



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

JOAS SILVA DOS SANTOS

APRIMORAMENTO DE TESTES DE HIPÓTESES PARA MODELOS DE
SUPERDISPERSÃO E MODELOS BETA PRIME

Recife

2024

JOAS SILVA DOS SANTOS

**APRIMORAMENTO DE TESTES DE HIPÓTESES PARA MODELOS DE
SUPERDISPERSÃO E MODELOS BETA PRIME**

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística. Área de concentração: Estatística Matemática.

Orientadora: Prof^a Dr^a. Audrey Helen Mariz De Aquino Cysneiros

Recife

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Santos, Joas Silva dos.

Aprimoramento de testes de hipóteses para modelos de superdispersão e modelos beta prime / Joas Silva dos Santos. - Recife, 2024.

81p.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-graduação em Estatística.

Orientação: Audrey Helen Mariz de Aquino Cysneiros.

1. Modelos de regressão; 2. Parâmetro de dispersão; 3. Função de ligação; 4. Teste da razão de verossimilhanças; 5. Verossimilhança perfilada; 6. Correção de Bartlett. I. Cysneiros, Audrey Helen Mariz de Aquino - Orientador. II. Título.

UFPE-Central

JOAS SILVA DOS SANTOS

**APRIMORAMENTO DE TESTES DE HIPÓTESES PARA MODELOS DE
SUPERDISPERSÃO E MODELOS BETA PRIME**

Tese apresentada ao Programa de Pós-Graduação
em Estatística da Universidade Federal de
Pernambuco, como requisito parcial para a obtenção
do título de Doutor em Estatística. Área de
concentração: Estatística Matemática.

Aprovado em: 26 de abril de 2024.

BANCA EXAMINADORA

Prof^a Dr^a. Audrey Helen Mariz De Aquino Cysneiros
UFPE

Prof. Dr. Miguel Angel Uribe Opazo
UNIOESTE

Prof^a Dr^a. Mariana Correia De Araújo
UFRN

Prof. Dr. Roberto Ferreira Manghi
UFPE

Prof. Dr. Abraão David Costa do Nascimento
UFPE

A minha mãe, Valmira Silva.

AGRADECIMENTOS

Agradeço a Deus, por me dar forças nos momentos difíceis, para que eu pudesse continuar e conquistar meus objetivos.

A minha mãe, Valmira Silva, pelo seu amor, dedicação, apoio e pelo suporte durante a minha trajetória na pós-graduação.

À professora Audrey Cysneiros, pela orientação e disponibilidade.

Aos professores do Programa de Pós-Graduação em Estatística da UFPE.

Aos professores do Departamento de Estatística e Ciências Atuariais da UFS, em especial a José Rodrigo, pela confiança, amizade e incentivo.

Aos membros da banca examinadora, pela atenção e sugestões.

À CAPES, pelo apoio financeiro.

RESUMO

Nesta tese, tratamos de aprimoramentos de testes de hipóteses em modelos de regressão não lineares generalizados superdispersados e no modelo de regressão linear beta prime. Duas linhas de pesquisa são abordadas. A primeira, referente aos Capítulos 1 e 2, trata da obtenção do ajuste proposto por Skovgaard para a estatística da razão de verossimilhanças, assim como o ajuste para a função de verossimilhança perfilada modificada proposta por Cox e Reid nos modelos de regressão não lineares generalizados superdispersados. A segunda linha de pesquisa, referente ao Capítulo 3, trata da obtenção de uma correção de Bartlett para o teste da razão de verossimilhanças no modelo de regressão beta prime. Adicionalmente, consideramos o teste da razão de verossimilhanças bootstrap. Os desempenhos dos testes de hipóteses baseados nos refinamentos propostos foram avaliados numericamente e comparados às suas contrapartidas usuais através de estudos de simulação de Monte Carlo. Por fim, a utilidade dos refinamentos foi ilustrada através de aplicações a conjuntos de dados reais.

Palavras-chave: Modelos de regressão. Parâmetro de dispersão. Função de ligação. Teste da razão de verossimilhanças. Verossimilhança perfilada. Correção de Bartlett.

ABSTRACT

In this thesis, we deal with improvements for hypothesis tests in the overdispersed generalized nonlinear regression models and the beta prime linear regression model. Two lines of research are addressed. The first, referring to Chapters 1 and 2, deals with obtaining the adjustment proposed by Skovgaard for the likelihood ratio statistic, as well as the adjustment for the modified profile likelihood function proposed by Cox and Reid in the overdispersed generalized nonlinear regression models. The second line of research, referring to Chapter 3, deals with obtaining a Bartlett correction for the likelihood ratio test in the beta prime regression model. Additionally, we considered the bootstrap likelihood ratio test. The performances of hypothesis tests based on the proposed refinements were numerically evaluated and compared to their usual counterparts through Monte Carlo simulation studies. Finally, the usefulness of the refinements was illustrated through applications to real data sets.

Keywords: Regression models. Dispersion parameter. Link function. Likelihood ratio test. Profile likelihood. Bartlett correction.

LISTA DE FIGURAS

Figura 1 – Gráficos quantil-quantil das estatísticas da razão de verossimilhanças (ω) , razão de verossimilhanças corrigida (ω_b) , razão de verossimilhanças ajustada (ω_s) e razão de verossimilhanças corrigida bootstrap (ω_r) , $\mathcal{H}_0 : \beta_5 = 0$	26
Figura 2 – Gráfico half-normal residual com envelopes simulados para o modelo Poisson duplo, dados de progênie de <i>Sitophilus zeamais</i>	32
Figura 3 – Gráficos quantil-quantil das estatísticas da razão de verossimilhanças (ω) e razão de verossimilhanças corrigida (ω_{b3}) , $\mathcal{H}_0 : \beta_2 = 0$	52
Figura 4 – Gráfico half-normal residual com envelopes simulados para o Modelo 3, dados de gordura corporal.	57

LISTA DE TABELAS

Tabela 1 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = 0$	25
Tabela 2 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_4 = \beta_5 = 0$	26
Tabela 3 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \gamma_5 = 0$	27
Tabela 4 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = \gamma_4 = 0$	27
Tabela 5 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_3 = 0$, $n = 40$ e diferentes valores para p	28
Tabela 6 – Taxas de rejeição não nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = 0$ e $n = 40$	28
Tabela 7 – Estimativas dos parâmetros, dados de incidentes com danos em navios.	30
Tabela 8 – Estimativas dos parâmetros e valores de AIC e BIC, dados de progênie de <i>Sitophilus zeamais</i>	32

Tabela 9 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $q = 2$, $n = 40$ e diferentes valores para p	41
Tabela 10 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 4$, $n = 40$ e diferentes valores para q	41
Tabela 11 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 3$ e $q = 2$	42
Tabela 12 – Taxas de rejeição não nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 3$, $q = 2$ e $n = 30$	42
Tabela 13 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = 0$	51
Tabela 14 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω), dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}) e do teste da razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$	53
Tabela 15 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \gamma_4 = 0$	54
Tabela 16 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = \beta_3 = \gamma_4 = 0$	54

Tabela 17 – Taxas de rejeição não nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = 0$ e $n = 20$	55
Tabela 18 – Estimativas dos parâmetros e valores de AIC e BIC para os Modelos 3 (beta prime) e 3b (gama), dados de gordura corporal.	57
Tabela 19 – Número de incidentes com danos (y) e total de meses em serviço por tipo de navio (A-E), ano de construção (1:1960-64, 2:1965-69, 3:1970-74 e 4:1975-79) e período de operação (1:1960-74 e 2:1975-79).	78
Tabela 20 – Número de insetos emergidos (y) em placas de Petri contendo 10g de milho tratadas com extratos produzidos a partir de diferentes partes de <i>Annona mucosa</i> (folhas, ramos e sementes) ou apenas água (controle).	79
Tabela 21 – Massa gorda corporal (y) em kg estimada por DXA, circunferência do quadril (x_2), largura do joelho (x_4), soma (x_3) do log da dobra cutânea do queixo, do tríceps e do subescapular, e soma do logaritmo de três medidas antropométricas (x_5), para 71 mulheres alemãs saudáveis.	80

SUMÁRIO

1	AJUSTE DE SKOVGAARD PARA O TESTE DA RAZÃO DE VEROSSIMILHANÇAS EM MODELOS DE REGRESSÃO NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS	14
1.1	INTRODUÇÃO	14
1.2	OS MODELOS NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS	16
1.3	AJUSTE DE SKOVGAARD	19
1.4	RESULTADOS NUMÉRICOS	23
1.5	APLICAÇÕES	29
1.5.1	Primeira Aplicação	29
1.5.2	Segunda Aplicação	31
1.6	CONCLUSÃO	32
2	TESTE DA RAZÃO DE VEROSSIMILHANÇAS PERFILADAS MODIFICADAS CORRIGIDO EM MODELOS DE REGRESSÃO NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS	34
2.1	INTRODUÇÃO	34
2.2	TESTE DA RAZÃO DE VEROSSIMILHANÇAS PERFILADAS MODIFICADAS	35
2.3	FATOR DE CORREÇÃO DE BARTLETT	36
2.4	RESULTADOS NUMÉRICOS	39
2.5	APLICAÇÃO	42
2.6	CONCLUSÃO	43
3	TESTE DA RAZÃO DE VEROSSIMILHANÇAS APRIMORADO NO MODELO DE REGRESSÃO BETA PRIME COM PRECISÃO VARIÁVEL	44
3.1	INTRODUÇÃO	44
3.2	O MODELO DE REGRESSÃO BETA PRIME	45
3.3	TESTE DA RAZÃO DE VEROSSIMILHANÇAS CORRIGIDO	47

3.4	RESULTADOS NUMÉRICOS	49
3.5	APLICAÇÃO	55
3.6	CONCLUSÃO	57
	REFERÊNCIAS	59
	APÊNDICE A – DERIVADAS DA FUNÇÃO $\psi(\mu, \phi)$	63
	APÊNDICE B – OBTENÇÃO DAS QUANTIDADES PRO- POSTAS POR SKOVGAARD	65
	APÊNDICE C – OBTENÇÃO DA CORREÇÃO DE BAR- TLETT PARA O TESTE DE VEROSSIMI- LHANÇA PERFILADA MODIFICADA	68
	APÊNDICE D – CUMULANTES NO MODELO DE REGRES- SÃO BETA PRIME	73
	APÊNDICE E – CONJUNTOS DE DADOS	78

1 AJUSTE DE SKOVGAARD PARA O TESTE DA RAZÃO DE VEROSSIMILHANÇAS EM MODELOS DE REGRESSÃO NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS

1.1 INTRODUÇÃO

Quando a variância observada nos dados é maior do que a variância nominal predita pelo modelo linear generalizado padrão (MCCULLAGH; NELDER, 1989), dizemos que os dados apresentam superdispersão. Esse fenômeno da superdispersão ocorre principalmente em dados de contagem e proporções, tipicamente analisados usando os modelos de Poisson e binomial, respectivamente. Não considerar a superdispersão pode resultar em uma subestimação de erros-padrão e inferência enganosa para os parâmetros de regressão, ver Hinde e Demétrio (1998). Consequentemente, vários modelos foram propostos para lidar com dados desse tipo. Dey et al. (1997) introduziram a classe de modelos lineares generalizados superdispersados (MLGS). Essa classe assume uma estrutura de regressão adicional para um parâmetro de escala que é incorporado na função de variância, tornando mais flexível a modelagem da variância e, portanto, da superdispersão. Além disso, abrange modelos importantes, como a família exponencial dupla de Efron (1986).

No que segue, apresentamos alguns estudos envolvendo os modelos lineares generalizados superdispersados. Cordeiro e Botter (2001) obtiveram estimadores de máxima verossimilhança melhorados. Cordeiro et al. (2006) derivaram um fator de correção de Bartlett para o teste da razão de verossimilhanças considerando o teste conjunto de parâmetros da média e da dispersão. Andrade (2013) obteve correções semelhantes para um teste baseado em verossimilhança perfilada modificada. Rodrigues (2013) desenvolveu técnicas de diagnóstico. Campos (2014) derivou uma correção tipo-Bartlett para testar efeitos da dispersão. Santos (2019) derivou ajustes semelhantes, porém, para testar efeitos da média e da dispersão simultaneamente, generalizando os resultados de Campos. Cordeiro et al. (2008) introduziram os modelos não lineares generalizados superdispersados (MNLGS), uma extensão dos MLGS, e obtiveram estimadores de máxima verossimilhança melhorados para o modelo proposto. Terra (2013) derivou testes corrigidos e também propôs técnicas de diagnóstico para MNLGS.

O teste da razão de verossimilhanças é comumente usado para realizar inferências sobre os parâmetros dos MNLGS. No entanto, esse teste utiliza uma aproximação qui-

quadrado para a distribuição nula¹ desconhecida da estatística de teste; o erro dessa aproximação é $O(n^{-1})^2$, em que n é o tamanho da amostra. Como uma consequência, é provável que o teste apresente distorções de tamanho em pequenas amostras. Skovgaard (2001) propôs um ajuste para a estatística da razão de verossimilhanças que visa melhorar seu desempenho em amostras de tamanho pequeno. A estatística ajustada tem distribuição qui-quadrado com alto grau de acurácia, ver Skovgaard (2001). Para obter tal ajuste são necessárias, além das matrizes de informação observada e esperada, duas quantidades de complexidade análoga à informação de Fisher. Isso tem atraído vários pesquisadores que se concentraram em obter expressões para o ajuste de Skovgaard em problemas específicos. Ferrari e Cysneiros (2008) derivaram uma expressão geral para a correção de Skovgaard nos modelos não lineares da família exponencial. Correções semelhantes foram obtidas por Ferrari e Pinheiro (2011) em regressão beta. Melo et al. (2014) desenvolveram testes da razão de verossimilhanças ajustados de Skovgaard nos modelos de regressão multivariados heteroscedásticos com erros nas variáveis. Rauber et al. (2020) obtiveram a correção de Skovgaard no modelo de regressão beta com função de ligação média paramétrica. Guedes et al. (2020) propuseram estatísticas da razão de verossimilhanças ajustadas no modelo de regressão gama unitária. Melo et al. (2022) obtiveram refinamentos para a estatística da razão de verossimilhanças nos modelos de regressão multivariados Dirichlet.

Neste capítulo, nosso objetivo principal é derivar o ajuste de Skovgaard para o teste da razão de verossimilhanças na classe dos modelos não lineares generalizados superdispersados. O capítulo está organizado da seguinte forma. Na Seção 1.2, apresentamos os MNLGS. Na Seção 1.3, derivamos uma expressão do ajuste de Skovgaard para a estatística da razão de verossimilhanças nos MNLGS. Estes ajustes são bastante gerais, pois são válidos para testar qualquer hipótese de interesse. Também apresentamos um teste corrigido alternativo e sua versão bootstrap. Os resultados de simulação de Monte Carlo comparando o desempenho de amostra finita dos quatro testes, a saber, teste da razão de verossimilhanças usual, teste corrigido de Skovgaard, e os dois testes alternativos, são fornecidos na Seção 1.4. Na Seção 1.5, apresentamos e discutimos duas aplicações empíricas. Finalmente, algumas considerações finais são dadas na Seção 1.6.

¹ A distribuição nula é a distribuição da estatística de teste obtida sob a hipótese nula.

² Sejam $\{a_n\}$ e $\{b_n\}$ sequências de números reais. Dizemos que a_n tem ordem no máximo b_n e escrevemos $a_n = O(b_n)$, se a razão $|a_n/b_n|$ é limitada para todo n suficientemente grande, isto é, se existirem $k > 0$ e n_0 tal que $|a_n/b_n| \leq k$, $\forall n \geq n_0$.

1.2 OS MODELOS NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS

Seja Y uma variável aleatória tendo uma função densidade na família exponencial de dois parâmetros da forma

$$\pi(y; \mu, \phi) = A(y) \exp \left[(y - \mu) \psi^{(1,0)}(\mu, \phi) + \phi T(y) + \psi(\mu, \phi) \right], \quad (1.1)$$

em que $A(\cdot)$, $T(\cdot)$, $\psi(\cdot, \cdot)$ são funções conhecidas e $\phi > 0$ é um parâmetro de dispersão. Em (1.1), seja $\psi^{(r,s)} = \partial^{r+s} \psi(\mu, \phi) / \partial \mu^r \partial \phi^s$ ($r, s \geq 0$). Aqui, $E(Y) = \mu$ e $\text{Var}(Y) = \psi^{(2,0)} \phi^{-1}$. Temos também que, $E[T(Y)] = -\psi^{(0,1)}$, $\text{Var}[T(Y)] = \psi^{(1,1)^2} \psi^{(2,0)} \phi^{-1} - \psi^{(0,2)}$ e $E[(Y - \mu)T(Y)] = -\psi^{(1,1)} \psi^{(2,0)} \phi^{-1}$.

Um caso especial da família (1.1) é a distribuição passeio aleatório (WISE, 1966; WASAN, 1968). Esta é a distribuição do recíproco de uma variável aleatória tendo uma distribuição Gaussiana inversa (JOHNSON et al., 1994). Se Y tem distribuição passeio aleatório $\mathcal{PA}(\theta, \delta)$, sua função densidade é dada por

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

O valor esperado e a variância de Y são $\mu = 1/\theta + 1/\delta$ e $1/(\theta\delta) + 2/\delta^2$, respectivamente. Definindo $\phi = -\delta/(2\theta^2)$ e escrevendo a densidade acima na forma da Equação (1.1), através de uma parametrização em termos da média, Cordeiro et al. (2006) mostraram que

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

Outro caso especial é a distribuição Poisson dupla proposta por Efron (1986). A distribuição Poisson dupla, denotada por $Y \sim \mathcal{PD}(\mu, \phi)$, tem função de probabilidade dada por

$$\pi(y; \mu, \phi) = c(\mu, \phi) \phi^{1/2} e^{-\phi\mu} \left(\frac{e^{-y} y^y}{y!} \right) \left(\frac{e\mu}{y} \right)^{\phi y}, \quad y = 0, 1, 2, \dots, \quad (1.2)$$

em que $\mu, \phi > 0$ e $c(\mu, \phi)$ é uma constante de normalização que garante que (1.2) soma 1. Efron (1986) mostrou que essa constante é quase igual a 1. Essa distribuição tem média aproximadamente μ e variância aproximadamente μ/ϕ . O modelo Poisson duplo permite superdispersão ($0 < \phi < 1$), bem como subdispersão ($\phi > 1$) e reduz-se ao modelo de

Poisson se $\phi = 1$. Aqui,

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

Um outro modelo de superdispersão que também pertence à família (1.1) é a distribuição binomial dupla, ver Efron (1986). Suponha que Y representa a proporção de sucessos em m tentativas. Então, Y segue distribuição binomial dupla (escalonada), isto é, $Y \sim \mathcal{BD}(\mu, \phi)/m$, se sua função de probabilidade é dada por

$$\pi(y; \mu, \phi) = c(\mu, \phi)\phi^{1/2} \binom{m}{my} \frac{\mu^{my\phi}(1-\mu)^{m(1-y)\phi}}{y^{-my(1-\phi)}(1-y)^{-m(1-y)(1-\phi)}}, \quad 0 < y < 1,$$

sendo $0 < \mu < 1$ e $\phi > 0$. Novamente, $c(\mu, \phi)$ é uma constante de normalização aproximadamente igual a 1. A média e variância dessa distribuição são aproximadamente μ e $\mu(1-\mu)/(m\phi)$, respectivamente. A distribuição será superdispersada se $0 < \phi < 1$. Quando $\phi = 1$, obtemos o modelo binomial. Adicionalmente,

$$\psi(\mu, \phi) = m\phi[\mu \log \mu + (1-\mu) \log(1-\mu)] + \frac{\log \phi}{2}.$$

Sejam $Y_1, \dots, Y_n$ variáveis aleatórias independentes, cada Y_l tendo uma densidade na família (1.1) com parâmetros μ_l (média) e ϕ_l (dispersão), $l = 1, \dots, n$. Os modelos de regressão não lineares generalizados superdispersados são definidos como

$$g(\mu_l) = \eta_l = f_1(\mathbf{x}_l; \boldsymbol{\beta}) \quad \text{e} \quad h(\phi_l) = \tau_l = f_2(\mathbf{s}_l; \boldsymbol{\gamma}),$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ e $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^\top$ são vetores de parâmetros desconhecidos a serem estimados ($p + q < n$), $\mathbf{x}_l = (x_{l1}, \dots, x_{lb})^\top$ ($b \leq p$) e $\mathbf{s}_l = (s_{l1}, \dots, s_{lc})^\top$ ($c \leq q$) são os respectivos vetores de covariáveis da média e da dispersão. Além disso, g e h são funções de ligação estritamente crescentes e duas vezes diferenciáveis. As funções f_1 e f_2 são continuamente diferenciáveis e possivelmente não lineares, de modo que as matrizes de derivadas $\mathbf{X}^* = \mathbf{X}^*(\boldsymbol{\beta}) = \partial \boldsymbol{\eta} / \partial \boldsymbol{\beta}^\top$ e $\mathbf{S}^* = \mathbf{S}^*(\boldsymbol{\gamma}) = \partial \boldsymbol{\tau} / \partial \boldsymbol{\gamma}^\top$, com $\boldsymbol{\eta} = (\eta_1, \dots, \eta_n)^\top$ e $\boldsymbol{\tau} = (\tau_1, \dots, \tau_n)^\top$, têm postos completos para todo $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, respectivamente.

O logaritmo da função de verossimilhança (log-verossimilhança) para um dado MNLGS é

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{l=1}^n \left[(y_l - \mu_l) \psi^{(1,0)}(\mu_l, \phi_l) + \phi_l T(y_l) + \psi(\mu_l, \phi_l) + \log A(y_l) \right]. \quad (1.3)$$

Agora, vamos introduzir algumas notações. Denotamos por $m_{il} = d^i \mu_l / d\eta_l^i$ e $\phi_{il} = d^i \phi_l / d\tau_l^i$ as i -ésimas derivadas das funções de ligação inversas $\mu = g^{-1}(\eta)$ e $\phi = h^{-1}(\tau)$,

para $i = 1, 2$ e $l = 1, \dots, n$. Definimos as seguintes matrizes diagonais de dimensão $n \times n$:

$$\mathbf{M}_i = \text{diag}\{m_{i1}, \dots, m_{in}\}, \quad \mathbf{\Phi}_i = \text{diag}\{\phi_{i1}, \dots, \phi_{in}\} \text{ e } \mathbf{\Psi}^{(r,s)} = \text{diag}\{\psi_1^{(r,s)}, \dots, \psi_n^{(r,s)}\}.$$

A função escore é dada por $\mathbf{U} = \mathbf{U}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (\mathbf{U}_\beta^\top, \mathbf{U}_\gamma^\top)^\top$, em que

$$\mathbf{U}_\beta = \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} = \mathbf{X}^{*\top} \mathbf{\Psi}^{(2,0)} \mathbf{M}_1 (\mathbf{y} - \boldsymbol{\mu}) \quad \text{e} \quad \mathbf{U}_\gamma = \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = \mathbf{S}^{*\top} \mathbf{\Phi}_1 \mathbf{v},$$

sendo $(\mathbf{y} - \boldsymbol{\mu}) = (y_1 - \mu_1, \dots, y_n - \mu_n)^\top$ e $\mathbf{v} = (v_1, \dots, v_n)^\top$, com $v_l = \psi_l^{(1,1)}(y_l - \mu_l) + T(y_l) + \psi_l^{(0,1)}$. A matriz de informação observada, isto é, o negativo da matriz de segundas derivadas da log-verossimilhança, é dada por

$$\mathbf{J} = \mathbf{J}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{bmatrix} \mathbf{J}_{\beta\beta} & \mathbf{J}_{\beta\gamma} \\ \mathbf{J}_{\gamma\beta} & \mathbf{J}_{\gamma\gamma} \end{bmatrix}, \quad (1.4)$$

em que

$$\begin{aligned} \mathbf{J}_{\beta\beta} &= \mathbf{X}^{*\top} \left[\mathbf{\Psi}^{(2,0)} (\mathbf{M}_1^2 - \mathbf{M}_2 \mathbf{N}) - \mathbf{\Psi}^{(3,0)} \mathbf{M}_1^2 \mathbf{N} \right] \mathbf{X}^* - [(\mathbf{y} - \boldsymbol{\mu})^\top \mathbf{\Psi}^{(2,0)} \mathbf{M}_1] [\mathbf{X}^{**}], \\ \mathbf{J}_{\gamma\gamma} &= -\mathbf{S}^{*\top} \left[(\mathbf{\Psi}^{(1,2)} \mathbf{N} + \mathbf{\Psi}^{(0,2)}) \mathbf{\Phi}_1^2 + \mathbf{V} \mathbf{\Phi}_2 \right] \mathbf{S}^* - [\mathbf{v}^\top \mathbf{\Phi}_1] [\mathbf{S}^{**}], \\ \mathbf{J}_{\beta\gamma} &= \mathbf{J}_{\gamma\beta}^\top = -\mathbf{X}^{*\top} \mathbf{\Psi}^{(2,1)} \mathbf{M}_1 \mathbf{\Phi}_1 \mathbf{N} \mathbf{S}^*, \end{aligned}$$

com $\mathbf{N} = \text{diag}\{y_1 - \mu_1, \dots, y_n - \mu_n\}$ e $\mathbf{V} = \text{diag}\{v_1, \dots, v_n\}$. Além disso, $\mathbf{X}^{**} = \partial^2 \boldsymbol{\eta} / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top$ e $\mathbf{S}^{**} = \partial^2 \boldsymbol{\tau} / \partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^\top$ são arrays $n \times p \times p$ e $n \times q \times q$, respectivamente. Aqui, o símbolo $[\cdot][\cdot]$ denota o produto colchete de uma matriz e um array, conforme definido em Wei (1998, p. 188). Por exemplo, sejam $\mathbf{A}$ uma matriz $m \times n$ e $\mathbf{B}$ um array $n \times p \times q$, então $\mathbf{C} = [\mathbf{A}][\mathbf{B}]$ é um array $m \times p \times q$ com elementos $C_{tij} = \sum_{l=1}^n A_{tl} B_{lij}$, para $t = 1, \dots, m$, $i = 1, \dots, p$ e $j = 1, \dots, q$.

A matriz de informação esperada de Fisher $\mathbf{K} = \mathbf{K}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \text{diag}\{\mathbf{K}_{\beta\beta}, \mathbf{K}_{\gamma\gamma}\}$ é bloco-diagonal, em que $\mathbf{K}_{\beta\beta} = \mathbf{X}^{*\top} \mathbf{\Psi}^{(2,0)} \mathbf{M}_1^2 \mathbf{X}^*$ e $\mathbf{K}_{\gamma\gamma} = -\mathbf{S}^{*\top} \mathbf{\Psi}^{(0,2)} \mathbf{\Phi}_1^2 \mathbf{S}^*$ são as matrizes de informação esperada para $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, respectivamente. Consequentemente, os parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são ortogonais (COX; REID, 1987) e suas estimativas de máxima verossimilhança (E.M.V.) $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$ são assintoticamente independentes. Essas estimativas satisfazem as equações não lineares $\mathbf{U}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}) = 0$. É possível usar o método escore de Fisher para estimar $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ iterativamente solucionando as seguintes equações:

$$\begin{aligned} \mathbf{X}^{*(m)\top} \mathbf{\Psi}^{(2,0)(m)} \mathbf{M}_1^{(m)2} \mathbf{X}^{*(m)} (\boldsymbol{\beta}^{(m+1)} - \boldsymbol{\beta}^{(m)}) &= \mathbf{X}^{*(m)\top} \mathbf{\Psi}^{(2,0)(m)} \mathbf{M}_1^{(m)2} \boldsymbol{\xi}_1^{(m)} \\ \left(-\mathbf{S}^{*(m)\top} \mathbf{\Psi}^{(0,2)(m)} \mathbf{\Phi}_1^{(m)2} \mathbf{S}^{*(m)} \right) (\boldsymbol{\gamma}^{(m+1)} - \boldsymbol{\gamma}^{(m)}) &= \mathbf{S}^{*(m)\top} \left(-\mathbf{\Psi}^{(0,2)(m)} \mathbf{\Phi}_1^{(m)2} \right) \boldsymbol{\xi}_2^{(m)}, \end{aligned}$$

em que $\xi_1 = \mathbf{M}_1^{-1}(\mathbf{y} - \boldsymbol{\mu})$ e $\xi_2 = -(\Psi^{(0,2)}\Phi_1)^{-1}\mathbf{v}$ são vetores $n \times 1$. Ou seja, podemos obter $\hat{\beta}$ e $\hat{\gamma}$ as E.M.V. de β e γ , respectivamente, por mínimos quadrados ponderados iterativamente.

