VIESES E CORREÇÕES DE BARTLETT E TIPO-BARTLETT EM MODELOS DE REGRESSÃO DE VALORES EXTREMOS

Edson Marcos RAMOS[1]
Gauss Moutinho CORDEIRO[2]

▪ RESUMO: Neste artigo foram obtidas estimativas de máxima verossimilhança do viés de segunda órdem nos modelos de regressão de valores extremos.Também foram consideradas correções de amostras finitas na matriz notação para a razão de verossimilhança e estatísticas de escores nesses modelos. Algumas simulações foram realizadas para comparar as estatísticas da razão de verossimilhança e de escores com suas versões modificadas com a distribuição qui-quadrado. Também foi ilustrada a correção do viés.

▪ PALAVRAS-CHAVE: Correrção de viés; teste *bootstrap*; distribuição qui-quadrado; modelos de valores extremos; estimatimação de máxima verossimilhança.

1 Introdução

A correção do viés de segunda ordem do estimador de máxima verossimilhança é apresentada por Cox e Snell (1968) no caso multiparamétrico. Alguns artigos tratam da aplicação desta fórmula em modelos especiais, tais como: modelos lineares generalizados (Cordeiro e McCullagh, 1991) e modelos lineares generalizados superdispersados (Cordeiro e Botter, 2001).

Recentemente, vários artigos têm sido publicados apresentando estatísticas escore corrigidas em classes amplas de modelos de regressão. Cordeiro, Ferrari e Paula (1993) e Cribari-Neto e Ferrari (1995a) obtiveram correções tipo-Bartlett para testes escore em modelos lineares generalizados com parâmetro de dispersão conhecido e desconhecido, respectivamente. Correções similares para testes escore em modelos lineares heterocedásticos foram obtidos por Cribari-Neto e Ferrari

[1]Departamento de Estatística, Universidade Federal do Pará – UFPA, Caixa postal 479, CEP 66075-110, Belém, PA, Brasil. E-mail: *ramosedson@yahoo.com.br*
[2]Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco – UFRPE, CEP 50171-900, Recife, PE, Brasil. E-mail: *gauss@ufrpe.br*

(1995b). Estudos de simulação mostram que as estatísticas escore modificadas propostas por Cordeiro e Ferrari (1991) são mais bem aproximadas pela distribuição χ^2 de referência do que as estatísticas escore usuais. Uma revisão da literatura dos testes escores aperfeiçoados é dada por Cribari-Neto e Cordeiro (1996). A tese de doutorado de Ferrari (1991) é uma referência importante sobre o melhoramento dos testes escore.

Neste artigo, corrigimos o viés de segunda-ordem no modelo linear de valor extremo. Deduzimos uma fórmula para os vieses de segunda ordem dos estimadores de máxima verossimilhança e apresentamos um estudo de simulação que confirma a superioridade dos estimadores corrigidos pelo viés sobre os estimadores clássicos em relação à teoria assintótica de segunda ordem.

Deduzimos, também, correções para os testes da razão de verossimilhanças e escore em modelos lineares de valor extremo. Destacamos três testes escore aperfeiçoados com melhor *performance* do que o teste escore usual. Os ajustamentos em amostras finitas, foram obtidos usando os resultados descritos em Lawley (1956), Cordeiro e Ferrari (1991), Kakizawa (1996) e Cordeiro; Ferrari e Cysneiros (1998). As fórmulas para as correções foram deduzidas em forma matricial, o que possibilita a obtenção das estatísticas da razão de verossimilhanças e escore corrigidas em uma variedade de situações. Por meio de comparações numéricas das correções analíticas e dos testes *bootstrap*, concluímos que existe uma equivalência de segunda-ordem entre as três estatísticas escore corrigidas.

Sejam $y_1, \ldots, y_n$ variáveis aleatórias, com cada y tendo uma distribuição de valor extremo com parâmetro de locação $\rho \in \Re$ e parâmetro de dispersão $\sigma > 0$. A função densidade da distribuição de valor extremo é dada por

$$f(y; \rho, \sigma) = \sigma^{-1} \exp\left\{-\frac{(y-\rho)}{\sigma} - \exp\left[-\frac{(y-\rho)}{\sigma}\right]\right\}. \tag{1}$$

A esperança e a variância de y são dadas por $E(y) = \rho - \psi(1)\sigma$ e $\mathrm{Var}(y) = \pi^2\sigma^2/6$, onde $\psi(1) = -\gamma$ e γ é a constante de Euler. A média de y é função não apenas de ρ mas também de σ. Mudanças em σ não alteram apenas a dispersão, mas também a média da distribuição. Assim, o parâmetro σ mede algumas incertezas à respeito da dispersão da distribuição. Em outras palavras, aumentos ou diminuições na locação ρ, fazem que a distribuição seja deslocada para a direita ou para a esquerda, enquanto aumentos ou diminuições em σ fazem que a distribuição se torne mais ou menos concentrada em torno de ρ. Verifica-se que y tem uma distribuição de valor extremo (1) se, e somente se, e^y tiver uma distribuição de Weibull, ou $e^{\frac{y}{\sigma}}$ tiver uma distribuição exponencial. Assim, alguns teoremas válidos para distribuições exponenciais podem ser usados na distribuição de valor extremo aplicando-os em $e^{\frac{y}{\sigma}}$.

A distribuição de valor extremo foi originada, principalmente, da necessidade dos astrônomos em aceitarem ou rejeitarem observações longínquas. Um grande número de publicações envolvendo aplicações práticas da distribuição de valor extremo estão detalhadas no livro de Gumbel (1958). O uso dessa distribuição para modelos relacionados ao tempo de falha em estudos de confiabilidade foi citado em

Canfield e Bordman (1975). A denominação valor extremo deve-se ao fato de que a distribuição é obtida a partir do limite de distribuições quando $n \to \infty$.

Suponhamos que o vetor de parâmetros de locação $\rho = (\rho_1, \ldots, \rho_n)^T$ possui uma estrutura linear da forma

$$\rho = X\theta, \tag{2}$$

em que $X = (x_1, \ldots, x_p)$ é uma matriz modelo $n \times p, (n > p)$, de constantes conhecidas e posto completo p e $\theta = (\theta_1, \ldots, \theta_p)^T$ é um conjunto de parâmetros de regressão desconhecidos a serem estimados. A inferência sobre θ pode ser feita via métodos de verossimilhança analogamente àqueles para modelos lineares ou normais. Assumimos no artigo que o parâmetro de dispersão σ é conhecido.

A função de log-verossimilhança para o parâmetro θ no modelo de regressão linear de valor extremo (1)-(2) com relação a amostra é dada por

$$\ell(\theta) = -n \log \sigma - \frac{1}{\sigma} \sum_{i=1}^{n} (y_i - \rho_i) - \sum_{i=1}^{n} \exp\left[-\frac{(y_i - \rho_i)}{\sigma} \right].$$

As condições de regularidade usuais para os estimadores de máxima verossimilhança são válidas nesse modelo. Essas condições de regularidade são as mesmas exigidas para a expansão de Edgeworth. Discussões detalhadas de tais condições podem ser encontradas em Bhattacharya e Rao (1976) e Bhattacharya e Ghosh (1978).

