COMPARAÇÃO DAS EXPANSÕES DE EDGEWORTH, LUGANNANI-RICE, DANIELS E CORDEIRO-FERRARI COM APLICAÇÕES NA ESTATÍSTICA

Rejane dos Santos BRITO[1]
Gauss Moutinho CORDEIRO[1]

▪ RESUMO: Neste artigo revisa-se as expansões de Laplace, ponto de sela e de Edgeworth, e apresenta-se exemplos que mostram a aplicação destas expansões assintóticas na estatística. Usando a expansão de Laplace para integrais e a expansão de Edgeworth, mostra-se como calcular a expansão ponto de sela para a função densidade de uma única variável aleatória e, ainda, como obter a função densidade da média amostral de variáveis aleatórias independentes e identicamente distribuídas. Cordeiro e Ferrari (1998) propuseram uma estatística que aproxima a soma padronizada de variáveis aleatórias independentes e identicamente distribuídas até ordem $O(n^{-1})$, em que n é o tamanho da amostra. Verifica-se, também, a eficiência desta estatística comparando-a com as expansões de Edgeworth, Lugannani e Rice (1980) e Daniels (1987).

▪ PALAVRAS-CHAVE: Aproximação de Edgeworth; aproximação de Laplace; aproximação ponto de sela; correção de Bartlett generalizada; identidade de Bartlett.

1 Introdução

O estudo das expansões ponto de sela teve início com Esscher (1932), mas Daniels (1954) foi o pioneiro no assunto, pois aplicou a expansão ponto de sela na área da estatística. Lugannani e Rice (1980) apresentaram a expansão ponto de sela e sua aplicabilidade referente à soma estocástica de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.).

Em muitas aplicações estatísticas, as expansões têm sua importância no que se refere, por exemplo, à construção de testes e intervalos de confiança, além,

[1] Departamento de Estatística e Informática – DEINFO, Universidade Federal Rural de Pernambuco – UFRPE, CEP: 50171-900, Recife, Pernambuco, Brasil. E-mail: *janesbrito@gmail.com* *gausscordeiro@uol.com.br*

também, do cálculo dos p-valores. Em muitos casos, a metodologia ponto de sela é utilizada para determinar limites uniformes com erros relativos sobre o intervalo da distribuição (Kolassa, 1997). Sabe-se ainda que as expansões ponto de sela são muito importantes na teoria assintótica, pois aproximam com grande precisão as funções densidade e de distribuição da soma e da média de variáveis aleatórias i.i.d. (Goutis e Casella, 1999). Uma escolha da aproximação de forma ampla é baseada na expansão de Edgeworth e na expansão inversa de Cornish-Fisher, mas, apesar da importância no estudo teórico das funções densidade e distribuição e aproximação dos quantis, estas expansões pondem dar estimativas negativas da função densidade e, também, aproximações não-monótonas da função distribuição e das funções dos quantis quando aplicadas a situações reais. Um alternativa natural é o uso das ferramentas de expansão assintótica para integrais, tais como, aproximações de Laplace e aproximações ponto de sela (Davison, 2001).

A obtenção das expansões ponto de sela pode ser feita através do uso de ferramentas como exponencial "inclinada", expansões de Edgeworth, polinômios de Hermite, integrações complexas e outras noções avançadas. As expansões ponto de sela são facilmente deduzidas da função geratriz de cumulantes da variável aleatória de interesse.

Os testes estatísticos tradicionais, como aqueles da razão de verossimilhanças, Wald e escore, são baseados em grandes amostras e o comportamento das estatísticas só pode ser estudado quando o tamando da amostra n tende para infinito. Para aplicar esses testes em amostras pequenas e moderadas, existem melhoramentos propostos através das correções de Bartlett e tipo-Bartlett. Uma revisão dessas correções pode ser encontrada em Cribari-Neto e Cordeiro (1996).

Na Seção 2, apresenta-se de forma resumida a aproximação de Laplace. Na Seção 3, considera-se a expansão ponto de sela em formas distintas. A Seção 4 trata das identidades de Barttlet, das estatísticas clássicas usuais e suas versões melhoradas no contexto estritamente uniparamétrico, com alguns estudos de simulação comparando as estatísticas usuais e suas modificações. A Seção 5 aborda a correção de Bartlett generalizada proposta por Cordeiro e Ferrari (1998) e a compara com as expansões de Edgeworth, Lugannani e Rice (1980) e Daniels (1987), na aproximação da distribuição t de Student pela distribuição normal. Na Seção 6, revisa-se a quadratura de Gauss-Hermite. Ao longo do artigo apresentam-se diversas aplicações à Estatística. Finalmente, a Seção 7 apresenta algumas conclusões.