1.3 AJUSTE DE SKOVGAARD

Seja $\boldsymbol{\theta} = (\beta^\top, \gamma^\top)^\top$ o vetor de parâmetros k -dimensional para um dado MNLGS. Particionando-o como $\boldsymbol{\theta} = (\boldsymbol{\nu}^\top, \boldsymbol{\lambda}^\top)^\top$, em que $\boldsymbol{\nu} = (\nu_1, \dots, \nu_r)^\top$ e $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_t)^\top$ são, respectivamente, os parâmetros de interesse e de perturbação, desejamos testar a hipótese nula $\mathcal{H}_0 : \boldsymbol{\nu} = \boldsymbol{\nu}^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1 : \boldsymbol{\nu} \neq \boldsymbol{\nu}^{(0)}$, em que $\boldsymbol{\nu}^{(0)}$ é um vetor especificado r -dimensional. A estatística da razão de verossimilhanças é

$$\omega = 2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}})\},$$

em que $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\nu}}^\top, \hat{\boldsymbol{\lambda}}^\top)^\top$ e $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\nu}^{(0)\top}, \tilde{\boldsymbol{\lambda}}^\top)^\top$ são as E.M.V. irrestrita e restrita de $\boldsymbol{\theta}$, respectivamente. Sob a hipótese nula $\mathcal{H}_0$, ω é assintoticamente distribuída como χ_r^2 , em que r é o número de restrições impostas sob $\mathcal{H}_0$. Uma vez que essa aproximação pela distribuição qui-quadrado se dá em grandes amostras, é provável que ocorram distorções de tamanho em pequenas amostras. Skovgaard (2001) propôs, para pequenas amostras, a seguinte estatística da razão de verossimilhanças ajustada:

$$\omega_s = \omega - 2 \log \xi,$$

em que a expressão geral para a quantidade ξ é dada por

$$\xi = \frac{|\tilde{\mathbf{K}}|^{1/2} |\hat{\mathbf{K}}|^{1/2} |\bar{\mathbf{\Upsilon}}|^{-1} |\tilde{\mathbf{J}}_{\lambda\lambda}|^{1/2}}{|\tilde{\mathbf{K}} \bar{\mathbf{\Upsilon}}^{-1} \hat{\mathbf{J}} \hat{\mathbf{K}}^{-1} \bar{\mathbf{\Upsilon}}|_{\lambda\lambda}^{1/2}} \frac{\{\tilde{\mathbf{U}}^\top \bar{\mathbf{\Upsilon}}^{-1} \hat{\mathbf{K}} \hat{\mathbf{J}}^{-1} \bar{\mathbf{\Upsilon}} \tilde{\mathbf{K}}^{-1} \tilde{\mathbf{U}}\}^{r/2}}{\omega^{r/2-1} \tilde{\mathbf{U}}^\top \bar{\mathbf{\Upsilon}}^{-1} \bar{\boldsymbol{\rho}}}, \quad (1.5)$$

com $\bar{\boldsymbol{\rho}}$ e $\bar{\mathbf{\Upsilon}}$ obtidos a partir de

$$\boldsymbol{\rho} = \mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}(\boldsymbol{\theta}_0)(\ell(\boldsymbol{\theta}_0) - \ell(\boldsymbol{\theta}))] \quad \text{e} \quad \mathbf{\Upsilon} = \mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}(\boldsymbol{\theta}_0)\mathbf{U}^\top(\boldsymbol{\theta})],$$

substituindo $\boldsymbol{\theta}_0$ por $\hat{\boldsymbol{\theta}}$ e $\boldsymbol{\theta}$ por $\tilde{\boldsymbol{\theta}}$ depois que os valores esperados são calculados. Note que $\boldsymbol{\rho}$ é um vetor k -dimensional e $\mathbf{\Upsilon}$ é uma matriz $k \times k$. Em (1.5), a notação $|\mathbf{A}|$ denota o determinante da matriz $\mathbf{A}$, enquanto $[\mathbf{A}]_{\lambda\lambda}$ denota a submatriz obtida de $\mathbf{A}$ correspondente a $\boldsymbol{\lambda}$. Além disso, $\mathbf{J}_{\lambda\lambda}$ representa a matriz de informação observada para o parâmetro de perturbação $\boldsymbol{\lambda}$. Adicionalmente, usamos abreviações: $\tilde{\mathbf{U}} = \mathbf{U}(\tilde{\boldsymbol{\theta}})$, $\hat{\mathbf{J}} = \mathbf{J}(\hat{\boldsymbol{\theta}})$, $\tilde{\mathbf{J}}_{\lambda\lambda} = \mathbf{J}_{\lambda\lambda}(\tilde{\boldsymbol{\theta}})$, $\hat{\mathbf{K}} = \mathbf{K}(\hat{\boldsymbol{\theta}})$ e $\tilde{\mathbf{K}} = \mathbf{K}(\tilde{\boldsymbol{\theta}})$.

A estatística ajustada ω_s é aproximadamente distribuída como χ_r^2 com alto grau de precisão, ver Skovgaard (2001). Essa estatística corrigida é uma generalização da estatística R^* de Barndorff-Nielsen (1986, 1991). R^* é restrita ao teste de uma hipótese unidimensional, enquanto ω_s pode ser utilizada para testar uma hipótese multidimensional.

Nosso objetivo aqui é derivar expressões gerais para as quantidades $\bar{\rho}$ e $\bar{\Upsilon}$ nos MNLGS. E assim obter um teste da razão de verossimilhanças ajustado para essa classe de modelos de superdispersão. Então, a partir da log-verossimilhança e da função escore definidas na Seção 1.2, obtivemos, após longa álgebra (ver Apêndice B), que $\bar{\rho}$ e $\bar{\Upsilon}$ têm a forma

$$\bar{\rho} = \begin{bmatrix} \hat{\mathbf{X}}^{*\top} \hat{\mathbf{M}}_1 \left\{ \hat{\Psi}^{(1,0)} - \tilde{\Psi}^{(1,0)} - \hat{\Psi}^{(1,1)}(\hat{\Phi} - \tilde{\Phi}) \right\} \mathbf{1} \\ \hat{\mathbf{S}}^{*\top} \hat{\Phi}_1 \hat{\Psi}^{(0,2)}(\tilde{\Phi} - \hat{\Phi}) \mathbf{1} \end{bmatrix} \quad (1.6)$$

e

$$\bar{\Upsilon} = \begin{bmatrix} \hat{\mathbf{X}}^{*\top} \hat{\mathbf{M}}_1 \tilde{\Psi}^{(2,0)} \tilde{\mathbf{M}}_1 \tilde{\mathbf{X}}^* & \hat{\mathbf{X}}^{*\top} \hat{\mathbf{M}}_1 \left\{ \tilde{\Psi}^{(1,1)} - \hat{\Psi}^{(1,1)} \right\} \tilde{\Phi}_1 \tilde{\mathbf{S}}^* \\ \mathbf{0} & -\hat{\mathbf{S}}^{*\top} \hat{\Phi}_1 \hat{\Psi}^{(0,2)} \tilde{\Phi}_1 \tilde{\mathbf{S}}^* \end{bmatrix}, \quad (1.7)$$

em que $\mathbf{1}$ denota um vetor n -dimensional de uns e $\Phi = \text{diag}\{\phi_1, \dots, \phi_n\}$. Adicionalmente, as matrizes com circunflexo são avaliadas em $\hat{\theta}$, enquanto que aquelas com til são avaliadas em $\tilde{\theta}$. É importante ressaltar que a estatística corrigida ω_s também pode ser utilizada com os MLGS; nesse caso, basta notar que $\eta_l = \mathbf{x}_l^\top \beta$ e $\tau_l = \mathbf{s}_l^\top \gamma$, $l = 1, \dots, n$, e substituir $\mathbf{X}^*$ por $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$ e $\mathbf{S}^*$ por $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_n)^\top$ nas quantidades em (1.5). Além disso, $\mathbf{X}^{**}$ e $\mathbf{S}^{**}$ são arrays de zeros, de modo que há uma simplificação da matriz de informação observada (1.4).

Agora, discutimos o teste da razão de verossimilhanças corrigido via Bartlett (1937) para a classe dos MNLGS obtido por Terra (2013), bem como sua versão bootstrap; ver Rocke (1989).

Inicialmente, considere uma partição dos parâmetros como $\beta = (\beta_1^\top, \beta_2^\top)^\top$ e $\gamma = (\gamma_1^\top, \gamma_2^\top)^\top$, em que $\beta_1 = (\beta_1, \dots, \beta_{p_1})^\top$ ($p_1 \leq p$), $\beta_2 = (\beta_{p_1+1}, \dots, \beta_p)^\top$, $\gamma_1 = (\gamma_1, \dots, \gamma_{q_1})^\top$ ($q_1 \leq q$) e $\gamma_2 = (\gamma_{q_1+1}, \dots, \gamma_q)^\top$. O interesse consiste em testar a hipótese nula $\mathcal{H}_0$: $\beta_1 = \beta_1^{(0)}$, $\gamma_1 = \gamma_1^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1$: violação de pelo menos uma igualdade, $\beta_1^{(0)}$ e $\gamma_1^{(0)}$ sendo vetores especificados de dimensões p_1 e q_1 , respectivamente. A partição dos parâmetros β e γ induz partições correspondentes das matrizes modelo como $\mathbf{X}^* = [\mathbf{X}_1^* \mathbf{X}_2^*]$ e $\mathbf{S}^* = [\mathbf{S}_1^* \mathbf{S}_2^*]$. A estatística da razão de verossimilhanças é $\omega = 2\{\ell(\hat{\beta}, \hat{\gamma}) - \ell(\tilde{\beta}, \tilde{\gamma})\}$, em que $\tilde{\beta} = (\beta_1^{(0)\top}, \tilde{\beta}_2^\top)^\top$ e $\tilde{\gamma} = (\gamma_1^{(0)\top}, \tilde{\gamma}_2^\top)^\top$ são as estimativas de

máxima verossimilhança restritas de β e γ , respectivamente. A estatística corrigida é

$$\omega_b = \frac{\omega}{c_1},$$

em que $c_1 = 1 + (\epsilon_{p+q} - \epsilon_{p-p_1+q-q_1})/(p_1 + q_1)$ é o fator de correção de Bartlett. Em grandes amostras, a distribuição nula de ω_b é aproximadamente $\chi^2_{p_1+q_1}$ com erro de ordem $O(n^{-2})$. Os termos ϵ_{p+q} e $\epsilon_{p-p_1+q-q_1}$ dependem de cumulantes de derivadas da log-verossimilhança até quarta ordem. As expressões gerais para esses termos são dadas em Lawley (1956) e Cordeiro e Cribari-Neto (2014). Antes de apresentar as expressões para os ϵ 's nos MNLGS obtidas por Terra (2013), é necessário definir algumas matrizes. Temos

$$\begin{aligned} \mathbf{Z}_\beta &= \mathbf{X}^*(\mathbf{X}^{*\top} \Psi^{(2,0)} \mathbf{M}_1^2 \mathbf{X}^*)^{-1} \mathbf{X}^{*\top}, & \mathbf{Z}_{\beta_2} &= \mathbf{X}_2^*(\mathbf{X}_2^{*\top} \Psi^{(2,0)} \mathbf{M}_1^2 \mathbf{X}_2^*)^{-1} \mathbf{X}_2^{*\top}, \\ \mathbf{Z}_\gamma &= \mathbf{S}^*(-\mathbf{S}^{*\top} \Psi^{(0,2)} \Phi_1^2 \mathbf{S}^*)^{-1} \mathbf{S}^{*\top}, & \mathbf{Z}_{\gamma_2} &= \mathbf{S}_2^*(-\mathbf{S}_2^{*\top} \Psi^{(0,2)} \Phi_1^2 \mathbf{S}_2^*)^{-1} \mathbf{S}_2^{*\top}. \end{aligned}$$

Adicionalmente,

$$\begin{aligned} \mathbf{D}_\beta &= \text{diag}\{d_{\beta 1}, \dots, d_{\beta n}\}, & \mathbf{D}_{\beta_2} &= \text{diag}\{d_{\beta_2 1}, \dots, d_{\beta_2 n}\}, \\ \mathbf{D}_\gamma &= \text{diag}\{d_{\gamma 1}, \dots, d_{\gamma n}\}, & \mathbf{D}_{\gamma_2} &= \text{diag}\{d_{\gamma_2 1}, \dots, d_{\gamma_2 n}\}, \end{aligned}$$

em que

$$\begin{aligned} d_{\beta l} &= \text{tr}\{(\mathbf{X}^{*\top} \Psi^{(2,0)} \mathbf{M}_1^2 \mathbf{X}^*)^{-1} \mathbf{X}_l^{**}\}, & d_{\beta_2 l} &= \text{tr}\{(\mathbf{X}_2^{*\top} \Psi^{(2,0)} \mathbf{M}_1^2 \mathbf{X}_2^*)^{-1} \mathbf{X}_{2l}^{**}\}, \\ d_{\gamma l} &= \text{tr}\{(-\mathbf{S}^{*\top} \Psi^{(0,2)} \Phi_1^2 \mathbf{S}^*)^{-1} \mathbf{S}_l^{**}\}, & d_{\gamma_2 l} &= \text{tr}\{(-\mathbf{S}_2^{*\top} \Psi^{(0,2)} \Phi_1^2 \mathbf{S}_2^*)^{-1} \mathbf{S}_{2l}^{**}\}, \end{aligned}$$

com $\mathbf{X}_l^{**} = \partial^2 \eta_l / \partial \beta \partial \beta^\top$, $\mathbf{X}_{2l}^{**} = \partial^2 \eta_l / \partial \beta_2 \partial \beta_2^\top$, $\mathbf{S}_l^{**} = \partial^2 \tau_l / \partial \gamma \partial \gamma^\top$ e $\mathbf{S}_{2l}^{**} = \partial^2 \tau_l / \partial \gamma_2 \partial \gamma_2^\top$.

Terra (2013) mostrou então que $\epsilon_{p+q} = \epsilon_p(\beta) + \epsilon_q(\gamma) + \epsilon_{p,q}(\beta, \gamma)$, em que

$$\begin{aligned} \epsilon_p(\beta) &= \frac{1}{4} \mathbf{1}^\top (\Psi^{(4,0)} \mathbf{M}_1^4 + 4 \Psi^{(3,0)} \mathbf{M}_1^2 \mathbf{M}_2 + \Psi^{(2,0)} \mathbf{M}_2^2) \mathbf{Z}_{\beta d}^{(2)} \mathbf{1} \\ &\quad - \frac{1}{3} \mathbf{1}^\top \Psi^{(3,0)} \mathbf{M}_1^3 \mathbf{Z}_\beta^{(3)} (\mathbf{M}_1^3 \Psi^{(3,0)} + 3 \mathbf{M}_1 \mathbf{M}_2 \Psi^{(2,0)}) \mathbf{1} \\ &\quad - \frac{1}{4} \mathbf{1}^\top \Psi^{(2,0)} \mathbf{M}_1 \mathbf{M}_2 (2 \mathbf{Z}_\beta^{(3)} - \mathbf{Z}_{\beta d} \mathbf{Z}_\beta \mathbf{Z}_{\beta d}) \mathbf{M}_1 \mathbf{M}_2 \Psi^{(2,0)} \mathbf{1} \\ &\quad - \frac{1}{4} \mathbf{1}^\top (2 \Psi^{(2,0)} \mathbf{M}_1 \mathbf{M}_2 \mathbf{Z}_{\beta d} + \Psi^{(2,0)} \mathbf{M}_1^2 \mathbf{D}_\beta) \\ &\quad \times (\mathbf{Z}_\beta \mathbf{D}_\beta \mathbf{M}_1^2 \Psi^{(2,0)} + \mathbf{D}_\beta) \mathbf{1}, \end{aligned} \tag{1.8}$$

$$\begin{aligned}
\epsilon_q(\gamma) = & \frac{1}{4} \mathbf{1}^\top (\Psi^{(0,4)} \Phi_1^4 + 2\Psi^{(0,3)} \Phi_1^2 \Phi_2 - \Psi^{(0,2)} \Phi_2^2) \mathbf{Z}_{\gamma d}^{(2)} \mathbf{1} \\
& + \frac{1}{12} \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (2\mathbf{Z}_\gamma^{(3)} + 3\mathbf{Z}_{\gamma d} \mathbf{Z}_\gamma \mathbf{Z}_{\gamma d}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} \\
& - \frac{1}{4} \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 (2\mathbf{Z}_\gamma^{(3)} - \mathbf{Z}_{\gamma d} \mathbf{Z}_\gamma \mathbf{Z}_{\gamma d}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} \\
& + \frac{1}{2} \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 (D_\gamma \mathbf{Z}_\gamma \mathbf{Z}_{\gamma d} \Phi_1^2 \Psi^{(0,2)} + \mathbf{Z}_{\gamma d} \mathbf{Z}_\gamma \mathbf{Z}_{\gamma d} \Phi_1^3 \Psi^{(0,3)}) \mathbf{1} \\
& + \frac{1}{4} \mathbf{1}^\top (2\Psi^{(0,3)} \Phi_1^3 \mathbf{Z}_{\gamma d} + \Psi^{(0,2)} \Phi_1^2 D_\gamma) (\mathbf{Z}_\gamma D_\gamma \Phi_1^2 \Psi^{(0,2)} + D_\gamma) \mathbf{1} \\
& + \frac{1}{2} \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 \mathbf{Z}_{\gamma d} D_\gamma \mathbf{1}
\end{aligned} \tag{1.9}$$

e

$$\begin{aligned}
\epsilon_{p,q}(\beta, \gamma) = & \frac{1}{2} \mathbf{1}^\top (\Psi^{(2,2)} \Phi_1^2 + \Psi^{(2,1)} \Phi_2) M_1^2 \mathbf{Z}_{\beta d} \mathbf{Z}_{\gamma d} \mathbf{1} \\
& - \frac{1}{4} \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 (2\mathbf{Z}_\beta^{(2)} \odot \mathbf{Z}_\gamma - \mathbf{Z}_{\beta d} \mathbf{Z}_\gamma \mathbf{Z}_{\beta d}) M_1^2 \Phi_1 \Psi^{(2,1)} \mathbf{1} \\
& + \frac{1}{2} \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 \mathbf{Z}_{\beta d} \mathbf{Z}_\gamma \mathbf{Z}_{\gamma d} (\Phi_1^3 \Psi^{(0,3)} + \Phi_1 \Phi_2 \Psi^{(0,2)}) \mathbf{1} \\
& + \frac{1}{2} \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_1 D_\gamma \mathbf{Z}_\gamma \mathbf{Z}_{\beta d} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} \\
& - \frac{1}{4} \mathbf{1}^\top \Psi^{(2,1)} M_1^2 \Phi_2 \mathbf{Z}_{\beta d} D_\gamma \mathbf{1}.
\end{aligned} \tag{1.10}$$

Como antes, $\mathbf{1}$ denota um vetor n -dimensional de uns. O símbolo $\odot$ representa o produto de Hadamard de duas matrizes (RAO, 1973, p. 30), $Z^{(2)} = Z \odot Z$ e $Z^{(3)} = Z^{(2)} \odot Z$, em que Z é uma matriz dada. Além disso, o subscrito d indica que uma matriz diagonal foi formada a partir da matriz original. Adicionalmente, $\epsilon_{p-p_1+q-q_1} = \epsilon_{p-p_1}(\beta_2) + \epsilon_{q-q_1}(\gamma_2) + \epsilon_{p-p_1,q-q_1}(\beta_2, \gamma_2)$, em que as quantidades $\epsilon_{p-p_1}(\beta_2)$, $\epsilon_{q-q_1}(\gamma_2)$ e $\epsilon_{p-p_1,q-q_1}(\beta_2, \gamma_2)$ seguem de (1.8)–(1.10), com $\mathbf{Z}_\beta$, $\mathbf{Z}_\gamma$, $\mathbf{Z}_{\beta d}$, $\mathbf{Z}_{\gamma d}$, $\mathbf{D}_\beta$ e $\mathbf{D}_\gamma$ substituídos por $\mathbf{Z}_{\beta_2}$, $\mathbf{Z}_{\gamma_2}$, $\mathbf{Z}_{\beta_2 d}$, $\mathbf{Z}_{\gamma_2 d}$, $\mathbf{D}_{\beta_2}$ e $\mathbf{D}_{\gamma_2}$, respectivamente. Os termos ϵ_{p+q} e $\epsilon_{p-p_1+q-q_1}$ devem ser avaliados na estimativa de máxima verossimilhança restrita $(\beta_1^{(0)}, \tilde{\beta}_2, \gamma_1^{(0)}, \tilde{\gamma}_2)$.

Para testar restrições sobre os parâmetros da média, isto é, $\mathcal{H}_0 : \beta_1 = \beta_1^{(0)}$ contra $\mathcal{H}_1 : \beta_1 \neq \beta_1^{(0)}$, o fator de correção de Bartlett é

$$c_2 = 1 + \{\epsilon_p(\beta) - \epsilon_{p-p_1}(\beta_2) + \epsilon_{p,q}(\beta, \gamma) - \epsilon_{p-p_1,q}(\beta_2, \gamma)\} / p_1,$$

em que $\epsilon_{p-p_1,q}(\beta_2, \gamma)$ é obtido de (1.10) com $\mathbf{Z}_\beta$ e $\mathbf{Z}_{\beta d}$ substituídos por $\mathbf{Z}_{\beta_2}$ e $\mathbf{Z}_{\beta_2 d}$, respectivamente. O fator de correção c_2 deve ser avaliado na E.M.V. restrita $(\beta_1^{(0)}, \tilde{\beta}_2, \tilde{\gamma})$.

Finalmente, se o interesse está em testar parâmetros que modelam a dispersão, $\mathcal{H}_0 : \gamma_1 = \gamma_1^{(0)}$ contra $\mathcal{H}_1 : \gamma_1 \neq \gamma_1^{(0)}$, então o fator de correção de Bartlett é dado por

$$c_3 = 1 + \{\epsilon_q(\gamma) - \epsilon_{q-q_1}(\gamma_2) + \epsilon_{p,q}(\beta, \gamma) - \epsilon_{p,q-q_1}(\beta, \gamma_2)\}/q_1,$$

em que $\epsilon_{p,q-q_1}(\beta, \gamma_2)$ segue de (1.10) com $\mathbf{Z}_\gamma$, $\mathbf{Z}_{\gamma_d}$ e $\mathbf{D}_\gamma$ substituídos por $\mathbf{Z}_{\gamma_2}$, $\mathbf{Z}_{\gamma_2d}$ e $\mathbf{D}_{\gamma_2}$, respectivamente. O fator c_3 deve ser avaliado na E.M.V. restrita $(\tilde{\beta}, \gamma_1^{(0)}, \tilde{\gamma}_2)$.

Alternativamente, Rocke (1989) propôs o fator de correção de Bartlett bootstrap. Sua abordagem difere da abordagem usual na qual o método bootstrap de Efron (1979) é utilizado para estimar a distribuição nula da estatística de teste, computando a partir dela um valor crítico bootstrap. A estatística da razão de verossimilhanças corrigida pode ser expressa como ω/c , em que $c = \mathbb{E}(\omega)/m$, m sendo o número de restrições impostas pela hipótese nula. A ideia de Rocke (1989) é estimar c usando bootstrap paramétrico. Especificamente, B amostras bootstrap são geradas utilizando o bootstrap paramétrico e impondo a hipótese nula e, para cada pseudo-amostra, é calculada a estatística da razão de verossimilhanças (ω^*). A estatística corrigida bootstrap é dada por

$$\omega_r = \frac{\omega m}{\bar{\omega}^*},$$

em que $\bar{\omega}^*$ é a média das estatísticas bootstrap, isto é,

$$\bar{\omega}^* = \frac{1}{B} \sum_{b=1}^B \omega_b^*,$$

ω_b^* sendo a estatística da razão de verossimilhanças calculada a partir da b -ésima amostra artificial, $b = 1, \dots, B$.

1.4 RESULTADOS NUMÉRICOS

Nesta seção, apresentamos resultados de simulações comparando o desempenho de amostra finita do teste da razão de verossimilhanças (ω), teste da razão de verossimilhanças ajustado via Skovgaard (ω_s), teste da razão de verossimilhanças corrigido (ω_b) e teste corrigido bootstrap (ω_r) na classe de MNLGS. Os resultados são baseados no modelo de regressão não linear passeio aleatório com

$$\log \mu_l = \beta_1 + \exp\{\beta_2 x_{l2}\} + \sum_{i=3}^p \beta_i x_{li} \quad \text{e} \quad \log \phi_l = \gamma_1 + \exp\{\gamma_2 s_{l2}\} + \sum_{j=3}^q \gamma_j s_{lj},$$

$l = 1, \dots, n$. A distribuição passeio aleatório com média μ e parâmetro de dispersão ϕ pode ser gerada a partir da distribuição χ_1^2 como $Y = 4(\mu\sqrt{Z\phi^{-1}} + \sqrt{\mu^2\phi^{-1}Z + 4\mu})^{-2}$, em que $Z \sim \chi_1^2$; ver Cordeiro et al. (2006). As derivadas da função $\psi(\mu, \phi)$ no modelo passeio aleatório são dadas no Apêndice A.

As hipóteses nulas consideradas são: (i) $\mathcal{H}_0 : \beta_5 = 0$ ($p = 5, q = 4$), (ii) $\mathcal{H}_0 : \beta_4 = \beta_5 = 0$ ($p = 5, q = 4$), (iii) $\mathcal{H}_0 : \gamma_5 = 0$ ($p = 5, q = 5$) e (iv) $\mathcal{H}_0 : \beta_5 = \gamma_4 = 0$ ($p = 5, q = 4$). Os parâmetros não fixados na hipótese nula são iguais a 1, exceto o intercepto do sub-modelo da média que é igual a -1 . Os valores das covariáveis foram obtidos como uma amostra aleatória da distribuição uniforme $\mathcal{U}(-0,5, 0,5)$, e mantidos constantes ao longo das simulações. Para cada tamanho de amostra $n = 30, 40, 50, 60, 70, 80, 90$, computamos as taxas de rejeição nulas dos testes nos níveis de significância $\alpha = 10\%, 5\%, 1\%$. As simulações foram realizadas usando a linguagem de programação matricial 0x (DOORNIK, 2007). A maximização da log-verossimilhança foi realizada utilizando o algoritmo BFGS. Utilizamos 10000 réplicas de Monte Carlo e $B = 400$ réplicas bootstrap. Os resultados são apresentados nas Tabelas 1-4; todas as entradas das tabelas são porcentagens.

A Tabela 1 contém os resultados de simulação correspondentes à hipótese nula $\mathcal{H}_0 : \beta_5 = 0$, com $p = 5, q = 4$. Notamos que o teste da razão de verossimilhanças usual ω é consideravelmente liberal, isto é, o teste rejeita a hipótese nula com mais frequência do que se espera com base nos níveis de significância, especialmente quando n é pequeno. Por exemplo, para $n = 30$ e $\alpha = 10\%$ a taxa de rejeição nula do teste ω é 24,4%, ou seja, maior do que duas vezes o nível de significância. No entanto, à medida que o tamanho da amostra n aumenta, as taxas de rejeição do teste se aproximam dos níveis de significância considerados. A correção de Bartlett atenua as distorções de tamanho do teste da razão de verossimilhanças, no entanto o teste corrigido ω_b ainda tem comportamento liberal em pequenas amostras. Por exemplo, no mesmo caso anterior, isto é, $n = 30$ e $\alpha = 10\%$, a taxa de rejeição nula do teste ω_b é igual a 17,6%. O teste da razão de verossimilhanças ajustado ω_s tem desempenho muito melhor que os dos testes original e corrigido de Bartlett. Ainda para $n = 30$ e $\alpha = 10\%$, a taxa de rejeição nula do teste ω_s é 11,9%. O teste corrigido bootstrap ω_r , por sua vez, também apresenta taxas de rejeição nulas muito próximas aos níveis de significância. Porém, diferente do teste ajustado ω_s , ω_r é ligeiramente conservativo em pequenas amostras. Por exemplo, quando $n = 30$, a taxa de rejeição nula do teste ω_r em $\alpha = 10\%$ é 9,0%. O teste ω_r tem um desempenho melhor que o teste ω_s em pequenas amostras, no entanto, possui um maior custo computacional.

A Figura 1 apresenta gráficos quantil-quantil das diferentes estatísticas de teste. Consideramos o primeiro cenário, isto é, a hipótese nula $\mathcal{H}_0 : \beta_5 = 0$, com tamanhos amostrais $n = 30, 40, 50$. Plotamos os quantis exatos das estatísticas de teste contra os respectivos quantis assintóticos obtidos da distribuição limite χ_1^2 . A linha sólida indica concordância perfeita entre as distribuições nulas exata e assintótica. Note que as linhas correspondentes às estatísticas corrigidas, especialmente aquelas para ω_s e ω_r , estão mais próximas da linha sólida que a da razão de verossimilhanças padrão ω . Isso ocorre nos três gráficos e sugere que as distribuições nulas das estatísticas corrigidas, principalmente ω_s e ω_r , são melhor aproximadas pela distribuição qui-quadrado em comparação com a razão de verossimilhanças usual.

Tabela 1 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
30	24,4	17,6	11,9	9,0	16,7	10,8	6,5	4,2	6,7	3,2	1,9	0,7
40	18,1	13,3	11,0	9,4	11,1	7,6	5,7	4,6	3,8	2,1	1,6	0,8
50	16,6	12,6	10,8	9,8	9,7	7,0	6,0	5,0	2,8	1,8	1,3	1,2
60	14,5	11,5	10,3	9,4	8,2	5,9	5,0	4,6	2,1	1,2	0,9	0,7
70	14,1	11,5	10,7	10,2	7,7	5,8	5,4	4,9	1,9	1,1	1,0	0,9
80	13,1	10,8	10,2	9,7	7,1	5,7	5,5	5,1	1,8	1,2	1,1	1,0
90	12,3	10,4	10,0	9,6	6,7	5,4	5,1	4,8	1,6	1,1	1,1	0,9

A Tabela 2 contém os resultados de simulação associados à hipótese nula $\mathcal{H}_0 : \beta_4 = \beta_5 = 0$, com $p = 5, q = 4$. Nesse caso, testamos duas restrições sobre o vetor de parâmetros de regressão da média. Como antes, o teste da razão de verossimilhanças ajustado de Skovgaard e o teste corrigido bootstrap tiveram os melhores desempenhos. Por exemplo, quando $n = 50$, as distorções de tamanho (isto é, tamanho empírico menos nível de significância) absolutas dos testes ω_r , ω_s , ω_b e ω em $\alpha = 5\%$ são 0,2%, 0,5%, 1,6%, 5,3%, respectivamente.

