi_\ell^{(4,0)} m_{1\ell}^4 + 12\psi_\ell^{(3,0)} m_{1\ell}^2 m_{2\ell} + 3\psi_\ell^{(2,0)} m_{2\ell} + 4\psi_\ell^{(2,0)} m_{1\ell} m_{3\ell}\right\}\right.$$
$$\times\ (r)_\ell(s)_\ell(t)_\ell(u)_\ell + \left\{2\psi_\ell^{(3,0)} m_{1\ell}^3 + 3\psi_\ell^{(2,0)} m_{1\ell} m_{2\ell}\right\} \times \left\{(st)_\ell(r)_\ell(u)_\ell + (rt)_\ell(s)_\ell(u)_\ell\right.$$
$$\left. +\ (su)_\ell(r)_\ell(t)_\ell + (ru)_\ell(s)_\ell(t)_\ell + (tu)_\ell(s)_\ell(r)_\ell + (sr)_\ell(t)_\ell(u)_\ell\right\} + \psi_\ell^{(2,0)} m_{1\ell}^2 \left\{(srt)_\ell(u)_\ell\right.$$
$$\left.\left. +\ (sru)_\ell(t)_\ell + (sut)_\ell(r)_\ell + (rut)_\ell(s)_\ell\right\} + \psi_\ell^{(2,0)} m_{1\ell}^2 \left\{(sr)_\ell(tu)_\ell + (st)_\ell(ru)_\ell + (rt)_\ell(su)_\ell\right\}\right]$$