2 Aproximação de Laplace

Goutis e Casella (1999) explicitam a obtenção da expansão de Laplace mediante o uso dos primeiros termos da expansão em série de Taylor. Utiliza-se a função $h(y) = \log f(y)$ tal que $f(y) = \exp\{h(y)\}$. Escolhendo $\hat{y}$ como um ponto a ser expandido em torno de y, tem-se

$$f(y) \approx \exp\left\{h(\hat{y}) + \frac{(y-\hat{y})^2}{2} h^{''}(\hat{y})\right\}. \qquad (1)$$

A aproximação (1) pode ser usada para calcular integrais de funções positivas, tais como $\int f(y)dy$ dentro do intervalo (a, b). Ao expandir o integrando como em (1), obtém-se

$$\int_a^b f(y)dy \approx \int_a^b \exp\left\{h(\hat{y}) + \frac{(y-\hat{y})^2}{2} h^{''}(\hat{y})\right\} dy. \tag{2}$$

Se $\hat{y}$ é um ponto de máximo, $h^{''}(\hat{y})$ é negativo e o lado direito de (2) pode ser calculado, pois a parte principal da equação representa o integrando da distribuição normal com média $\hat{y}$ e variância $-1/h^{''}(\hat{y})$. Dessa forma, sendo agora o intervalo (a, b) equivalente ao intervalo $(-\infty, +\infty)$, obtém-se

$$\int_{-\infty}^{+\infty} f(y)dy \approx \exp\{h(\hat{y})\} \left\{-\frac{2\pi}{h^{''}(\hat{y})}\right\}^{1/2}. \tag{3}$$

A equação (3) é conhecida como aproximação de Laplace para integrais.

2.1 Fórmula da inversão e função geradora de cumulantes

Considere que Y é uma variável aleatória tendo função densidade ou função de probabilidade $f(y)$. As diferentes maneiras de aproximar estas funções são obtidas calculando a função geradora de momentos (f.g.m.)

$$\phi(\lambda) = \int_{-\infty}^{+\infty} \exp(\lambda y) f(y) dy, \tag{4}$$

em que $f(y)$ é a função densidade de probabilidade para o caso de uma variável aleatória contínua. De modo análogo, faz-se o uso do somatório para o caso de uma variável aleatória discreta e, assim, $f(y)$ é uma função de probabilidade. A fórmula da transformação inversa implica

$$\begin{aligned} f(y) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(i\lambda) \exp(-i\lambda y) dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp\{K(i\lambda) - i\lambda y\} d\lambda, \end{aligned} \tag{5}$$

em que $i = \sqrt{-1}$, $\phi(i\lambda)$ é a função característica de Y e $K(\lambda) = \log \phi(\lambda)$. A função $K(\lambda)$ é conhecida como função geradora de cumulantes (f.g.c.) de Y e desempenha um papel importante na teoria assintótica.

3 Expansões ponto de sela

A primeira aplicação estatística da expansão ponto de sela foi feita por Daniels (1954) através da determinação de uma aproximação da função densidade utilizando a transformada de Fourier. A aproximação pode ser obtida através da família exponencial conjugada e, também, através de uma relação entre o ponto de sela e o estimador de máxima verossimilhança (E.M.V.). Uma dificuldade em se usar esta expansão vem do fato de ser necessário o uso de integrações complexas.

Segundo Daniels (1954), existem formas distintas de se obter (analiticamente) a aproximação ponto de sela e suas expansões associadas em potências de ordem $O(n^{-1})$ para a função densidade da média amostral $\bar{Y}$ de uma amostra aleatória i.i.d.

3.1 Expansão ponto de sela através do método de Laplace

Considera-se a técnica da expansão de Laplace apresentada na Seção 2. Realizando uma mudança de variável na fórmula da transformação inversa (5), $\lambda' = i\lambda$, obtém-se uma função densidade em termos da integral complexa

$$f(y) = \frac{1}{2\pi i} \int_{t-i\infty}^{t+i\infty} \exp\{K(\lambda) - \lambda y\} d\lambda \tag{6}$$

para t em torno de zero. Segundo um teorema de análise complexa (teorema da curva fechada), pode-se aproximar um valor de t de modo a obter a integração.

Aproximando a integral de $\exp\{K(\lambda)\}$ com relação à variável λ através da expansão de Laplace, obtém-se a equação

$$f(y) \approx \int_{-\infty}^{+\infty} \exp\left\{ K(\hat{\lambda}) + \frac{(\lambda-\hat{\lambda})^2}{2} \left.\frac{\partial^2 K(\lambda)}{\partial \lambda^2}\right|_{\hat{\lambda}} \right\} d\lambda$$

ou

$$f(y) \approx \exp\{K(\hat{\lambda})\} \left\{ - \frac{2\pi}{\left.\frac{\partial^2 K(\lambda)}{\partial \lambda^2}\right|_{\hat{\lambda}}} \right\}^{1/2}, \tag{7}$$