Consideramos agora o teste dos parâmetros do submodelo da dispersão. Os resultados relacionados à hipótese nula $\mathcal{H}_0 : \gamma_5 = 0$, com $p = 5, q = 5$, são fornecidos na Tabela 3. Notamos que o teste da razão de verossimilhanças também apresenta distorções de tamanho consideráveis, quando utilizado para testar parâmetros da dispersão. Por exemplo,

Figura 1 – Gráficos quantil-quantil das estatísticas da razão de verossimilhanças (ω), razão de verossimilhanças corrigida (ω_b), razão de verossimilhanças ajustada (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r), $\mathcal{H}_0 : \beta_5 = 0$.

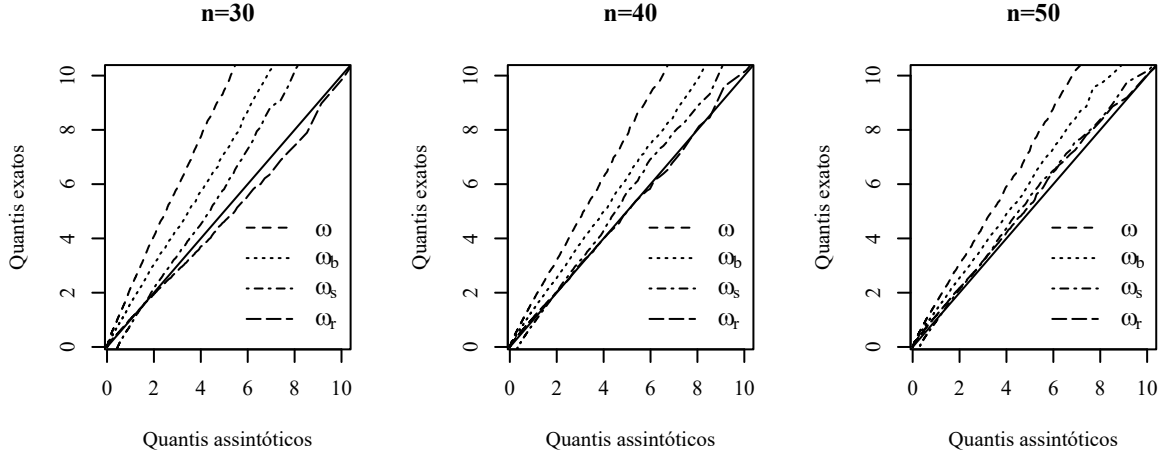


Tabela 2 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_4 = \beta_5 = 0$.

	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
n	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
30	27,6	18,5	12,5	9,7	18,9	11,2	7,4	4,9	7,7	3,6	2,1	0,9
40	20,0	13,9	11,4	9,8	12,5	8,0	6,3	4,9	4,1	2,2	1,6	1,0
50	17,5	12,8	10,9	9,7	10,3	6,6	5,5	4,8	2,9	1,4	1,3	0,9
60	15,2	11,1	10,4	9,3	8,4	5,7	5,4	4,6	2,2	1,3	1,2	1,0
70	14,4	11,3	10,4	9,8	8,0	5,8	5,2	4,8	2,1	1,2	1,0	0,9
80	14,0	11,2	10,9	9,9	7,7	6,0	5,7	5,3	2,1	1,3	1,3	1,1
90	12,7	10,3	10,0	9,5	6,9	5,2	5,0	4,7	1,7	1,0	0,9	0,9

quando $n = 40$ e $\alpha = 5\%$, o teste rejeita a hipótese nula em 11,0% dos casos, mais que o dobro do nível de significância. Aqui, a correção de Bartlett também melhora o desempenho do teste original em pequenas amostras, porém não o suficiente. Por exemplo, para $n = 40$ e $\alpha = 5\%$, a taxa de rejeição nula do teste ω_b é 7,5%. O teste ajustado, por outro lado, exibe um bom controle da frequência do erro tipo I, superando os testes usual e corrigido. Por exemplo, no mesmo caso anterior, $n = 40$ e $\alpha = 5\%$, a taxa de rejeição nula do teste ω_s é 5,9%. O teste de corrigido bootstrap foi ligeiramente melhor que o teste ajustado. Ainda para $n = 40$ e $\alpha = 5\%$, a taxa de rejeição nula para o teste ω_r é 5,2%.

Também consideramos o teste simultâneo de parâmetros dos submodelos da média e da dispersão. A hipótese nula sob teste é $\mathcal{H}_0 : \beta_5 = \gamma_4 = 0$, com $p = 5, q = 4$. As taxas de

Tabela 3 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \gamma_5 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
30	24,3	17,2	13,3	11,1	16,1	10,2	8,2	5,7	6,2	3,5	3,1	1,3
40	18,0	13,4	11,1	10,4	11,0	7,5	5,9	5,2	3,5	1,9	1,4	1,1
50	16,2	12,1	10,2	10,1	9,2	6,3	5,0	4,9	2,2	1,3	1,1	1,0
60	14,3	11,1	9,8	10,0	8,0	5,5	4,9	4,9	2,1	1,1	1,0	0,9
70	13,6	10,7	9,7	10,3	7,5	5,5	5,0	5,1	1,9	1,1	1,1	1,1
80	12,9	10,5	9,7	10,3	7,2	5,5	5,1	5,3	1,9	1,1	1,0	1,2
90	12,8	10,7	10,1	10,4	7,1	5,6	5,3	5,5	1,7	1,1	1,0	1,1

rejeição nulas dos testes são fornecidas na Tabela 4. Os testes da razão de verossimilhanças ajustado e corrigido bootstrap apresentam desempenhos semelhantes e ambos superam os demais teste.

Tabela 4 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = \gamma_4 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
30	24,5	15,2	11,3	10,0	16,2	8,8	6,4	5,3	6,1	2,6	2,1	1,0
40	19,9	13,5	10,3	9,5	12,1	7,3	5,2	4,8	3,9	1,8	1,2	1,0
50	17,4	12,1	10,0	9,9	10,1	6,1	5,2	4,7	3,0	1,6	1,2	1,2
60	15,6	11,6	9,9	9,9	8,7	5,4	4,8	4,6	2,2	1,2	0,9	0,9
70	14,1	10,6	9,8	9,6	7,8	5,5	4,9	4,9	2,1	1,3	1,1	1,1
80	13,8	10,5	9,7	9,8	7,5	5,5	5,0	4,8	2,0	1,3	1,1	1,1
90	13,5	10,7	9,9	10,0	7,3	5,5	4,7	5,0	1,9	1,2	1,0	1,1

Investigamos ainda o impacto do número de parâmetros de perturbação no desempenho dos testes. Para isso, fixamos o tamanho da amostra $n = 40$ e o número de covariáveis de dispersão $q = 3$, então testamos a hipótese nula $\mathcal{H}_0 : \beta_3 = 0$, para $p = 4, 5, 6, 7, 8$. Os resultados são dados na Tabela 5. Note que as distorções de tamanho do teste da razão de verossimilhanças ficam mais evidentes conforme o número de parâmetros de perturbação aumenta. Por exemplo, para $\alpha = 5\%$ as taxas de rejeição nulas do teste são 8,2% ($p = 4$) e 12,3% ($p = 8$). Esse efeito também pode ser observado no teste corrigido, porém de

forma menos marcante. O teste ajustado, em contrapartida, é muito menos sensível ao número de parâmetros de perturbação. Note que suas taxas de rejeição nulas continuam muito próximas dos níveis de significância quando aumentamos o número de parâmetros de perturbação. O teste corrigido bootstrap também é pouco afetado pelos parâmetros de perturbação, e apresenta um comportamento conservativo nos cenários considerados.

Tabela 5 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_3 = 0$, $n = 40$ e diferentes valores para p .

p	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
4	14,3	11,3	10,6	9,3	8,2	5,6	5,5	4,4	2,1	1,3	1,1	0,9
5	16,1	12,3	10,2	9,4	9,6	6,7	5,3	4,7	3,0	1,7	1,2	1,0
6	17,2	12,9	10,6	9,6	10,4	7,5	5,7	4,9	3,6	2,1	1,5	1,2
7	16,9	12,2	10,0	9,5	10,4	6,7	5,3	4,8	3,5	1,7	1,3	0,9
8	19,0	13,8	10,8	9,9	12,3	7,8	5,9	4,9	4,4	2,0	1,6	1,0

Adicionalmente, realizamos simulações comparando os poderes dos diferentes testes. A hipótese nula sob teste é $\mathcal{H}_0 : \beta_5 = 0$, com $p = 5$, $q = 4$, que corresponde ao primeiro cenário. Na geração da resposta, assumimos que $\beta_5 = \delta$, com δ diferente de zero. Logo, a hipótese nula é falsa. Uma vez que alguns testes tiveram comportamento liberal, utilizamos

Tabela 6 – Taxas de rejeição não nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças corrigido (ω_b), razão de verossimilhanças ajustado (ω_s) e razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_5 = 0$ e $n = 40$.

δ	$\alpha = 10\%$				$\alpha = 5\%$			
	ω	ω_b	ω_s	ω_r	ω	ω_b	ω_s	ω_r
-1,3	100,0	99,9	100,0	100,0	99,9	99,8	100,0	99,9
-1,1	99,6	99,5	99,7	99,5	99,3	99,1	99,4	99,1
-0,9	98,0	97,7	98,2	97,8	95,8	95,5	96,7	95,1
-0,7	89,4	89,4	90,2	88,8	82,3	81,9	84,4	81,0
-0,5	67,3	67,6	68,0	67,2	54,7	54,4	56,2	53,3
0,5	68,7	69,0	69,6	68,6	55,8	55,5	57,3	54,5
0,7	90,2	90,2	90,8	89,9	83,4	83,0	84,2	82,0
0,9	97,9	97,8	98,2	97,7	96,0	95,6	96,6	95,4
1,1	99,7	99,6	99,9	99,7	99,3	99,0	99,5	99,1
1,3	100,0	99,7	100,0	100,0	99,9	99,6	100,0	99,9

valores críticos exatos (simulados). Usamos $n = 40$ e $\alpha = 10\%, 5\%$. A Tabela 6 fornece os poderes estimados dos testes para diferentes valores de δ . Notamos que os poderes estimados dos testes foram bastante semelhantes. Além disso, não houve perda de poder devido ao uso do ajuste de Skovgaard para a estatística da razão de verossimilhanças.

1.5 APLICAÇÕES

1.5.1 Primeira Aplicação

McCullagh e Nelder (1989, p. 204) consideram um conjunto de dados em que a variável resposta (Y) é o número de incidentes com danos causados por ondas a navios de carga. Para cada um dos $n = 34$ navios foram registrados o período total de meses em serviço, bem como o número de incidentes com danos nesse período. Os regressores qualitativos são: tipo de navio (A, B, C, D e E), ano de construção (AC1, AC2, AC3 e AC4) e período de operação (PO1 e PO2). Os dados são fornecidos no Apêndice E, Tabela 19. O número médio de incidentes é igual a 10,47 e a variância é 247,59. Ou seja, a variância amostral das contagens de incidentes é substancialmente maior que a média, indicando a presença de superdispersão.

McCullagh e Nelder (1989) realizam uma análise desses dados usando um modelo de regressão de Poisson com ligação canônica dado por

$$\begin{aligned} \log \mu = & \log(\text{número de meses em serviço}) + \beta_1 + \text{efeito devido ao tipo de navio} \\ & + \text{efeito devido ao ano de construção} + \text{efeito devido ao período de operação,} \end{aligned} \quad (1.11)$$

em que μ é o número esperado de incidentes com danos, a média da distribuição de Poisson. O termo anterior ao intercepto β_1 é uma variável quantitativa cujo coeficiente é fixado em 1, esse termo é conhecido como *offset*. Para cada variável qualitativa, a categoria-base é tomada como sendo o primeiro nível. As estimativas dos parâmetros e erros-padrão para o modelo (1.11) são dados na primeira coluna da Tabela 7.

Aqui vamos assumir uma distribuição Poisson dupla para o número de incidentes, uma vez que essa distribuição é útil na modelagem de contagens superdispersas. Sendo assim, consideramos a mesma especificação dada na Equação (1.11) para a média da distribuição Poisson dupla e, seguindo Dey et al. (1997), modelamos o parâmetro de dispersão como

$\phi = \gamma_1 + \gamma_2 \log(\text{número de meses em serviço})$ e usamos a função $T(y) = (y+1) \log(y+1)$. As estimativas dos parâmetros e erros-padrão para esse modelo são dados na segunda coluna da Tabela 7.

Tabela 7 – Estimativas dos parâmetros, dados de incidentes com danos em navios.

Parâmetro	Modelo de Poisson	Modelo Poisson duplo
Intercepto	−6,406 (0,217)	−6,408 (0,623)
B	−0,543 (0,178)	−0,528 (0,439)
C	−0,687 (0,329)	−0,622 (0,717)
D	−0,076 (0,291)	−0,071 (0,650)
E	0,326 (0,236)	0,371 (0,531)
AC2	0,697 (0,150)	0,717 (0,512)
AC3	0,818 (0,170)	0,844 (0,526)
AC4	0,453 (0,233)	0,446 (0,657)
PO2	0,384 (0,118)	0,348 (0,341)
γ_1	—	0,501 (0,188)
γ_2	—	−0,042 (0,021)

Nota: Erros-padrão são dados entre parênteses.

A partir da especificação do parâmetro de dispersão em Dey et al. (1997), nota-se que os autores sugerem que a superdispersão varia com o número total de meses em serviço. Nosso interesse consiste em testar a hipótese nula $\mathcal{H}_0 : \gamma_2 = 0$ (dispersão constante) contra a hipótese alternativa $\mathcal{H}_1 : \gamma_2 \neq 0$ (dispersão variável). As derivadas de $\psi(\mu, \phi)$ no modelo Poisson duplo são dadas no Apêndice A. Os valores das estatísticas, com p -valores entre parênteses, são: $\omega = 3,5094$ (0,0610), $\omega_s = 5,5246$ (0,0188) e $\omega_b = 1,8051$ (0,1791). Note que, o teste da razão de verossimilhanças usual ω e o teste corrigido ω_b não rejeitam a hipótese nula de dispersão constante no nível de significância 5%. Por outro lado, o teste ajustado ω_s toma uma decisão diferente, a saber, rejeição da hipótese nula de dispersão constante no nível 5%. É importante ressaltar que a decisão do teste ω_s está de acordo com os resultados de Dey et al. (1997), que concluíram que a superdispersão presente nos dados varia com o número de meses em serviço.

1.5.2 Segunda Aplicação

Ribeiro et al. (2013) conduziram um experimento para avaliar a atividade inseticida de extratos de *Annona mucosa* (Annonaceae) sobre o gorgulho *Sitophilus zeamais*, que é uma das principais pragas de cereais armazenados em condições tropicais. No ensaio, placas de Petri contendo 10g de milho foram tratadas com extratos preparados com diferentes partes da planta (folhas, ramos e sementes) em uma concentração de 1500mg/kg ou somente água (controle), com 10 repetições. Então, cada placa de Petri foi infestada com 20 insetos, e, após 60 dias, contou-se o número (Y_{ij}) de insetos emergidos (progênie). Os dados são apresentados no Apêndice E, Tabela 20. A média e variância do número de insetos emergidos são 23,45 e 228,15, respectivamente. Temos portanto uma variância amostral consideravelmente maior que a média, sugerindo a existência de superdispersão nas contagens de gorgulhos emergidos.

Vamos analisar esses dados usando um modelo de regressão Poisson duplo com

$$\log \mu_{ij} = \beta_1 + \beta_{2i},$$

$$\log \phi_{ij} = \gamma_1,$$

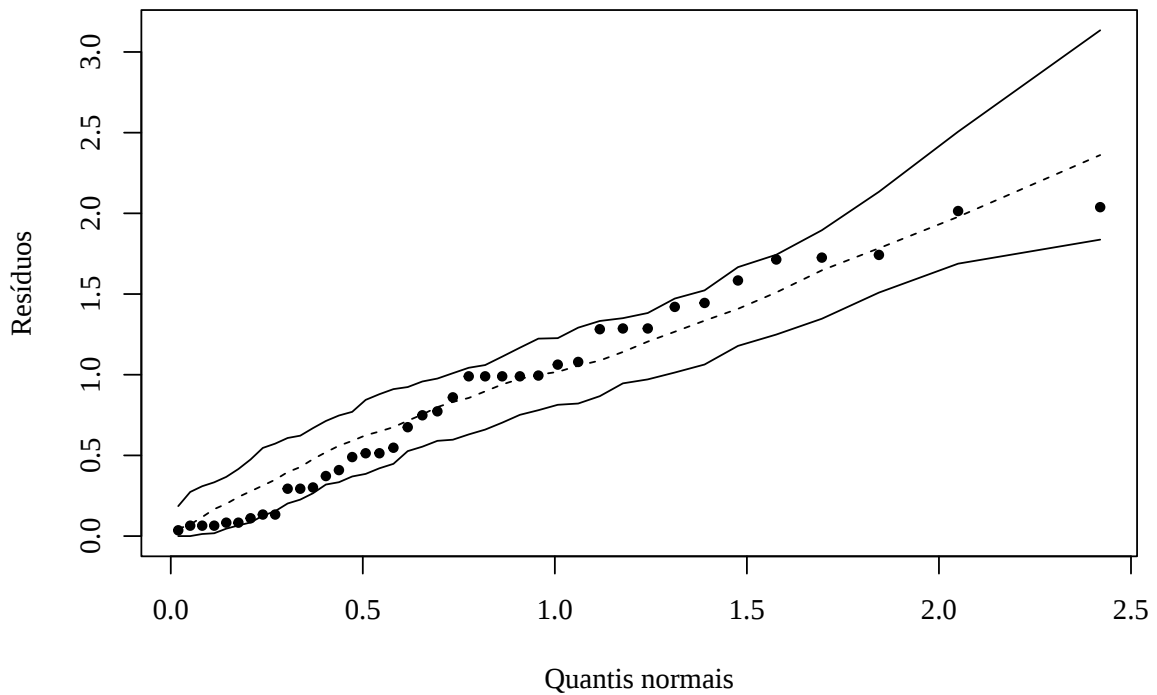
em que β_{2i} é o efeito do i -ésimo extrato, $i = 1, \dots, 4$. As estimativas dos parâmetros, erros-padrão e valores de AIC e BIC para esse modelo são dados na Tabela 8. Testamos então o modelo de Poisson contra o modelo Poisson duplo, ou seja, testamos $\mathcal{H}_0 : \gamma_1 = 0 (\phi = 1)$ contra $\mathcal{H}_1 : \gamma_1 \neq 0 (\phi \neq 1)$. Os valores das estatísticas, com p -valores entre parênteses, são: $\omega = 17,4339$ (0,0001), $\omega_s = 21,2416$ (0,0001) e $\omega_b = 13,3253$ (0,0003). Todos os três testes ω , ω_s e ω_b rejeitam o modelo de Poisson em favor do modelo Poisson duplo nos níveis de significância usuais.

Note que os critérios de seleção de modelos AIC e BIC também favorecem o modelo Poisson duplo (ver Tabela 8); uma vez que tal modelo apresenta os menores valores para ambos os critérios. A Figura 2 mostra o gráfico half-normal residual com envelopes simulados para o modelo Poisson duplo. Uma vez que apenas um resíduo ficou fora dos envelopes, concluímos que o modelo apresenta um ajuste adequado aos dados.

Tabela 8 – Estimativas dos parâmetros e valores de AIC e BIC, dados de progênie de *Sitophilus zeamais*.

Parâmetro	Modelo de Poisson	Modelo Poisson duplo
β_1	3,450 (0,056)	3,450 (0,084)
β_{22}	-0,006 (0,080)	-0,006 (0,120)
β_{23}	-0,052 (0,081)	-0,052 (0,121)
β_{24}	-3,355 (0,307)	-3,355 (0,460)
γ_1	—	-0,808 (0,224)
AIC	269,166	253,732
BIC	275,921	262,176

Nota: Erros-padrão são dados entre parênteses.

Figura 2 – Gráfico half-normal residual com envelopes simulados para o modelo Poisson duplo, dados de progênie de *Sitophilus zeamais*.

1.6 CONCLUSÃO

Os modelos não lineares generalizados superdispersados são úteis para análise de dados com superdispersão, isto é, dados em que a variância da resposta é maior do que a variância predita pelo modelo linear generalizado padrão. A superdispersão é comumente

observada em dados na forma de contagem e de proporção, que são tipicamente modelados usando as distribuições de Poisson e binomial, respectivamente. Os MNLGS incluem um parâmetro de dispersão que pode ser especificado como uma função das covariáveis. Esse parâmetro é incorporado à função de variância, e isso dá maior flexibilidade à modelagem da variância e, portanto, da superdispersão. O teste da razão de verossimilhanças é comumente utilizado para realizar inferências sobre os parâmetros de regressão nos MNLGS. Uma vez que esse teste usa uma aproximação qui-quadrado de primeira ordem, e que esta aproximação pode não ser precisa em amostras de tamanho pequeno ou moderado, é provável que o teste apresente distorções de tamanho.

Neste capítulo, derivamos um ajuste para a estatística da razão de verossimilhanças nos MNLGS, usando a proposta de Skovgaard (2001). Os resultados de simulação de Monte Carlo mostram que o teste ajustado é muito superior ao teste da razão de verossimilhanças usual no controle da frequência do erro tipo I em amostras de tamanho pequeno. Além disso, o teste ajustado que propomos também superou o teste corrigido de Bartlett e teve um desempenho tão bom quanto aquele do teste corrigido bootstrap. Por fim, apresentamos e discutimos duas aplicações empíricas.

2 TESTE DA RAZÃO DE VEROSSIMILHANÇAS PERFILADAS MODIFICADAS CORRIGIDO EM MODELOS DE REGRESSÃO NÃO LINEARES GENERALIZADOS SUPERDISPERSADOS

2.1 INTRODUÇÃO

A função de verossimilhança perfilada é comumente utilizada quando a inferência de interesse envolve parâmetros de perturbação. Tal função é obtida maximizando a verossimilhança usual com respeito ao parâmetro de perturbação, para dado valor do parâmetro de interesse. Contudo, a verossimilhança perfilada não se comporta como uma função de verossimilhança verdadeira, por exemplo, ela pode apresentar vícios na função escore e na informação; ver Ferrari et al. (2005). Consequentemente, vários ajustes foram propostos para a função de verossimilhança perfilada. Barndorff-Nielsen (1983) propôs uma verossimilhança perfilada modificada que pode ser difícil de obter em modelos gerais, pois depende da especificação de uma estatística ancilar e envolve derivadas com respeito ao espaço amostral. Cox e Reid (1987) propuseram um ajuste simples para a função de verossimilhança perfilada, mas que requer ortogonalidade entre os parâmetros de interesse e de perturbação. Outras modificações para a função de verossimilhança perfilada também estão disponíveis; ver, por exemplo, McCullagh e Tibshirani (1990), Cox e Reid (1993) e Stern (1997). Uma consequência importante do uso dessas versões ajustadas é o fato de que a esperança da função escore obtida com base nelas é de ordem $O(n^{-1})$, isto é, aproximadamente igual a zero, sendo n o tamanho da amostra. Uma revisão detalhada sobre funções de verossimilhanças perfiladas modificadas pode ser encontrada em Severini (2000, Capítulo 9), ver também Fraser e Reid (1989).

A estatística da razão de verossimilhanças perfiladas modificadas, assim como a estatística da razão de verossimilhanças original, tem uma distribuição assintótica qui-quadrado (χ^2) até um erro de ordem $O(n^{-1})$. Contudo, espera-se que ela forneça inferências mais precisas do que a estatística original, uma vez que atenua o efeito dos parâmetros de perturbação. Além disso, DiCiccio e Stern (1994) propuseram um fator de correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas que reduz o erro da aproximação pela distribuição χ^2 para ordem $O(n^{-2})$. Tal abordagem tem sido empregada em diversos estudos na literatura. Ferrari e Cribari-Neto (2002) desenvolveram um teste da razão de verossimilhanças perfiladas modificadas corrigido no modelo

de regressão normal linear válido somente quando o parâmetro que define o comportamento da heteroscedasticidade é escalar. Ferrari et al. (2004) generalizaram esse resultado considerando um vetor de parâmetros na estrutura de heteroscedasticidade. Ferrari et al. (2005) obtiveram um teste da razão de verossimilhanças perfiladas modificadas corrigido para os modelos lineares generalizados. Cysneiros e Ferrari (2006) derivaram a correção de Bartlett para um teste de dispersão variável na classe dos modelos não lineares da família exponencial. Araújo et al. (2020) derivaram ajustes semelhantes, porém na classe dos modelos não lineares simétricos heteroscedásticos.

Neste capítulo, nosso objetivo é derivar uma correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas de Cox e Reid (1987) nos modelos não lineares generalizados superdispersados. No que segue, o capítulo está estruturado da seguinte forma. Na Seção 2.2, apresentamos o teste da razão de verossimilhanças perfiladas modificadas através da metodologia de Cox e Reid (1987). Na Seção 2.3, derivamos a correção de Bartlett para a estatística da razão de verossimilhanças perfiladas modificadas de Cox e Reid para testar os efeitos da dispersão nos MNLGS. Na Seção 2.4, apresentamos resultados de simulação de Monte Carlo comparando os desempenhos do teste da razão de verossimilhanças perfiladas modificadas e sua versão corrigida. Uma aplicação empírica é discutida na Seção 2.5. Por fim, algumas considerações finais são fornecidas na Seção 2.6.

2.2 TESTE DA RAZÃO DE VEROSSIMILHANÇAS PERFILADAS MODIFICADAS

Seja $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ um vetor aleatório tendo uma função densidade que depende do parâmetro $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{q+p})^\top$. Denotamos a função de verossimilhança para $\boldsymbol{\theta}$ por $L(\boldsymbol{\theta})$, e o logaritmo da função de verossimilhança (log-verossimilhança) por $\ell = \ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$. Considere uma partição do parâmetro como $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\zeta}^\top)^\top$, em que $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)^\top$ é o parâmetro de interesse e $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_p)^\top$ é o parâmetro de perturbação. A função de log-verossimilhança perfilada é

$$\ell_p(\boldsymbol{\alpha}) = \ell(\boldsymbol{\alpha}, \tilde{\boldsymbol{\zeta}}_\alpha), \quad (2.1)$$

em que $\tilde{\boldsymbol{\zeta}}_\alpha$ é a estimativa de máxima verossimilhança (E.M.V.) restrita de $\boldsymbol{\zeta}$, dado o valor de $\boldsymbol{\alpha}$. A função escore perfilada é dada por

$$\mathbf{U}_p(\boldsymbol{\alpha}) = \frac{\partial \ell_p(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}.$$

A E.M.V. de $\boldsymbol{\alpha}$, denotada por $\hat{\boldsymbol{\alpha}}$, é obtida solucionando $\mathbf{U}_p(\boldsymbol{\alpha}) = \mathbf{0}$. A estatística da razão de verossimilhanças usual para testar $\mathcal{H}_0 : \boldsymbol{\alpha} = \boldsymbol{\alpha}^{(0)}$ contra $\mathcal{H}_1 : \boldsymbol{\alpha} \neq \boldsymbol{\alpha}^{(0)}$, em que $\boldsymbol{\alpha}^{(0)}$ é um vetor q -dimensional especificado, é dada por

$$\omega = 2\{\ell_p(\hat{\boldsymbol{\alpha}}) - \ell_p(\boldsymbol{\alpha}^{(0)})\}.$$

Em grandes amostras, a distribuição nula da estatística ω é aproximadamente χ_q^2 com um erro de ordem $O(n^{-1})$.

Cox e Reid (1987) introduziram a função de log-verossimilhança perfilada modificada

$$\ell_m(\boldsymbol{\alpha}) = \ell_p(\boldsymbol{\alpha}) - \frac{1}{2} \log \det\{\mathbf{J}_{\zeta\zeta}(\boldsymbol{\alpha}, \tilde{\boldsymbol{\zeta}}_{\boldsymbol{\alpha}})\}, \quad (2.2)$$

em que $\mathbf{J}_{\zeta\zeta}$ é o bloco da matriz de informação observada que corresponde ao parâmetro de perturbação $\boldsymbol{\zeta}$ avaliado em $(\boldsymbol{\alpha}^\top, \tilde{\boldsymbol{\zeta}}_{\boldsymbol{\alpha}}^\top)^\top$. Aqui, o parâmetro de interesse $\boldsymbol{\alpha}$ deve ser ortogonal ao parâmetro de perturbação $\boldsymbol{\zeta}$.

A estatística da razão de verossimilhanças perfiladas modificadas de Cox e Reid para testar $\mathcal{H}_0$ contra $\mathcal{H}_1$ é dada por

$$\omega_m = 2\{\ell_m(\bar{\boldsymbol{\alpha}}) - \ell_m(\boldsymbol{\alpha}^{(0)})\}, \quad (2.3)$$

em que $\bar{\boldsymbol{\alpha}}$ é o ponto em que $\ell_m(\boldsymbol{\alpha})$ é maximizada. Em grandes amostras, a distribuição nula da estatística ω_m é aproximadamente χ_q^2 com um erro de ordem $O(n^{-1})$. Uma vez que o teste da razão de verossimilhanças perfiladas modificadas atenua o impacto dos parâmetros de perturbação, espera-se que este forneça inferências mais precisas do que o teste da razão de verossimilhanças usual. Além disso, é possível melhorar ainda mais o desempenho da estatística modificada de Cox e Reid usando um fator de correção de Bartlett. Essa questão será abordada na próxima seção.