Sejam $U = U(\theta) = \partial\ell(\theta)/\partial\theta$ e $K = K(\theta)$, a função escore e a informação total de Fisher, respectivamente. Podemos mostrar que $U = -\sigma^{-1} X^T g$ e $K = \sigma^{-2} X^T X$, onde g é um vetor $n \times 1$ com componentes da forma $\exp[-(y_i - \theta_i)/\sigma] - 1$. O estimador de máxima verossimilhança (EMV) $\hat{\theta}$ pode ser obtido pelo método dos mínimos quadrados reponderados pela relação

$$\theta^{(m+1)} = \theta^{(m)} - \sigma (X^T X)^{-1} X^T g^{(m)},$$

a qual é facilmente implementada no sistema S-Plus (Becker; Chambers; Wilks, 1988).

Na seção seguinte é obtida uma fórmula geral em notação matricial para o viés de segunda ordem dos estimadores de máxima verossimilhança de θ no modelo de regressão linear de valor extremo. Correções de vieses para outras classes importantes de modelos de regressão são encontradas em Cordeiro e McCullagh (1991), Cordeiro, Vasconcellos e Santos (1998) e Cordeiro et al. (2000). Correções de Bartlett para modelos de regressão são dadas por Cordeiro (1983, 1987) e por Cordeiro; Paula e Botter (1994) e correções tipo-Bartlett são obtidas em Cordeiro; Ferrari e Paula (1993), Cribari-Neto e Ferrari (1995a) e Cysneiros e Cordeiro (2002). As correções reduzem o erro da aproximação da distribuição da estatística por χ^2 para as estatísticas escore e da razão de verossimilhanças, e, portanto, os testes corrigidos resultam em inferências mais precisas em problemas práticos com amostras reduzidas.

2 Viés de segunda ordem do estimador $\hat{\theta}$

O objetivo dessa seção é usar a fórmula assintótica de Cox e Snell (1968) para o viés até ordem n^{-1} dos estimadores de máxima verossimilhança para calcular o viés de segunda ordem do estimador $\hat{\theta}$ no modelo (1)-(2). As notações usadas para obtenção do conjunto de cumulantes de derivadas da log-verossimilhança são iguais àquelas dadas em Cordeiro e McCullagh (1991) e Cordeiro e Botter (2001). Os cumulantes relativos aos modelos (1)-(2) podem ser encontrados sem dificuldade. Por exemplo,

$$\kappa_{r,s} = \sigma^{-2}\sum_i x_{ir}x_{is},\ \kappa_{rst} = \sigma^{-3}\sum_i x_{ir}x_{is}x_{it},\ \kappa_{rs,t} = -\sigma^{-3}\sum_i x_{ir}x_{is}x_{it},$$
$$\kappa_{rs,t} = -\sigma^{-3}\sum_i x_{ir}x_{is}x_{it}, \kappa_{r,s,t} = 2\sigma^{-3}\sum_i x_{ir}x_{is}x_{it}, \kappa_{rs,tu} = \sigma^{-4}\sum_i x_{ir}x_{is}x_{it}x_{iu},$$
$$\kappa_{r,s,t,u} = 6\sigma^{-4}\sum_i x_{ir}x_{is}x_{it}x_{iu} \quad \text{e} \quad \kappa_{r,s,tu} = -2\sigma^{-4}\sum_i x_{ir}x_{is}x_{it}x_{iu}.$$

A partir da expressão geral para os vieses até ordem n^{-1} dos estimadores de máxima verossimilhança dada por Cox e Snell (1968), podemos escrever os vieses de $\hat{\theta}_a$ até ordem n^{-1} como

$$B(\hat{\theta}_a) = \sum{}' \kappa^{ar}\kappa^{tu}\left(\kappa_{rt}^{(u)} - \frac{1}{2}\kappa_{rtu}\right),$$

onde os índices r, t e u varrem as combinações dos p parâmetros em θ. Inserindo os cumulantes acima e rearrumando os termos, obtemos

$$B(\hat{\theta}_a) = -\frac{1}{2}\sigma^{-3}\sum_i \left(\sum_r{}' \kappa^{ar}x_{ir}\right)\left(\sum_{t,u}{}' \kappa^{tu}x_{it}x_{iu}\right).$$

Com o objetivo de expressar a correção de viés $B(\hat{\theta})$, usando notação matricial, definimos a seguinte matriz e vetor ambos de ordem n:

$$Z = \{z_{ij}\} = X(X^TX)^{-1}X^T \quad \text{e} \quad \varepsilon = (z_{11}, \ldots, z_{nn})^T.$$

A matriz Z independe do parâmetro de dispersão σ e o vetor ε contêm os elementos da diagonal da matriz Z. Após reagrupamento dos termos acima, a correção de viés $B(\hat{\theta})$ pode ser escrita como

$$B(\hat{\theta}) = -\frac{\sigma}{2}(X^TX)^{-1}X^T\varepsilon. \tag{3}$$

O vetor de viés $B(\hat{\theta})$ pode ser obtido a partir de uma regressão de mínimos quadrados de ε sobre as colunas de X. A expressão (3) sugere que $B(\hat{\theta})$ seja pequeno quando todos os elementos de ε estão suficientemente próximos de zero, isto é, ε está próximo ao espaço ortogonal das colunas de X. É claro que o viés de $\hat{\theta}$ é dependente do parâmetro de dispersão σ. Valores grandes de σ podem resultar em estimadores $\hat{\theta}$ com vieses apreciáveis, enquanto para valores suficientemente pequenos de σ os vieses de $\hat{\theta}$ serão insignificantes. É previsto que o estimador corrigido $\tilde{\theta} = \hat{\theta} + \frac{\sigma}{2}(X^TX)^{-1}X^T\varepsilon$ tenha melhores propriedades amostrais do que o estimador não corrigido $\hat{\theta}$. Estudos de simulação são feitos na Seção 5 que confirmam essa previsão, mostrando que $\tilde{\theta}$ tem vieses menores que o correspondente $\hat{\theta}$. Assim, a correção de vieses de segunda ordem direciona o estimador corrigido para ficar mais próximo dos parâmetros verdadeiros.