em que, para cada y, $\hat{\lambda}$ satisfaz $\partial K(\lambda)/\partial\lambda = 0$ e $\partial^2 K(\lambda)/\partial\lambda^2 < 0$ e, assim, maximiza $K(\lambda)$. Seja $\hat{\lambda}$ obtido pela solução de $K'(\hat{\lambda}) = y$. Pode-se imaginar que esta solução se refere ao E.M.V. no modelo da família exponencial baseado na observação y. Expandindo $K(\lambda) - \lambda y$, referente ao expoente da equação (6), em série de Taylor, tem-se

$$K(\lambda) - \lambda y \approx K(\hat{\lambda}) - \hat{\lambda} y + \frac{(\lambda-\hat{\lambda})^2}{2} K''(\hat{\lambda})$$

e associando esta aproximação com a equação (7), obtém-se

$$f(y) \approx \left\{ \frac{1}{2\pi K''(\hat{\lambda})} \right\}^{1/2} \exp\{K(\hat{\lambda}) - \hat{\lambda} y\}. \tag{8}$$

A equação (8) é denominada *expansão ponto de sela* para uma função densidade ou função de probabilidade e seu erro de aproximação é bem menor que o da função obtida através da aproximação em série de Taylor. A sua aplicabilidade depende apenas do conhecimento da f.g.c. $K(\lambda)$ e do cálculo de $\hat{\lambda}$ em $K'(\hat{\lambda}) = y$. Esta equação é, em geral, não-linear.

Em muitas aplicações, a equação ponto de sela ($K'(\hat{\lambda}) = y$) pode não ser resolvida analiticamente, mesmo quando a solução de $\hat{\lambda}$ existe. Mesmo assim, os métodos de ponto de sela podem ser aplicados resolvendo a equação numericamente. Usa-se o algoritmo de Newton-Raphson para calcular o ponto de sela, tendo este em geral bom desempenho desde que a função $K(\lambda) - \lambda y$, que é minimizada, seja convexa. Há ainda o método da secante para encontrar o ponto de sela. Esse método, geralmente, produz a resposta correta em uma iteração se $K'(\lambda)$ é linear, como é o caso do método de Newton-Raphson. Como no algoritmo de Newton-Raphson, se $K''(\lambda)$ varia muito rapidamente, a convergência pode ser muito lenta, então, não existe a possibilidade de divergência para o método secante (Kolassa, 1997).

Uma aplicação importante da expansão ponto de sela é a aproximação da função densidade ou função de probabilidade de uma soma ou média de variáveis aleatórias i.i.d. Por exemplo, é possível determinar uma aproximação para a função densidade de uma soma ou média de variáveis aleatórias com distribuição exponencial de parâmetro α. Como a equação (8) é facilmente obtida através da f.g.c., tem-se que a f.g.m. da média amostral ($\bar{Y} = \sum_{i=1}^{n} Y_i/n$) é dada por $\phi_{\bar{Y}}(\lambda) = \phi(\lambda/n)^n$ e, portanto, a f.g.c. de $\bar{Y}$ é $K_{\bar{Y}}(\lambda) = nK(\lambda/n)$. Assim, uma aplicação direta de (8) produz a expansão ponto de sela da função densidade de $\bar{Y}$

$$f_{\bar{Y}}(\bar{y}) \approx \left\{ \frac{n}{2\pi K''(\hat{\lambda})} \right\}^{1/2} \exp[n\{K(\hat{\lambda}) - \hat{\lambda}\bar{y}\}]. \tag{9}$$

A qualidade da expansão ponto de sela pode ser, frequentemente, obtida pela multiplicação da aproximação da função densidade por uma constante de forma que sua integração resulte um. Uma versão da renormalização é exata para as funções densidades da normal, gama e normal inversa (Daniels, 1980).

Exemplo 1. A aplicação da equação (8) é mostrada a seguir por meio da obtenção da função densidade da distribuição $\chi^2_{p,\alpha}$ não-central. Essa distribuição não apresenta forma fechada e sua função densidade é escrita como

$$f(y) = \sum_{k=0}^{\infty} \frac{y^{p/2+k-1} e^{-y/2}}{\Gamma(p/2+k) 2^{p/2+k}} \frac{\alpha^k e^{-\alpha}}{k!},$$

em que p são os graus de liberdade e α é o parâmetro de não-centralidade. A f.g.m. é expressa em forma fechada por

$$\phi_Y(\lambda) = \frac{\exp\{2\alpha\lambda/(1-2\lambda)\}}{(1-2\lambda)^{p/2}}.$$