2.3 FATOR DE CORREÇÃO DE BARTLETT

DiCiccio e Stern (1994) mostraram que a estatística da razão de verossimilhanças perfiladas modificadas é corrigível por um fator de ajuste de Bartlett. A estatística corrigida é dada por

$$\omega_m^* = \frac{\omega_m}{1 + c_m/q}.$$

Aqui, temos $P(\omega_m^* > z_\delta) = \delta + O(n^{-2})$, enquanto $P(\omega_m > z_\delta) = \delta + O(n^{-1})$, em que z_δ é tal que $P(\chi_q^2 > z_\delta) = \delta$. Ou seja, a inferência baseada na estatística corrigida ω_m^* tem

erro de ordem $O(n^{-2})$ sob a hipótese nula, enquanto que testes baseados em ω_m tem erro de ordem $O(n^{-1})$. Antes de definir o termo c_m , é necessário introduzir algumas notações. Os índices $r, s, t, \dots$ variam sobre $1, \dots, q + p$; $a, b, c, \dots$ variam sobre $1, \dots, q$ e $i, j, k, \dots$ variam sobre $q + 1, \dots, q + p$. Seja θ_r denotando o r -ésimo elemento de $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\zeta}^\top)^\top$. Adotamos a seguinte notação para os cumulantes de derivadas da log-verossimilhança: $\lambda_{rs} = \mathbb{E}(\partial^2 \ell / \partial \theta_r \partial \theta_s)$, $\lambda_{rst} = \mathbb{E}(\partial^3 \ell / \partial \theta_r \partial \theta_s \partial \theta_t)$ e assim por diante. As derivadas dos cumulantes são denotadas por $(\lambda_{rs})_t = \partial \lambda_{rs} / \partial \theta_t$, $(\lambda_{rs})_{tu} = \partial^2 \lambda_{rs} / \partial \theta_t \partial \theta_u$ etc. A matriz $(q + p) \times (q + p)$ inversa de (λ_{rs}) é representada por (λ^{rs}) . O elemento (r, s) da matriz de informação de Fisher é dado por $-\lambda_{rs}$ e o elemento correspondente de sua inversa é $-\lambda^{rs}$. Definimos $\nu^{rs} = \lambda^{rs} - \tau^{rs}$, em que $\tau^{rs} = \lambda^{ra} \lambda^{sb} \sigma_{ab}$, com (σ_{ab}) sendo a inversa da matriz (λ^{ab}) ($q \times q$). Note que as entradas da matriz (ν^{rs}) são todas zero exceto para a submatriz $p \times p$ inferior direita, que é a inversa da matriz (λ_{ij}) . Qualquer índice repetido como um sobrescrito e um subscrito indica uma soma implícita sobre todos os valores desse índice.

A expressão para o termo c_m definido em DiCiccio e Stern (1994, p. 404) pode ser escrita como

$$\begin{aligned} c_m = & \frac{1}{4} \tau^{ru} \tau^{st} \lambda_{rstu} - \lambda^{ru} \tau^{st} (\lambda_{rst})_u + (\lambda^{ru} \lambda^{st} - \nu^{ru} \nu^{st}) (\lambda_{rs})_{tu} \\ & - \left(\frac{1}{4} \lambda^{ru} \tau^{st} \tau^{vw} + \frac{1}{2} \lambda^{ru} \tau^{sw} \tau^{tv} - \frac{1}{3} \tau^{ru} \tau^{sw} \tau^{tv} \right) \lambda_{rst} \lambda_{uvw} \\ & + (\lambda^{ru} \tau^{st} \lambda^{vw} + \lambda^{ru} \lambda^{sw} \lambda^{tv} - \nu^{ru} \lambda^{sw} \nu^{tv}) \lambda_{rst} (\lambda_{uv})_w \\ & - (\lambda^{ru} \lambda^{st} \lambda^{vw} - \nu^{ru} \nu^{st} \nu^{vw} + \lambda^{ru} \lambda^{sw} \lambda^{tv} - \nu^{ru} \nu^{sw} \nu^{tv}) (\lambda_{rs})_t (\lambda_{uv})_w. \end{aligned}$$

Uma vez que assumimos a ortogonalidade dos parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\zeta}$, é possível mostrar que, ver Ferrari et al. (2004) para mais detalhes,

$$\begin{aligned} c_m = & \frac{1}{4} \lambda^{ab} \lambda^{cd} \lambda_{abcd} - \lambda^{ab} \lambda^{cd} (\lambda_{acd})_b + \lambda^{ab} \lambda^{cd} (\lambda_{ac})_{db} - \lambda^{ij} \lambda^{ab} (\lambda_{iab})_j \\ & - \left(\frac{1}{4} \lambda^{ab} \lambda^{cd} \lambda^{ef} + \frac{1}{2} \lambda^{ab} \lambda^{cf} \lambda^{de} - \frac{1}{3} \lambda^{ab} \lambda^{cf} \lambda^{de} \right) \lambda_{acd} \lambda_{bef} \\ & + (\lambda^{ab} \lambda^{cd} \lambda^{ef} + \lambda^{ab} \lambda^{cf} \lambda^{de}) \lambda_{acd} (\lambda_{be})_f - (\lambda^{ab} \lambda^{cd} \lambda^{ef} + \lambda^{ab} \lambda^{cf} \lambda^{de}) (\lambda_{ac})_d (\lambda_{be})_f \\ & - \left(\frac{1}{4} \lambda^{ij} \lambda^{ab} \lambda^{cd} + \frac{1}{2} \lambda^{ij} \lambda^{ad} \lambda^{bc} \right) \lambda_{iab} \lambda_{jcd} + \lambda^{ij} \lambda^{ab} \lambda^{kl} \lambda_{iab} (\lambda_{jk})_l. \end{aligned} \quad (2.4)$$

No que segue, derivamos uma expressão matricial para o ajuste c_m considerando o teste dos efeitos da dispersão nos modelos não lineares generalizados superdispersados, definidos na Seção 1.2. Nosso interesse consiste em testar a hipótese nula $\mathcal{H}_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1 : \boldsymbol{\gamma} \neq \boldsymbol{\gamma}^{(0)}$, em que $\boldsymbol{\gamma}^{(0)}$ é um vetor q -dimensional especificado. A

estatística da razão de verossimilhanças perfiladas modificadas (2.3) para testar $\mathcal{H}_0$ contra $\mathcal{H}_1$ pode ser facilmente computada a partir da log-verossimilhança (1.3) e da matriz de informação observada (1.4), conforme descrito na seção anterior. Note que, de acordo com a notação utilizada, temos $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\zeta}) = (\boldsymbol{\gamma}, \boldsymbol{\beta})$.

Uma vez que os cumulantes λ_{iab} e $(\lambda_{iab})_j$ são iguais a zero para a classe de MNLGS, segue que $c_m = c_{m1} - c_{m2} - c_{m3}$ com

$$\begin{aligned} c_{m1} &= \lambda^{ab} \lambda^{cd} \left(\frac{1}{4} \lambda_{abcd} - (\lambda_{acd})_b + (\lambda_{ac})_{db} \right), \\ c_{m2} &= \lambda^{ab} \lambda^{cd} \lambda^{ef} \left(\lambda_{acd} \left\{ \frac{1}{4} \lambda_{bef} - (\lambda_{be})_f \right\} + (\lambda_{ac})_d (\lambda_{be})_f \right), \\ c_{m3} &= \lambda^{ab} \lambda^{cf} \lambda^{de} \left(\lambda_{acd} \left\{ \frac{1}{6} \lambda_{bef} - (\lambda_{be})_f \right\} + (\lambda_{ac})_d (\lambda_{be})_f \right). \end{aligned}$$

Substituindo as expressões para os cumulantes nas equações acima, obtivemos, após álgebra tediosa (ver Apêndice C),

$$\begin{aligned} c_{m1} &= \frac{1}{4} \text{tr} \{ (\boldsymbol{\Psi}^{(0,4)} \boldsymbol{\Phi}_1^4 - \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_2^2) \mathbf{Z}_d^2 \} + \frac{1}{2} \text{tr} \{ \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^2 (\boldsymbol{\Phi}_2 \mathbf{Z}_d + \boldsymbol{\Phi}_1 \mathbf{D}) \mathbf{Z}_d \} \\ &\quad - \frac{1}{2} \text{tr} \{ \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 (2\mathbf{C}_d - \mathbf{D} \mathbf{Z}_d) \} + \frac{1}{4} \text{tr} \{ \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 (\mathbf{D}^2 - 2\mathbf{B}_d) \}, \end{aligned} \quad (2.5)$$

$$\begin{aligned} c_{m2} &= -\frac{1}{4} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \mathbf{Z}_d \mathbf{Z} (\mathbf{Z}_d \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} + 2\mathbf{D} \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(0,2)}) \mathbf{1} \\ &\quad - \frac{1}{4} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \mathbf{Z}_d \mathbf{Z} (\mathbf{Z}_d \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} + 2\mathbf{D} \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(0,2)}) \mathbf{1} \\ &\quad - \frac{1}{2} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \mathbf{Z}_d \mathbf{Z} \mathbf{Z}_d \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1} - \frac{1}{4} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{D} \mathbf{Z} \mathbf{D} \boldsymbol{\Phi}_1^2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1} \end{aligned} \quad (2.6)$$

e

$$\begin{aligned} c_{m3} &= -\frac{1}{6} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,3)} \boldsymbol{\Phi}_1^3 \mathbf{Z}^{(3)} \boldsymbol{\Phi}_1^3 \boldsymbol{\Psi}^{(0,3)} \mathbf{1} + \frac{1}{2} \mathbf{1}^\top \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \mathbf{Z}^{(3)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \boldsymbol{\Psi}^{(0,2)} \mathbf{1} \\ &\quad + \text{tr} \{ \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1 \boldsymbol{\Phi}_2 \mathbf{C} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{Z} \} + \frac{1}{2} \text{tr} \{ \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{B} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{Z} \}, \end{aligned} \quad (2.7)$$

em que $\mathbf{1}$ é um vetor $n \times 1$ de uns, $\mathbf{Z} = \mathbf{S}^* (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1} \mathbf{S}^{*\top}$ é uma matriz $n \times n$ com elementos $z_{lm} = \mathbf{s}_l^{*\top} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1} \mathbf{s}_m^*$, $\mathbf{s}_l^*$ sendo a l -ésima linha de $\mathbf{S}^*$, $\mathbf{Z}_d = \text{diag}\{z_{11}, \dots, z_{nn}\}$ e $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$, $d_l = \text{tr}\{\mathbf{S}_l^{*} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1}\}$, $\mathbf{S}_l^*$ é uma matriz $q \times q$ com elementos $\partial^2 \tau_l / \partial \gamma_a \partial \gamma_b = (ab)_l$. Além disso, $\mathbf{Z}^{(3)} = \{z_{lm}^3\}$, $\mathbf{B} = \{b_{lm}\}$, $b_{lm} = \text{tr}\{\mathbf{S}_l^{*} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1} \mathbf{S}_m^{*} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1}\}$, $\mathbf{C}$ é uma matriz $n \times n$ com elemento (l, m) dado por $a_{ll}^{(m)} = \mathbf{s}_l^{*\top} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1} \mathbf{S}_m^{*} (-\mathbf{S}^{*\top} \boldsymbol{\Psi}^{(0,2)} \boldsymbol{\Phi}_1^2 \mathbf{S}^*)^{-1} \mathbf{s}_l^*$, $\mathbf{B}_d = \text{diag}\{b_{11}, \dots, b_{nn}\}$ e $\mathbf{C}_d = \text{diag}\{a_{11}^{(1)}, \dots, a_{nn}^{(n)}\}$.

Utilizando as expressões acima podemos reescrever c_m como a soma de duas componentes, $c_m = c_m^L + c_m^{NL}$ com

$$\begin{aligned} c_m^L = & \frac{1}{12} \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 (3\mathbf{Z}_d \mathbf{Z} \mathbf{Z}_d + 2\mathbf{Z}^{(3)}) \Phi_1^3 \Psi^{(0,3)} \mathbf{1} + \frac{1}{2} \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 \mathbf{Z}_d \mathbf{Z} \mathbf{Z}_d \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} \\ & + \frac{1}{4} \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 \Phi_2 (\mathbf{Z}_d \mathbf{Z} \mathbf{Z}_d - 2\mathbf{Z}^{(3)}) \Phi_1 \Phi_2 \Psi^{(0,2)} \mathbf{1} \\ & + \frac{1}{4} \text{tr} \{ (\Psi^{(0,4)} \Phi_1^4 + 2\Psi^{(0,3)} \Phi_1^2 \Phi_2 - \Psi^{(0,2)} \Phi_2^2) \mathbf{Z}_d^2 \} \end{aligned}$$

e

$$\begin{aligned} c_m^{NL} = & \frac{1}{2} \mathbf{1}^\top \Psi^{(0,3)} \Phi_1^3 \mathbf{Z}_d \mathbf{Z} \mathbf{D} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} + \frac{1}{4} \mathbf{1}^\top \Psi^{(0,2)} \Phi_1 (2\Phi_2 \mathbf{Z}_d \mathbf{Z} + \Phi_1 \mathbf{D} \mathbf{Z}) \mathbf{D} \Phi_1^2 \Psi^{(0,2)} \mathbf{1} \\ & - \frac{1}{4} \text{tr} \{ \Psi^{(0,2)} \Phi_1 [\Phi_2 (4\mathbf{C}_d - 2\mathbf{D} \mathbf{Z}_d) + \Phi_1 (2\mathbf{B}_d - \mathbf{D}^2)] - 2\Psi^{(0,3)} \Phi_1^3 \mathbf{D} \mathbf{Z}_d \} \\ & - \frac{1}{2} \text{tr} \{ \Psi^{(0,2)} (\Phi_1^2 \mathbf{B} + 2\Phi_1 \Phi_2 \mathbf{C}) \Psi^{(0,2)} \Phi_1^2 \mathbf{Z} \}. \end{aligned}$$

Quando o ajuste c_m envolve parâmetros desconhecidos, eles devem ser substituídos por suas estimativas de máxima verossimilhança sob a hipótese nula.

2.4 RESULTADOS NUMÉRICOS

Nesta seção, apresentamos resultados de simulações comparando os desempenhos de amostra finita dos testes da razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças usual (ω) e suas versões corrigidas (ω_s e ω_b). Consideramos o modelo de regressão não linear Poisson duplo com

$$\begin{aligned} \log \mu_l &= \beta_2 + e^{\beta_1 x_{l1}} + \sum_{i=3}^p \beta_i x_{li}, \\ \log \phi_l &= \gamma_2 + e^{\gamma_1 s_{l1}} + \sum_{j=3}^q \gamma_j s_{lj}, \end{aligned}$$

$l = 1, \dots, n$. Testamos a hipótese nula $\mathcal{H}_0 : \gamma_1 = 1, \gamma_2 = \dots = \gamma_q = 0$ contra uma hipótese alternativa bilateral. Na geração da resposta, assumimos que $\beta_1 = \dots = \beta_p = 1$. Os valores das covariáveis foram obtidos como uma amostra aleatória da distribuição uniforme no intervalo $(-0,5, 0,5)$. Os níveis de significância são $\alpha = 10\%, 5\%, 1\%$. Já os tamanhos de amostra são $n = 20, 30, 40, 50, 60$. Todas as simulações foram realizadas utilizando a linguagem de programação matricial 0x (DOORNIK, 2007). Os resultados são baseados em 10000 réplicas de Monte Carlo.

A Tabela 9 apresenta as taxas de rejeição nulas dos testes para o caso em que $n = 40$, $q = 2$ e $p = 3, 4, 5, 6$. Variamos p para avaliar o efeito dos parâmetros de perturbação sobre os testes. O teste da razão de verossimilhanças usual ω é bastante afetado pelo número de parâmetros de perturbação no modelo. Notamos que o comportamento liberal do teste ω torna-se mais evidente conforme o número de parâmetros de perturbação aumenta. Por exemplo, para $p = 3$ e $p = 6$ as taxas de rejeição estimadas do teste ω em $\alpha = 5\%$ são 8,2% e 20,0%, respectivamente. O teste da razão de verossimilhanças corrigido ω_b é melhor que o teste usual, mas ainda é afetado pelo número de parâmetros de perturbação. Por exemplo, para $p = 3$ e $p = 6$, as taxas de rejeição estimadas do teste ω_b em $\alpha = 5\%$ são 4,2% e 8,6%, respectivamente. O teste ω_s supera os testes ω e ω_b . Na mesma situação anterior, as taxas de rejeição estimadas do teste ω_s são 5,5% e 4,8%. Por sua vez, o teste da razão de verossimilhanças perfiladas modificadas ω_m atenua o impacto do número de parâmetros de perturbação. Notamos que suas taxas de rejeição estão apenas ligeiramente acima dos níveis de significância considerados. Por exemplo, para $p = 3$ e $p = 6$, as taxas de rejeição estimadas do teste ω_m em $\alpha = 10\%$ são 10,7% e 10,9%, respectivamente. A correção de Bartlett para o teste da razão de verossimilhanças perfiladas modificadas traz suas taxas de rejeição para ainda mais próximo dos níveis de significância. Na mesma situação anterior, as taxas de rejeição estimadas do teste da razão de verossimilhanças perfiladas modificadas corrigido ω_m^* são 10,2% e 9,9%.

A Tabela 10 contém as taxas de rejeição nulas dos testes para o caso em que $n = 40$, $p = 4$ e $q = 2, 3, 4, 5$. Note que, embora o teste da razão de verossimilhanças tenha um comportamento liberal, suas taxas de rejeição não variam muito com o aumento do número de parâmetros de interesse. Analogamente, os testes da razão de verossimilhanças corrigidos apresentam taxas de rejeição estáveis conforme o número de parâmetros de interesse aumenta. O mesmo vale para o teste da razão de verossimilhanças perfiladas modificadas e sua versão corrigida. Agora, comparando os desempenhos dos testes, notamos que o teste da razão de verossimilhanças ajustado ω_s tem desempenho semelhante ao do teste da razão de verossimilhanças perfiladas modificadas, e ambos superam o teste usual ω e sua versão corrigida ω_b . Além disso, está claro a partir dos resultados que o teste de melhor desempenho é o teste da razão de verossimilhanças perfiladas modificadas corrigido ω_m^* .

Na Tabela 11 apresentamos os resultados para o caso em que $p = 3$, $q = 2$ e $n = 20, 30, 50, 60$. Notamos que o teste ajustado ω_s apresentou um comportamento conservativo

Tabela 9 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $q = 2$, $n = 40$ e diferentes valores para p .

p	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	3	4	5	6	3	4	5	6	3	4	5	6
ω	14,0	21,5	25,4	29,1	8,2	13,6	17,5	20,0	2,2	5,0	8,4	8,9
ω_m	10,7	11,2	11,3	10,9	5,5	5,6	6,0	5,3	1,2	1,1	1,2	1,1
ω_m^*	10,2	10,4	10,3	9,9	5,1	5,2	5,4	4,8	1,0	0,9	0,9	0,9
ω_b	8,9	11,0	14,0	14,9	4,2	6,0	8,6	8,6	0,8	1,5	2,9	2,8
ω_s	11,7	11,1	12,5	10,3	5,5	6,0	6,4	4,8	1,5	1,1	1,0	1,0

Tabela 10 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 4$, $n = 40$ e diferentes valores para q .

q	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	2	3	4	5	2	3	4	5	2	3	4	5
ω	21,5	22,0	23,0	24,1	13,6	14,1	14,8	15,3	5,0	5,3	5,4	5,2
ω_m	11,2	10,6	10,8	11,0	5,6	5,5	5,4	5,3	1,1	1,2	1,2	1,0
ω_m^*	10,4	9,7	9,6	9,8	5,2	4,9	4,7	4,7	0,9	1,0	0,9	0,9
ω_b	11,0	11,3	11,3	11,7	6,0	6,3	6,2	6,0	1,5	1,5	1,5	1,6
ω_s	11,1	9,5	10,2	10,0	6,0	5,1	5,8	5,8	1,1	1,4	1,6	1,8

neste cenário, mas suas taxas de rejeição tendem rapidamente aos níveis de significância à medida que n aumenta. Vale destacar o bom desempenho do teste da razão de verossimilhanças perfiladas modificadas corrigido em pequenas amostras. Note que suas taxas de rejeição estimadas foram ligeiramente distorcidas, mesmo para n pequeno. Por exemplo, quando $n = 20$, a taxa de rejeição estimada do teste ω_m^* em $\alpha = 10\%$ é 10,4%.

Também realizamos simulação de poder para o caso $p = 3$, $q = 2$ e $n = 30$. Na geração da resposta, assumimos que $\gamma_1 = 1 + \delta$ e $\gamma_2 = \delta$ com $\delta \neq 0$. Os poderes estimados dos testes nos níveis de significância $\alpha = 10\%$ e 5% , para diferentes valores de δ , são dados na Tabela 12. Notamos que os poderes estimados dos testes não diferiram muito. Além disso, o teste da razão de verossimilhanças perfiladas modificadas apresentou, em geral, poderes simulados ligeiramente maiores do que sua versão corrigida.

Tabela 11 – Taxas de rejeição nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 3$ e $q = 2$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	20	30	50	60	20	30	50	60	20	30	50	60
ω	22,6	16,7	14,4	13,1	14,4	9,8	7,7	7,1	5,5	2,9	1,8	1,7
ω_m	11,1	11,1	11,2	11,1	5,8	5,7	5,6	5,9	1,1	1,3	1,1	1,3
ω_m^*	10,4	10,4	10,8	10,8	5,2	5,4	5,3	5,7	0,9	1,1	1,1	1,2
ω_b	11,2	9,2	9,5	9,5	6,0	4,7	4,7	4,6	1,3	1,1	0,9	1,0
ω_s	6,8	8,1	9,8	10,3	3,1	4,1	4,9	5,4	0,6	0,9	0,9	1,2

Tabela 12 – Taxas de rejeição não nulas (%) dos testes da razão de verossimilhanças (ω), razão de verossimilhanças perfiladas modificadas (ω_m), razão de verossimilhanças perfiladas modificadas corrigido (ω_m^*), razão de verossimilhanças corrigido (ω_b) e razão de verossimilhanças ajustado (ω_s): $p = 3$, $q = 2$ e $n = 30$.

δ	$\alpha = 10\%$					$\alpha = 5\%$				
	ω	ω_m	ω_m^*	ω_b	ω_s	ω	ω_m	ω_m^*	ω_b	ω_s
-1,1	98,4	99,1	99,1	97,2	96,6	97,3	98,5	98,4	95,5	95,0
-0,9	90,4	93,8	93,6	86,4	88,3	86,4	90,5	90,3	80,1	84,1
-0,7	79,5	85,7	85,4	71,9	79,0	72,0	79,6	78,9	62,4	72,7
-0,5	56,1	65,7	65,0	44,6	62,0	44,9	54,8	53,9	33,1	51,7
0,5	70,6	47,4	46,4	55,3	65,9	57,7	33,1	32,1	38,8	47,7
0,7	92,2	76,2	75,4	82,5	96,5	84,4	63,2	61,8	68,9	86,1
0,9	98,8	93,9	93,3	96,0	100,0	96,7	86,2	85,4	89,4	99,6
1,1	100,0	99,3	99,2	99,6	100,0	99,8	97,7	97,4	98,5	100,0

2.5 APLICAÇÃO

Na Seção 1.5.2 analisamos os dados de um experimento realizado para avaliar a atividade inseticida de extratos de *Annona mucosa* (Annonaceae) sobre o gorgulho *Sitophilus zeamais* em amostras de milho; ver Ribeiro et al. (2013). A variável de interesse é o número (Y_{ij}) de insetos emergidos (progênie) após 60 dias. Ajustamos o seguinte modelo de regressão Poisson duplo:

$$\log \mu_{ij} = \beta_1 + \beta_{2i},$$

$$\log \phi_{ij} = \gamma_1,$$

em que β_{2i} é o efeito do i -ésimo extrato, $i = 1, \dots, 4$. A Seção 1.5.2 contém detalhes sobre a estimação e avaliação da bondade de ajuste desse modelo. Aqui, também realizamos o teste da hipótese nula $\mathcal{H}_0 : \gamma_1 = 0 (\phi = 1)$ contra a hipótese alternativa $\mathcal{H}_1 : \gamma_1 \neq 0 (\phi \neq 1)$. Ou seja, testamos o modelo de Poisson contra o modelo Poisson duplo. Os valores das estatísticas de teste (p -valores entre parênteses) são: $\omega = 17,4339 (0,0001)$, $\omega_m = 20,6673 (0,0001)$, $\omega_m^* = 20,4965 (0,0001)$, $\omega_s = 21,2416 (0,0001)$ e $\omega_b = 13,3253 (0,0003)$. Todos os cinco testes rejeitam a hipótese nula nos níveis de significância usuais. Em outras palavras, os testes considerados favorecem fortemente o modelo Poisson duplo.

2.6 CONCLUSÃO

A superdispersão é caracterizada por uma variabilidade extra observada nos dados e não capturada pelo modelo padrão. Esse problema está mais relacionado aos modelos de Poisson e binomial, pois a heterogeneidade presente nos dados tende a violar a suposição subjacente de que a variância é uma função específica da média. Neste capítulo, consideramos os modelos não lineares generalizados superdispersados para análise de dados com superdispersão, e abordamos o problema de realizar inferências na presença de parâmetros de perturbação. Derivamos uma correção de Bartlett para um teste de dispersão de verossimilhança perfilada modificada de Cox e Reid (1987) nos MNLGS, usando a abordagem de DiCiccio e Stern (1994). Nossos resultados numéricos mostram que o número de parâmetros de perturbação afeta o desempenho do teste da razão de verossimilhanças usual, no sentido de que seu comportamento liberal torna-se mais evidente à medida que o número de parâmetros de perturbação aumenta. O teste da razão de verossimilhanças perfiladas modificadas é muito menos afetado pelos parâmetros de perturbação, superando o teste padrão, e sendo superado por sua versão corrigida que apresentou um excelente desempenho com taxas de rejeição muito próximas dos níveis de significância. Uma aplicação empírica também é apresentada e discutida.

3 TESTE DA RAZÃO DE VEROSSIMILHANÇAS APRIMORADO NO MODELO DE REGRESSÃO BETA PRIME COM PRECISÃO VARIÁVEL

3.1 INTRODUÇÃO

O modelo de regressão beta prime, introduzido por Bourguignon et al. (2021), é útil para a análise de variáveis respostas positivas. Ele é baseado na suposição de que a resposta tem distribuição beta prime indexada pelos parâmetros de média e de precisão. E ambos os parâmetros podem ser especificados como funções das variáveis explicativas.

Keeping (1962) discutiu algumas propriedades da distribuição beta prime; ver também McDonald (1984) e McDonald (1987). Essa distribuição tem sido considerada em várias pesquisas recentemente desenvolvidas. Biazatti et al. (2022) definiram uma generalização do modelo beta prime considerando também a modelagem de regressão. Freitas et al. (2022) propuseram uma abordagem para análise de variáveis aleatórias positivas com medidas repetidas baseada na distribuição beta prime. Outros trabalhos envolvendo tal distribuição são Leão et al. (2021), Kamyari et al. (2022), Medeiros et al. (2023) e Santos et al. (2023).

O método de máxima verossimilhança é empregado para estimação dos parâmetros nos submodelos da média e da precisão na regressão beta prime; e o teste da razão de verossimilhanças é tipicamente usado para realizar inferências quanto a significância estatística de tais parâmetros. No entanto, o teste da razão de verossimilhanças usa valores críticos assintóticos obtidos da distribuição limite qui-quadrado (χ^2) da estatística de teste. E essa aproximação qui-quadrado pode ser pobre em pequenas amostras, de modo que o teste pode apresentar distorções de tamanho consideráveis. Em particular, nossos resultados de simulação mostram que o teste tem comportamento liberal em pequenas amostras, isto é, ele rejeita excessivamente hipóteses nulas verdadeiras.

Uma estratégia amplamente utilizada na literatura para melhorar o desempenho de amostra finita do teste em problemas específicos consiste na obtenção do fator de correção de Bartlett para a estatística da razão de verossimilhanças (BARTLETT, 1937). A correção de Bartlett reduz o erro na aproximação qui-quadrado para a distribuição nula da estatística da razão de verossimilhanças de ordem $O(n^{-1})$ para ordem $O(n^{-2})$, de forma que o teste corrigido tende a apresentar comportamento superior em pequenas amostras.

Várias pesquisas têm focado na obtenção de tais correções em modelos de regressão específicos. Botter e Cordeiro (1997) derivaram um fator de correção de Bartlett para a estatística da razão de verossimilhanças nos modelos lineares generalizados com covariáveis de dispersão. Cordeiro et al. (2006) obtiveram resultados semelhantes para uma classe de modelos de superdispersão. Araújo et al. (2020) derivaram correções de Bartlett para modelos não lineares simétricos heteroscedásticos. Guedes et al. (2021) desenvolveram um fator de correção de Bartlett para a estatística da razão de verossimilhanças no modelo de regressão beta com precisão variável. Uma introdução ao tema pode ser encontrada em Cordeiro e Cribari-Neto (2014).