3 Correções de Bartlett

Considere agora uma partição do vetor $\hat{\theta}$, $p \times 1$, com $\theta = (\theta_1^T, \theta_2^T)^T$ onde $\theta_1 = (\theta_1, \ldots, \theta_q)^T$ e $\theta_2 = (\theta_{q+1}, \ldots, \theta_p)^T$ para $q \leq p$, e uma partição induzida correspondente da matriz modelo $X = (X_1, X_2)$. Nosso interesse é testar a hipótese nula composta $H : \theta_1 = \theta_1^{(0)}$ *versus* a hipótese alternativa bilateral $A : \theta_1 \neq \theta_1^{(0)}$. Geralmente, o vetor $\theta_1^{(0)} = 0$. Seja $\hat{\theta} = (\hat{\theta}_1^T, \hat{\theta}_2^T)^T$ o estimador de máxima verossimilhança irrestrito de θ e $\tilde{\theta}_2$ o estimador de máxima verossimilhança restrito de θ_2 sob $H : \theta_1 = \theta_1^{(0)}$. Funções avaliadas em $\tilde{\theta} = (\theta_1^{(0)T}, \tilde{\theta}_2^T)^T$ são distinguidas pela adição de um til. A estatística da razão de verossimilhanças para testar a hipótese nula é dada por $w = 2\{\ell(\hat{\theta}) - \ell(\tilde{\theta})\}$ e sob H, mediante algumas condições de regularidade, temos que $P(w \leq x) = P(\chi_q^2 \leq x) + O(n^{-1})$. Testes da razão de verossimilhanças corrigidos em amostras finitas podem ser calculados usando uma estatística corrigida por uma correção de Bartlett dada por $w^* = w/\tilde{c}$, onde $\tilde{c}$ é o estimador de máxima verossimilhança (EMV) de $c = E(w)/q$. Então, sob H temos, $P(w^* \leq x) = P(\chi_q^2 \leq x) + O(n^{-2})$. O valor esperado da estatística da razão de verossimilhanças é obtido da expansão de Lawley (1956) dada por $E(w) = q + \varepsilon_p - \varepsilon_{p-q} + O(n^{-2})$, onde ε_p e ε_{p-q} são termos de ordem n^{-1} expressos como função de cumulantes de derivadas da log-verossimilhança. O método usado para computar ε_p e ε_{p-q}, usando notação matricial, não é complicado e segue de cálculos algébricos similares àqueles desenvolvidos por Cordeiro (1983, 1987) e por Cordeiro, Paula e Botter (1994). Todas essas publicações mostram os detalhes dos cálculos dos $\varepsilon's$ que embora elementar é muito cansativo. Depois de alguma álgebra, obtemos ε_p dado como

$$\varepsilon_p = -\frac{1}{4}tr(Z_d^2) + \frac{1}{12}1^T(2Z^{(3)} + 3Z_dZZ_d)1, \tag{4}$$

onde $Z_d = \mathrm{diag}\{z_{11}, \ldots, z_{nn}\}$, 1 é um vetor $n \times 1$ de uns e $Z = \{z_{ij}^3\}$ é uma matriz $n \times n$ obtida diretamente dos elementos da matriz Z. O ajustamento ε_{p-q} vem diretamente de (4) substituindo a matriz X pela matriz X_2 cujas colunas correspondem aos parâmetros de perturbação θ_2.

A expressão (4) envolve apenas operações simples com matrizes e vetores e pode ser computada usando programas de computação algébrica como MAPLE ou MATHEMATICA. A grande vantagem no uso da equação (4) na prática deve-se ao fato da matriz modelo X ser facilmente adaptada a vários casos especiais. Citaremos agora dois modelos especiais: o modelo de regressão simples e o modelo de classificação de um fator.

3.1 O modelo de regressão linear simples

Suponhamos que estamos interessados em um modelo de regressão linear de dois parâmetros definido por $\rho_i = \alpha + \theta x_i$, onde x_i denota o valor de uma variável explicativa, para $i = 1, \ldots, n$. Para testar a hipótese nula composta $H : \theta = 0$ *versus* a hipótese alternativa $A : \theta \neq 0$, podemos calcular a correção de Bartlett examinando (4) sob as hipóteses nula H e a alternativa A. Sob $H : \theta = 0$, $X = 1$, $Z = n^{-1}11^T$, $Z_d = n^{-1}I$ e $Z^{(3)} = Z_d Z Z_d = n^{-1}11^T$, onde I é a matriz identidade, resultando $\varepsilon_1 = \frac{1}{6n}$. Usaremos a notação $s_a = \sum_i (x_i - \overline{x})^a/n$ para o a-ésimo momento central da covariável x para $a = 2, 3, 4$. Sob a hipótese alternativa $A : \theta \neq 0$, $X = (1, x)$ com $x = (x_1, \ldots, x_n)^T$. Um elemento típico de Z sob a hipótese alternativa $A : \theta \neq 0$ tem a forma $z_{ij} = (ns_2)^{-1}\{s_2 + (x_i - \overline{x})(x_j - \overline{x})\}$. Então, temos

$$\mathrm{tr}(Z_d^2) = (\gamma_{2x} + 6)/n \quad \text{e} \quad 1^T Z^{(3)} 1 = 1^T Z_d Z Z_d 1 = (\gamma_{1x}^2 + 4)/n,$$

onde $\gamma_{1x} = s_3/s_2^{\frac{3}{2}}$ e $\gamma_{2x} = (s_4 - 3s_2^2)/s_2^2$ são medidas de assimetria e curtose de x, respectivamente.

Depois de alguma álgebra, obtemos a partir de (4), $\varepsilon_2 = \frac{1}{12n}\{5\gamma_{1x}^2 - 3(\gamma_{2x}+6)\}$. Assim, a correção de Bartlett para testar $H : \theta = 0$ pode ser escrita como

$$c = 1 + \frac{1}{12n}(5\gamma_{1x}^2 - 3\gamma_{2x}). \tag{5}$$

Desse modo, a equação (5) sugere que a aproximação χ_1^2 é particularmente sensível às mudanças nas medidas amostrais de assimetria e curtose da variável explanatória x. O termo $\lambda_x = 5\gamma_{1x}^2 - 3\gamma_{2x}$ da correção em (5) possui uma forma elegante. Os valores de λ_x estão limitados por $\gamma_{1x}^2 + \gamma_{2x} + 2 \geq 0$ e $\gamma_{2x} - \gamma_{1x}^2 \geq 1$, o que implica $-3\gamma_{2x} \leq \lambda_x \leq 2\gamma_{1x}^2 - 3$. É claro que λ_x diminui quando γ_{2x} aumenta. Observa-se, também, que λ_x exibe um comportamento quadrático como função de γ_{1x}.

3.2 O modelo de classificação de um fator

Consideramos agora o modelo de análise de variâncias com um fator fixo. Cada observação é classificada dentro de p categorias diferentes, onde existem n_i observações na i-ésima categoria com correspondentes parâmetros de locação ρ_i fixado como: $\rho_i = \theta + \theta_i$ para $i = 1, \ldots, p$, com $\theta_p = 0$, onde θ é a média geral e θ_i é o efeito da i-ésima população sobre o parâmetro de locação ρ. No modelo de classificação de um fator temos $X^T X = \mathrm{diag}\{n_i\}$ e posto $(X^T X) = p$. A matriz Z tem ordem $n = \sum_i n_i$ com elemento genérico $\delta_{ij} n^{-1}$, onde $\delta_{ij} = 1$ se i e j são observações na mesma população e, caso contrário, zero. Observemos que $tr(Z_d^2) = 1^T Z^{(3)} 1 = 1^T Z_d Z Z_d 1 = \sum_i n_i^{-1}$. Estamos interessados em testar $H : \theta_1 = \cdots = \theta_{p-1} = 0$ *versus* A : a violação de pelo menos uma dessas igualdades. A correção de Bartlett para o teste da razão de verossimilhanças de homogeneidade dos parâmetros de locação no modelo de classificação de um fator é obtido de (4) como $c = 1 + (\varepsilon_p - \varepsilon_1)/(p-1)$, produzido

$$c = 1 + \frac{1}{6(p-1)} \left(\sum\nolimits_i n_i^{-1} - n^{-1} \right). \tag{6}$$