O ponto de sela da equação é obtido por $K'(\hat{\lambda}) = y$ de forma que

$$\hat{\lambda}(y) = \frac{-p + 2y - \sqrt{p^2 + 8\alpha y}}{4y}$$

e, mediante a equação (8), é calculada a função densidade aproximada da distribuição $\chi^2_{p,\alpha}$ não-central por

$$f(y) \approx \left\{ \frac{-4\pi(4\alpha + p - 2p\lambda)}{(-1+2\lambda)^3} \right\}^{-1/2} \exp\left\{ \frac{2\alpha\lambda}{1-2\lambda} - \frac{1}{2} p \log(1-2\lambda) + \frac{p}{4} - \frac{y}{2} \right. \\ \left. + \frac{1}{4}\sqrt{p^2 + 8\alpha y} \right\}.$$

Exemplo 2. Tendo por base a equação (9), obtém-se a expansão ponto de sela para a média amostral de uma distribuição exponencial de parâmetro um. A f.g.c. desta distribuição é dada por $K(\lambda) = -\log(1-\lambda)$. O ponto de sela $\hat{\lambda}$ é obtido de $K'(\hat{\lambda}) = \bar{y}$. Dessa maneira, $\hat{\lambda}$ é obtido de $\hat{\lambda} = 1 - 1/\bar{y}$ e, tem-se ainda, que $K(\hat{\lambda}) = \log(\bar{y})$ e $K''(\hat{\lambda}) = \bar{y}^2$. Através da equação (9) a aproximação ponto de sela para a média amostral da distribuição exponencial com parâmetro um é

$$f_{\bar{Y}}(\bar{y}) \approx \left(\frac{n}{2\pi}\right)^{1/2} \bar{y}^{n-1} \exp\{-n(\bar{y}-1)\}.$$

Exemplo 3. A aplicação da equação (9) é agora introduzida para a média amostral da distribuição de Poisson. Sejam $Y_1, \ldots, Y_n$ variáveis aleatórias i.i.d. que seguem uma distribuição de Poisson com média α. A f.g.c. de Y_i é $K(\lambda) = \alpha\{\exp(\lambda) - 1\}$ cujo ponto de sela iguala $\hat{\lambda} = \log(\bar{y}/\alpha)$. A equação (9) pode ser usada diretamente, sendo que a média $\bar{y} = s/n$ admite agora somente valores inteiros. Dessa forma, tem-se

$$\begin{aligned} f_{\bar{Y}}(\bar{y}) &\approx \left(\frac{n}{2\pi}\right)^{1/2} \exp\left[n \left\{ \alpha\left(\frac{\bar{y}}{\alpha} - 1\right) - \left(\log\frac{\bar{y}}{\alpha}\right)\bar{y} \right\} \right] \\ &= \left(\frac{n}{2\pi}\right)^{1/2} \frac{\alpha^{n\bar{y}} e^{n\{\bar{y}-\alpha\}}}{\bar{y}^{n\bar{y}+1/2}}. \end{aligned}$$

3.2 Expansão ponto de sela através da expansão de Edgeworth

Segundo Daniels (1987), um enfoque para o estudo da expansão de Edgeworth utiliza a idéia pioneira de Esscher (1932) e que foi explorada por Cramér (1938), Blackwell e Hodges (1959), Bahadur e Ranga Rao (1960), Barndorff-Nielsen e Cox (1979), Robinson (1982), entre outros.

Apesar da derivação original de ponto de sela de Daniels (1954) ter sido baseada na inversa da função característica, existe uma derivação alternativa que Reid (1988) denominou de "versão mais estatística" baseada na expansão de Edgeworth. A expansão de Edgeworth de uma distribuição é obtida expandindo a f.g.c. através da série de Taylor em torno de zero e invertendo-a em seguida.

Sejam $Y_1, \ldots, Y_n$ variáveis aleatórias i.i.d. de realizações de Y com função densidade $f(y)$ se Y for contínua ou função de probabilidade se Y for discreta. Define-se a família exponencial conjugada uniparamétrica, em que λ é o parâmetro canônico, por

$$f(y;\lambda) = \exp\{\lambda y - K(\lambda)\} f(y), \tag{10}$$

em que $K(t)$ é a f.g.c. da variável aleatória Y. A família exponencial conjugada (10) reproduz exatamente a função densidade $f(y)$ postulada para os dados quando $\lambda = 0$. O divisor necessário para normalizar a expressão $\exp(\lambda y) f(y)$ é igual à f.g.m. dada na expressão (4). Pode-se, facilmente, verificar que a f.g.c. $K(t;\lambda)$ da família exponencial (10) é expressa em termos daquela $K(t)$ de Y por $K(t;\lambda) = K(t+\lambda) - K(\lambda)$.

Usaremos a distribuição gama denotada por $G(\alpha, \beta)$ para exemplificar o uso da expressão (10) cuja função densidade conjugada é expressa por

$$f(y;\alpha,\beta,\lambda) = \frac{(\beta-\lambda)^\alpha}{\Gamma(\alpha)} \exp\{-(\beta-\lambda)y\} y^{\alpha-1},$$

em que a f.g.c. da função densidade conjugada é $K(t) = \alpha \log\left(\frac{\beta-\lambda}{\beta-\lambda-t}\right)$.