$$
\begin{aligned}
\kappa_{rs}^{(t)} &= -\sum_{\ell}\Big[\Big\{\psi_\ell^{(3,0)}m_{1\ell}^3 + 2\psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}(r)_\ell(s)_\ell(t)_\ell + \psi_\ell^{(2,0)}m_{1\ell}^2\{(st)_\ell(r)_\ell + (rt)_\ell(s)_\ell\}\Big]\\
\kappa_{tu}^{(rs)} &= -\sum_{\ell}\Big[\Big\{\psi_\ell^{(4,0)}m_{1\ell}^4 + 5\psi_\ell^{(3,0)}m_{1\ell}^2m_{2\ell} + 2\psi_\ell^{(2,0)}m_{2\ell} + 2\psi_\ell^{(2,0)}m_{1\ell}m_{3\ell}\Big\}\times(r)_\ell(s)_\ell(t)_\ell(u)_\ell\\
&+\Big\{\psi_\ell^{(3,0)}m_{1\ell}^3 + 2\psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}\times\{(st)_\ell(r)_\ell(u)_\ell + (rt)_\ell(s)_\ell(u)_\ell + (su)_\ell(r)_\ell(t)_\ell + (rs)_\ell(u)_\ell(t)_\ell\\
&+(ur)_\ell(s)_\ell(t)_\ell\} + \psi_\ell^{(2,0)}m_{1\ell}^2\{(trs)_\ell(u)_\ell + (urs)_\ell(t)_\ell\} + \psi_\ell^{(2,0)}m_{1\ell}^2\{(ur)_\ell(ts)_\ell + (tr)_\ell(us)_\ell\}\Big]\\
\kappa_{rtu}^{(s)} &= -\sum_{\ell}\Big[\Big\{2\psi_\ell^{(4,0)}m_{1\ell}^4 + 9\psi_\ell^{(3,0)}m_{1\ell}^2m_{2\ell} + 3\psi_\ell^{(2,0)}m_{2\ell} + 3\psi_\ell^{(2,0)}m_{1\ell}m_{3\ell}\Big\}\times(r)_\ell(s)_\ell(t)_\ell(u)_\ell\\
&+\Big\{2\psi_\ell^{(3,0)}m_{1\ell}^3 + 3\psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}\times\{(sr)_\ell(t)_\ell(u)_\ell + (ts)_\ell(r)_\ell(u)_\ell(su)_\ell(r)_\ell(t)_\ell\}\\
&+\Big\{\psi_\ell^{(3,0)}m_{1\ell}^3 + 2\psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}\times\{(tr)_\ell(s)_\ell(u)_\ell + (tu)_\ell(r)_\ell(s)_\ell(ru)_\ell(t)_\ell(s)_\ell\}\\
&+\psi_\ell^{(2,0)}m_{1\ell}^2\{(sru)_\ell(t)_\ell + (sut)_\ell(r)_\ell(trs)_\ell(u)_\ell\}\\
&+\psi_\ell^{(2,0)}m_{1\ell}^2\{(sr)_\ell(tu)_\ell + (st)_\ell(ru)_\ell + (rt)_\ell(su)_\ell\}\Big]\\
\kappa_{rs,t} &= \sum_{\ell}\Big[\Big\{\psi_\ell^{(3,0)}m_{1\ell}^3 + 2\psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}(r)_\ell(s)_\ell(t)_\ell + \psi_\ell^{(2,0)}m_{1\ell}^2(sr)_\ell(t)_\ell\Big]\\
\kappa_{r,s,t} &= -\sum_{\ell}\psi_\ell^{(3,0)}m_{1\ell}^3(r)_\ell(s)_\ell(t)_\ell\\
\kappa_{rs,tu} &= \sum_{\ell}\left[\left\{\frac{\left(\psi_\ell^{(3,0)}\right)^2}{\psi_\ell^{(2,0)}}m_{1\ell}^4 + 2\psi_\ell^{(3,0)}m_{1\ell}^2m_{2\ell} + \psi_\ell^{(2,0)}m_{2\ell}^2\right\}\times(r)_\ell(s)_\ell(t)_\ell(u)_\ell\right.\\
&\left.+\Big\{\psi_\ell^{(3,0)}m_{1\ell}^3 + \psi_\ell^{(2,0)}m_{1\ell}m_{2\ell}\Big\}\times\{(tu)_\ell(r)_\ell(s)_\ell + (rs)_\ell(u)_\ell(t)_\ell\} + \psi_\ell^{(2,0)}m_{1\ell}^2(sr)_\ell(tu)_\ell\right.\\
\kappa_{r,s,tu} &= -\sum_{\ell}\left[\left\{\frac{\left(\psi_\ell^{(3,0)}\right)^2}{\psi_\ell^{(2,0)}}m_{1\ell}^4 + \psi_\ell^{(3,0)}m_{1\ell}^2m_{2\ell}\right\}\times(r)_\ell(s)_\ell(t)_\ell(u)_\ell + \psi_\ell^{(3,0)}m_{1\ell}^3(tu)_\ell(r)_\ell(s)_\ell\right]\\
\kappa_{r,s,t,u} &= \sum_{\ell}\left[\left\{\frac{\left(\psi_\ell^{(3,0)}\right)^2}{\psi_\ell^{(2,0)}}m_{1\ell}^4 - \psi_\ell^{(4,0)}m_{1\ell}^4\right\}\times(r)_\ell(s)_\ell(t)_\ell(u)_\ell\right]
\end{aligned}
$$