4 Correções tipo-Bartlett

Cordeiro e Ferrari (1991) mostraram que, sob condições de regularidade gerais, a estatística escore S_R pode ser modificada por um fator de correção que envolve constantes de ordem n^{-1} e um polinômio de segundo grau na própria estatística S_R produzindo uma estatística escore modificada ajustada com distribuição χ^2 até ordem n^{-1}, segundo a hipótese nula, cujo erro é reduzido de $O(n^{-1})$ para $O(n^{-3/2})$. A estatística escore modificada proposta por esses autores é dada por

$$S_R^* = S_R \left\{ 1 - (c + bS_R + aS_R^2) \right\}, \tag{7}$$

onde o fator multiplicativo entre chaves é uma correção tipo-Bartlett como função da própria estatística S_R e os coeficientes a, b, c são de ordem n^{-1}. Estes coeficientes produzem uma estatística modificada (7) com distribuição χ_q^2, sob a hipótese nula, quando termos de ordens inferiores a n^{-1} são omitidos. No artigo de Cordeiro e Ferrari são dadas as relações que ligam a, b e c aos coeficientes A_1, A_2 e A_3 de Harris (1985).

Testes escore aperfeiçoados podem ser baseados na estatística S_R^* proposta por Cordeiro e Ferrari (1991). Espera-se obter inferência mais refinada do que àquela baseada em S_R, tomando como referência a distribuição χ^2, ou seja para S_R temos $P(S_R \leq x) = P(\chi_q^2 \leq x) + o(n^{-1})$, enquanto que para testes baseados na estatística S_R^*, temos $P(S_R^* \leq x) = P(\chi_q^2 \leq x) + o(n^{-2})$, isto é, o erro da aproximação por χ^2 cái da ordem o(n^{-1}) para ordem $o(n^{-2})$ nos testes relativos a S_R e S_R^*, respectivamente. No entanto, a estatística aperfeiçoada S_R^* nem sempre é uma transformação monótona. Para solucionar esse problema, Kakizawa (1996) sugeriu uma transformação monótona dada por $K(S_R) = S_R^* + P(S_R)$ envolvendo a própria estatística S_R e os coeficientes a, b e c, onde $P(S_R)$ é um polinômio de quinto grau na estatística escore original S_R e é de ordem $O(n^{-2})$. $P(S_R)$ é dado por

$$P(S_R) = \frac{1}{4} \left\{ c^2 S_R + 2bcS_R^2 + \left(2ac + \frac{4}{3}b^2 \right) S_R^3 + 3abS_R^4 + \frac{9}{5}a^2 S_R^5 \right\}. \tag{8}$$

Cordeiro, Ferrari e Cysneiros (1998) também apresentaram uma fórmula alternativa para a estatística escore aperfeiçoada que é uma transformação monótona de S_R. A estatística alternativa $\tilde{S}_R$ é expressa em termos da função de distribuição da normal padrão $\Phi(\cdot)$ por

$$\tilde{S}_R = \sqrt{\frac{\pi}{3a}} \exp\left(\frac{b^2}{3a} - c \right) \left\{ \Phi\left(\sqrt{6a} S_R + \sqrt{\frac{2}{3a}} b \right) - \Phi\left(\frac{2}{3a} b \right) \right\} \tag{9}$$

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

$$\tilde{S}_R = \frac{1}{2b}\exp(-c)\{1 - \exp(-2bS)\}$$

se $a = 0$ e $b \neq 0$. Note que se $a = b = 0$, temos que S_R^* é uma transformação monótona de S_R, com isso, não há necessidade de definir uma estatística aperfeiçoada alternativa. As três estatísticas modificadas S_R^*, $K(S_R)$ e $\tilde{S}_R$ são equivalentes até segunda ordem, isto é, elas diferem por termos de ordem $O(n^{-\frac{3}{2}})$. Nesta seção, deduzimos fórmulas gerais para os ajustamentos a, b e c em amostras finitas usando os resultados de Cordeiro e Ferrari (1991) e Harris (1985), para as estatísticas escore para testar a hipótese nula no modelo de regressão linear de valor extremo dado em (1)-(2).

A estatística escore $S_R = \tilde{U}^T\tilde{K}^{-1}\tilde{U}$ para testar a hipótese nula $H : \theta_1 = \theta_1^{(0)}$ no modelo (1)-(2) é dada por

$$S_R = g^T X_1(R^T R)^{-1}X_1^T g, \quad \text{onde} \quad R = X_1 - X_2(X_2^T X_2)^{-1}X_2^T X_1$$

e g está definido na Seção 1.

Como espera-se que o teste aperfeiçoado S_R^* produza uma inferência mais refinada em amostras finitas, ou seja segundo a hipótese nula H, o erro da aproximação χ^2 é reduzido de $O(n^{-1})$ para $O(n^{-3/2})$, tem-se $P(S_R^* \leq x) = P(\chi_q^2 \leq x) + O(n^{-3/2})$. Com isso, calcularemos as quantidades A_1, A_2 e A_3 dos coeficientes obtidos a partir das fórmulas de Harris (1985).

Inserindo os cumulantes $\kappa's$ de derivadas da log-verossimilhança, como discutido na Seção 2, nas equações de Harris (1985) e usando uma aproximação similar àquela discutida em Cordeiro, Ferrari e Paula (1993) e Cordeiro e Ferrari (1996), obtêm-se expressões similares simples em notação matricial para A_1, A_2 e A_3. Os detalhes algébricos usados para obtenção dos $A's$ são exaustivos, porém eles seguem desenvolvimentos similares aos de Cordeiro, Ferrari e Paula (1993). Esses detalhes matemáticos podem ser obtidos dos autores deste artigo. Depois de alguns cálculos, obtemos

$$\begin{aligned}
A_1 &= 3 1^T Z_{2d}(Z_{-}Z_2)z_{2d}1 + 12\ 1^T\ Z_{2d}Z_2(Z - Z_2)d \\
&\quad +18\ 1^T\ \{Z_2^{(2)} \otimes (Z - Z_2)\}1 - 24\ 1^T\ Z_{2d}(Z - Z_2)_d\,1, \qquad (10)\\
A_2 &= -12\ 1^T\ (Z - Z_2)_d Z_{2d}(Z - Z_2)_d\,1 - 12\ 1^T\ Z_{2d}(Z - Z_2)(Z - Z_2)_d \\
&\quad -24\ 1^T\ \{(Z - Z_2)^{(2)} \otimes Z_2\}\,1\ + 18\ 1^T\ (Z - Z_2)_d^{(2)}\,1, \qquad (11)\\
A_3 &= 12\ 1^T\ (Z - Z_2)_d(Z - Z_2)(Z - Z_2)_d\,1\ + 8\ 1^T\ (Z - Z_2)_d^{(3)}\,1, \qquad (12)
\end{aligned}$$