Um outro exemplo é ilustrado a partir da distribuição de Weibull com parâmetros $\alpha > 0$ e $\beta > 0$, sendo sua função densidade conjugada dada por

$$f(y;\alpha,\beta,\lambda) = \alpha^{\frac{\lambda}{\beta}+1} \beta y^{\beta-1} e^{-y(\alpha y^{\beta-1} - \lambda)} \left\{\Gamma\left(1 + \frac{\lambda}{\beta}\right)\right\}^{-1}.$$

Definem-se $S = \sum_{i=1}^n Y_i$ e $S^* = (S - n\mu)/n^{1/2}\sigma$, em que μ e σ^2 são a média e variância de Y_i. Sejam $f_S(s;\lambda)$ e $K_S(t;\lambda)$ as funções densidade e geratriz de cumulantes de S relativas à família (10). Dessa forma, tem-se $K_S(t;\lambda) = nK(t + \lambda) - nK(\lambda)$ e, por inversão, vem

$$f_S(s;\lambda) = \exp\{s\lambda - nK(\lambda)\} f_S(s) \tag{11}$$

sendo $f_S(s) = f_S(s;0)$.

As funções densidades de S e S^* correspondentes à família (11) estão relacionadas por

$$f_S(s;\lambda) = f_{S^*}(y;\lambda) \frac{1}{\sqrt{nK''(\lambda)}}, \tag{12}$$

em que $y = \{s - nK'(\lambda)\}/\sqrt{nK''(\lambda)}$. Aproxima-se $f_{S^*}(y;\lambda)$ pela expansão de Edgeworth escolhendo convenientemente $y = 0$ para anular o termo de ordem $O(n^{-\frac{1}{2}})$. Esta escolha equivale a considerar a distribuição em (10) definida por $\hat{\lambda}$ satisfazendo a equação $K'(\hat{\lambda}) = s/n$. Pode-se interpretar $\hat{\lambda}$ como a estimativa de máxima verossimilhança de λ baseada numa única observação s de (11).

A expansão de Edgeworth da soma padronizada é dada por

$$f_{S^*}(y) = \phi(y)\left\{1 + \frac{\rho_3}{6\sqrt{n}} H_3(y) + \frac{\rho_4}{24n} H_4(y) + \frac{\rho_3^2}{72n} H_6(y)\right\} + O(n^{-3/2}), \tag{13}$$

sendo $\phi(y)$ a f.d.p. da distribuição normal $N(0,1)$, $f_{S^*}(y)$ a função densidade da soma padronizada S^*, $H_r(y)$ o polinômio de Hermite de ordem r, $\rho_j = \rho_j(\lambda) = K^{(j)}(\lambda)/K''(\lambda)^{j/2}$ para $j = 3$ e 4, $K^{(j)}(\lambda) = d^jK(\lambda)/d\lambda^j$ e $\rho_3(\lambda)$ e $\rho_4(\lambda)$ sendo os cumulantes padronizados que medem a assimetria e a curtose da distribuição de Y. Aqui, $H_3(y) = y^3 - 3y$, $H_4(y) = y^4 - 6y^2 + 3$ e $H_6(y) = y^6 - 15y^4 + 45y^2 - 15$.

Logo, $f_S(s;\lambda) = f_{S^*}(0;\lambda)\frac{1}{\sqrt{nK''(\lambda)}}$. Agora, $f_{S^*}(0;\lambda)$ segue de (13), observando que os cumulantes referentes a (10) são iguais a n vezes as derivadas de $K(\lambda)$

$$f_{S^*}(0;\hat{\lambda}) = \frac{1}{\sqrt{2\pi}}\{1 + M(\hat{\lambda}) + O(n^{-2})\}, \tag{14}$$

em que $M(\lambda)$ é um termo de ordem $O(n^{-1})$ dado por

$$M(\lambda) = \frac{3\rho_4(\lambda) - 5\rho_3(\lambda)^2}{24n}.$$

Fazendo $\lambda = \hat{\lambda}$ em (11), explicitando $f_S(s)$ e usando (12) e (14), tem-se

$$f_S(s) = \frac{\exp\{nK(\hat{\lambda}) - s\hat{\lambda}\}}{\sqrt{2n\pi K''(\hat{\lambda})}}\{1 + M(\hat{\lambda}) + O(n^{-2})\}. \tag{15}$$

A fórmula (15) para aproximar a função densidade de S é denominada *expansão ponto de sela da soma estocástica* e produz aproximações precisas para funções densidades baseadas nas suas funções geratrizes de cumulantes. O termo principal da equação (15) é chamado *expansão ponto de sela* para a função densidade da soma estocástica S proveniente de Y. Observe-se que o termo principal (15) só depende da f.g.c. de $K(\lambda)$. Esta fórmula é bem diferente da expansão de Edgeworth (13). Primeiro, para usar (15) é necessário calcular, além de $\hat{\lambda}$, a f.g.c. de $K(\lambda)$ e não somente os seus 4 primeiros cumulantes. Entretanto, nas aplicações isso não apresenta grandes dificuldades.