A.2 Cumulantes para o componente γ

$$
\begin{aligned}
\kappa_{RS} &= \sum_{\ell} \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} (R)_{\ell}(S)_{\ell} \\
\kappa_{RST} &= \sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(0,3)} \phi_{1\ell}^{3} + 3\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \Big\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (ST)_{\ell}(R)_{\ell} + (RT)_{\ell}(S)_{\ell} \right. \\
&\quad \left. + (SR)_{\ell}(T)_{\ell} \right\} \Big] \\
\kappa_{RS}^{(T)} &= \sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(0,3)} \phi_{1\ell}^{3} + 2\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \Big\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (ST)_{\ell}(R)_{\ell} + (RT)_{\ell}(S)_{\ell} \right\} \Big] \\
\kappa_{RSTU} &= \sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(0,4)} \phi_{1\ell}^{4} + 6\psi_{\ell}^{(0,3)} \phi_{1\ell}^{2}\phi_{2\ell} + 3\psi_{\ell}^{(0,2)} \phi_{2\ell}^{2} + 4\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{3\ell} \Big\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell} \\
&\quad + \Big\{ \psi_{\ell}^{(0,3)} \phi_{1\ell}^{3} + 3\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \Big\} \left\{ (SU)_{\ell}(R)_{\ell}(T)_{\ell} + (RU)_{\ell}(S)_{\ell}(T)_{\ell} + (TU)_{\ell}(R)_{\ell}(S)_{\ell} \right. \\
&\quad \left. + (ST)_{\ell}(R)_{\ell}(U)_{\ell} + (RT)_{\ell}(S)_{\ell}(U)_{\ell} + (SR)_{\ell}(U)_{\ell}(T)_{\ell} \right\} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (STU)_{\ell}(R)_{\ell} \right. \\
&\quad \left. + (RTU)_{\ell}(S)_{\ell} + (SRU)_{\ell}(T)_{\ell} + (RST)_{\ell}(U)_{\ell} \right\} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (ST)_{\ell}(RU)_{\ell} + (SU)_{\ell}(RT)_{\ell} \right. \\
&\quad \left. + (SR)_{\ell}(TU)_{\ell} \right\} \Big] \\
\kappa_{RS,T} &= -\sum_{\ell} \Big[\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} (R)_{\ell}(S)_{\ell}(T)_{\ell} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} (SR)_{\ell}(T)_{\ell} \Big] \\
\kappa_{R,S,T} &= -\sum_{\ell} \psi_{\ell}^{(0,3)} \phi_{1\ell}^{3} (R)_{\ell}(S)_{\ell}(T)_{\ell} \\
\kappa_{RTU}^{(S)} &= \sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(0,4)} \phi_{1\ell}^{4} + 6\psi_{\ell}^{(0,3)} \phi_{1\ell}^{2}\phi_{2\ell} + 3\psi_{\ell}^{(0,2)} \phi_{2\ell}^{2} + 3\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{3\ell} \Big\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell} \\
&\quad + \psi_{\ell}^{(0,3)} \phi_{1\ell}^{3} \left\{ (SU)_{\ell}(R)_{\ell}(T)_{\ell} + (RU)_{\ell}(S)_{\ell}(T)_{\ell} + (TU)_{\ell}(R)_{\ell}(S)_{\ell} + (ST)_{\ell}(R)_{\ell}(U)_{\ell} \right. \\
&\quad \left. + (RT)_{\ell}(S)_{\ell}(U)_{\ell} + (SR)_{\ell}(U)_{\ell}(T)_{\ell} \right\} + 3\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \left\{ (SR)_{\ell}(U)_{\ell}(T)_{\ell} \right. \\
&\quad \left. + (ST)_{\ell}(R)_{\ell}(U)_{\ell} + (SU)_{\ell}(R)_{\ell}(T)_{\ell} \right\} + 2\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \left\{ (TU)_{\ell}(R)_{\ell}(S)_{\ell} \right. \\
&\quad \left. + (RU)_{\ell}(S)_{\ell}(T)_{\ell} + (RT)_{\ell}(S)_{\ell}(U)_{\ell} \right\} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (STU)_{\ell}(R)_{\ell} + (SRU)_{\ell}(T)_{\ell} \right. \\
&\quad \left. + (RST)_{\ell}(U)_{\ell} \right\} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^{2} \left\{ (ST)_{\ell}(RU)_{\ell} + (SU)_{\ell}(RT)_{\ell} + (SR)_{\ell}(TU)_{\ell} \right\} \Big]
\end{aligned}
$$