onde $Z = X(X^T X)^{-1}X^T$ e $Z_2 = (X_2^T X_2)^{-1}$ e o símbolo $\otimes$ denota o produto (direto) de matrizes de mesma dimensão e usamos a notação $Z^{(2)} = Z \otimes Z$, $Z^{(3)} = Z^{(2)} \otimes Z$ e $Z_d = \text{diag}\{z_{11}, \ldots, z_{nn}\}$ que contém os elementos da diagonal da matriz Z. A substituição das equações (10)-(12) nos coeficientes a, b e c, produzem as estatísticas escore aperfeiçoadas descritas em (7)-(9). Para uma hipótese nula simples $H : \theta = \theta^{(0)}$, temos $Z_2 = 0$, e que implica $A_1 = 0$. As quantidades A_1, A_2 e A_3 são fáceis de serem obtidas sob $H : \theta_1 = \theta_1^{(0)}$, pois as equações (10)-(12) envolvem operações simples com matrizes e vetores. Entretanto, estas equações não possuem uma interpretação simples. Alguns sistemas algébricos computacionais tais como MAPLE e MATHEMATICA podem ser de grande utilidade no cálculo com estas equações. Também é possível obtermos consideráveis simplificações quando os modelos possuem forma fechada para as expressões Z_2 e $Z - Z_2$. O uso da equação $Z - Z_2 = R(R^T R)^{-1} R^T$, onde R foi definido associado com a estatística escore S_R, evita o cálculo de uma matriz inversa de dimensão p, porém exige o cálculo de uma matriz de ordem inferior a q. Apresentamos acima correções tipo-Bartlett para o teste escore em modelos lineares de valor extremo com parâmetro de dispersão σ conhecido. Mostraremos, na Seção 5, que se este parâmetro for substituído pelo seu EMV sob a hipótese nula H, mesmo assim, a correção tipo-Bartlett é ainda eficaz para melhorar a aproximação χ^2 para a estatística escore modificada.

4.1 O modelo de regressão linear simples

Para o modelo de regressão linear descrito na Seção 3.1, a estatística escore para testar a hipótese nula $H : \theta_1 = 0$ é dada por

$$S_R = \frac{1}{ns_2} \left\{ \sum_{i=1}^{n} \left[\exp\left(\frac{\overline{y} - y_i}{\sigma} \right) - 1 \right] x_i \right\}^2 .$$

Usando a mesma notação, obtemos, $A_1 = -\frac{24}{n}$, $A_2 = -18(1+\gamma_{2x})/n$ e $A_3 = 20\gamma_{1x}^2$.

4.2 O modelo de classificação de um fator

Para o modelo de classificação descrito na Seção 3.2, as matrizes $Z - Z_2$ e Z_2 têm como elementos $n\delta_{ij}/n_i - 1$ e n^{-1}, respectivamente, onde $\delta_{ij} = 1$ se i e j indicam observações da mesma população, e zero caso contrário. A estatística escore para testar $H : \theta = \cdots = \theta_{p-1} = 0$ pode ser escrita como

$$S_R = \sum_{i=11}^{p} n_i^{-1} \left\{ \sum_{j=1}^{n_i} \left(\frac{\overline{y} - y_{ij}}{\sigma} \right) - 1 \right\}^2 ,$$

onde $\overline{y} = \sum_{i=1}^{p} \sum_{j=1}^{n_i} y_{ij} / \sum_{j=1}^{p} n_i$ é a média amostral. Após alguma álgebra mostra-se que

$$A_1 = -\frac{24(p-1)}{n}, \quad A_2 = \frac{18}{n} \left\{ n \sum_{i}^{p} n_i^{-1} - (p^2 - 4p - 3) \right\},$$

e

$$A_3 = \frac{1}{n} \left\{ 20n - \sum_{i=1}^{p} n_i^{-1} - (3p^2 - 6p - 4) \right\}.$$

5 Resultados e discussões

Para uma comparação da *performance* dos estimadores de máxima verossimilhança $\hat{\theta}$ em contra partida com os estimadores de viés corrigidos, simulamos um modelo de regressão linear de valor extremo com parâmetro de locação fixado como $\rho_i = \theta_0 + \sum_{r=1}^{3} \theta_r x_{ir}$ e parâmetro de dispersão $\sigma = 1$ e $\sigma = 2$. As covariáveis $x_{i1},\ x_{i2}$ e x_{i3} foram escolhidas como quantidades de uma distribuição uniforme $U(0,1)$ e seus valores foram mantidos fixos para um mesmo tamanho de amostra. Sem perda de generalidade, tomamos os seguintes valores para os parâmetros da regressão: $\theta_0 = 2,\ \theta_1 = 1,\ \theta_2 = 3$ e $\theta_3 = -2$. O número de observações foi fixado em $n = 30$ e $n = 50$. Usamos um método conveniente de amostragem a partir de uma distribuição de valor extremo baseado na equação $y_i = \rho_i - \sigma \log(-\log u_i)$, onde a variável aleatória u_i tem uma distribuição uniforme $U(0,1)$.

Para cada conjunto de dados os estimadores de máxima verossimilhança $\hat{\theta}'s$ foram obtidos iterativamente a partir das equações dadas na Seção 1. Na simulação, foram geradas 10.000 amostras em cada caso e para análise do desempenho os cálculos foram feitos usando o pacote estatístico S-PLUS. Para cada uma das 10.000 replicações, calculamos o estimador $\hat{\theta}$ e o estimador corrigido $\tilde{\theta} = \hat{\theta} + \frac{\sigma}{2}(X^T X)^{-1} X^T \varepsilon$.

Tabela 1 – Estimadores não corrigidos (EMV) e estimadores corrigidos pelo viés (ECV)

	n		$\theta_0 = 2$	$\theta_1 = 1$	$\theta_2 = 3$	$\theta_3 = -2$
		EMV	1.61	1.34	2.28	−1.73
	30		(0.31)	(0.23)	(0.31)	(0.40)
		ECV	1.73	1.21	2.47	−1.84
$\sigma = 2$			(0.29)	(0.17)	(0.27)	(0.33)
		EMV	1.75	1.25	2.63	−1.81
	50		(0.19)	(0.16)	(0.24)	(0.31)
		ECV	1.83	1.14	2.79	−1.95
			(0.17)	(0.12)	(0.17)	(0.24)
		EMV	1.84	1.15	2.67	−1.85
	30		(0.27)	(0.19)	(0.27)	(0.31)
		ECV	1.89	1.12	2.70	−1.90
$\sigma = 1$			(0.21)	(0.13)	(0.18)	(0.27)
		EMV	1.88	1.20	2.84	−1.94
	50		(0.15)	(0.11)	(0.21)	(0.27)
		ECV	1.92	1.09	2.91	−1.97
			(0.11)	(0.08)	(0.14)	(0.21)

A Tabela 1 apresenta as médias das amostras de ambos estimadores não corrigidos (EMV) e os corrigidos pelo viés (ECV) com seus respectivos erros padrões entre parênteses. Esta tabela nos mostra que os vieses dos estimadores de máxima verossimilhança $\hat{\theta}$ são significantes mesmo para amostras de tamanho $n = 50$ e que o estimador corrigido pelo viés $\tilde{\theta}$ possui valores mais próximos dos parâmetros verdadeiros do que àqueles do estimador tradicional $\hat{\theta}$. É óbvio que quando n aumenta a correção de viés tem um menor impacto e que o estimador corrigido tende a ter erros padrões cada vez menores do que os estimadores usuais em amostras de tamanhos pequenos a moderados. Entretanto, pelo menos neste caso, a correção de viés produz uma redução no erro médio quadrático de segunda ordem do estimador aperfeiçoado $\tilde{\theta}$ por causa da redução dos erros padrões. Note que os erros padrões dos estimadores corrigidos pelo viés são inferiores aos correspondentes erros padrões dos estimadores usuais.