O termo principal em (15) não é a função densidade da distribuição normal $N(0,1)$ e, embora seja sempre positivo, nem sempre integra um. Assim, uma desvantagem de (15) é que nem sempre é fácil integrar o seu lado direito para obter uma aproximação para a função de distribuição de S. Entretanto, este termo pode ser normalizado. A expansão (15) é expressa em potências de ordem $O(n^{-1})$, enquanto a expansão de Edgeworth é dada em potências de ordem $O(n^{-1/2})$.

A expansão para a função densidade da média amostral $\bar{Y} = S/n$ segue diretamente de (15) como

$$f_{\bar{Y}}(y) = \left\{\frac{n}{2\pi K''(\hat{\lambda})}\right\}^{\frac{1}{2}} \exp[n\{K(\hat{\lambda}) - \hat{\lambda}y\}]\{1 + M(\hat{\lambda}) + O(n^{-2})\}. \tag{16}$$

O termo principal em (16) é denominado *expansão ponto de sela para a função densidade (ou de probabilidade) da média amostral.*

Segundo Hinkley *et al.* (1990) e Cordeiro (1999), o interesse maior na inferência está na obtenção de aproximações precisas para probabilidades do tipo $P(S \geq s)$ (ou $P(\bar{Y} \geq y)$) de uma amostra i.i.d. de n observações. Uma maneira de aproximar $P(S \geq s)$ é integrar numericamente a expansão ponto de sela representada pelo termo principal em (15), ou seja, calcular

$$P(S \leq s) = \int_{-\infty}^{s} \frac{\exp\{nK(\hat{\lambda}) - y\hat{\lambda}\}}{\sqrt{2n\pi K''(\hat{\lambda})}} dy.$$

A expansão de Edgeworth correspondente à função de distribuição de S^* é obtida de (13) por integração, sendo expressa por

$$F_{S^*}(y) = \Phi(y) - \phi(y) \left\{ \frac{\rho_3}{6\sqrt{n}} H_2(y) + \frac{\rho_4}{24n} H_3(y) + \frac{\rho_3^2}{72n} H_5(y) \right\} + O(n^{-3/2}).$$

Aqui, $\Phi(\cdot)$ é a função de distribuição acumulada (f.d.a.) da distribuição normal $N(0,1)$ e os respectivos polinômios de Hermite de ordem r são: $H_2(y) = y^2 - 1$, $H_3(y) = y^3 - 3y$ e $H_5(y) = y^5 - 10y^3 + 15y$.

3.3 Expansão ponto de sela através de Lugannani-Rice

Uma forma alternativa simples de obter $P(S \leq s)$ até ordem $O(n^{-1})$, válida sobre todo o intervalo de variação de s, é devida a Lugannani e Rice (1980).

Integra-se a equação (15), em que $nK'(\hat{\lambda}) = s$, e, obtém-se, invertendo-a $\hat{\lambda} = \hat{\lambda}(x)$. Desde que $dx/d\hat{\lambda} = nK''(\hat{\lambda})$, vem

$$P(S \leq s) = \int_{-\infty}^{\hat{\lambda}(s)} \sqrt{\frac{nK''(\hat{\lambda})}{2\pi}} \exp[n\{K(\hat{\lambda}) - \hat{\lambda}K'(\hat{\lambda})\}] d\hat{\lambda}.$$

Considere ainda a seguinte mudança de variável

$$q = q(\hat{\lambda}) = \mathsf{sinal}(\hat{\lambda}) \sqrt{2\{\hat{\lambda}K'(\hat{\lambda}) - K(\hat{\lambda})\}}.$$

A função $q(\hat{\lambda})$ é estritamente crescente e contínua. Com a mudança de variável $q = q(\hat{\lambda})$, a integral $P(S \leq s)$ é reescrita como

$$P(S \leq s) = \int_{-\infty}^{q_s} \sqrt{\frac{n}{2\pi K''(\hat{\lambda})}} \frac{q}{\hat{\lambda}} \exp(-nq^2/2) dq = \int_{-\infty}^{q_s} f(q)\phi(q; n^{-1}) dq,$$

em que $q_s = q(\hat{\lambda}(s))$, enquanto $\phi(u; n^{-1})$ indica a função densidade da distribuição normal $N(0, n^{-1})$. Além disso,

$$f(q) = \frac{q}{\hat{\lambda}\sqrt{K''(\hat{\lambda})}},$$

com $\hat{\lambda}$ expresso como uma função de q. Dessa forma, a integral $P(S \leq s)$ pode ser reescrita na forma que possibilita a aplicação da expansão de Laplace. Entretanto, esse método oferece três expressões distintas para a aproximar $P(S \leq s)$ dependendo de q_s ser negativo, igual a zero, ou positivo. Tem-se,

$$\begin{aligned} P(S \leq s) &= \int_{-\infty}^{q_s} \{1 + f(q) - 1\}\phi(q; n^{-1})dq \\ &= \Phi(\sqrt{n}q_s) + \int_{-\infty}^{q_s} \frac{f(q) - 1}{q} q\phi(q; n^{-1})dq. \end{aligned}$$