$$\kappa_{TU}^{(RS)} = \sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(0,4)} \phi_{1\ell}^4 + 5\psi_{\ell}^{(0,3)} \phi_{1\ell}^2 \phi_{2\ell} + 2\psi_{\ell}^{(0,2)} \phi_{2\ell}^2 + 2\psi_{\ell}^{(0,2)} \phi_{1\ell} \phi_{3\ell} \Big\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell}$$
$$+ \Big\{ \psi_{\ell}^{(0,3)} \phi_{1\ell}^3 + 2\psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell} \Big\} \{ (SU)_{\ell}(R)_{\ell}(T)_{\ell} + (RU)_{\ell}(S)_{\ell}(T)_{\ell} + (ST)_{\ell}(R)_{\ell}(U)_{\ell}$$
$$+ (TR)_{\ell}(U)_{\ell}(S)_{\ell} + (RS)_{\ell}(U)_{\ell}(T)_{\ell} \} + \psi_{\ell}^{(0,2)} \phi_{1\ell}^2 \{ (TRS)_{\ell}(U)_{\ell} + (SRU)_{\ell}(T)_{\ell} \}$$
$$+ \psi_{\ell}^{(0,2)} \phi_{1\ell}^2 \{ (ST)_{\ell}(RU)_{\ell} + (SU)_{\ell}(RT)_{\ell} \} \Big]$$

$$\kappa_{TU,RS} = \sum_{\ell} \left[\left\{ \frac{\left(\psi_{\ell}^{(2,1)}\right)^2}{\psi_{\ell}^{(2,0)}} \phi_{1\ell}^4 - \psi_{\ell}^{(0,2)} \phi_{2\ell}^2 \right\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell} - \psi_{\ell}^{(0,2)} \phi_{1\ell}\phi_{2\ell}(SR)_{\ell}(U)_{\ell}(T)_{\ell} \right.$$
$$\left. - \psi_{\ell}^{(0,2)} \phi_{1\ell}^2 (SR)_{\ell}(TU)_{\ell} \right]$$

$$\kappa_{R,S,TU} = -\sum_{\ell} \left[\left\{ -\frac{\left(\psi_{\ell}^{(2,1)}\right)^2}{\psi_{\ell}^{(2,0)}} \phi_{1\ell}^4 + \psi_{\ell}^{(0,3)} \phi_{1\ell}^2 \phi_{2\ell} \right\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell} + \psi_{\ell}^{(0,3)} \phi_{1\ell}^3 (TU)_{\ell}(R)_{\ell}(S)_{\ell} \right]$$

$$\kappa_{TU,RS} = \sum_{\ell} \left\{ 3\frac{\left(\psi_{\ell}^{(2,1)}\right)^2}{\psi_{\ell}^{(2,0)}} \phi_{1\ell}^4 + \psi_{\ell}^{(0,4)} \phi_{1\ell}^4 \right\} \times (R)_{\ell}(S)_{\ell}(T)_{\ell}(U)_{\ell}$$

A.3 Cumulantes mistos

$$\kappa_{rS} = 0, \quad \kappa_{rST} = 0, \quad \kappa_{rS}^{(t)} = 0, \quad \kappa_{RSTu} = 0, \quad \kappa_{rST}^{(u)} = 0, \quad \kappa_{rST}^{(U)} = 0, \quad \kappa_{rS}^{(TU)} = 0,$$
$$\kappa_{rS}^{(tU)} = 0, \quad \kappa_{rs,T} = 0, \quad \kappa_{rS,T} = 0,$$

$$\kappa_{rsT} = \kappa_{sr}^{(T)} = -\sum_{\ell} \psi_{\ell}^{(2,1)} m_{1\ell}^2 \phi_{1\ell} (r)_{\ell}(s)_{\ell}(T)_{\ell}, \quad \kappa_{RS}^{(t)} = \sum_{\ell} \psi_{\ell}^{(1,2)} m_{1\ell} \phi_{1\ell}^2 (R)_{\ell}(S)_{\ell}(t)_{\ell}$$