Faremos agora um estudo de simulação Monte Carlo num modelo de regressão linear de três parâmetros

$$\rho = \alpha + \theta_1 x_1 + \theta_2 x_2 \tag{13}$$

cujo objetivo é testar a hipótese nula $H : \theta_1 = 0$ *versus* a hipótese alternativa $A : \theta_1 \neq 0$. A estatística da razão de verossimilhanças e a estatística escore são dadas, respectivamente, por

$$\omega = 2\left\{\ell(\hat{\theta}) - \ell(\tilde{\theta})\right\} = \frac{2}{\sigma}\sum_{i=1}^{n}(\hat{\rho}_i - \tilde{\rho}_i) + 2\sum_{i=1}^{n}\left[\exp\left\{-\frac{(y_i - \tilde{\rho}_i}{\sigma}\right\} - \exp\left\{-\frac{(y_i - \hat{\rho}_i)}{\sigma}\right\}\right]$$

e $S_R = \tilde{g}^T X_1 (R^T R)^{-1} X_1^T \tilde{g}$, onde g tem uma componente típica da forma $1 + e^{-(y-\rho)/\sigma}$. A correção de Bartlett é obtida de (4) e as correções tipo Bartlett são dadas pelas equações (7)-(12). Nas simulações, os parâmetros foram fixados em $\alpha = 1$, $\theta_1 = 0$, $\theta_2 = 1.5$ e $\sigma = 4$, as variáveis independentes x_1 e x_2 foram escolhidas de maneira aleatória a partir de uma distribuição uniforme nos intervalos (0,1) e (0,5), respectivamente. Seus valores foram mantidos constantes durante as simulações realizadas com amostras do mesmo tamanho. As simulações foram baseadas em 10.000 amostras em cada caso para $n = 20, 30, 40, 50, 60$ e 80. Foram comparadas as estatística da razão de verossimilhanças (ω) e sua versão aperfeiçoada (ω^*), a estatística escore (S_R) e suas três versões aperfeiçoadas: a estatística escore corrigida (S_R^*) dada em (7) e as estatísticas escore corrigidas monótonas $K(S_R)$ e $\tilde{S}_R$ determinadas pelas equações (8)-(9), respectivamente, com testes da razão de verossimilhanças e escore baseados em *bootstrap*. As estatísticas aperfeiçoadas ω^*, S_R^*, $K(S_R)$ e $\tilde{S}_R$ seguem de (4) e (7)-(12), juntamente com as quantidades a, b e c.

O algoritmo usado para estimação das taxas de rejeição dos testes da razão de verossimilhanças e escore usando o *bootstrap* é dado a seguir:

i. a partir do modelo (13) simulamos os dados $y_1, \ldots, y_n$. Depois, obtém-se os estimadores de máxima verossimilhança, usando o algoritmo descrito na Seção 1;

ii. em seguida aplicamos um esquema de reamostragem substituindo os dados $y_1, \ldots, y_n$ por uma pseudo-amostra $y_1^*, \ldots, y_n^*$;

iii. calculamos os estimadores de *bootstrap* $\hat{\theta}^*$ e $\tilde{\theta}^*$ das equações dadas em i) aplicadas para o conjunto de dados de *bootstrap* $y_1^*, \ldots, y_n^*$. Então, são calculadas as estatísticas de *bootstrap* da razão de verossimilhanças ω_B e escore S_{R_B};

iv. usamos as etapas ii) e iii) para obter um total de $B = 500$ valores da estatística da razão verossimilhanças ω_B e da estatística escore S_{R_B}. O *bootstrap* baseado nos testes da razão de verossimilhanças e escore rejeitam a hipótese nula $H : \theta_1 = 0$ $100\alpha\%$ se $(\#\{\omega_B \geq \omega\}+1)/(B+1) \leq \alpha$ e $(\#\{S_{R_B} \geq S_R\}+1)/(B+1) \leq \alpha$, respectivamente;

v. os tamanhos empíricos de *bootstrap* w_B e S_{R_B} correspondentes ao nível nominal $\alpha\%$ igualam às proporções de casos nas 10.000 simulações correspondentes à rejeição da hipótese nula $H : \theta_1 = 0$.

Tabela 2 – Taxas de rejeição estimadas da razão de verossimilhanças, da razão de verossimilhanças aperfeiçoada, escore, dos três testes escore aperfeiçoados e testes *bootstrap*

Tam.da Amostra	Tam. Nom.	w	w^*	S_R	S_R^*	$K(S_R)$	$\tilde{S}_R$	w_B	S_{R_B}
20	10	35.8	17.2	15.8	10.9	11.4	10.9	12.7	10.0
	5	29.7	10.6	9.7	5.6	5.6	5.7	5.3	5.6
30	10	26.4	13.1	15.2	10.3	10.5	10.8	10.5	11.3
	5	17.2	7.9	8.8	5.2	5.4	5.3	5.2	5.5
40	10	15.9	11.8	12.1	10.2	10.3	10.3	10.2	11.5
	5	10.5	6.4	6.2	5.1	5.1	5.1	5.1	5.2
50	10	15.1	11.3	11.5	10.0	10.1	10.2	10.1	10.8
	5	8.7	5.9	5.8	4.8	5.2	4.9	5.0	5.3
60	10	13.9	10.8	11.2	9.9	10.3	10.0	10.1	10.5
	5	7.8	5.2	5.7	5.1	5.1	5.1	4.9	5.1
80	10	12.1	10.3	10.9	9.8	10.2	10.0	10.0	10.3
	5	6.7	5.1	5.2	5.0	4.9	5.1	5.0	5.1

A Tabela 2 apresenta os tamanhos empíricos dos oito testes (w, w^*, S_R, S_R^*, $K(S_R)$, $\tilde{S}_R$, w_B e S_{R_B}), correspondentes aos níveis nominais de 10% e 5%. As seis primeiras colunas correspondem às taxas de rejeição das estatísticas. As duas últimas colunas se referem aos tamanhos empíricos de *bootstrap* descritos na etapa (v) do algoritmo acima. A Tabela 2, também, nos revela algumas informações relevantes. Primeiro, os testes da razão de verossimilhanças e escore possuem taxas de rejeição bem superiores aos níveis fixados. A aproximação assintótica por χ_1^2 para estatísticas não modificadas funciona bem para n grande em concordância com os resultados assintóticos de grandes amostras. Observamos que a estatística escore S_R produz um teste mais preciso quando comparado com a distribuição χ_1^2 do que o teste baseado na estatística da razão de verossimilhanças. As correções de Bartlett para a estatística w e as correções tipo-Bartlett para a estatística S_R são muito eficazes, pois empurram as taxas de rejeição na direção dos níveis nominais, embora convém salientar que isso possui menor impacto quando o tamanho da amostra aumenta. Vale destacar que o teste w^* corrigido e o teste w_B baseado no *bootstrap* possuem valores simulados mais próximos dos níveis nominais do que o teste w não modificado. O mesmo ocorre para os três testes escore corrigidos e o teste escore baseado no *bootstrap*. Assim, a correção analítica e aquela de reamostagem de *bootstrap* são eficazes em empurrar as taxas de rejeição das estatísticas modificadas na direção dos níveis nominais correspondentes. Em outras palavras, as simulações mostram que em amostras finitas, a *performance* das correções de Bartlett e das correções tipo-Bartlett são equivalentes aos testes *bootstrap*. Embora consideramos apenas um conjunto particular de parâmetros no modelo (13), para várias variações dos parâmetros α, θ_1, θ_2 e σ, foram obtidos (aproximadamente) os mesmos resultados descritos na Tabela 2.