Neste capítulo, nosso objetivo é obter uma expressão para o fator de correção de Bartlett à estatística da razão de verossimilhanças no modelo de regressão beta prime com precisão variável. A partir desse resultado, vamos definir três testes baseados nas estatísticas da razão de verossimilhanças corrigidas para a classe de regressão beta prime. Os testes baseados nas estatísticas corrigidas podem ser utilizados para testar conjuntamente os parâmetros dos submodelos da média e da precisão.

O capítulo está organizado da seguinte forma. Na Seção 3.2, apresentamos o modelo de regressão beta prime com precisão variável. Na Seção 3.3, derivamos a correção de Bartlett para o teste da razão de verossimilhanças no modelo de regressão beta prime. Os resultados de simulação comparando os comportamentos de amostra finita do teste da razão de verossimilhanças usual e suas versões corrigidas são apresentados na Seção 3.4. Uma aplicação empírica é discutida na Seção 3.5. Finalmente, a Seção 3.6 contém algumas considerações finais.

3.2 O MODELO DE REGRESSÃO BETA PRIME

Seja Y uma variável aleatória que tem distribuição beta prime com parâmetros $\mu > 0$ (média) e $\phi > 0$ (precisão). A função densidade de Y é

$$f(y; \mu, \phi) = \frac{y^{\mu(1+\phi)-1}(1+y)^{-(\mu(1+\phi)+\phi+2)}}{B(\mu(1+\phi), \phi+2)}, \quad y > 0, \quad (3.1)$$

em que $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ é a função beta e $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$ é a função gama. A variância de Y é dada por $\mu(1+\mu)/\phi$. A parametrização alternativa em (3.1) para a distribuição beta prime foi proposta por Bourguignon et al. (2021).

Sejam $Y_1, \dots, Y_n$ variáveis aleatórias independentes, em que Y_l , $l = 1, \dots, n$, é beta prime distribuído com média μ_l e precisão ϕ_l . O modelo de regressão beta prime é definido pelos submodelos da média e da precisão

$$g(\mu_l) = \eta_l = \sum_{i=1}^p \beta_i x_{li} \quad \text{e} \quad h(\phi_l) = \tau_l = \sum_{j=1}^q \gamma_j s_{lj},$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ e $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_q)^\top$ são vetores de parâmetros de regressão desconhecidos a serem estimados ($p + q < n$), $\mathbf{x}_l = (x_{l1}, \dots, x_{lp})^\top$ e $\mathbf{s}_l = (s_{l1}, \dots, s_{lq})^\top$ são as respectivas covariáveis da média e da precisão, $x_{l1} = s_{l1} = 1 \forall l$. Adicionalmente, $g: \mathbb{R}_+ \mapsto \mathbb{R}$ e $h: \mathbb{R}_+ \mapsto \mathbb{R}$ são funções de ligação estritamente crescentes e diferenciáveis.

A função de log-verossimilhança para $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$, o vetor de parâmetros da regressão beta prime, é dada por

$$\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{l=1}^n \ell(\mu_l, \phi_l),$$

em que $\ell(\mu_l, \phi_l) = (\mu_l(1 + \phi_l) - 1) \log y_l - (\mu_l(1 + \phi_l) + \phi_l + 2) \log(1 + y_l) - \log \Gamma(\mu_l(1 + \phi_l)) - \log \Gamma(\phi_l + 2) + \log \Gamma(\mu_l(1 + \phi_l) + \phi_l + 2)$.

Diferenciando a função de log-verossimilhança com respeito aos parâmetros de regressão obtivemos a função escore $\mathbf{U}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (\mathbf{U}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})^\top, \mathbf{U}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})^\top)^\top$, com

$$\mathbf{U}_\beta = \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} = \mathbf{X}^\top (\boldsymbol{\Phi} + \mathbf{I}) \mathbf{D}_1 (\mathbf{y}^* - \boldsymbol{\mu}^*) \quad \text{e} \quad \mathbf{U}_\gamma = \frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} = \mathbf{S}^\top \mathbf{D}_2 (\mathbf{y}^* - \boldsymbol{\mu}^*),$$

em que $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$ ($n \times p$) e $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_n)^\top$ ($n \times q$) são as matrizes de covariáveis da média e da precisão, $\mathbf{I}$ é a matriz identidade $n \times n$, $\boldsymbol{\Phi}$, $\mathbf{D}_1$ e $\mathbf{D}_2$ são matrizes diagonais $n \times n$ com elementos ϕ_l , $1/g'(\mu_l)$ e $1/h'(\phi_l)$, respectivamente, $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$, $\mathbf{y}_l^* = (y_1^*, \dots, y_n^*)^\top$, $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^\top$ e $\boldsymbol{\mu}_l^* = (\mu_1^*, \dots, \mu_n^*)^\top$, com $y_l^* = \log y_l - \log(1 + y_l)$, $y_l^* = \mu_l \log y_l - (1 + \mu_l) \log(1 + y_l)$, $\mu_l^* = \psi^{(0)}(\mu_l(1 + \phi_l)) - \psi^{(0)}(\mu_l(1 + \phi_l) + \phi_l + 2)$ e $\mu_l^* = \mu_l \mu_l^* - \psi^{(0)}(\mu_l(1 + \phi_l) + \phi_l + 2) + \psi^{(0)}(\phi_l + 2)$, sendo $\psi^{(0)}(z) = d \log \Gamma(z)/dz$ a função digama.

Os estimadores de máxima verossimilhança (E.M.V.) dos parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, denotados por $\hat{\boldsymbol{\beta}}$ e $\hat{\boldsymbol{\gamma}}$, não possuem forma fechada, as estimativas são obtidas por meio da maximização da log-verossimilhança usando um algoritmo de otimização não linear, como, por exemplo, o algoritmo BFGS; ver Nocedal e Wright (2006). Os E.M.V. de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ solucionam o sistema de equações não lineares $\mathbf{U}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$.

A matriz de informação de Fisher é

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

em que $\mathbf{K}_{\beta\beta} = \mathbf{X}^\top \mathbf{H}_1 \mathbf{X}$, $\mathbf{K}_{\beta\gamma} = \mathbf{K}_{\gamma\beta}^\top = \mathbf{X}^\top \mathbf{H}_2 \mathbf{S}$ e $\mathbf{K}_{\gamma\gamma} = \mathbf{S}^\top \mathbf{H}_3 \mathbf{S}$, com $\mathbf{H}_1 = \text{diag}\{h_{11}, \dots, h_{1n}\}$, $h_{1l} = a_l[(1+\phi_l)/g'(\mu_l)]^2$, $\mathbf{H}_2 = \text{diag}\{h_{21}, \dots, h_{2n}\}$, $h_{2l} = b_l/(g'(\mu_l)h'(\phi_l))$, e $\mathbf{H}_3 = \text{diag}\{h_{31}, \dots, h_{3n}\}$, $h_{3l} = c_l/(h'(\phi_l))^2$. As quantidades a_l , b_l e c_l são definidas no Apêndice D.

3.3 TESTE DA RAZÃO DE VEROSSIMILHANÇAS CORRIGIDO

Em geral, o interesse consiste em testar restrições sobre um subconjunto dos componentes do vetor de parâmetros k -dimensional $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$. Então, particionamos $\boldsymbol{\theta}$ na forma $\boldsymbol{\theta} = (\boldsymbol{\lambda}^\top, \boldsymbol{\zeta}^\top)^\top$, em que $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)^\top$ e $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_r)^\top$ são, respectivamente, os parâmetros de interesse e de perturbação, com $m + r = p + q$. Logo, desejamos testar a hipótese nula $\mathcal{H}_0 : \boldsymbol{\lambda} = \boldsymbol{\lambda}^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1 : \boldsymbol{\lambda} \neq \boldsymbol{\lambda}^{(0)}$, em que $\boldsymbol{\lambda}^{(0)}$ é um vetor m -dimensional especificado. A estatística da razão de verossimilhanças é

$$\omega = 2\{\ell(\hat{\boldsymbol{\lambda}}, \hat{\boldsymbol{\zeta}}) - \ell(\boldsymbol{\lambda}^{(0)}, \tilde{\boldsymbol{\zeta}})\},$$

em que $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\lambda}}^\top, \hat{\boldsymbol{\zeta}}^\top)^\top$ e $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\lambda}^{(0)\top}, \tilde{\boldsymbol{\zeta}}^\top)^\top$ são, respectivamente, os estimadores de máxima verossimilhança irrestrito e restrito (sob $\mathcal{H}_0$) de $\boldsymbol{\theta} = (\boldsymbol{\lambda}^\top, \boldsymbol{\zeta}^\top)^\top$. Sob a hipótese nula, ω é assintoticamente distribuído como χ_m^2 , em que m é o número de restrições impostas por $\mathcal{H}_0$. Os valores críticos para o teste da razão de verossimilhanças são, portanto, obtidos dessa distribuição limite qui-quadrado. Como uma consequência, o teste pode apresentar distorções de tamanho em pequenas amostras.

Uma abordagem tipicamente utilizada para melhorar o desempenho de amostra finita do teste em problemas específicos é a obtenção da correção de Bartlett para a estatística da razão de verossimilhanças. A estatística corrigida é $\omega_{b1} = \omega/c$, em que $c = 1 + b/m$ tornou-se conhecido como o fator de correção de Bartlett. Aqui, a constante $b = b(\boldsymbol{\theta})$, de ordem $O(n^{-1})$, segue da seguinte expansão para o valor esperado de ω sob a hipótese nula obtida por Bartlett (1937): $\mathbb{E}(\omega) = m + b + O(n^{-2})$. Adicionalmente, $b(\boldsymbol{\theta})$ deve ser avaliado na estimativa restrita $\tilde{\boldsymbol{\theta}}$. A estatística ajustada ω_{b1} é melhor aproximada pela distribuição χ_m^2 do que a estatística usual ω , já que a correção de Bartlett reduz o

erro na aproximação qui-quadrado para ordem n^{-2} . Lawley (1956) obteve uma fórmula geral para o ajuste de Bartlett como uma função de cumulantes de derivadas da log-verossimilhança. Ele mostrou que $\mathbb{E}(\omega) = m + \varepsilon_k - \varepsilon_{k-m} + O(n^{-2})$, em que ambos ε_k e ε_{k-m} são de ordem n^{-1} . Consequentemente, o fator de correção de Bartlett pode ser escrito como $c = 1 + (\varepsilon_k - \varepsilon_{k-m})/m$, em que

$$\varepsilon_k = \sum_{\theta} (l_{rstu} - l_{rstuvw}), \quad (3.2)$$

com

$$l_{rstu} = \kappa^{rs} \kappa^{tu} \left\{ \frac{1}{4} \kappa_{rstu} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right\}, \quad (3.3)$$

$$l_{rstuvw} = \kappa^{rs} \kappa^{tu} \kappa^{vw} \left\{ \frac{1}{6} \kappa_{rtv} \kappa_{suw} + \frac{1}{4} \kappa_{rtu} \kappa_{svw} - \kappa_{rtv} \kappa_{sw}^{(u)} - \kappa_{rtu} \kappa_{sv}^{(v)} + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right\}. \quad (3.4)$$

Os cumulantes (κ 's) acima são definidos no Apêndice D. No somatório em (3.2) os índices r, s, t, u, v, w variam de 1 a k . Além disso, $-\kappa^{rs}$ é o elemento (r, s) da matriz inversa da informação de Fisher $\mathbf{K}(\boldsymbol{\beta}, \boldsymbol{\gamma})^{-1}$. O termo ε_{k-m} também é obtido das Equações (3.2)-(3.4), entretanto o somatório varia somente sobre os $k - m$ parâmetros de perturbação.

Lawley (1956) demonstrou ainda que os cumulantes de ω_{b1} concordam com os da distribuição χ_m^2 até um erro de ordem $O(n^{-2})$. Uma expansão assintótica para a distribuição nula de ω foi obtida por Hayakawa (1977). Mais detalhes sobre as correções de Bartlett podem ser encontrados em Cordeiro e Cribari-Neto (2014); ver também Cribari-Neto e Cordeiro (1996).

No que segue, derivamos o fator de correção de Bartlett à estatística da razão de verossimilhanças no modelo de regressão beta prime com precisão variável. É importante notar que os parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ não são ortogonais, isto é, a matriz de informação de Fisher não é bloco-diagonal. Isso torna o processo de obtenção da correção de Bartlett mais longo e também mais tedioso. Utilizaremos o principal resultado em Cordeiro (1993), que forneceu uma expressão matricial para o termo de ajuste ε_k .

Sejam $\mathbf{A}^{(tu)}$, $\mathbf{P}^{(t)}$ e $\mathbf{Q}^{(u)}$ matrizes $k \times k$ com elementos (r, s) da forma

$$\mathbf{A}^{(tu)} = \left\{ \frac{1}{4} \kappa_{rstu} - \kappa_{rst}^{(u)} + \kappa_{rt}^{(su)} \right\}, \quad \mathbf{P}^{(t)} = \{\kappa_{rst}\}, \quad \mathbf{Q}^{(u)} = \{\kappa_{su}^{(r)}\},$$

$t, u = 1, \dots, k$. Avaliamos as matrizes acima usando os cumulantes que derivamos para o modelo de regressão beta prime com precisão variável (ver Apêndice D). Então, definimos

as matrizes $\mathbf{L}$, $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$, $\mathbf{N}_1$, $\mathbf{N}_2$ e $\mathbf{N}_3$ cujos elementos (r, s) são

$$\begin{aligned} L_{rs} &= \text{tr}(\mathbf{K}^{-1} \mathbf{A}^{(rs)}), \\ M_{1rs} &= \text{tr}(\mathbf{K}^{-1} \mathbf{P}^{(r)} \mathbf{K}^{-1} \mathbf{P}^{(s)}), \\ M_{2rs} &= \text{tr}(\mathbf{K}^{-1} \mathbf{P}^{(r)} \mathbf{K}^{-1} \mathbf{Q}^{(s)\top}), \\ M_{3rs} &= \text{tr}(\mathbf{K}^{-1} \mathbf{Q}^{(r)} \mathbf{K}^{-1} \mathbf{Q}^{(s)}), \\ N_{1rs} &= \text{tr}(\mathbf{P}^{(r)} \mathbf{K}^{-1}) \text{tr}(\mathbf{P}^{(s)} \mathbf{K}^{-1}), \\ N_{2rs} &= \text{tr}(\mathbf{P}^{(r)} \mathbf{K}^{-1}) \text{tr}(\mathbf{Q}^{(s)} \mathbf{K}^{-1}), \\ N_{3rs} &= \text{tr}(\mathbf{Q}^{(r)} \mathbf{K}^{-1}) \text{tr}(\mathbf{Q}^{(s)} \mathbf{K}^{-1}), \end{aligned}$$

$r, s = 1, \dots, k$. A partir desses resultados, segue que, em forma matricial,

$$\begin{aligned} \sum_{\theta} l_{rstu} &= \text{tr}(\mathbf{K}^{-1} \mathbf{L}), \\ \sum_{\theta} \kappa^{rs} \kappa^{tu} \kappa^{vw} &\left\{ \frac{1}{6} \kappa_{rtv} \kappa_{suw} - \kappa_{rtv} \kappa_{sw}^{(u)} + \kappa_{rt}^{(v)} \kappa_{sw}^{(u)} \right\} \\ &= -\frac{1}{6} \text{tr}(\mathbf{K}^{-1} \mathbf{M}_1) + \text{tr}(\mathbf{K}^{-1} \mathbf{M}_2) - \text{tr}(\mathbf{K}^{-1} \mathbf{M}_3), \\ \sum_{\theta} \kappa^{rs} \kappa^{tu} \kappa^{vw} &\left\{ \frac{1}{4} \kappa_{rtu} \kappa_{svw} - \kappa_{rtu} \kappa_{sw}^{(v)} + \kappa_{rt}^{(u)} \kappa_{sw}^{(v)} \right\} \\ &= -\frac{1}{4} \text{tr}(\mathbf{K}^{-1} \mathbf{N}_1) + \text{tr}(\mathbf{K}^{-1} \mathbf{N}_2) - \text{tr}(\mathbf{K}^{-1} \mathbf{N}_3). \end{aligned}$$

Assim, ε_k pode ser expresso como

$$\varepsilon_k = \text{tr}[\mathbf{K}^{-1}(\mathbf{L} - \mathbf{M} - \mathbf{N})],$$

em que $\mathbf{M} = -\frac{1}{6} \mathbf{M}_1 + \mathbf{M}_2 - \mathbf{M}_3$ e $\mathbf{N} = -\frac{1}{4} \mathbf{N}_1 + \mathbf{N}_2 - \mathbf{N}_3$. A quantidade ε_{k-m} é obtida de modo semelhante, todavia considerando somente os parâmetros de perturbação.

Consideramos também as seguintes estatísticas corrigidas que são assintoticamente equivalentes a ω_{b1} até ordem $O(n^{-1})$: $\omega_{b2} = \omega \exp[-(\varepsilon_k - \varepsilon_{k-m})/m]$ e $\omega_{b3} = \omega[1 - (\varepsilon_k - \varepsilon_{k-m})/m]$.

3.4 RESULTADOS NUMÉRICOS

Realizamos simulações de Monte Carlo para comparar o desempenho de amostra finita do teste da razão de verossimilhanças original (ω) com os de suas versões corrigidas de Bartlett (ω_{b1} , ω_{b2} e ω_{b3}) no modelo de regressão beta prime com precisão variável. Todas

as simulações foram realizadas usando a linguagem de programação matricial **0x** (DO-ORNIK, 2007). A maximização da função de log-verossimilhança foi realizada utilizando o algoritmo BFGS. Todos os resultados são baseados no modelo de regressão beta prime com $\log \mu_l = \beta_1 + \beta_2 x_{l2} + \dots + \beta_p x_{lp}$ e $\log \phi_l = \gamma_1 + \gamma_2 s_{l2} + \dots + \gamma_q s_{lq}$, $l = 1, \dots, n$. Os valores das covariáveis foram obtidos como amostras aleatórias da distribuição $\mathcal{U}(0, 1)$. Consideramos quatro hipóteses nulas distintas. Testamos (i) $\mathcal{H}_0 : \beta_2 = 0$ com $(p = 4, q = 2)$, $(p = 6, q = 2)$ e $(p = 4, q = 4)$, (ii) $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$ $(p = 4, q = 4)$, (iii) $\mathcal{H}_0 : \gamma_4 = 0$ $(p = 4, q = 4)$, e (iv) $\mathcal{H}_0 : \beta_2 = \beta_3 = \gamma_4 = 0$ $(p = 4, q = 4)$. No primeiro caso, os correspondentes valores dos parâmetros são $\beta = (1, 0, 0, 0, 1, 8, 2, 5)$ e $\gamma = (3, 0, 1, 5)$, $\beta = (1, 0, 0, 0, 1, 8, 2, 5, 0, 9, 1, 4)$ e $\gamma = (3, 0, 1, 5)$, $\beta = (1, 0, 0, 0, 1, 8, 2, 5)$ e $\gamma = (3, 0, 1, 5, 0, 7, -1, 1)$. No segundo, configuramos $\beta = (1, 0, 0, 0, 0, 0, 2, 5)$ e $\gamma = (3, 0, 1, 5, 0, 7, -1, 1)$. No terceiro, temos $\beta = (1, 0, 0, 3, 1, 8, 2, 5)$ e $\gamma = (3, 0, 1, 5, 0, 7, 0, 0)$. No quarto, $\beta = (1, 0, 0, 0, 0, 0, 2, 5)$ e $\gamma = (3, 0, 1, 5, 0, 7, 0, 0)$. Computamos as taxas de rejeição nulas dos testes, para cada valor de $n = 20, 30, 40, 50, 60, 80$, nos níveis de significância $\alpha = 10\%, 5\%, 1\%$. O número de réplicas é 10000. Os resultados são apresentados nas Tabelas 13-16; entradas são porcentagens.

A Tabela 13 contém os resultados de simulação para a hipótese nula $\mathcal{H}_0 : \beta_2 = 0$. Nesse caso, um único parâmetro de regressão da média é fixado na hipótese nula. Notamos que o teste da razão de verossimilhanças original ω tem comportamento não confiável em pequenas amostras. Por exemplo, quando $n = 20$, $p = 4$, $q = 2$ e $\alpha = 5\%$, a taxa de rejeição nula do teste ω é 10,7%, ou seja, mais que o dobro do nível de significância. Além disso, o desempenho pobre do teste da razão de verossimilhanças em amostras de tamanho pequeno fica mais evidente com o aumento do número de parâmetros de perturbação. Por exemplo, quando $n = 20$, $p = 4$ e $\alpha = 10\%$, as taxas de rejeição nulas do teste ω para $q = 2$ e $q = 4$ são, respectivamente, 17,6% e 27,1%. Com respeito aos testes da razão de verossimilhanças ajustados de Bartlett, é fácil ver que todos os três testes têm desempenho superior ao do teste original, especialmente quando n é pequeno. Por exemplo, para $n = 40$, $p = 4$, $q = 4$ e $\alpha = 5\%$, as correspondentes taxas de rejeição nulas dos testes ω , ω_{b1} , ω_{b2} e ω_{b3} são 9,3%, 6,1%, 5,6% e 5,2%. Também verificamos que o impacto dos parâmetros de perturbação sobre o desempenho dos testes corrigidos é menos pronunciado em relação ao teste não modificado, especialmente no caso do teste ω_{b3} . Por exemplo, quando $n = 20$, $p = 4$ e $\alpha = 10\%$, o teste ω_{b3} apresenta taxas de rejeição nulas para $q = 2$ e $q = 4$ iguais a 11,0% e 10,7%, respectivamente. Claramente, o teste com

melhor desempenho é o teste ω_{b3} .

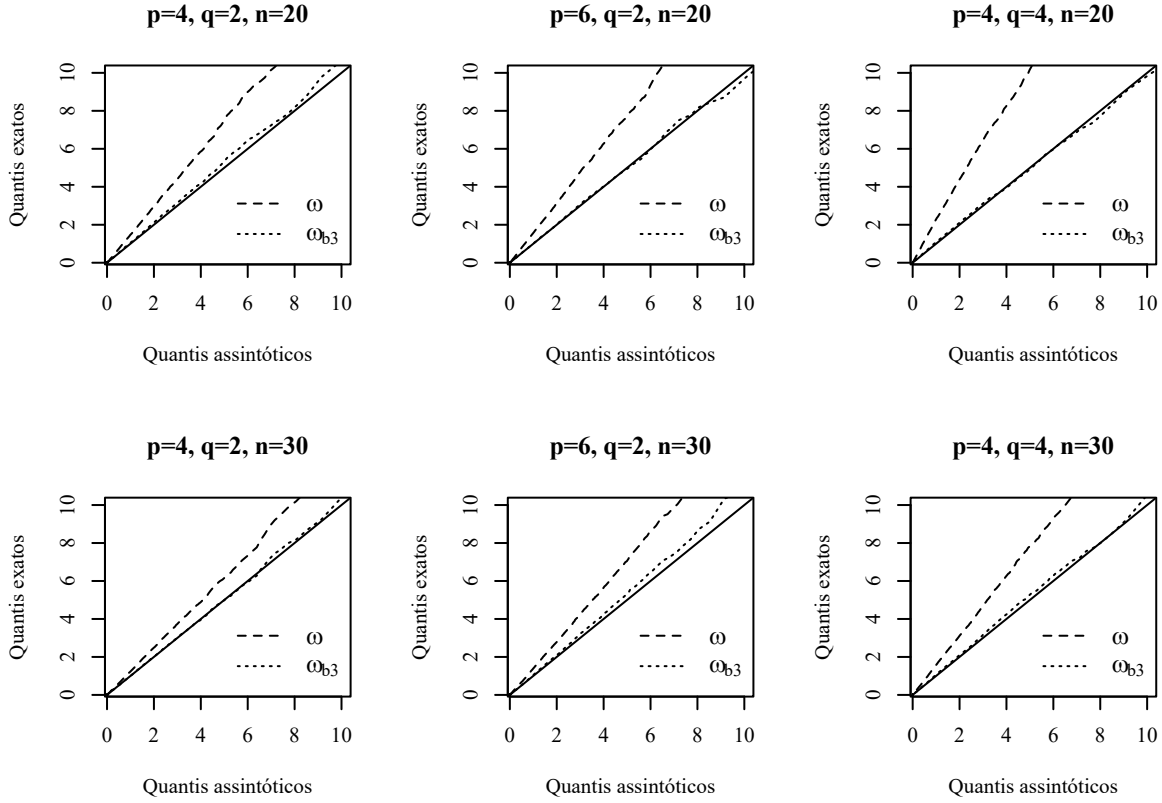
Tabela 13 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}
$p = 4, q = 2$												
20	17,6	12,7	11,9	11,0	10,7	6,8	6,3	5,6	3,3	1,7	1,5	1,2
30	14,0	10,8	10,5	10,1	7,9	5,5	5,3	5,0	2,0	1,1	1,1	1,0
40	12,9	10,4	10,3	10,1	7,1	5,4	5,3	5,2	1,7	1,1	1,0	0,9
50	12,6	10,8	10,6	10,5	6,9	5,6	5,5	5,5	1,8	1,2	1,2	1,2
60	11,7	10,3	10,3	10,2	6,2	5,0	5,0	5,0	1,4	1,0	1,0	1,0
80	10,8	9,5	9,5	9,5	5,5	5,0	4,9	4,9	1,0	0,8	0,8	0,8
$p = 6, q = 2$												
20	18,7	12,7	11,8	10,3	11,9	7,1	6,2	5,1	4,0	1,6	1,4	1,1
30	16,4	12,0	11,6	11,0	9,9	6,6	6,2	5,7	3,1	1,6	1,5	1,3
40	14,1	10,8	10,5	10,2	7,8	5,7	5,5	5,3	2,0	1,2	1,1	1,1
50	12,9	10,2	10,1	9,8	7,0	5,2	5,1	5,0	1,8	1,2	1,1	1,1
60	12,5	10,4	10,3	10,1	6,8	5,5	5,4	5,2	1,8	1,2	1,2	1,1
80	12,0	10,2	10,1	10,1	6,2	5,2	5,1	5,1	1,4	1,0	1,0	1,0
$p = 4, q = 4$												
20	27,1	17,2	15,0	10,7	18,7	10,1	8,4	4,9	7,9	2,6	2,0	0,9
30	18,7	12,8	11,9	10,7	11,4	7,2	6,6	5,8	3,9	1,7	1,5	1,2
40	15,6	11,7	11,2	10,5	9,3	6,1	5,6	5,2	2,5	1,2	1,2	1,0
50	14,1	10,7	10,4	10,1	7,9	5,4	5,2	4,9	2,2	1,2	1,2	1,1
60	13,1	10,5	10,3	10,0	7,2	5,2	5,1	4,9	1,9	1,1	1,0	0,9
80	12,3	10,2	10,1	10,0	6,6	5,2	5,1	5,0	1,6	1,0	0,9	0,9

A Figura 3 mostra gráficos quantil-quantil das estatísticas da razão de verossimilhanças ω e razão de verossimilhanças corrigida ω_{b3} . Tais gráficos estão associados ao primeiro cenário, isto é, à hipótese nula $\mathcal{H}_0 : \beta_2 = 0$. Plotamos os quantis exatos das estatísticas de teste contra os quantis assintóticos correspondentes obtidos da distribuição limite χ_1^2 . A linha sólida indica que as distribuições nulas exata e assintótica concordam perfeitamente. Note que a distribuição nula da estatística da razão de verossimilhanças é pobremente aproximada pela distribuição qui-quadrado de referência, uma vez que, em todos os seis gráficos, a linha associada a ω está afastada da linha sólida. Em contrapartida, a linha correspondente a ω_{b3} está muito próxima da linha sólida, sugerindo, portanto, que a distribuição nula dessa estatística de teste é bem aproximada pela distribuição assintótica

qui-quadrado.

Figura 3 – Gráficos quantil-quantil das estatísticas da razão de verossimilhanças (ω) e razão de verossimilhanças corrigida (ω_{b3}), $\mathcal{H}_0 : \beta_2 = 0$.



A Tabela 14 contém os resultados de simulação para a hipótese nula $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$ ($p = 4, q = 4$). Aqui, também consideramos o fator de correção de Bartlett bootstrap proposto por Rocke (1989). Podemos computá-lo da seguinte forma. Usando o método bootstrap paramétrico e impondo a hipótese nula obtemos B amostras artificiais. Para cada uma das B amostras bootstrap, calculamos a estatística da razão de verossimilhanças, cujo valor denotamos por ω_j^* , $j = 1, \dots, B$. Seja $\bar{\omega}^*$ a média dos valores ω_j^* , a estatística da razão de verossimilhanças corrigida bootstrap é então dada por

$$\omega_r = \frac{\omega m}{\bar{\omega}^*},$$

em que m é o número de restrições sob a hipótese nula. Note que aqui $m = 2$. Utilizamos $B = 1000$ réplicas bootstrap. Os resultados para os testes ω , ω_{b1} , ω_{b2} e ω_{b3} são semelhantes aos apresentados na Tabela 13 para o caso $p = 4, q = 4$; o teste ω_{b3} tendo o melhor desempenho. Com relação ao teste corrigido bootstrap ω_r , notamos que ele é conservativo em pequenas amostras. Seu desempenho é ligeiramente melhor que o do teste corrigido

ω_{b3} . Todavia, tendo em vista o maior custo computacional envolvido no cômputo de ω_r , e que seu desempenho não é significativamente superior ao de ω_{b3} , recomendamos o uso do teste ajustado ω_{b3} em análises empíricas baseadas no modelo de regressão beta prime.