$$\kappa_{rstU} = -\sum_{\ell} \Big[\Big\{ 2\psi_{\ell}^{(3,1)} m_{1\ell}^3 \phi_{1\ell} + 3\psi_{\ell}^{(2,1)} m_{1\ell} m_{2\ell} \phi_{1\ell} \Big\} \times (r)_{\ell}(s)_{\ell}(t)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)} m_{1\ell}^2 \phi_{1\ell} \{ (sr)_{\ell}(t)_{\ell}$$
$$\cdot (U)_{\ell} + (st)_{\ell}(r)_{\ell}(U)_{\ell} + (rt)_{\ell}(s)_{\ell}(U)_{\ell} \} \Big]$$

$$\kappa_{Rst}^{(u)} = -\sum_{\ell} \Big[\Big\{ \psi_{\ell}^{(3,1)} m_{1\ell}^3 \phi_{1\ell} + 2\psi_{\ell}^{(2,1)} m_{1\ell} m_{2\ell} \phi_{1\ell} \Big\} \times (R)_{\ell}(s)_{\ell}(t)_{\ell}(u)_{\ell} + \psi_{\ell}^{(2,1)} m_{1\ell}^2 \phi_{1\ell} \{ (su)_{\ell}(t)_{\ell}$$
$$\cdot (R)_{\ell} + (tu)_{\ell}(s)_{\ell}(R)_{\ell} \} \Big]$$

$$\kappa_{rst}^{(U)} = -\sum_{\ell}\Big[\Big\{2\psi_{\ell}^{(3,1)}m_{1\ell}^{3}\phi_{1\ell} + 3\psi_{\ell}^{(2,1)}m_{1\ell}m_{2\ell}\phi_{1\ell}\Big\}\times(r)_{\ell}(s)_{\ell}(t)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\{(sr)_{\ell}(t)_{\ell}$$
$$\cdot (U)_{\ell} + (st)_{\ell}(r)_{\ell}(U)_{\ell} + (rt)_{\ell}(s)_{\ell}(U)_{\ell}\}\Big]$$

$$\kappa_{rs}^{(tU)} = -\sum_{\ell}\Big[\Big\{\psi_{\ell}^{(3,1)}m_{1\ell}^{3}\phi_{1\ell} + 2\psi_{\ell}^{(2,1)}m_{1\ell}m_{2\ell}\phi_{1\ell}\Big\}\times(r)_{\ell}(s)_{\ell}(t)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\{(st)_{\ell}(r)_{\ell}$$
$$\cdot (U)_{\ell} + (rt)_{\ell}(s)_{\ell}(U)_{\ell}\}\Big]$$

$$\kappa_{rsTU} = -\sum_{\ell}\Big[\Big\{\psi_{\ell}^{(2,2)}m_{1\ell}^{2}\phi_{1\ell}^{2} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\Big\}\times(r)_{\ell}(s)_{\ell}(T)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}(TU)_{\ell}(r)_{\ell}(s)_{\ell}\Big]$$

$$\kappa_{rsT}^{(U)} = -\sum_{\ell}\Big[\Big\{\psi_{\ell}^{(2,2)}m_{1\ell}^{2}\phi_{1\ell}^{2} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\Big\}\times(r)_{\ell}(s)_{\ell}(T)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}(TU)_{\ell}(r)_{\ell}(s)_{\ell}\Big]$$

$$\kappa_{rs}^{(TU)} = -\sum_{\ell}\Big[\Big\{\psi_{\ell}^{(2,2)}m_{1\ell}^{2}\phi_{1\ell}^{2} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\Big\}\times(r)_{\ell}(s)_{\ell}(T)_{\ell}(U)_{\ell} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}(TU)_{\ell}(r)_{\ell}(s)_{\ell}\Big]$$