Investigaremos agora como os tamanhos e os poderes das estatísticas da razão de verossimilhanças, escore e suas versões corrigidas são sensíveis à falta de especificação do parâmetro de dispersão. Para isso, consideramos o modelo (13) e calculamos as taxas de rejeição e poderes das estatísticas da razão de verossimilhanças não corrigida, corrigida, escore não corrigida e escore corrigida (8) considerando valores incorretos para o parâmetro de dispersão a um nível nominal de 5% da distribuição χ_1^2 de referência. Suponha que $\sigma_T = 4$ seja o valor verdadeiro do parâmetro de dispersão. Usamos um sobrescrito (i) para $i = 1, 2, 3, 4$ sobre as estatísticas corrigidas para indicar que elas correspondem à falta de especificação do parâmetro de dispersão quando σ está fixo em $\sigma_T - 2,\ \sigma_T - 1,\ \sigma_T + 1,\ \sigma_T + 2$, respectivamente. As simulações dos poderes para os testes não modificados e para os testes modificados foram feitos fixando os tamanhos das amostras em $n = 30$ e assumindo que $\alpha = 1$ e $\theta_2 = 1.5$ e $\theta_1 = 0.2\,,\, 0.4\,,\, 0.6$ e 0.8. As simulações de poder e tamanho foram baseadas em 10.000 replicações. Sob a hipótese nula foram obtidas as taxas de rejeição de todas as estatísticas e suas formas corrigidas para $i = 1, \ldots, 4$ ao nível nominal de 5% da distribuição χ_1^2 de referência.

A Tabela 3 apresenta os tamanhos empíricos (taxas de rejeição) de todas as estatísticas (sob verdade e sob a não-especificação do parâmetro de dispersão) para $n = 20, 30, 40$ e 50 e a Tabela 4 mostra seus poderes estimados para $n = 30$. Analisando as celas da Tabela 3, observamos que: (i) as taxas de rejeição das estatísticas ω não corrigida e escore S_R estão superestimadas e subestimadas, respectivamente, em todos os casos, especialmente quando n é pequeno; (ii) os tamanhos empíricos de todas as estatísticas corrigidas estão mais próximos de 5% do que o tamanho empírico das estatísticas ω e escore S_R não corrigidas.

Tabela 3 – Taxas de rejeição estimadas da razão de verossimilhanças, da razão de verossimilhanças aperfeiçoada, escore, escore aperfeiçoada e das estatísticas melhoradas à falta de especificação do parâmetro de dispersão

n	ω	ω^*	${\omega^*}^{(1)}$	${\omega^*}^{(2)}$	${\omega^*}^{(3)}$	${\omega^*}^{(4)}$
20	17.8	7.5	13.4	8.6	10.1	14.4
30	13.2	6.4	11.3	7.9	8.7	13.1
40	9.4	5.7	8.4	7.7	8.1	8.8
50	7.6	5.2	6.5	6.3	6.8	7.1
	S_R	S_R^*	${S_R^*}^{(1)}$	${S_R^*}^{(2)}$	${S_R^*}^{(3)}$	${S_R^*}^{(4)}$
20	2.3	3.9	3.5	3.7	3.8	3.6
30	3.2	4.5	4.2	4.4	4.1	3.9
40	3.5	4.6	4.1	4.3	4.4	4.2
50	4.1	4.8	4.3	4.5	4.6	4.4
	S_R	$K(S_R)$	$K(S_R)^{(1)}$	$K(S_R)^{(2)}$	$K(S_R)^{(3)}$	$K(S_R)^{(4)}$
20	2.3	4.1	3.6	3.8	3.7	3.4
30	3.2	4.2	3.7	4.0	3.8	3.5
40	3.5	4.4	3.8	3.9	4.0	3.7
50	4.1	4.7	4.0	4.2	4.3	3.8
	S_R	$\tilde{S}_R$	$\tilde{S}_R^{(1)}$	$\tilde{S}_R^{(2)}$	$\tilde{S}_R^{(3)}$	$\tilde{S}_R^{(4)}$
20	2.3	4.0	3.7	3.9	3.8	3.4
30	3.2	4.3	4.0	4.1	3.9	3.8
40	3.5	4.5	4.2	4.3	4.1	3.9
50	4.1	4.7	4.4	4.5	4.3	4.1

Além disso, o tamanho dos testes corrigidos não são substancialmente afetados pela não-especificação do parâmetro de dispersão, pelo menos quando σ pertence ao intervalo (2,6) e, até mesmo, quando σ não está especificado corretamente, as estatísticas modificadas induzem a alguns melhoramentos em termos de aproximação pela distribuição qui-quadrado; (iii) como era de se esperar, a aproximação χ_1^2 para as estatísticas corrigidas deteriora-se quando a distância entre σ e o valor verdadeiro 4 aumenta; (iv) as taxas de rejeição de todas as estatísticas aproximam-se do nível nominal 5% quando $n \to \infty$.

As celas da Tabela 4 mostram que as estatísticas modificadas ω^*, S_R^*, $K(S_R)$, e $\tilde{S}_R$ têm poderes simulados maiores do que os correspondentes das estatísticas não modificadas ω e S_R e das estatísticas modificadas ${\omega^*}^{(i)}$, ${S_R^*}^{(i)}$, $K(S_R)^{(i)}$ e $\tilde{S}_R^{(i)}$ sob a não especificação do parâmetro de dispersão. A não especificação do parâmetro de dispersão induz, portanto, a uma redução no poder das estatísticas corrigidas de Bartlett e tipo-Bartlett, conforme pode ser comprovado na Tabela 4.

Encontraram-se expressões matriciais para os vieses de ordem n^{-1} das estimativas de máxima verossimilhança dos parâmetros no modelo de regressão linear de valor extremo. Os resultados obtidos na forma matricial possuem grande vantagem nos cálculos operacionais. Através de estudos de simulação os estimadores corrigidos têm melhor desempenho do que àqueles não-corrigidos, pelo menos em termos de proximidade dos parâmetros verdadeiros.