Tabela 14 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω), dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}) e do teste da razão de verossimilhanças corrigido bootstrap (ω_r): $\mathcal{H}_0 : \beta_2 = \beta_3 = 0$.

n	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω_r
$\alpha = 10\%$					
20	32,6	19,1	16,4	11,6	9,5
30	21,5	13,4	12,3	11,0	9,8
40	17,9	12,1	11,6	10,8	10,3
50	15,6	11,1	10,8	10,4	9,9
60	15,2	10,9	10,7	10,4	10,1
80	13,5	10,5	10,4	10,3	10,2
$\alpha = 5\%$					
20	22,9	11,2	9,1	5,8	4,4
30	13,5	7,2	6,5	5,4	4,6
40	10,8	6,7	6,3	5,7	5,2
50	9,0	5,8	5,5	5,1	5,0
60	8,2	5,5	5,3	5,1	5,0
80	7,5	5,4	5,3	5,2	5,2
$\alpha = 1\%$					
20	10,1	3,3	2,3	1,2	0,9
30	4,5	1,8	1,5	1,2	0,9
40	3,4	1,5	1,4	1,1	1,0
50	2,5	1,2	1,1	1,0	0,9
60	2,1	1,1	1,0	0,9	0,9
80	2,0	1,3	1,3	1,2	1,2

Agora, estamos interessados em considerar testes para os coeficientes de regressão da precisão. A Tabela 15 contém os resultados de simulação correspondentes à hipótese nula $\mathcal{H}_0 : \gamma_4 = 0$ ($p = 4, q = 4$). Como antes, a frequência com que o teste da razão de verossimilhanças rejeita a hipótese nula é muito maior que o esperado com base nos níveis de significância selecionados. Os testes corrigidos controlam muito melhor a frequência do erro tipo I relativamente ao teste padrão. Por exemplo, quando $n = 30$, as taxas de rejeição nulas dos testes ω_{b1} , ω_{b2} e ω_{b3} em $\alpha = 1\%$ são 1,8%, 1,7% e 1,5%, respectivamente, enquanto que a do teste ω é 3,1%.

Tabela 15 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \gamma_4 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}
20	24,3	16,8	15,3	13,1	16,4	9,9	8,8	7,0	6,4	2,7	2,3	1,5
30	17,4	13,1	12,4	11,7	10,8	6,8	6,4	5,8	3,1	1,8	1,7	1,5
40	15,1	11,8	11,5	11,1	8,8	6,3	6,0	5,7	2,6	1,4	1,4	1,3
50	13,9	11,1	10,8	10,5	7,7	5,8	5,7	5,5	2,0	1,3	1,3	1,2
60	12,8	10,6	10,5	10,3	6,9	5,1	5,1	5,0	1,7	1,1	1,1	1,0
80	11,7	10,0	9,9	9,7	5,8	4,8	4,8	4,7	1,4	1,0	1,0	1,0

Também abordamos o teste simultâneo de parâmetros dos submodelos da média e da precisão. A Tabela 16 apresenta as taxas de rejeição nulas para o teste da hipótese nula $\mathcal{H}_0 : \beta_2 = \beta_3 = \gamma_4 = 0$ ($p = 4, q = 4$). Os resultados não diferem muito daqueles nas tabelas acima. O teste da razão de verossimilhanças tem comportamento destacadamente liberal. Os testes corrigidos apresentam taxas de rejeição nulas muito mais próximas dos níveis de significância que as do teste padrão, de modo que este é superado por todos os testes corrigidos.

Tabela 16 – Taxas de rejeição nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = \beta_3 = \gamma_4 = 0$.

n	$\alpha = 10\%$				$\alpha = 5\%$				$\alpha = 1\%$			
	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}
20	31,8	18,5	16,4	13,1	22,0	10,9	9,0	6,8	8,9	2,9	2,3	1,5
30	20,3	12,7	11,8	10,7	12,3	6,6	6,2	5,4	4,0	1,7	1,4	1,2
40	17,2	11,6	11,2	10,7	10,1	6,2	5,8	5,6	2,9	1,5	1,4	1,2
50	15,3	11,2	10,8	10,5	8,7	5,7	5,5	5,3	2,2	1,0	1,0	0,9
60	14,5	11,1	10,8	10,6	8,0	5,6	5,5	5,3	1,9	1,1	1,1	1,1
80	12,3	9,6	9,5	9,4	6,6	5,0	4,9	4,9	1,5	1,0	1,0	0,9

Por fim, realizamos algumas simulações sob a hipótese alternativa para avaliar o desempenho de poder dos testes. Consideramos a mesma hipótese nula do primeiro cenário, $\mathcal{H}_0 : \beta_2 = 0$, com $p = 6$ e $q = 2$. A geração dos dados foi realizada usando $\beta_2 = \delta \neq 0$. Como alguns testes apresentaram um comportamento liberal, usamos valores críticos exatos obtidos da simulação de tamanho. Utilizamos $n = 20$ e $\alpha = 10\%, 5\%$. Os poderes estimados dos testes para diferentes valores de δ são dados na Tabela 17. Notamos que

não houve perda de poder dos testes baseados nas estatísticas corrigidas via correção de Bartlett.

Tabela 17 – Taxas de rejeição não nulas (%) do teste da razão de verossimilhanças (ω) e dos três testes da razão de verossimilhanças corrigidos (ω_{b1} , ω_{b2} e ω_{b3}): $\mathcal{H}_0 : \beta_2 = 0$ e $n = 20$.

δ	$\alpha = 10\%$				$\alpha = 5\%$			
	ω	ω_{b1}	ω_{b2}	ω_{b3}	ω	ω_{b1}	ω_{b2}	ω_{b3}
-0,6	92,8	92,4	92,1	91,7	88,0	87,0	86,6	85,7
-0,4	74,1	73,5	73,3	72,7	61,6	60,6	60,2	59,6
-0,2	31,3	31,4	31,3	31,1	20,9	20,6	20,4	20,5
-0,1	15,3	15,2	15,2	15,2	8,8	8,7	8,7	8,8
0,1	14,8	14,9	14,8	14,9	8,5	8,5	8,4	8,5
0,2	28,8	28,9	28,8	28,8	18,3	18,1	18,1	18,1
0,4	69,9	69,8	69,6	69,4	56,7	56,1	56,0	56,1
0,6	93,1	93,0	92,9	92,7	87,6	87,3	87,0	86,8

3.5 APLICAÇÃO

Consideramos os dados de um estudo sobre a predição do conteúdo de gordura corporal a partir de medidas antropométricas comuns usando modelos de regressão linear (GARCIA et al., 2005). Essas medidas antropométricas, bem como a estimativa (Y) de massa gorda corporal (kg) por DXA (*Dual-Energy X-Ray Absorptiometry*), foram obtidas para $n = 71$ mulheres alemãs saudáveis. O modelo de regressão linear final ajustado por Garcia et al. (2005) contém os seguintes regressores: circunferência do quadril (x_2), largura do joelho (x_4), e uma variável (x_3) definida como a soma do log da dobra cutânea do queixo, do tríceps e do subescapular.

Aqui, vamos analisar esses dados usando o modelo de regressão beta prime. Inicialmente, consideramos o seguinte modelo:

$$\log \mu_l = \beta_1 + \beta_2 x_{l2} + \beta_3 x_{l3} + \beta_4 x_{l4},$$

$$\log \phi_l = \gamma_1 + \gamma_2 x_{l5},$$

em que x_5 é a soma do logaritmo de três medidas antropométricas. Esse modelo será chamado de Modelo 1. Primeiro, testamos a hipótese nula $\mathcal{H}_0 : \gamma_2 = 0$ (precisão constante). Os valores das estatísticas da razão de verossimilhanças e suas versões corrigidas,

com p -valores entre parênteses, são: $\omega = 4,7365$ (0,0295), $\omega_{b1} = 4,2243$ (0,0398), $\omega_{b2} = 4,1956$ (0,0405) e $\omega_{b3} = 4,1622$ (0,0413). Logo, tanto o teste não modificado como os testes corrigidos rejeitam a hipótese nula de precisão fixa no nível de significância 5%. Em seguida, testamos a exclusão da covariável x_4 , isto é, testamos $\mathcal{H}_0 : \beta_4 = 0$. Os valores das estatísticas são: $\omega = 2,8854$ (0,0894), $\omega_{b1} = 2,6700$ (0,1023), $\omega_{b2} = 2,6618$ (0,1028) e $\omega_{b3} = 2,6526$ (0,1034). Ou seja, os quatro testes não rejeitam a hipótese nula no nível de significância 5%. Assim, consideramos um modelo sem o regressor x_4 , a saber:

$$\log \mu_l = \beta_1 + \beta_2 x_{l2} + \beta_3 x_{l3},$$

$$\log \phi_l = \gamma_1 + \gamma_2 x_{l5}.$$

Esse modelo será chamado de Modelo 2. Testamos então a hipótese nula $\mathcal{H}_0 : \gamma_2 = 0$ no Modelo 2. Os valores das estatísticas são: $\omega = 4,1080$ (0,0427), $\omega_{b1} = 3,7032$ (0,0543), $\omega_{b2} = 3,6826$ (0,0550) e $\omega_{b3} = 3,6589$ (0,0558). Note que o teste da razão de verossimilhanças produz rejeição, enquanto que os testes corrigidos produzem não rejeição, isso no nível 5%. Seguindo a inferência dos testes corrigidos chegamos ao nosso modelo final, chamado de Modelo 3:

$$\log \mu_l = \beta_1 + \beta_2 x_{l2} + \beta_3 x_{l3},$$

$$\log \phi_l = \gamma_1.$$

Vale ressaltar que testamos exclusões individuais dos regressores x_2 e x_3 do Modelo 3, isto é, testamos $\mathcal{H}_0 : \beta_2 = 0$ e $\mathcal{H}_0 : \beta_3 = 0$. Em ambos os casos, os p -valores dos testes ω , ω_{b1} , ω_{b2} e ω_{b3} são menores que 0,0001. As estimativas dos parâmetros, erros-padrão e valores de AIC e BIC para o Modelo 3 são dados na Tabela 18.

A Figura 4 apresenta o gráfico half-normal com envelopes simulados para os resíduos quantílicos do Modelo 3. Todos os resíduos estão dentro dos envelopes. Dessa forma, concluímos que o modelo representa adequadamente os dados.

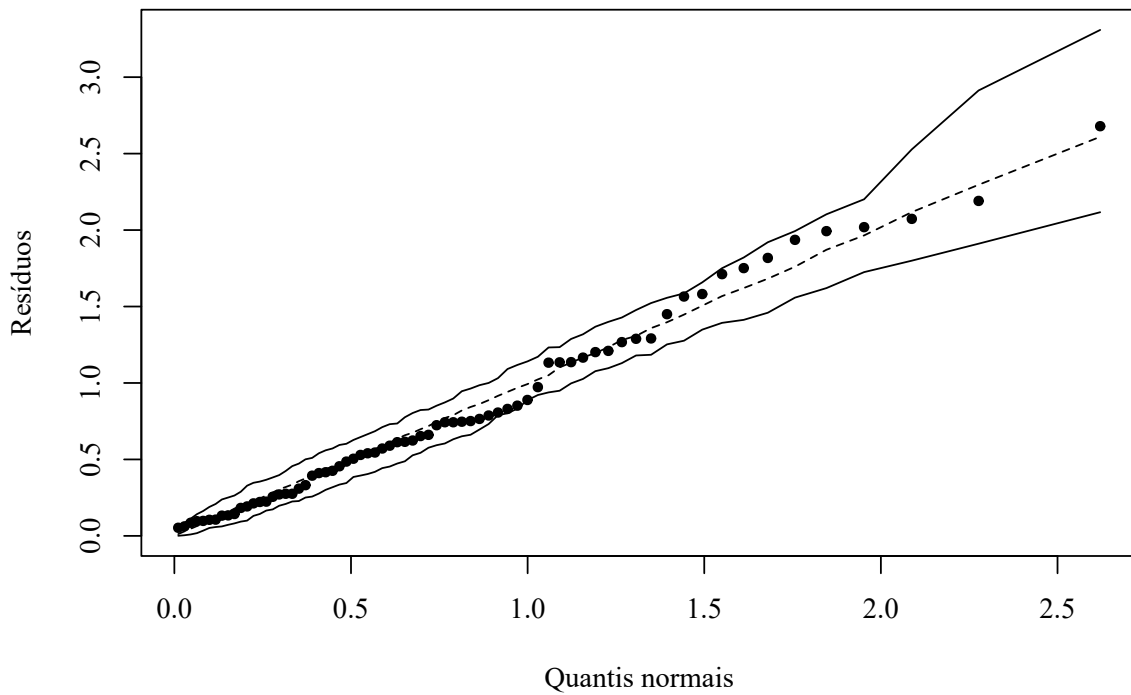
Ajustamos ainda um modelo de regressão gama (MCCULLAGH; NELDER, 1989) usando a mesma especificação para a média e precisão como no Modelo 3; esse modelo é chamado de Modelo 3b (ver Tabela 18). É interessante notar que os critérios de informação AIC e BIC favorecem o Modelo 3, uma vez que os valores desses critérios são menores para este modelo.

Tabela 18 – Estimativas dos parâmetros e valores de AIC e BIC para os Modelos 3 (beta prime) e 3b (gama), dados de gordura corporal.

Parâmetro	Modelo 3	Modelo 3b
β_1	-0,0641 (0,1412)	-0,0589 (0,1424)
β_2	0,0168 (0,0019)	0,0172 (0,0019)
β_3	0,4287 (0,0436)	0,4182 (0,0437)
γ_1	4,2402 (0,1723)	4,2121 (0,1674)
AIC	386,4081	387,8469
BIC	395,4589	396,8977

Nota: Erros-padrão são dados entre parêntesis.

Figura 4 – Gráfico half-normal residual com envelopes simulados para o Modelo 3, dados de gordura corporal.



3.6 CONCLUSÃO

O modelo de regressão beta prime com precisão variável, introduzido por Bourguignon et al. (2021), é útil para análise de variáveis respostas que assumem valores positivos. A estimação dos parâmetros de regressão nos submodelos da média e da precisão é realizada

usando o método de máxima verossimilhança, e o teste da razão de verossimilhanças é comumente empregado para realizar inferências sobre tais parâmetros. No entanto, as inferências do teste da razão de verossimilhanças podem ser enganosas quando o tamanho da amostra é pequeno. Isso ocorre porque a qualidade da aproximação qui-quadrado para a estatística de teste pode ser pobre em pequenas amostras. Neste capítulo, obtivemos uma correção de Bartlett para o teste da razão de verossimilhanças no modelo de regressão beta prime com precisão variável. Com base nesse resultado definimos três testes corrigidos para uso em análises de regressão beta prime. Os resultados de simulação de Monte Carlo mostram que o teste da razão de verossimilhanças é consideravelmente liberal, rejeitando excessivamente hipóteses nulas verdadeiras; que os três testes corrigidos superam o teste padrão; e que um dos três testes corrigidos tem desempenho tão bom quanto o do teste corrigido bootstrap, sendo que este último tem um custo computacional consideravelmente maior. Por fim, apresentamos e discutimos uma aplicação empírica.

REFERÊNCIAS

- ANDRADE, T. A. N. *Verossimilhança perfilada nos modelos lineares generalizados com superdispersão*. Dissertação (Mestrado em Estatística) — Universidade Federal de Pernambuco, Recife, 2013.
- ARAÚJO, M. C.; CYSNEIROS, A. H. M. A.; MONTENEGRO, L. C. Improved heteroskedasticity likelihood ratio tests in symmetric nonlinear regression models. *Statistical Papers*, v. 61, p. 167–188, 2020.
- BARNDORFF-NIELSEN, O. E. On a formula for the distribution of the maximum likelihood estimator. *Biometrika*, v. 70, p. 343–365, 1983.
- BARNDORFF-NIELSEN, O. E. Inference on full or partial parameters based on the standardized signed log likelihood ratio. *Biometrika*, v. 73, p. 307–322, 1986.
- BARNDORFF-NIELSEN, O. E. Modified signed log likelihood ratio. *Biometrika*, v. 78, p. 557–563, 1991.
- BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London A*, v. 160, p. 268–282, 1937.
- BIAZATTI, E. C.; CORDEIRO, G. M.; RODRIGUES, G. M.; ORTEGA, E. M. M.; DE SANTANA, L. H. A weibull-beta prime distribution to model covid-19 data with the presence of covariates and censored data. *Stats*, v. 5, p. 1159–1173, 2022.
- BOTTER, D. A.; CORDEIRO, G. M. Bartlett corrections for generalized linear models with dispersion covariates. *Communications in Statistics - Theory and Methods*, v. 26, n. 2, p. 279–307, 1997.
- BOURGUIGNON, M.; SANTOS-NETO, M.; de Castro, M. A new regression model for positive random variables with skewed and long tail. *METRON*, v. 79, p. 33–55, 2021.
- CAMPOS, R. P. S. *Aperfeiçoamento de teste em modelos lineares generalizados superdispersados*. Dissertação (Mestrado em Estatística) — Universidade Federal de Pernambuco, Recife, 2014.
- CORDEIRO, G. M. General matrix formulae for computing Bartlett corrections. *Statistics & Probability Letters*, v. 16, p. 11–18, 1993.
- CORDEIRO, G. M.; BOTTER, D. A. Second-order biases of maximum likelihood estimates in overdispersed generalized linear models. *Statistics & Probability Letters*, v. 55, p. 269–280, 2001.
- CORDEIRO, G. M.; CRIBARI-NETO, F. *An Introduction to Bartlett Correction and Bias Reduction*. New York: Springer, 2014.
- CORDEIRO, G. M.; CYSNEIROS, A. H. M. A.; CYSNEIROS, F. J. A. Bartlett adjustments for overdispersed generalized linear models. *Communications in Statistics - Theory and Methods*, v. 35, n. 5, p. 937–952, 2006.
- CORDEIRO, G. M.; PREVIDELLI, I.; SAMOBYL, R. W. Bias corrected maximum likelihood estimators in nonlinear overdispersed models. *Brazilian Journal of Probability and Statistics*, v. 22, p. 105–118, 2008.

- COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, v. 49, p. 1–39, 1987.
- COX, D. R.; REID, N. A note on the calculation of adjusted profile likelihood. *Journal of the Royal Statistical Society B*, v. 55, p. 467–471, 1993.
- CRIBARI-NETO, F.; CORDEIRO, G. M. On Bartlett and Bartlett-type corrections. *Econometric Reviews*, v. 15, p. 339–367, 1996.
- CYSNEIROS, A. H. M. A.; FERRARI, S. L. P. An improved likelihood ratio test for varying dispersion in exponential family nonlinear models. *Statistics & Probability Letters*, v. 76, n. 3, p. 255–265, 2006.
- DEY, D. K.; GELFAND, A. E.; PENG, F. Overdispersed generalized linear models. *Journal of Statistical Planning and Inference*, v. 64, p. 93–107, 1997.
- DICICCIO, T. J.; STERN, S. E. Frequentist and bayesian Bartlett correction of test statistics based on adjusted profile likelihoods. *Journal of the Royal Statistical Society B*, v. 56, p. 397–408, 1994.
- DOORNIK, J. A. *Object-oriented matrix programming using Ox*. 3rd ed. London: Timberlake Consultants Press and Oxford, 2007.
- EFRON, B. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, v. 7, p. 1 – 26, 1979.
- EFRON, B. Double exponential families and their use in generalized linear regression. *Journal of the American Statistical Association*, v. 81, p. 709–721, 1986.
- FERRARI, S. L. P.; CRIBARI-NETO, F. Corrected modified profile likelihood heteroskedasticity tests. *Statistics & Probability Letters*, v. 57, p. 353–361, 2002.
- FERRARI, S. L. P.; CYSNEIROS, A. H. M. A. Skovgaard’s adjustment to likelihood ratio tests in exponential family nonlinear models. *Statistics & Probability Letters*, v. 78, p. 3047–3055, 2008.
- FERRARI, S. L. P.; CYSNEIROS, A. H. M. A.; CRIBARI-NETO, F. An improved test for heteroskedasticity using adjusted modified profile likelihood inference. *Journal of Statistical Planning and Inference*, v. 124, p. 423–437, 2004.
- FERRARI, S. L. P.; LUCAMBIO, F.; CRIBARI-NETO, F. Improved profile likelihood inference. *Journal of Statistical Planning and Inference*, v. 134, p. 373–391, 2005.
- FERRARI, S. L. P.; PINHEIRO, E. C. Improved likelihood inference in beta regression. *Journal of Statistical Computation and Simulation*, v. 81, p. 431–443, 2011.
- FRASER, D. A. S.; REID, N. Adjustments to profile likelihood. *Biometrika*, v. 76, p. 477–488, 1989.
- FREITAS, J. V. B. D.; NOBRE, J. S.; BOURGUIGNON, M.; SANTOS-NETO, M. A new approach to modeling positive random variables with repeated measures. *Journal of Applied Statistics*, v. 49, n. 15, p. 3784–3803, 2022.

- GARCIA, A. L.; WAGNER, K.; HOTHORN, T.; KOEBNICK, C.; ZUNFT, H.-J. F.; TRIPPO, U. Improved prediction of body fat by measuring skinfold thickness, circumferences, and bone breadths. *Obesity research*, v. 13, p. 626–634, 2005.
- GUEDES, A. C.; CRIBARI-NETO, F.; ESPINHEIRA, P. L. Modified likelihood ratio tests for unit gamma regressions. *Journal of Applied Statistics*, v. 47, p. 1562–1586, 2020.
- GUEDES, A. C.; CRIBARI-NETO, F.; ESPINHEIRA, P. L. Bartlett-corrected tests for varying precision beta regressions with application to environmental biometrics. *PLOS ONE*, v. 16, n. 6, p. e0253349, 2021.
- HAYAKAWA, T. The likelihood ratio criterion and the asymptotic expansion of its distribution. *Annals of the Institute of Statistical Mathematics*, v. 29, p. 359–378, 1977.
- HINDE, J.; DEMÉTRIO, C. G. B. Overdispersion: models and estimation. *Computational Statistics & Data Analysis*, v. 27, p. 151–170, 1998.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous Univariate Distributions*. 2nd. ed. New York: Wiley, 1994.
- KAMYARI, N.; SOLTANIAN, A. R.; MAHJUB, H.; MOGHIMBEIGI, A.; SEYEDTABIB, M. Zero-augmented beta-prime model for multilevel semi-continuous data: a bayesian inference. *BMC Medical Research Methodology*, v. 22, n. 283, 2022.
- KEEPING, E. S. *Introduction to Statistical Inference*. Princeton, New Jersey: Van Nostrand, 1962.
- LAWLEY, D. N. A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika*, v. 43, p. 295–303, 1956.
- LEÃO, J.; BOURGUIGNON, M.; SAULO, H.; SANTOS-NETO, M.; CALSAVARA, V. The negative binomial beta prime regression model with cure rate: application with a melanoma dataset. *Journal of Statistical Theory and Practice*, v. 15, n. 63, 2021.
- MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. London: Chapman & Hall, 1989.
- MCCULLAGH, P.; TIBSHIRANI, R. A simple method for the adjustment of profile likelihoods. *Journal of the Royal Statistical Society B*, v. 52, p. 325–344, 1990.
- MCDONALD, J. B. Some generalized functions for the size distribution of income. *Econometrica*, v. 52, p. 647–663, 1984.
- MCDONALD, J. B. Model selection: some generalized distributions. *Communications in Statistics - Theory and Methods*, v. 16, p. 1049–1074, 1987.
- MEDEIROS, F. M. C.; ARAÚJO, M. C.; BOURGUIGNON, M. Improved estimators in beta prime regression models. *Communications in Statistics - Simulation and Computation*, v. 52, p. 5125–5138, 2023.
- MELO, T. F. N.; FERRARI, S. L. P.; PATRIOTA, A. G. Modified likelihood ratio tests in heteroskedastic multivariate regression models with measurement error. *Journal of Statistical Computation and Simulation*, v. 84, p. 2233–2247, 2014.

- MELO, T. F. N.; VARGAS, T. M.; LEMONTE, A. J.; MORENO-ARENAS, G. Higher-order asymptotic refinements in the multivariate Dirichlet regression model. *Communications in Statistics - Simulation and Computation*, v. 51, p. 53–71, 2022.
- NOCEDAL, J.; WRIGHT, S. J. *Numerical Optimization*. New York: Springer, 2006.
- RAO, C. R. *Linear Statistical Inference and its Applications*. 2nd. ed. New York: Wiley, 1973.
- RAUBER, C.; CRIBARI-NETO, F.; BAYER, F. M. Improved testing inferences for beta regressions with parametric mean link function. *AStA Advances in Statistical Analysis*, v. 104, p. 687–717, 2020.
- RIBEIRO, L. P.; VENDRAMIM, J. D.; BICALHO, K. U.; ANDRADE, M. S.; FERNANDES, J. B.; MORAL, R. A.; DEMÉTRIO, C. G. B. *Annona mucosa* jacq. (annonaceae): A promising source of bioactive compounds against sitophilus zeamais mots. (coleoptera: Curculionidae). *Journal of Stored Products Research*, v. 55, p. 6–14, 2013.
- ROCKE, D. M. Bootstrap Bartlett adjustment in seemingly unrelated regression. *Journal of the American Statistical Association*, v. 84, p. 598–601, 1989.
- RODRIGUES, H. M. *Técnicas de diagnóstico nos modelos lineares generalizados com superdispersão*. Dissertação (Mestrado em Estatística) — Universidade Federal de Pernambuco, Recife, 2013.
- SANTOS, J. S. *Correção tipo-Bartlett à estatística gradiente nos modelos lineares generalizados superdispersados*. Dissertação (Mestrado em Estatística) — Universidade Federal de Pernambuco, Recife, 2019.
- SANTOS, K. H. dos; PEREIRA, T. L.; SOUZA, T. C. de; BOURGUIGNON, M. A misspecification test for beta prime regression models. *Communications in Statistics - Simulation and Computation*, v. 52, p. 4561–4574, 2023.
- SEVERINI, T. A. *Likelihood Methods in Statistics*. Oxford University Press, 2000.
- SKOVGAARD, I. M. Likelihood asymptotics. *Scandinavian Journal of Statistics*, v. 28, p. 3–32, 2001.
- STERN, S. E. A second-order adjustment to the profile likelihood in the case of a multidimensional parameter of interest. *Journal of the Royal Statistical Society B*, v. 59, p. 653–665, 1997.
- TERRA, M. L. C. *Modelos não lineares generalizados com superdispersão*. Tese (Doutorado em Estatística) — Universidade Federal de Pernambuco, Recife, 2013.
- WASAN, M. T. *First Passage Time Distribution of Brownian Motion*. Monograph, Department of Mathematics, Kingston, Ontario: Queen's University, 1968.
- WEI, B. C. *Exponential Family Nonlinear Models*. Singapore: Springer-Verlag, 1998.
- WISE, M. E. Tracer dilution curves in cardiology and random walk and lognormal distributions. *Acta Physiologica Pharmacologica Neerlandica*, v. 14, p. 175–204, 1966.