$$\kappa_{RST}^{(u)} = \sum_{\ell}\Big[\Big\{\psi_{\ell}^{(1,3)}m_{1\ell}\phi_{1\ell}^{3} + 3\psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}\phi_{2\ell}\Big\}\times(R)_{\ell}(S)_{\ell}(T)_{\ell}(u)_{\ell} + \psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}\{(ST)_{\ell}(R)_{\ell}$$
$$\cdot (u)_{\ell} + (RT)_{\ell}(S)_{\ell}(u)_{\ell} + (RS)_{\ell}(T)_{\ell}(u)_{\ell}\}\Big]$$

$$\kappa_{RS}^{(tU)} = \sum_{\ell}\Big[\Big\{\psi_{\ell}^{(1,3)}m_{1\ell}\phi_{1\ell}^{3} + 2\psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}\phi_{2\ell}\Big\}\times(R)_{\ell}(S)_{\ell}(U)_{\ell}(t)_{\ell} + \psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}\{(SU)_{\ell}(R)_{\ell}$$
$$\cdot (t)_{\ell} + (RU)_{\ell}(S)_{\ell}(t)_{\ell}\}\Big]$$

$$\kappa_{RS}^{(tu)} = -\sum_{\ell}\Big[\Big\{\psi_{\ell}^{(2,2)}m_{1\ell}^{2}\phi_{1\ell}^{2} + \psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}\Big\}\times(R)_{\ell}(S)_{\ell}(t)_{\ell}(u)_{\ell} + \psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}(tu)_{\ell}(R)_{\ell}(S)_{\ell}\Big]$$

$$\kappa_{Rs,t} = \sum_{\ell}\psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}(R)_{\ell}(s)_{\ell}(t)_{\ell}$$

$$\kappa_{R,s,t} = -\sum_{\ell}\psi_{\ell}^{(2,1)}m_{1\ell}^{2}\phi_{1\ell}(R)_{\ell}(s)_{\ell}(t)_{\ell}$$

$$\kappa_{RS,t} = \sum_{\ell}\psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}(R)_{\ell}(S)_{\ell}(t)_{\ell}$$

$$\kappa_{R,S,t} = -\sum_{\ell}\psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}(R)_{\ell}(S)_{\ell}(t)_{\ell}$$

$$\kappa_{rs,TU} = \sum_{\ell}\Bigg[\Bigg\{\frac{\psi_{\ell}^{(3,0)}\psi_{\ell}^{(1,2)}}{\psi_{\ell}^{(2,0)}}m_{1\ell}^{2}\phi_{1\ell}^{2} + \psi_{\ell}^{(1,2)}m_{2\ell}\phi_{1\ell}^{2} - \psi_{\ell}^{(2,0)}\psi_{\ell}^{(0,2)}m_{1\ell}^{2}\phi_{1\ell}^{2}\Bigg\}(r)_{\ell}(s)_{\ell}(T)_{\ell}(U)_{\ell}$$
$$+ \psi_{\ell}^{(1,2)}m_{1\ell}\phi_{1\ell}^{2}(rs)_{\ell}(T)_{\ell}(U)_{\ell}\Bigg]$$

$$\kappa_{rS,Tu} = \sum_{\ell} \frac{(\psi_\ell^{(2,1)})^2}{\psi_\ell^{(2,0)}} m_{1\ell}^2 \phi_{1\ell}^2 (r)_\ell (s)_\ell (T)_\ell (u)_\ell$$

$$\kappa_{r,s,T,U} = \sum_{\ell} \left\{ \frac{\psi_\ell^{(3,0)}\psi_\ell^{(1,2)}}{\psi_\ell^{(2,0)}} + 2\frac{(\psi_\ell^{(2,1)})^2}{\psi_\ell^{(2,0)}} - \psi_\ell^{(2,2)}\psi_\ell^{(2,0)}\psi_\ell^{(0,2)} \right\} m_{1\ell}^2 \phi_{1\ell}^2 (r)_\ell (s)_\ell (T)_\ell (U)_\ell$$