Integrando por partes o último termo da soma, tomando $q\phi(q; n^{-1})$ como o fator diferencial, sendo ainda

$$\lim_{q \to -\infty} f(q)\phi(q; n^{-1})/q = 0,$$

tem-se

$$\begin{aligned} P(S \leq s) = \Phi(\sqrt{n}q_s) &- \frac{1}{n}\phi(q_s; n^{-1})\frac{f(q_s) - 1}{q_s} \\ &+ \frac{1}{n}\int_{-\infty}^{q_s} \frac{d}{dq}\left(\frac{f(q) - 1}{q}\right)\phi(q; n^{-1})dq, \end{aligned}$$

em que a última integral é ainda da forma $\int_{-\infty}^{q_s} f(q)\phi(q; n^{-1})dq$, com uma especificação diferente da função $f(\cdot)$. Essa integral é multiplicada por n^{-1}, de forma que sua contribuição é da ordem $O(n^{-1})$. Assim, tem-se a expansão assintótica

$$P(S \leq s) = \left\{\Phi(\sqrt{n}q_s) + \phi(q_s; n^{-1})\frac{1 - f(q_s)}{nq_s}\right\}\{1 + O(n^{-1})\}.$$

Escrevendo $\hat{r} = \sqrt{n}q_s$, obtém-se a aproximação conhecida como a expansão de Lugannani-Rice expressa como

$$P(S \leq s) = \Phi(\hat{r}) + \left(\frac{1}{\hat{r}} - \frac{1}{\hat{\nu}}\right)\phi(\hat{r}) + O(n^{-1}), \tag{17}$$

em que $\hat{r} = \mathsf{sinal}(\hat{\lambda})[2n\{\hat{\lambda}K'(\hat{\lambda}) - K(\hat{\lambda})\}]^{1/2}$ e $\hat{\nu} = \hat{\lambda}\{nK''(\hat{\lambda})\}^{1/2}$.

As quantidades $\hat{r}$ e $\hat{\nu}$ podem ser interpretadas como a razão sinalizada de verossimilhanças e a estatística escore, respectivamente, para testar $\lambda = 0$ no modelo exponencial (11) definido para S.

A aproximação (17) é boa em quase todo intervalo de variação de s, exceto próximo ao ponto $s = E(S)$ ou $r = 0$, em que deve ser substituída pelo seu limite, quando $r \to 0$, dado por

$$P(S \leq s) = \frac{1}{2} + \frac{\hat{\rho}_3}{6\sqrt{2\pi n}} + O(n^{-1}).$$

3.4 Expansão ponto de sela através da expansão de Daniels

Daniels (1954) afirma que é, frequentemente, necessário aproximar as distribuições de algumas estatísticas cujas distribuições exatas não podem ser obtidas convenientemente. Segundo Reid (1988), a aproximação obtida por Daniels em 1954 é mais precisa que a aproximação de Edgeworth especialmente para n pequeno. Com o propósito de obter aproximações precisas para as distribuições de probabilidade da média amostral $\bar{Y}$ de n observações de uma distribuição com média $E(Y)$, Daniels (1987) propõe uma aproximação cujo erro relativo é controlado sobre todo o campo de variação de $\bar{Y}$.

Quando $\hat{\lambda} > 0$, a sua expansão de $P(S \geq s)$ até termos de ordem $O(n^{-1})$ é dada por

$$\begin{aligned} P(S \geq s) = \exp(n\hat{K} - s\hat{\lambda} + \hat{\nu}^2/2) \Bigg[\{1 - \Phi(\hat{\nu})\} \left\{ 1 - \frac{\hat{\rho}_3 \hat{\nu}^3}{6\sqrt{n}} + \frac{1}{n} \left(\frac{\hat{\rho}_4 \hat{\nu}^4}{24} + \frac{\hat{\rho}_3^2 \hat{\nu}^6}{72} \right) \right\} \\ + \phi(\hat{\nu}) \left\{ \frac{\hat{\rho}_3(\hat{\nu}^2 - 1)}{6\sqrt{n}} - \frac{1}{n} \left(\frac{\hat{\rho}_4(\hat{\nu}^3 - \hat{\nu})}{24} + \frac{\hat{\rho}_3^2(\hat{\nu}^5 - \hat{\nu}^3 + 3\hat{\nu})}{72} \right) \right\} \Bigg], \quad (18) \end{aligned}$$

em que $\hat{\rho}_3 = \rho_3(\hat{\lambda})$, $\hat{\rho}_4 = \rho_4(\hat{\lambda})$, $\hat{K} = K(\hat{\lambda})$ e $\hat{\nu} = \hat{\lambda}\{nK''(\hat{\lambda})\}^{1/2}$. A aproximação de (18) com apenas os termos de ordem $O(\sqrt{n})$ fornece, em geral, bons resultados.