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

- Modelo passeio aleatório

$$\begin{aligned}
 \psi^{(2,0)} &= \frac{2}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^3 (2\mu - \phi)^{1/2}}, \\
 \psi^{(3,0)} &= \frac{-2[4(2\mu - \phi)^{1/2} - (-\phi)^{1/2}]}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^4 (2\mu - \phi)^{3/2}}, \\
 \psi^{(4,0)} &= \frac{-54\phi + 96\mu - 30(-\phi)^{1/2}(2\mu - \phi)^{1/2}}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^5 (2\mu - \phi)^{5/2}}, \\
 \psi^{(0,2)} &= \frac{1}{2\phi(\phi^2 - 2\mu\phi)^{1/2}}, \quad \psi^{(0,3)} = -\frac{2\phi - 3\mu}{2\phi^2(\phi - 2\mu)(\phi^2 - 2\mu\phi)^{1/2}}, \\
 \psi^{(0,4)} &= \frac{6\phi^2 - 18\mu\phi + 15\mu^2}{2\phi^3(-2\mu + \phi)^2(\phi^2 - 2\mu\phi)^{1/2}}, \\
 \psi^{(1,1)} &= \frac{1}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^2 (-\phi)^{1/2} (2\mu - \phi)^{1/2}}, \\
 \psi^{(1,2)} &= \frac{1}{2(2\mu - \phi)^{3/2} (-\phi)^{3/2}}, \quad \psi^{(2,2)} = \frac{-3}{2(-\phi)^{3/2} (2\mu - \phi)^{5/2}}, \\
 \psi^{(2,1)} &= \frac{-2[2\phi - 3\mu + 2(\phi^2 - 2\mu\phi)^{1/2}]}{(\phi - 2\mu)\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^4 (\phi^2 - 2\mu\phi)^{1/2}}, \\
 \psi^{(0,1)} &= \frac{2[(-\phi)^{1/2}(2\mu - \phi)^{1/2} - \mu + \phi]}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^3 (-\phi)^{1/2}}, \\
 \psi^{(1,0)} &= -\frac{1}{\{(2\mu - \phi)^{1/2} - (-\phi)^{1/2}\}^2}.
 \end{aligned}$$

- Modelo Poisson duplo

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

- Modelo binomial duplo

$$\psi^{(1,2)} = 0, \quad \psi^{(2,2)} = 0, \quad \psi^{(0,2)} = -\frac{1}{2\phi^2}, \quad \psi^{(0,3)} = \frac{1}{\phi^3}, \quad \psi^{(0,4)} = -\frac{3}{\phi^4},$$

$$\psi^{(1,0)} = m\phi \log\left(\frac{\mu}{1-\mu}\right), \quad \psi^{(2,0)} = \frac{m\phi}{\mu(1-\mu)}, \quad \psi^{(1,1)} = m \log\left(\frac{\mu}{1-\mu}\right),$$

$$\psi^{(2,1)} = \frac{m}{\mu(1-\mu)}, \quad \psi^{(0,1)} = m[\mu \log \mu + (1-\mu) \log(1-\mu)] + \frac{1}{2\phi},$$

$$\psi^{(3,0)} = \frac{m\phi(2\mu-1)}{(\mu(1-\mu))^2}, \quad \psi^{(4,0)} = \frac{m\phi(6\mu^2-6\mu+2)}{(\mu(1-\mu))^3}.$$

APÊNDICE B – OBTENÇÃO DAS QUANTIDADES PROPOSTAS POR SKOVGAARD

Este apêndice contém detalhes sobre a derivação das expressões para $\bar{\rho}$ e $\bar{\Upsilon}$ nos MNLGS. Lembre-se que $\theta = (\beta^\top, \gamma^\top)^\top$. Começamos, calculando as seguintes quantidades:

$$\begin{aligned}\rho &= \mathbb{E}_{\theta_0}[\mathbf{U}(\theta_0)(\ell(\theta_0) - \ell(\theta))] \\ &= \begin{bmatrix} \mathbb{E}_{\theta_0}[\mathbf{U}_\beta(\theta_0)\ell(\theta_0)] - \mathbb{E}_{\theta_0}[\mathbf{U}_\beta(\theta_0)\ell(\theta)] \\ \mathbb{E}_{\theta_0}[\mathbf{U}_\gamma(\theta_0)\ell(\theta_0)] - \mathbb{E}_{\theta_0}[\mathbf{U}_\gamma(\theta_0)\ell(\theta)] \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix},\end{aligned}\quad (\text{B.1})$$

$$\begin{aligned}\Upsilon &= \mathbb{E}_{\theta_0}[\mathbf{U}(\theta_0)\mathbf{U}^\top(\theta)] \\ &= \begin{bmatrix} \mathbb{E}_{\theta_0}[\mathbf{U}_\beta(\theta_0)\mathbf{U}_\beta^\top(\theta)] & \mathbb{E}_{\theta_0}[\mathbf{U}_\beta(\theta_0)\mathbf{U}_\gamma^\top(\theta)] \\ \mathbb{E}_{\theta_0}[\mathbf{U}_\gamma(\theta_0)\mathbf{U}_\beta^\top(\theta)] & \mathbb{E}_{\theta_0}[\mathbf{U}_\gamma(\theta_0)\mathbf{U}_\gamma^\top(\theta)] \end{bmatrix}.\end{aligned}\quad (\text{B.2})$$

Em (B.1), o i -ésimo componente do vetor ρ_1 ($p \times 1$) é

$$\begin{aligned}\rho_{1i} &= \mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)\ell(\theta_0)] - \mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)\ell(\theta)] \\ &= \mathbb{E}_{\theta_0} \left[U_{\beta_i}(\theta_0) \sum_{l=1}^n \left\{ (Y_l - \mu_l^0) \psi_l^{(1,0)0} + \phi_l^0 T(Y_l) + \psi_l^0 + \log A(Y_l) \right\} \right] \\ &\quad - \mathbb{E}_{\theta_0} \left[U_{\beta_i}(\theta_0) \sum_{l=1}^n \left\{ (Y_l - \mu_l) \psi_l^{(1,0)} + \phi_l T(Y_l) + \psi_l + \log A(Y_l) \right\} \right] \\ &= \sum_{l=1}^n \left\{ \mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)Y_l](\psi_l^{(1,0)0} - \psi_l^{(1,0)}) + \mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)T(Y_l)](\phi_l^0 - \phi_l) \right\},\end{aligned}\quad (\text{B.3})$$

pois $\mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)] = 0$.

Por definição $\mathbb{E}_{\theta_0}(Y) = \int y \pi(y; \theta_0) dy$ e $\mathbb{E}_{\theta_0}[T(Y)] = \int T(y) \pi(y; \theta_0) dy$, em que $\pi(y; \theta_0)$ é a função densidade de Y . Além disso, sabemos que $\mathbb{E}_{\theta_0}(Y) = \mu^0$ e $\mathbb{E}_{\theta_0}[T(Y)] = -\psi^{(0,1)0}$. Assim,

$$\int y \pi(y; \theta_0) dy = \mu^0 \quad \text{e} \quad \int T(y) \pi(y; \theta_0) dy = -\psi^{(0,1)0}. \quad (\text{B.4})$$

Derivando ambos os lados das equações em (B.4) com respeito a β_i^0 , e permutando as operações de diferenciação e de integração, obtivemos, respectivamente,

$$\mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)Y_l] = \frac{d\mu_l^0}{d\eta_l^0} \frac{\partial \eta_l^0}{\partial \beta_i^0} = m_{1l}^0 \frac{\partial \eta_l^0}{\partial \beta_i^0} \quad \text{e} \quad \mathbb{E}_{\theta_0}[U_{\beta_i}(\theta_0)T(Y_l)] = -\psi_l^{(1,1)0} m_{1l}^0 \frac{\partial \eta_l^0}{\partial \beta_i^0}. \quad (\text{B.5})$$

Substituindo as expressões em (B.5) na Equação (B.3), segue que

$$\rho_{1_i} = \sum_{l=1}^n \frac{\partial \eta_l^0}{\partial \beta_i^0} m_{1l}^0 \left\{ \psi_l^{(1,0)0} - \psi_l^{(1,0)} - \psi_l^{(1,1)0} (\phi_l^0 - \phi_l) \right\}. \quad (\text{B.6})$$

Ainda em (B.1), considere o vetor $\boldsymbol{\rho}_2$ ($q \times 1$). Seu j -ésimo elemento é dado por

$$\begin{aligned} \rho_{2_j} &= \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)\ell(\boldsymbol{\theta}_0)] - \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)\ell(\boldsymbol{\theta})] \\ &= \sum_{l=1}^n \left\{ \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)Y_l](\psi_l^{(1,0)0} - \psi_l^{(1,0)}) + \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)T(Y_l)](\phi_l^0 - \phi_l) \right\}, \end{aligned} \quad (\text{B.7})$$

pois $\mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)] = 0$. Derivando as equações em (B.4) em relação a γ_j^0 , obtivemos

$$\mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)Y_l] = 0 \quad \text{e} \quad \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_j}(\boldsymbol{\theta}_0)T(Y_l)] = -\psi_l^{(0,2)0} \frac{d\phi_l^0}{d\tau_l^0} \frac{\partial \tau_l^0}{\partial \gamma_j^0} = -\psi_l^{(0,2)0} \phi_{1l}^0 \frac{\partial \tau_l^0}{\partial \gamma_j^0}. \quad (\text{B.8})$$

Agora, substituindo as expressões em (B.8) na Equação (B.7), temos

$$\rho_{2_j} = \sum_{l=1}^n \frac{\partial \tau_l^0}{\partial \gamma_j^0} \phi_{1l}^0 \psi_l^{(0,2)0} (\phi_l - \phi_l^0). \quad (\text{B.9})$$

Então, usando (B.6), (B.9), e notação matricial, obtivemos

$$\boldsymbol{\rho} = \begin{bmatrix} \mathbf{X}^{*\top 0} \mathbf{M}_1^0 \left\{ \boldsymbol{\Psi}^{(1,0)0} - \boldsymbol{\Psi}^{(1,0)} - \boldsymbol{\Psi}^{(1,1)0} (\boldsymbol{\Phi}^0 - \boldsymbol{\Phi}) \right\} \mathbf{1} \\ \mathbf{S}^{*\top 0} \boldsymbol{\Phi}_1^0 \boldsymbol{\Psi}^{(0,2)0} (\boldsymbol{\Phi} - \boldsymbol{\Phi}^0) \mathbf{1} \end{bmatrix}. \quad (\text{B.10})$$

Agora vamos considerar (B.2). O elemento (i, j) da matriz $\mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}_{\beta}(\boldsymbol{\theta}_0)\mathbf{U}_{\beta}^{\top}(\boldsymbol{\theta})]$ é dado por

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_0} \left[U_{\beta_i}(\boldsymbol{\theta}_0) \sum_{l=1}^n \frac{\partial \eta_l}{\partial \beta_j} m_{1l} \psi_l^{(2,0)}(Y_l - \mu_l) \right] \\ = \sum_{l=1}^n \frac{\partial \eta_l}{\partial \beta_j} m_{1l} \psi_l^{(2,0)} \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\beta_i}(\boldsymbol{\theta}_0)Y_l] = \sum_{l=1}^n \frac{\partial \eta_l^0}{\partial \beta_i^0} m_{1l}^0 \psi_l^{(2,0)} m_{1l} \frac{\partial \eta_l}{\partial \beta_j}, \end{aligned} \quad (\text{B.11})$$

para $i, j = 1, \dots, p$.

Já o elemento (a, b) da matriz $\mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}_{\gamma}(\boldsymbol{\theta}_0)\mathbf{U}_{\gamma}^{\top}(\boldsymbol{\theta})]$ de dimensão $q \times q$ é

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_0} \left[U_{\gamma_a}(\boldsymbol{\theta}_0) \sum_{l=1}^n \frac{\partial \tau_l}{\partial \gamma_b} \phi_{1l} \left\{ \psi_l^{(1,1)}(Y_l - \mu_l) + T(Y_l) + \psi_l^{(0,1)} \right\} \right] \\ = \sum_{l=1}^n \frac{\partial \tau_l}{\partial \gamma_b} \phi_{1l} \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\gamma_a}(\boldsymbol{\theta}_0)T(Y_l)] = - \sum_{l=1}^n \frac{\partial \tau_l^0}{\partial \gamma_a^0} \phi_{1l}^0 \psi_l^{(0,2)0} \phi_{1l} \frac{\partial \tau_l}{\partial \gamma_b}. \end{aligned} \quad (\text{B.12})$$

Usando o fato de que $\mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}_{\gamma}(\boldsymbol{\theta}_0)Y] = \mathbf{0}$, é possível mostrar que

$$\mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}_{\gamma}(\boldsymbol{\theta}_0)\mathbf{U}_{\beta}^{\top}(\boldsymbol{\theta})] = \mathbf{0}, \quad (\text{B.13})$$

cujas dimensões são $q \times p$.

Ainda temos que o elemento (i, a) da matriz $\mathbb{E}_{\boldsymbol{\theta}_0}[\mathbf{U}_{\beta}(\boldsymbol{\theta}_0)\mathbf{U}_{\gamma}^{\top}(\boldsymbol{\theta})]$ é dado por

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}_0} & \left[U_{\beta_i}(\boldsymbol{\theta}_0) \sum_{l=1}^n \frac{\partial \tau_l}{\partial \gamma_a} \phi_{1l} \left\{ \psi_l^{(1,1)}(Y_l - \mu_l) + T(Y_l) + \psi_l^{(0,1)} \right\} \right] \\ &= \sum_{l=1}^n \frac{\partial \tau_l}{\partial \gamma_a} \phi_{1l} \left\{ \psi_l^{(1,1)} \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\beta_i}(\boldsymbol{\theta}_0)Y_l] + \mathbb{E}_{\boldsymbol{\theta}_0}[U_{\beta_i}(\boldsymbol{\theta}_0)T(Y_l)] \right\} \\ &= \sum_{l=1}^n \frac{\partial \eta_l^0}{\partial \beta_i^0} m_{1l}^0 \left\{ \psi_l^{(1,1)} - \psi_l^{(1,1)0} \right\} \phi_{1l} \frac{\partial \tau_l}{\partial \gamma_a}, \end{aligned} \quad (\text{B.14})$$

para $i = 1, \dots, p$ e $a = 1, \dots, q$.

Assim, a partir de (B.11)–(B.14) e usando notação matricial, obtivemos

$$\boldsymbol{\Upsilon} = \begin{bmatrix} \mathbf{X}^{*\top 0} \mathbf{M}_1^0 \boldsymbol{\Psi}^{(2,0)} \mathbf{M}_1 \mathbf{X}^* & \mathbf{X}^{*\top 0} \mathbf{M}_1^0 \left\{ \boldsymbol{\Psi}^{(1,1)} - \boldsymbol{\Psi}^{(1,1)0} \right\} \boldsymbol{\Phi}_1 \mathbf{S}^* \\ \mathbf{0} & -\mathbf{S}^{*\top 0} \boldsymbol{\Phi}_1^0 \boldsymbol{\Psi}^{(0,2)0} \boldsymbol{\Phi}_1 \mathbf{S}^* \end{bmatrix}. \quad (\text{B.15})$$

Finalmente, em (B.10) e (B.15), avaliamos as quantidades com sobrescrito zero na E.M.V. irrestrita e as demais na E.M.V. restrita. Isso conduz às expressões (1.6) e (1.7).

APÊNDICE C – OBTENÇÃO DA CORREÇÃO DE BARTLETT PARA O TESTE DE VEROSSIMILHANÇA PERFILADA MODIFICADA

Este apêndice fornece alguns detalhes sobre a derivação da expressão para c_m nos MNLGS. O interesse consiste no teste da hipótese nula $\mathcal{H}_0 : \gamma = \gamma^{(0)}$ contra a hipótese alternativa $\mathcal{H}_1 : \gamma \neq \gamma^{(0)}$. Os cumulantes necessários são

$$\begin{aligned}
\lambda_{ab} &= \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l}^2(a)_l(b)_l, \\
\lambda_{abc} &= \sum_{l=1}^n \left[\left(\psi_l^{(0,3)} \phi_{1l}^3 + 3\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) (a)_l(b)_l(c)_l \right. \\
&\quad \left. + \psi_l^{(0,2)} \phi_{1l}^2 \{ (bc)_l(a)_l + (ac)_l(b)_l + (ab)_l(c)_l \} \right], \\
(\lambda_{ab})_c &= \sum_{l=1}^n \left[\left(\psi_l^{(0,3)} \phi_{1l}^3 + 2\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) (a)_l(b)_l(c)_l + \psi_l^{(0,2)} \phi_{1l}^2 \{ (bc)_l(a)_l + (ac)_l(b)_l \} \right], \\
\lambda_{abcd} &= \sum_{l=1}^n \left[\left(\psi_l^{(0,4)} \phi_{1l}^4 + 6\psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 3\psi_l^{(0,2)} \phi_{2l}^2 + 4\psi_l^{(0,2)} \phi_{1l} \phi_{3l} \right) (a)_l(b)_l(c)_l(d)_l \right. \\
&\quad + \left(\psi_l^{(0,3)} \phi_{1l}^3 + 3\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) \{ (cd)_l(a)_l(b)_l + (bd)_l(a)_l(c)_l + (bc)_l(a)_l(d)_l \\
&\quad + (ad)_l(b)_l(c)_l + (ac)_l(b)_l(d)_l + (ab)_l(c)_l(d)_l \} \\
&\quad + \psi_l^{(0,2)} \phi_{1l}^2 \{ (bcd)_l(a)_l + (acd)_l(b)_l + (abd)_l(c)_l + (abc)_l(d)_l + (ab)_l(cd)_l \\
&\quad + (ac)_l(bd)_l + (ad)_l(bc)_l \} \Big], \\
(\lambda_{abc})_d &= \sum_{l=1}^n \left[\left(\psi_l^{(0,4)} \phi_{1l}^4 + 6\psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 3\psi_l^{(0,2)} \phi_{2l}^2 + 3\psi_l^{(0,2)} \phi_{1l} \phi_{3l} \right) (a)_l(b)_l(c)_l(d)_l \right. \\
&\quad + \left(\psi_l^{(0,3)} \phi_{1l}^3 + 3\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) \{ (cd)_l(a)_l(b)_l + (bd)_l(a)_l(c)_l + (ad)_l(b)_l(c)_l \} \\
&\quad + \left(\psi_l^{(0,3)} \phi_{1l}^3 + 2\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) \{ (bc)_l(a)_l(d)_l + (ac)_l(b)_l(d)_l + (ab)_l(c)_l(d)_l \} \\
&\quad + \psi_l^{(0,2)} \phi_{1l}^2 \{ (bcd)_l(a)_l + (acd)_l(b)_l + (abd)_l(c)_l + (ab)_l(cd)_l + (ac)_l(bd)_l \\
&\quad + (bc)_l(ad)_l \} \Big], \\
(\lambda_{ab})_{cd} &= \sum_{l=1}^n \left[\left(\psi_l^{(0,4)} \phi_{1l}^4 + 5\psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} + 2\psi_l^{(0,2)} \phi_{2l}^2 + 2\psi_l^{(0,2)} \phi_{1l} \phi_{3l} \right) (a)_l(b)_l(c)_l(d)_l \right. \\
&\quad + \left(\psi_l^{(0,3)} \phi_{1l}^3 + 2\psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right) \{ (bd)_l(a)_l(c)_l + (bc)_l(a)_l(d)_l + (ad)_l(b)_l(c)_l \\
&\quad + (ac)_l(b)_l(d)_l + (cd)_l(a)_l(b)_l \} \\
&\quad + \psi_l^{(0,2)} \phi_{1l}^2 \{ (bcd)_l(a)_l + (acd)_l(b)_l + (ad)_l(bc)_l + (ac)_l(bd)_l \} \Big],
\end{aligned}$$

em que $(a)_l = \partial \tau_l / \partial \gamma_a$, $(ab)_l = \partial^2 \tau_l / \partial \gamma_a \partial \gamma_b$ e $(abc)_l = \partial^3 \tau_l / \partial \gamma_a \partial \gamma_b \partial \gamma_c$.

Uma vez que $c_m = c_{m1} - c_{m2} - c_{m3}$, precisamos calcular as quantidades c_{mj} , para $j = 1, 2, 3$. Temos que

$$c_{m1} = \lambda^{ab} \lambda^{cd} \left(\frac{1}{4} \lambda_{abcd} - (\lambda_{acd})_b + (\lambda_{ac})_{db} \right). \quad (\text{C.1})$$

Note que, a expressão acima possui um somatório implícito sobre os índices $a, b, c, d = 1, \dots, q$. É mais conveniente começar calculando o termo entre parênteses em (C.1), isto é,

$$\begin{aligned} & \frac{1}{4} \lambda_{abcd} - (\lambda_{acd})_b + (\lambda_{ac})_{db} \\ &= \frac{1}{4} \sum_{l=1}^n \left[\left(\psi_l^{(0,4)} \phi_{1l}^4 + 2\psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} - \psi_l^{(0,2)} \phi_{2l}^2 \right) (a)_l (b)_l (c)_l (d)_l \right. \\ & \quad + \psi_l^{(0,3)} \phi_{1l}^3 \left\{ (bd)_l (a)_l (c)_l + (bc)_l (a)_l (d)_l + (ab)_l (c)_l (d)_l + (cd)_l (a)_l (b)_l \right. \\ & \quad \left. + (ad)_l (b)_l (c)_l - 3(ac)_l (b)_l (d)_l \right\} \\ & \quad - \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \left\{ (bd)_l (a)_l (c)_l + (bc)_l (a)_l (d)_l + (ab)_l (c)_l (d)_l - 3(cd)_l (a)_l (b)_l \right. \\ & \quad \left. - 3(ad)_l (c)_l (b)_l + 5(ac)_l (b)_l (d)_l \right\} \\ & \quad \left. + \psi_l^{(0,2)} \phi_{1l}^2 \left\{ (bcd)_l (a)_l + (abd)_l (c)_l + (ad)_l (bc)_l + (ab)_l (cd)_l - 3(abc)_l (d)_l \right. \right. \\ & \quad \left. \left. - 3(ac)_l (bd)_l + (acd)_l (b)_l \right\} \right]. \quad (\text{C.2}) \end{aligned}$$

Substituindo (C.2) em (C.1), e rearranjando os somatórios, obtivemos, depois da simplificação de alguns termos,

$$\begin{aligned} c_{m1} &= \frac{1}{4} \sum_{l=1}^n \left\{ \psi_l^{(0,4)} \phi_{1l}^4 + 2\psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} - \psi_l^{(0,2)} \phi_{2l}^2 \right\} \left(\sum_{a,b=1}^q (a)_l \lambda^{ab} (b)_l \right)^2 \\ & \quad + \frac{1}{2} \sum_{l=1}^n \left\{ \psi_l^{(0,3)} \phi_{1l}^3 + \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right\} \left(\sum_{a,b=1}^q (ab)_l \lambda^{ab} \right) \left(\sum_{c,d=1}^q (c)_l \lambda^{cd} (d)_l \right) \\ & \quad - \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \left(\sum_{a,b,c,d=1}^q (a)_l \lambda^{ab} (bc)_l \lambda^{cd} (d)_l \right) \\ & \quad - \frac{1}{2} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \left(\sum_{a,b,c,d=1}^q (ad)_l \lambda^{dc} (cb)_l \lambda^{ba} \right) + \frac{1}{4} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \left(\sum_{a,b=1}^q (ab)_l \lambda^{ab} \right)^2. \end{aligned}$$

Notamos que, $\sum_{a,b=1}^q (a)_l \lambda^{ab} (b)_l = -z_l$, $\sum_{a,b=1}^q (ab)_l \lambda^{ab} = -d_l$, $\sum_{a,b,c,d=1}^q (ad)_l \lambda^{dc} (cb)_l \lambda^{ba} = b_l$ e $\sum_{a,b,c,d=1}^q (a)_l \lambda^{ab} (bc)_l \lambda^{cd} (d)_l = a_l^{(l)}$. Desse modo, segue que

$$\begin{aligned}
c_{m1} = & \frac{1}{4} \sum_{l=1}^n \psi_l^{(0,4)} \phi_{1l}^4 z_{ll}^2 + \frac{1}{2} \sum_{l=1}^n \psi_l^{(0,3)} \phi_{1l}^2 \phi_{2l} z_{ll}^2 - \frac{1}{4} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{2l}^2 z_{ll}^2 \\
& + \frac{1}{2} \sum_{l=1}^n \psi_l^{(0,3)} \phi_{1l}^3 d_l z_{ll} + \frac{1}{2} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} d_l z_{ll} \\
& - \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} a_{ll}^{(l)} - \frac{1}{2} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l}^2 b_{ll} + \frac{1}{4} \sum_{l=1}^n \psi_l^{(0,2)} \phi_{1l}^2 d_l^2.
\end{aligned}$$

Escrevendo a expressão acima em notação matricial, encontramos (2.5).

A quantidade c_{m2} é dada por

$$c_{m2} = \lambda^{ab} \lambda^{cd} \lambda^{ef} \left[\lambda_{acd} \left(\frac{1}{4} \lambda_{bef} - (\lambda_{be})_f \right) + (\lambda_{ac})_d (\lambda_{be})_f \right]. \quad (C.3)$$

Temos que

$$\begin{aligned}
& \lambda_{acd} \left(\frac{1}{4} \lambda_{bef} - (\lambda_{be})_f \right) + (\lambda_{ac})_d (\lambda_{be})_f \\
& = \frac{1}{4} \sum_{l,m=1}^n \left[\left(\psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 + 3 \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} - \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right. \right. \\
& \quad \times \left. \psi_m^{(0,3)} \phi_{1m}^3 + \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \right) (a)_l (c)_l (d)_l (b)_m (e)_m (f)_m \\
& \quad + \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m}^2 (a)_l (c)_l (d)_l \left\{ (ef)_m (b)_m + (bf)_m (e)_m + (be)_m (f)_m \right\} \\
& \quad - \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 (a)_l (c)_l (d)_l \left\{ (ef)_m (b)_m + (bf)_m (e)_m - 3(be)_m (f)_m \right\} \\
& \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,3)} \phi_{1m}^3 \left\{ (cd)_l (a)_l + (ad)_l (c)_l - 3(ac)_l (d)_l \right\} (b)_m (e)_m (f)_m \\
& \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \left\{ 3(cd)_l (a)_l + 3(ad)_l (c)_l - 5(ac)_l (d)_l \right\} (b)_m (e)_m (f)_m \\
& \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 \left\{ (cd)_l (a)_l + (ad)_l (c)_l \right\} \left\{ (ef)_m (b)_m + (bf)_m (e)_m + (be)_m (f)_m \right\} \\
& \quad \left. - \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 (ac)_l (d)_l \left\{ 3(ef)_m (b)_m + 3(bf)_m (e)_m - (be)_m (f)_m \right\} \right].
\end{aligned}$$

Substituindo a expressão acima em (C.3), rearranjando os somatórios e simplificando alguns termos, obtivemos

$$\begin{aligned}
c_{m2} = & \frac{1}{4} \sum_{l,m=1}^n \left\{ \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 + 2 \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} + \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \right\} \\
& \times \left(\sum_{a,b=1}^q (a)_l \lambda^{ab} (b)_m \right) \left(\sum_{c,d=1}^q (c)_l \lambda^{cd} (d)_l \right) \left(\sum_{e,f=1}^q (e)_m \lambda^{ef} (f)_m \right) \\
& + \frac{1}{2} \sum_{l,m=1}^n \left\{ \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m}^2 + \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 \right\} \\
& \times \left(\sum_{a,b=1}^q (a)_l \lambda^{ab} (b)_m \right) \left(\sum_{c,d=1}^q (c)_l \lambda^{cd} (d)_l \right) \left(\sum_{e,f=1}^q (ef)_m \lambda^{ef} \right) \\
& + \frac{1}{4} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 \left(\sum_{a,b=1}^q (a)_l \lambda^{ab} (b)_m \right) \left(\sum_{c,d=1}^q (cd)_l \lambda^{cd} \right) \left(\sum_{e,f=1}^q (ef)_m \lambda^{ef} \right).
\end{aligned}$$

Notamos que, $\sum_{a,b=1}^q (a)_l \lambda^{ab}(b)_m = -z_{lm}$. Assim,

$$\begin{aligned} c_{m2} = & -\frac{1}{4} \sum_{l,m=1}^n \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 z_{ll} z_{lm} z_{mm} - \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} z_{ll} z_{lm} z_{mm} \\ & - \frac{1}{4} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} z_{ll} z_{lm} z_{mm} - \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m}^2 z_{ll} z_{lm} d_m \\ & - \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 z_{ll} z_{lm} d_m - \frac{1}{4} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 d_l z_{lm} d_m. \end{aligned}$$

Expressando o resultado acima em notação matricial, encontramos (2.6).

A última quantidade é

$$c_{m3} = \lambda^{ab} \lambda^{cf} \lambda^{de} \left[\lambda_{acd} \left(\frac{1}{6} \lambda_{bef} - (\lambda_{be})_f \right) + (\lambda_{ac})_d (\lambda_{be})_f \right]. \quad (C.4)$$

Temos que

$$\begin{aligned} & \lambda_{acd} \left(\frac{1}{6} \lambda_{bef} - (\lambda_{be})_f \right) + (\lambda_{ac})_d (\lambda_{be})_f \\ &= \frac{1}{6} \sum_{l,m=1}^n \left[\left(\psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 + 3 \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} - 3 \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \right. \right. \\ & \quad \times \left. \left. \psi_m^{(0,3)} \phi_{1m}^3 - 3 \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \right) (a)_l (c)_l (d)_l (b)_m (e)_m (f)_m \right. \\ & \quad + \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,2)} \phi_{1m}^2 (a)_l (c)_l (d)_l \{ (ef)_m (b)_m + (bf)_m (e)_m + (be)_m (f)_m \} \\ & \quad - 3 \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 (a)_l (c)_l (d)_l \{ (ef)_m (b)_m + (bf)_m (e)_m - (be)_m (f)_m \} \\ & \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,3)} \phi_{1m}^3 \{ (cd)_l (a)_l + (ad)_l (c)_l - 5(ac)_l (d)_l \} (b)_m (e)_m (f)_m \\ & \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \{ 3(cd)_l (a)_l + 3(ad)_l (c)_l - 9(ac)_l (d)_l \} (b)_m (e)_m (f)_m \\ & \quad + \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 \{ (cd)_l (a)_l + (ad)_l (c)_l \} \{ (ef)_m (b)_m + (bf)_m (e)_m \\ & \quad + (be)_m (f)_m \} \\ & \quad \left. - \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 (ac)_l (d)_l \{ 5(ef)_m (b)_m + 5(bf)_m (e)_m - (be)_m (f)_m \} \right]. \end{aligned}$$

Agora, substituindo a expressão acima em (C.4), obtivemos

$$\begin{aligned} c_{m3} = & \frac{1}{6} \sum_{l,m=1}^n \left\{ \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 - 3 \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} \right\} \\ & \times \left(\sum_{a,b=1}^q (a)_l \lambda^{ab}(b)_m \right) \left(\sum_{c,f=1}^q (c)_l \lambda^{cf}(f)_m \right) \left(\sum_{d,e=1}^q (d)_l \lambda^{de}(e)_m \right) \\ & - \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 \left(\sum_{a,b=1}^q (a)_l \lambda^{ab}(b)_m \right) \left(\sum_{c,d,e,f=1}^q (c)_l \lambda^{cf}(fe)_m \lambda^{ed}(d)_l \right) \\ & - \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 \left(\sum_{a,b=1}^q (a)_l \lambda^{ab}(b)_m \right) \left(\sum_{c,d,e,f=1}^q (cd)_l \lambda^{de}(ef)_m \lambda^{fc} \right). \end{aligned}$$

Notamos que, $\sum_{c,d,e,f=1}^q (c)_l \lambda^{cf} (fe)_m \lambda^{ed} (d)_l = a_{ll}^{(m)}$ e $\sum_{c,d,e,f=1}^q (cd)_l \lambda^{de} (ef)_m \lambda^{fc} = b_{lm}$.