$$\begin{aligned}\kappa_{r,s,TU} = \sum_{\ell} \Bigg[\Bigg(\Bigg\{ -\frac{\psi_\ell^{(3,0)}\psi_\ell^{(1,2)}}{\psi_\ell^{(2,0)}} \psi_\ell^{(2,0)}\psi_\ell^{(0,2)} \Bigg\} m_{1\ell}^2 \phi_{1\ell}^2 - \psi_\ell^{(2,1)} m_{1\ell}^2 \phi_{2\ell} \Bigg) (r)_\ell (s)_\ell (T)_\ell (U)_\ell \\ - \psi_\ell^{(2,1)} m_{1\ell}^2 \phi_{1\ell} (r)_\ell (s)_\ell (TU)_\ell \Bigg]\end{aligned}$$

$$\begin{aligned}\kappa_{R,S,tu} = \sum_{\ell} \Bigg[\Bigg(\Bigg\{ -\frac{\psi_\ell^{(3,0)}\psi_\ell^{(1,2)}}{\psi_\ell^{(2,0)}} \psi_\ell^{(2,0)}\psi_\ell^{(0,2)} \Bigg\} m_{1\ell}^2 \phi_{1\ell}^2 - \psi_\ell^{(2,1)} m_{1\ell}^2 \phi_{2\ell} \Bigg) (R)_\ell (S)_\ell (t)_\ell (t)_\ell \\ - \psi_\ell^{(1,2)} m_{1\ell} \phi_{1\ell}^2 (R)_\ell (S)_\ell (tu)_\ell \Bigg]\end{aligned}$$

APÊNDICE B

Dados da Aplicação dos Métodos de Diagnósticos

Tabela B.1: Erros totais cometidos pelos ratos isquêmicos (lesionados).

Ratos	Bl_1	Bl_2	Bl_3	Bl_4	Bl_5	Ratos	Bl_1	Bl_2	Bl_3	Bl_4	Bl_5
1	5	12	3	5	3	14	32	26	27	8	3
2	5	10	6	0	3	15	22	19	3	0	3
3	3	5	2	6	3	16	19	3	4	2	3
4	3	4	0	0	3	17	9	18	11	10	3
5	8	1	0	0	3	18	18	19	8	2	3
6	13	10	5	0	3	19	32	26	27	8	3
7	8	1	2	0	3	20	22	19	3	0	3
8	23	4	0	0	3	21	19	3	4	2	3
9	6	8	0	0	3	22	9	18	11	10	3
10	4	0	0	0	3	23	18	19	8	2	3
11	1	2	6	0	3	24	8	10	0	3	3
12	3	2	8	0	3	25	2	4	2	2	3
13	2	0	9	1	3						

Tabela B.2: Erros totais cometidos pelos ratos não lesionados.

Ratos	Bl_1	Bl_2	Bl_3	Bl_4	Bl_5	Ratos	Bl_1	Bl_2	Bl_3	Bl_4	Bl_5
1	8	0	0	0	0	14	9	0	1	0	0
2	3	4	2	0	0	15	28	18	10	0	1
3	7	0	0	0	0	16	9	7	0	0	0
4	7	8	0	0	0	17	8	4	2	0	0
5	8	4	1	0	0	18	4	2	1	0	0
6	9	3	1	0	0	19	7	2	0	0	0
7	6	1	0	0	0	20	33	18	7	2	0
8	11	12	3	2	0	21	24	13	0	0	0
9	2	6	3	0	0	22	33	18	7	2	0
10	11	0	0	0	0	23	24	13	0	0	0
11	2	1	0	0	0	24	6	3	3	0	0
12	7	1	0	0	1	25	4	0	2	2	0
13	1	0	1	0	0	26	12	1	0	0	0