Tabela 4 – Simulações do poder para $n = 30$

θ_1	ω	ω^*	${\omega^*}^{(1)}$	${\omega^*}^{(2)}$	${\omega^*}^{(3)}$	${\omega^*}^{(4)}$
0.2	21.2	32.5	15.6	17.3	18.2	16.2
0.4	48.7	63.2	37.2	42.5	50.6	39.4
0.6	78.2	93.4	63.1	72.5	80.3	59.7
0.8	94.0	98.3	77.1	81.4	86.2	75.5
	S_R	S_R^*	${S_R^*}^{(1)}$	${S_R^*}^{(2)}$	${S_R^*}^{(3)}$	${S_R^*}^{(4)}$
0.2	35.4	42.6	28.5	33.7	35.9	38.2
0.4	41.2	51.8	37.3	40.4	42.3	35.6
0.6	68.6	83.2	57.8	69.3	70.1	73.2
0.8	91.2	94.5	81.3	84.6	82.7	77.9
	S_R	$K(S_R)$	$K(S_R)^{(1)}$	$K(S_R)^{(2)}$	$K(S_R)^{(3)}$	$K(S_R)^{(4)}$
0.2	35.4	45.8	38.2	40.3	41.7	35.2
0.4	41.2	62.7	53.7	58.2	56.4	51.3
0.6	68.6	89.3	66.9	77.2	70.3	64.1
0.8	91.2	97.2	87.6	93.8	90.2	85.1
	S_R	$\tilde{S}_R$	$\tilde{S}_R^{(1)}$	$\tilde{S}_R^{(2)}$	$\tilde{S}_R^{(3)}$	$\tilde{S}_R^{(4)}$
0.2	35.4	47.5	41.3	44.5	45.2	42.3
0.4	41.2	66.2	50.4	56.3	54.1	49.7
0.6	68.6	93.1	65.2	88.2	80.3	59.3
0.8	91.2	95.6	87.1	90.1	84.2	81.0

Conclusões

Há várias décadas, estudos são desenvolvidos no sentido de se obter estatísticas modificadas cujas distribuições são mais bem aproximadas por χ^2 no qual Bartlett foi o pioneiro.

Obteve-se estatísticas aperfeiçoadas no modelo de regressão linear de valor extremo cujas aproximações por χ^2 têm uma inferência mais refinada, principalmente, em pequenas amostras nas quais as estatísticas testes usuais dão resultados imprecisos em termos de aproximação assintótica. As estatísticas consideradas foram a razão de verossimilhanças e escore.

Resultados de simulação mostram, também, que as estatísticas corrigidas ω^* e S_R^* para o modelo de valor extremo, têm uma melhor performance do que as estatísticas não corrigidas, uma vez que suas taxas de rejeição são mais próximas dos níveis nominais da distribuição χ^2. Fez-se um estudo de sensibilidade das taxas de rejeição e dos poderes pela falta de especificação correta do parâmetro de dispersão.

RAMOS, E.M.; CORDEIRO, G.M. Biases and Bartlett and Bartlett-type corrections in extreme value linear regression model. *Rev. Mat. Estat.,* São Paulo, v.24, n.1, p.77-93, 2006.

- *ABSTRACT: In this paper, we obtain bias corrected maximum likelihood estimates in extreme value linear regression models. We also consider finite-sample corrections in matrix notation for the likelihood ratio and score statistics in these models. Some simulations are performed to compare the likelihood ratio and score statistics and their modified versions with the chi-squared distribution. We also illustrate the bias correction.*

- *KEYWORDS: Bias correction;* bootstrap *test; Chi-squared distribution; Extreme value model; Maximum likelihood estimate.*

Referências

BECKER, R.A.; CHAMBERS, J.M.; WILKS, A.R. *The New S Language: A Programming Environment for Data Analysis and Graphics*, Pacific Grove, CA: Wadsworth and Brooks/Cole, 1988.

BHATTACHARYA, R.N.;GHOSH, J.K. On the validity of the formal Edgeworth Expansion. *Ann. Statist.,* v.6, p.434-451, 1978.

BHATTACHARYA, R.N.; RAO, R.R. *Normal Approximation and Asymptotic Expansions*, New York: John Wiley, 1976.

CANFIELD, R.V.; BORGMAN, L.E. (1975). Some distributions of time to failure for reliability applications. *Technometrics*, v.17,, p.263-268, 1975.

CORDEIRO, G.M. Improved likelihood ratio statistics for generalized linear models. *J.R. Statist. Soc. B*, v.45, p.404-413, 1983.

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

CORDEIRO, G.M.; BOTTER, D.A. Second-order biases of maximum likelihood estimates in overdispersed generalized linear models. *Statist. Prob. Lett.*, v.55, p.269-280, 2001.

CORDEIRO, G.M.; PAULA, G.A.; BOTTER, D.A. Improved likelihood ratio tests for dipersion models. *Int. Statist. Rev.*, v.62, p.257-276, 1994.

CORDEIRO, G.M.; VASCONCELLOS, K.L.P.; SANTOS, M.L.F. On the second-order bias of parameter estimates in nonlinear regression models with Student t errors. *J. Statist. Comput. Simul.*, v.60, p.363-378, 1998.

CORDEIRO, G.M.; McCULLAGH, P. Bias correction in generalized linear models. *J.R. Statist. Soc. B*, v.53, p.629-643, 1991.

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

CORDEIRO, G.M.; FERRARI, S.L.P.; PAULA, G.A. Improved score tests for generalized linear models. *J. R. Statist. Soc. B*, v.53, p.629-643, 1993.

CORDEIRO, G.M.; FERRARI, S.L.P. Bartlett-type corrections for some score dispersion models. *Comm. Statist. Theory Meth.*, v.25, p.29-48, 1996.

CORDEIRO, G.M.; FERRARI, S.L.P.; CYSNEIROS, A.H.M.A. A formula to improve score test statistics. *J. Statist. Comput. Simul.*, v.62, p.123-136, 1998.

CORDEIRO, G.M.; FERRARI, S.L.P., URIBE-OPAZO, M.; VASCONCELLOS, K.L.P. Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models. *Statist. Prob. Lett.*, v.46, p.317-328, 2000.

COX, D.R.; SNELL, E.J. A general definition of residuals (with discussion). *J.R. Statist. Soc. B,* v.30, p.248-275, 1968.

CRIBARI-NETO, F.; CORDEIRO, G.M. On Bartlett and Bartlett-type corrections. *Econometric Reviews*, v.15, p.339-367, 1996.

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

CRIBARI-NETO, F.; FERRARI, S.L.P An improved Lagrange multiplier test for heteroskedasticity. *Comm. Statist. Simul. Comput.*, v.24, p.31-44, 1995b.

CYSNEIROS, A.H.M.A.; CORDEIRO, G.M. On improving the χ^2 approximation of score tests in location-scale nonlinear models. *Comm. Statist. Theory Meth.*, v.31, p.1709-1732, 2002.

FERRARI, S.L.P. *Aperfeiçoamento de Testes Escore, Aplicações e Extensões*, 1991. Tese (Doutorado) – Instituto de Matemática e Estatística da Universidade de São Paulo, 1991.

GUMBEL, E.J. *Statistics of Extremes*, New York: Columbia University Press, 1958.

HARRIS, P. An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika,* v.72, p.653-659, 1985.

KAKIZAWA, Y. Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika,* v.83, p.923-927, 1996.

LAWLEY, D.N. A general method for approximating to the distribution of the likelihood ratio criteria. *Biometrika,* v.71, p.233-244, 1956.

Recebido em 12.05.2004.

Aprovado após revisão em 15.03.2006.