Dessa forma, temos

$$\begin{aligned} c_{m3} = & -\frac{1}{6} \sum_{l,m=1}^n \psi_l^{(0,3)} \phi_{1l}^3 \psi_m^{(0,3)} \phi_{1m}^3 z_{lm}^3 + \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m} \phi_{2m} z_{lm}^3 \\ & + \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l} \phi_{2l} \psi_m^{(0,2)} \phi_{1m}^2 z_{lm} a_{ll}^{(m)} + \frac{1}{2} \sum_{l,m=1}^n \psi_l^{(0,2)} \phi_{1l}^2 \psi_m^{(0,2)} \phi_{1m}^2 z_{lm} b_{lm}. \end{aligned}$$

Em notação matricial, resulta em (2.7).

APÊNDICE D – CUMULANTES NO MODELO DE REGRESSÃO BETA PRIME

Neste apêndice, apresentamos os cumulantes para o modelo de regressão beta prime com precisão variável. Utilizamos a notação padrão para os cumulantes de derivadas da log-verossimilhança com respeito aos componentes de β e γ , com índices minúsculos e maiúsculos denotando elementos de β e γ , respectivamente: $\kappa_{rs} = \mathbb{E}(\partial^2 \ell(\beta, \gamma) / \partial \beta_r \partial \beta_s)$, $\kappa_{RS} = \mathbb{E}(\partial^2 \ell(\beta, \gamma) / \partial \gamma_R \partial \gamma_S)$, $\kappa_{rS} = \mathbb{E}(\partial^2 \ell(\beta, \gamma) / \partial \beta_r \partial \gamma_S)$, $\kappa_{rsT} = \mathbb{E}(\partial^3 \ell(\beta, \gamma) / \partial \beta_r \partial \beta_s \partial \gamma_T)$, $\kappa_{rsTUV} = \mathbb{E}(\partial^4 \ell(\beta, \gamma) / \partial \beta_r \partial \beta_s \partial \gamma_T \partial \gamma_U)$, etc. Adicionalmente, denotamos as derivadas dos cumulantes por: $\kappa_{rs}^{(t)} = \partial \kappa_{rs} / \partial \beta_t$, $\kappa_{RS}^{(T)} = \partial \kappa_{RS} / \partial \gamma_T$, $\kappa_{rS}^{(t)} = \partial \kappa_{rS} / \partial \beta_t$, $\kappa_{rS}^{(tU)} = \partial^2 \kappa_{rS} / \partial \beta_t \partial \gamma_U$, $\kappa_{rsT}^{(U)} = \partial \kappa_{rsT} / \partial \gamma_U$, etc.

Aqui, definimos algumas quantidades que foram utilizadas na obtenção dos cumulantes. Sejam

$$\begin{aligned} a_l &= \psi^{(1)}[\mu_l(1 + \phi_l)] - \psi^{(1)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ b_l &= (1 + \phi_l)\{\mu_l a_l - \psi^{(1)}[\mu_l(1 + \phi_l) + \phi_l + 2]\}, \\ c_l &= \mu_l^2 \psi^{(1)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^2 \psi^{(1)}[\mu_l(1 + \phi_l) + \phi_l + 2] + \psi^{(1)}(\phi_l + 2), \\ d_l &= \psi^{(2)}[\mu_l(1 + \phi_l)] - \psi^{(2)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ e_l &= -(1 + \phi_l) \left[2a_l + (1 + \phi_l) \frac{\partial a_l}{\partial \phi_l} \right]. \end{aligned}$$

Também utilizamos as seguintes derivadas: $\partial a_l / \partial \mu_l = (1 + \phi_l) d_l$,

$$\begin{aligned} \frac{\partial a_l}{\partial \phi_l} &= \mu_l \psi^{(2)}[\mu_l(1 + \phi_l)] - (1 + \mu_l) \psi^{(2)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ \frac{\partial \mu_l^*}{\partial \phi_l} &= \mu_l \psi^{(1)}[\mu_l(1 + \phi_l)] - (1 + \mu_l) \psi^{(1)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ \frac{\partial d_l}{\partial \mu_l} &= (1 + \phi_l)\{\psi^{(3)}[\mu_l(1 + \phi_l)] - \psi^{(3)}[\mu_l(1 + \phi_l) + \phi_l + 2]\}, \\ \frac{\partial d_l}{\partial \phi_l} &= \mu_l \psi^{(3)}[\mu_l(1 + \phi_l)] - (1 + \mu_l) \psi^{(3)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ \frac{\partial^2 a_l}{\partial \phi_l^2} &= \mu_l^2 \psi^{(3)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^2 \psi^{(3)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ \frac{\partial^2 \mu_l^*}{\partial \phi_l^2} &= \mu_l^2 \psi^{(2)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^2 \psi^{(2)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\ \frac{\partial^2 \mu_l^*}{\partial \mu_l \partial \phi_l} &= a_l + (1 + \phi_l) \frac{\partial a_l}{\partial \phi_l}, \quad \frac{\partial^3 \mu_l^*}{\partial \mu_l \partial \phi_l^2} = 2 \frac{\partial a_l}{\partial \phi_l} + (1 + \phi_l) \frac{\partial^2 a_l}{\partial \phi_l^2}, \\ \frac{\partial c_l}{\partial \mu_l} &= 2 \frac{\partial \mu_l^*}{\partial \phi_l} + (1 + \phi_l) \frac{\partial^2 \mu_l^*}{\partial \phi_l^2}, \quad \frac{\partial^2 c_l}{\partial \mu_l^2} = 2 \frac{\partial^2 \mu_l^*}{\partial \mu_l \partial \phi_l} + (1 + \phi_l) \frac{\partial^3 \mu_l^*}{\partial \mu_l \partial \phi_l^2}, \end{aligned}$$

$$\begin{aligned}
\frac{\partial^3 \mu_l^*}{\partial \phi_l^3} &= \mu_l^3 \psi^{(3)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^3 \psi^{(3)}[\mu_l(1 + \phi_l) + \phi_l + 2], \\
\frac{\partial^2 c_l}{\partial \mu_l \partial \phi_l} &= 3 \frac{\partial^2 \mu_l^*}{\partial \phi_l^2} + (1 + \phi_l) \frac{\partial^3 \mu_l^*}{\partial \phi_l^3}, \quad \frac{\partial b_l}{\partial \phi_l} = \frac{\partial \mu_l^*}{\partial \phi_l} + (1 + \phi_l) \frac{\partial^2 \mu_l^*}{\partial \phi_l^2}, \\
\frac{\partial c_l}{\partial \phi_l} &= \mu_l^3 \psi^{(2)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^3 \psi^{(2)}[\mu_l(1 + \phi_l) + \phi_l + 2] + \psi^{(2)}(\phi_l + 2), \\
\frac{\partial^2 c_l}{\partial \phi_l^2} &= \mu_l^4 \psi^{(3)}[\mu_l(1 + \phi_l)] - (1 + \mu_l)^4 \psi^{(3)}[\mu_l(1 + \phi_l) + \phi_l + 2] + \psi^{(3)}(\phi_l + 2), \\
\frac{\partial^2 a_l}{\partial \mu_l \partial \phi_l} &= d_l + (1 + \phi_l) \frac{\partial d_l}{\partial \phi_l}, \quad \frac{\partial e_l}{\partial \mu_l} = -(1 + \phi_l) \left[2 \frac{\partial a_l}{\partial \mu_l} + (1 + \phi_l) \frac{\partial^2 a_l}{\partial \mu_l \partial \phi_l} \right], \\
\frac{\partial b_l}{\partial \mu_l} &= (1 + \phi_l) \left\{ a_l + \mu_l \frac{\partial a_l}{\partial \mu_l} - (1 + \phi_l) \psi^{(2)}[\mu_l(1 + \phi_l) + \phi_l + 2] \right\}, \\
\frac{\partial e_l}{\partial \phi_l} &= -2a_l - (1 + \phi_l) \left[4 \frac{\partial a_l}{\partial \phi_l} + (1 + \phi_l) \frac{\partial^2 a_l}{\partial \phi_l^2} \right], \\
\frac{\partial^3 \mu_l^*}{\partial \mu_l^2 \partial \phi_l} &= \frac{\partial a_l}{\partial \mu_l} + (1 + \phi_l) \frac{\partial^2 a_l}{\partial \mu_l \partial \phi_l}, \quad \frac{\partial^2 b_l}{\partial \phi_l^2} = 2 \frac{\partial^2 \mu_l^*}{\partial \phi_l^2} + (1 + \phi_l) \frac{\partial^3 \mu_l^*}{\partial \phi_l^3}.
\end{aligned}$$

Os cumulantes da log-verossimilhança para o modelo de regressão beta prime são dados por

$$\begin{aligned}
\kappa_{rs} &= - \sum_{l=1}^n (1 + \phi_l)^2 a_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 x_{lr} x_{ls}, \quad \kappa_{rS} = - \sum_{l=1}^n b_l \frac{d\mu_l}{d\eta_l} \frac{d\phi_l}{d\tau_l} x_{lr} s_{lS}, \\
\kappa_{RS} &= - \sum_{l=1}^n c_l \left(\frac{d\phi_l}{d\tau_l} \right)^2 s_{lR} s_{lS}, \quad \kappa_{rst} = - \sum_{l=1}^n (1 + \phi_l)^2 \left\{ (1 + \phi_l) d_l \left(\frac{d\mu_l}{d\eta_l} \right)^3 + 3a_l \frac{d\mu_l}{d\eta_l} \frac{d^2 \mu_l}{d\eta_l^2} \right\} \\
&\quad \times x_{lr} x_{ls} x_{lt}, \quad \kappa_{rST} = \sum_{l=1}^n \left\{ e_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 - b_l \frac{d^2 \mu_l}{d\eta_l^2} \right\} \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} s_{lT}, \\
\kappa_{rST} &= - \sum_{l=1}^n \left\{ \frac{\partial c_l}{\partial \mu_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + b_l \frac{d^2 \phi_l}{d\tau_l^2} \right\} \frac{d\mu_l}{d\eta_l} x_{lr} s_{lS} s_{lT}, \\
\kappa_{RST} &= - \sum_{l=1}^n \left\{ \frac{\partial c_l}{\partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^3 + 3c_l \frac{d\phi_l}{d\tau_l} \frac{d^2 \phi_l}{d\tau_l^2} \right\} s_{lR} s_{lS} s_{lT}, \\
\kappa_{rstu} &= - \sum_{l=1}^n (1 + \phi_l)^2 \left\{ (1 + \phi_l) \left[\frac{\partial d_l}{\partial \mu_l} \left(\frac{d\mu_l}{d\eta_l} \right)^4 + 6d_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 \frac{d^2 \mu_l}{d\eta_l^2} \right] + a_l \left[3 \left(\frac{d^2 \mu_l}{d\eta_l^2} \right)^2 \right. \right. \\
&\quad \left. \left. + 4 \frac{d\mu_l}{d\eta_l} \frac{d^3 \mu_l}{d\eta_l^3} \right] \right\} x_{lr} x_{ls} x_{lt} x_{lu}, \\
\kappa_{rstU} &= - \sum_{l=1}^n \left\{ (1 + \phi_l)^2 \left[3d_l + (1 + \phi_l) \frac{\partial d_l}{\partial \phi_l} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^3 - 3e_l \frac{d\mu_l}{d\eta_l} \frac{d^2 \mu_l}{d\eta_l^2} + (1 + \phi_l) \frac{\partial \mu_l^*}{\partial \phi_l} \frac{d^3 \mu_l}{d\eta_l^3} \right\} \\
&\quad \times \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} x_{lt} s_{lU},
\end{aligned}$$

$$\begin{aligned}
\kappa_{rsTU} &= - \sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \mu_l^2} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \frac{\partial c_l}{\partial \mu_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 + \left[b_l \frac{d^2 \mu_l}{d\eta_l^2} - e_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 \right] \frac{d^2 \phi_l}{d\tau_l^2} \right\} \\
&\quad \times x_{lr} x_{ls} s_{lT} s_{lU}, \\
\kappa_{rSTU} &= - \sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \mu_l \partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^3 + b_l \frac{d^3 \phi_l}{d\tau_l^3} \right] + 3 \frac{\partial c_l}{\partial \mu_l} \frac{d\phi_l}{d\tau_l} \frac{d^2 \phi_l}{d\tau_l^2} \right\} \frac{d\mu_l}{d\eta_l} x_{lr} s_{lS} s_{lT} s_{lU}, \\
\kappa_{RSTU} &= - \sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \phi_l^2} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 6 \frac{\partial c_l}{\partial \phi_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 + c_l \left[3 \left(\frac{d^2 \phi_l}{d\tau_l^2} \right)^2 + 4 \frac{d\phi_l}{d\tau_l} \frac{d^3 \phi_l}{d\tau_l^3} \right] \right\} \\
&\quad \times s_{lR} s_{lS} s_{lT} s_{lU}.
\end{aligned}$$

As derivadas dos cumulantes são dadas por

$$\begin{aligned}
\kappa_{rs}^{(t)} &= - \sum_{l=1}^n (1 + \phi_l)^2 \left[(1 + \phi_l) d_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 2a_l \frac{d^2 \mu_l}{d\eta_l^2} \right] \frac{d\mu_l}{d\eta_l} x_{lr} x_{ls} x_{lt}, \\
\kappa_{rs}^{(T)} &= \sum_{l=1}^n e_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} s_{lT}, \\
\kappa_{rS}^{(t)} &= - \sum_{l=1}^n (1 + \phi_l) \left[\frac{\partial^2 \mu_l^*}{\partial \mu_l \partial \phi_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \frac{\partial \mu_l^*}{\partial \phi_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \frac{d\phi_l}{d\tau_l} x_{lr} x_{lt} s_{lS}, \\
\kappa_{rS}^{(T)} &= - \sum_{l=1}^n \left[\frac{\partial b_l}{\partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + b_l \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\mu_l}{d\eta_l} x_{lr} s_{lS} s_{lT}, \\
\kappa_{RS}^{(t)} &= - \sum_{l=1}^n \frac{\partial c_l}{\partial \mu_l} \frac{d\mu_l}{d\eta_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 x_{lt} s_{lR} s_{lS}, \\
\kappa_{RS}^{(T)} &= - \sum_{l=1}^n \left[\frac{\partial c_l}{\partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 2c_l \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\phi_l}{d\tau_l} s_{lR} s_{lS} s_{lT}, \\
\kappa_{rs}^{(tu)} &= - \sum_{l=1}^n (1 + \phi_l)^2 \left\{ \left[(1 + \phi_l) \frac{\partial d_l}{\partial \mu_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \left(3(1 + \phi_l) d_l + 2 \frac{\partial a_l}{\partial \mu_l} \right) \frac{d^2 \mu_l}{d\eta_l^2} \right] \right. \\
&\quad \times \left. \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 2a_l \left[\frac{d\mu_l}{d\eta_l} \frac{d^3 \mu_l}{d\eta_l^3} + \left(\frac{d^2 \mu_l}{d\eta_l^2} \right)^2 \right] \right\} x_{lr} x_{ls} x_{lt} x_{lu}, \\
\kappa_{rs}^{(tU)} &= - \sum_{l=1}^n (1 + \phi_l) \left\{ (1 + \phi_l) \left[3d_l + (1 + \phi_l) \frac{\partial d_l}{\partial \phi_l} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 2 \left[2a_l + (1 + \phi_l) \frac{\partial a_l}{\partial \phi_l} \right] \right. \\
&\quad \times \left. \frac{d^2 \mu_l}{d\eta_l^2} \right\} \frac{d\mu_l}{d\eta_l} \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} x_{lt} s_{lU}, \\
\kappa_{rs}^{(TU)} &= \sum_{l=1}^n \left[\frac{\partial e_l}{\partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + e_l \frac{d^2 \phi_l}{d\tau_l^2} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^2 x_{lr} x_{ls} s_{lT} s_{lU},
\end{aligned}$$

$$\begin{aligned}
\kappa_{rS}^{(tu)} &= -\sum_{l=1}^n (1 + \phi_l) \left\{ \left[\frac{\partial^3 \mu_l^*}{\partial \mu_l^2 \partial \phi_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 3 \frac{\partial^2 \mu_l^*}{\partial \mu_l \partial \phi_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \frac{d\mu_l}{d\eta_l} + \frac{\partial \mu_l^*}{\partial \phi_l} \frac{d^3 \mu_l}{d\eta_l^3} \right\} \\
&\quad \times \frac{d\phi_l}{d\tau_l} x_{lr} x_{lt} x_{lu} s_{lS}, \\
\kappa_{rS}^{(tU)} &= -\sum_{l=1}^n \left\{ \left[\frac{\partial^2 \mu_l^*}{\partial \mu_l \partial \phi_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \frac{\partial \mu_l^*}{\partial \phi_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \left[(1 + \phi_l) \frac{d^2 \phi_l}{d\tau_l^2} + \left(\frac{d\phi_l}{d\tau_l} \right)^2 \right] \right. \\
&\quad \left. + (1 + \phi_l) \left[\frac{\partial^3 \mu_l^*}{\partial \mu_l \partial \phi_l^2} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \frac{\partial^2 \mu_l^*}{\partial \phi_l^2} \frac{d^2 \mu_l}{d\eta_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 \right\} x_{lr} x_{lt} s_{lS} s_{lU}, \\
\kappa_{rS}^{(TU)} &= -\sum_{l=1}^n \left\{ \left[\frac{\partial^2 b_l}{\partial \phi_l^2} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 3 \frac{\partial b_l}{\partial \phi_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\phi_l}{d\tau_l} + b_l \frac{d^3 \phi_l}{d\tau_l^3} \right\} \frac{d\mu_l}{d\eta_l} x_{lr} s_{lS} s_{lT} s_{lU}, \\
\kappa_{RS}^{(tu)} &= -\sum_{l=1}^n \left[\frac{\partial^2 c_l}{\partial \mu_l^2} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \frac{\partial c_l}{\partial \mu_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 x_{lt} x_{lu} s_{lR} s_{lS}, \\
\kappa_{RS}^{(tU)} &= -\sum_{l=1}^n \left[\frac{\partial^2 c_l}{\partial \mu_l \partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 2 \frac{\partial c_l}{\partial \mu_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\mu_l}{d\eta_l} \frac{d\phi_l}{d\tau_l} x_{lt} s_{lR} s_{lS} s_{lU}, \\
\kappa_{RS}^{(TU)} &= -\sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \phi_l^2} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 5 \frac{\partial c_l}{\partial \phi_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 2c_l \left[\frac{d\phi_l}{d\tau_l} \frac{d^3 \phi_l}{d\tau_l^3} + \left(\frac{d^2 \phi_l}{d\tau_l^2} \right)^2 \right] \right\} \\
&\quad \times s_{lR} s_{lS} s_{lT} s_{lU}, \\
\kappa_{rst}^{(u)} &= -\sum_{l=1}^n \left\{ (1 + \phi_l)^3 \left[\frac{\partial d_l}{\partial \mu_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 3d_l \frac{d^2 \mu_l}{d\eta_l^2} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^2 + 3(1 + \phi_l)^2 \left[\frac{\partial a_l}{\partial \mu_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 \frac{d^2 \mu_l}{d\eta_l^2} \right. \right. \\
&\quad \left. \left. + a_l \left(\frac{d\mu_l}{d\eta_l} \frac{d^3 \mu_l}{d\eta_l^3} + \left(\frac{d^2 \mu_l}{d\eta_l^2} \right)^2 \right) \right] \right\} x_{lr} x_{ls} x_{lt} x_{lu}, \\
\kappa_{rst}^{(U)} &= -\sum_{l=1}^n \left\{ (1 + \phi_l)^2 \left[3d_l + (1 + \phi_l) \frac{\partial d_l}{\partial \phi_l} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^2 - 3e_l \frac{d^2 \mu_l}{d\eta_l^2} \right\} \frac{d\mu_l}{d\eta_l} \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} x_{lt} s_{lU}, \\
\kappa_{rsT}^{(u)} &= \sum_{l=1}^n \left[\frac{\partial e_l}{\partial \mu_l} \left(\frac{d\mu_l}{d\eta_l} \right)^3 + \left(2e_l - \frac{\partial b_l}{\partial \mu_l} \right) \frac{d\mu_l}{d\eta_l} \frac{d^2 \mu_l}{d\eta_l^2} - b_l \frac{d^3 \mu_l}{d\eta_l^3} \right] \frac{d\phi_l}{d\tau_l} x_{lr} x_{ls} x_{lu} s_{lT}, \\
\kappa_{rsT}^{(U)} &= \sum_{l=1}^n \left\{ \left[\frac{\partial e_l}{\partial \phi_l} \left(\frac{d\mu_l}{d\eta_l} \right)^2 - \frac{\partial b_l}{\partial \phi_l} \frac{d^2 \mu_l}{d\eta_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 + \left[e_l \left(\frac{d\mu_l}{d\eta_l} \right)^2 - b_l \frac{d^2 \mu_l}{d\eta_l^2} \right] \frac{d^2 \phi_l}{d\tau_l^2} \right\} \\
&\quad \times x_{lr} x_{ls} s_{lT} s_{lU}, \\
\kappa_{rST}^{(u)} &= -\sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \mu_l^2} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + \frac{\partial b_l}{\partial \mu_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \left(\frac{d\mu_l}{d\eta_l} \right)^2 + \left[\frac{\partial c_l}{\partial \mu_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + b_l \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d^2 \mu_l}{d\eta_l^2} \right\} \\
&\quad \times x_{lr} x_{lu} s_{lS} s_{lT}, \\
\kappa_{rST}^{(U)} &= -\sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \mu_l \partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + \left(2 \frac{\partial c_l}{\partial \mu_l} + \frac{\partial b_l}{\partial \phi_l} \right) \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\phi_l}{d\tau_l} + b_l \frac{d^3 \phi_l}{d\tau_l^3} \right\} \frac{d\mu_l}{d\eta_l} \\
&\quad \times x_{lr} s_{lS} s_{lT} s_{lU},
\end{aligned}$$

$$\begin{aligned}
\kappa_{RST}^{(u)} &= - \sum_{l=1}^n \left[\frac{\partial^2 c_l}{\partial \mu_l \partial \phi_l} \left(\frac{d\phi_l}{d\tau_l} \right)^3 + 3 \frac{\partial c_l}{\partial \mu_l} \frac{d\phi_l}{d\tau_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \frac{d\mu_l}{d\eta_l} x_{lu} s_{lR} s_{lS} s_{lT}, \\
\kappa_{RST}^{(U)} &= - \sum_{l=1}^n \left\{ \left[\frac{\partial^2 c_l}{\partial \phi_l^2} \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 6 \frac{\partial c_l}{\partial \phi_l} \frac{d^2 \phi_l}{d\tau_l^2} \right] \left(\frac{d\phi_l}{d\tau_l} \right)^2 + 3c_l \left[\frac{d\phi_l}{d\tau_l} \frac{d^3 \phi_l}{d\tau_l^3} + \left(\frac{d^2 \phi_l}{d\tau_l^2} \right)^2 \right] \right\} \\
&\quad \times s_{lR} s_{lS} s_{lT} s_{lU}.
\end{aligned}$$

APÊNDICE E – CONJUNTOS DE DADOS

- Incidentes com danos em navios de carga

Tabela 19 – Número de incidentes com danos (y) e total de meses em serviço por tipo de navio (A-E), ano de construção (1:1960-64, 2:1965-69, 3:1970-74 e 4:1975-79) e período de operação (1:1960-74 e 2:1975-79).

Tipo de navio	Ano de construção	Período de operação	Meses em serviço	Número de incidentes
A	1	1	127	0
A	1	2	63	0
A	2	1	1095	3
A	2	2	1095	4
A	3	1	1512	6
A	3	2	3353	18
A	4	2	2244	11
B	1	1	44882	39
B	1	2	17176	29
B	2	1	28609	58
B	2	2	20370	53
B	3	1	7064	12
B	3	2	13099	44
B	4	2	7117	18
C	1	1	1179	1
C	1	2	552	1
C	2	1	781	0
C	2	2	676	1
C	3	1	783	6
C	3	2	1948	2
C	4	2	274	1
D	1	1	251	0
D	1	2	105	0
D	2	1	288	0
D	2	2	192	0
D	3	1	349	2
D	3	2	1208	11
D	4	2	2051	4
E	1	1	45	0
E	2	1	789	7

Tipo de navio	Ano de construção	Período de operação	Meses em serviço	Número de incidentes
E	2	2	437	7
E	3	1	1157	5
E	3	2	2161	12
E	4	2	542	1

• **Progênie de *Sitophilus zeamais***

Tabela 20 – Número de insetos emergidos (y) em placas de Petri contendo 10g de milho tratadas com extratos produzidos a partir de diferentes partes de *Annona mucosa* (folhas, ramos e sementes) ou apenas água (controle).

Extrato	y	Extrato	y
Folha	19	Semente	4
Folha	20	Semente	0
Folha	36	Semente	1
Folha	32	Semente	1
Folha	18	Semente	1
Folha	47	Semente	0
Folha	38	Semente	2
Folha	31	Semente	0
Folha	32	Semente	2
Folha	40	Semente	0
Ramo	20	Controle	35
Ramo	34	Controle	26
Ramo	41	Controle	41
Ramo	29	Controle	34
Ramo	31	Controle	23
Ramo	15	Controle	29
Ramo	31	Controle	39
Ramo	33	Controle	34
Ramo	45	Controle	16
Ramo	20	Controle	38

• **Conteúdo de gordura corporal**

Tabela 21 – Massa gorda corporal (y) em kg estimada por DXA, circunferência do quadril (x_2), largura do joelho (x_4), soma (x_3) do log da dobra cutânea do queixo, do tríceps e do subescapular, e soma do logaritmo de três medidas antropométricas (x_5), para 71 mulheres alemãs saudáveis.

x_2	x_3	x_4	x_5	y
112,0	4,42	9,4	4,50	41,68
116,5	4,63	8,9	4,48	43,29
108,5	4,12	8,9	4,60	35,41
96,5	4,03	9,2	3,91	22,79
100,5	4,24	10,0	4,15	36,42
97,0	3,55	8,8	3,64	24,13
103,0	4,14	8,9	4,31	29,83
105,0	4,04	8,5	4,47	35,96
97,0	3,91	8,8	3,47	23,69
93,0	3,66	8,8	3,60	22,71
99,0	3,70	8,6	3,67	23,42
94,0	4,14	8,7	3,85	23,24
95,0	4,00	8,2	3,85	26,25
90,0	3,72	8,2	3,48	21,94
107,5	4,01	8,6	3,89	30,13
109,0	4,42	9,6	4,33	36,31
101,5	3,78	9,4	3,97	27,72
116,0	4,14	10,7	4,04	46,99
112,0	4,25	9,4	4,28	42,01
91,0	3,90	8,8	3,28	18,63
107,5	4,15	8,6	4,05	38,65
96,0	3,70	8,6	3,26	21,20
101,0	3,73	9,2	4,27	35,40
99,5	3,89	8,1	4,27	29,63
99,0	3,16	8,6	3,45	25,16
102,0	4,16	10,8	4,40	31,75
110,3	3,95	9,5	3,60	40,58
97,0	3,51	8,8	3,54	21,69
124,0	4,30	11,2	4,51	46,60
102,0	3,88	8,3	4,21	27,62
122,5	4,24	10,8	4,35	41,30
125,0	4,24	11,1	4,34	42,76
100,0	3,97	9,7	4,18	28,84

x_2	x_3	x_4	x_5	y
113,0	4,33	9,6	4,32	36,88
101,2	3,12	9,3	3,47	25,09
104,3	3,77	8,1	4,43	29,73
106,0	3,53	10,0	3,48	28,92
118,5	3,99	10,6	4,05	43,80
99,0	3,82	9,8	3,77	26,74
109,0	4,14	9,8	4,37	33,79
126,0	4,64	11,4	4,61	62,02
117,0	4,12	10,6	4,35	40,01
109,0	4,42	9,9	4,57	42,72
108,0	4,01	9,8	4,38	32,49
132,0	4,45	9,8	4,41	45,92
127,0	4,34	11,0	4,19	42,23
128,5	4,44	10,0	4,62	47,48
125,0	4,68	11,8	4,45	60,72
109,0	4,08	9,6	3,71	32,74
103,6	3,58	9,2	3,96	27,04
99,3	3,53	8,4	3,17	21,07
109,5	4,46	10,0	4,46	37,49
116,3	3,92	9,5	4,15	38,08
122,0	4,11	10,0	3,99	40,83
94,0	3,26	9,0	3,48	18,51
107,7	4,04	9,0	3,78	26,36
95,0	3,36	9,0	3,28	20,08
122,5	4,19	10,0	4,51	43,71
109,0	4,13	9,5	4,20	31,61
102,3	3,92	9,2	4,10	28,98
92,0	3,35	8,5	3,34	18,62
93,0	3,42	7,2	3,39	18,64
88,0	3,39	8,2	3,16	13,70
94,5	3,29	8,8	2,87	14,88
95,0	2,99	9,1	3,38	16,46
92,2	2,40	8,5	2,05	11,21
92,2	2,40	8,5	2,05	11,21
93,2	3,10	8,1	2,76	14,18
100,0	3,50	8,5	3,76	20,84
103,0	3,33	8,5	3,21	19,00
88,3	3,48	8,9	3,45	18,07