No caso em que $\hat{\lambda} < 0$, pode-se obter $P(S \geq s)$ até ordem $O(n^{-1/2})$ como

$$\begin{aligned} P(S \geq s) = H(-\hat{\nu}) + \exp(n\hat{K} - s\hat{\lambda} + \hat{\nu}^2/2) \times \\ \left[\{H(\hat{\nu}) - \Phi(\hat{\nu})\} \left(1 - \frac{\hat{\rho}_3 \hat{\nu}^3}{6\sqrt{n}} \right) + \phi(\hat{\nu}) \frac{\hat{\rho}_3(\hat{\nu}^2 - 1)}{6\sqrt{n}} \right], \quad (19) \end{aligned}$$

em que $H(w) = 0,\ 1/2$ e 1 quando $w < 0$, $w = 0$, $w > 0$, respectivamente.

4 Melhoramento de estatísticas no caso uniparamétrico

Segundo Cordeiro (1999), um problema natural que surge ao se usar os testes em grandes amostras é o de verificar se a aproximação de primeira ordem é adequada para a distribuição nula da estatística de teste em consideração. Para o caso em que se trata de pequenas amostras, a aproximação de primeira ordem pode não ser satisfatória, conduzindo a taxas de rejeição bastante distorcidas. Na existência dessas situações, Bartlett (1937) propôs uma primeira idéia para melhorar os testes estatísticos. Ele considerou apenas a razão de verossimilhanças, computando o seu valor esperado segundo H até ordem $O(n^{-1})$, sendo n o tamanho da amostra.

O desenvolvimento desta seção é baseado nas Seções 5.7 a 5.9 do livro de Cordeiro (1999), que apresenta um estudo mais amplo sobre o melhoramento dos testes da razão de verossimilhanças, escore e Wald. As expressões gerais dos coeficientes α_1, α_2 e α_3 descritos ao longo desta seção, são definidas nas seções acima do livro (num contexto multiparamétrico) em termos de cumulantes conjuntos de derivadas da log-verossimilhança. Apresentaremos aqui três propostas para corrigir as estatísticas da razão de verossimilhanças, escore e de Wald, seguindo os artigos de Cordeiro e Ferrari (1991), Kakizawa (1996) e Cordeiro et al. (1998), relativas apenas ao caso uniparamétrico. A originalidade aqui se restringe ao estudo de simulação para testar os graus de liberdade (g.l.) θ da distribuição t de Student para $\theta = 2$ e 6.

Considera-se um conjunto de n observações independentes e identicamente distribuídas $y_1, \ldots, y_n$, que seguem qualquer distribuição regular uniparamétrica indexada por um parâmetro escalar desconhecido θ supondo uma observação y. Seja $L(\theta; y)$ a verossimilhança total de um problema suposto regular e, seja ainda, $\ell(\theta) = \log L(\theta; y)$ a log-verossimilhança. Assume-se que $\ell(\theta)$ satisfaz as condições de regularidade padrão e ainda $U_\theta = d\ell(\theta)/d\theta, U_{\theta\theta} = d^2\ell(\theta)/d\theta^2$, etc. No que se segue, usa-se a notação padrão para os cumulantes conjuntos de derivadas da log-verossimilhança:

$$\kappa_{\theta\theta} = E(U_{\theta\theta}),\ \kappa_{\theta\theta\theta} = E(U_{\theta\theta\theta}),\ \kappa_{\theta,\theta} = E(U_\theta^2) = -\kappa_{\theta\theta},$$
$$\kappa_{\theta,\theta\theta} = E(U_\theta U_{\theta\theta}),\ \kappa_{\theta\theta,\theta\theta} = E(U_{\theta\theta}^2) - \kappa_{\theta\theta}^2,\ \kappa_{\theta\theta\theta\theta} = E(U_{\theta\theta\theta\theta}),$$
$$\kappa_{\theta,\theta,\theta\theta} = E(U_\theta^2 U_{\theta\theta}) - \kappa_{\theta,\theta}\,\kappa_{\theta\theta},\ \kappa_{\theta,\theta,\theta,\theta} = E(U_\theta^4) - 3\kappa_{\theta,\theta}^2 \quad \text{e} \quad \kappa_{\theta,\theta\theta\theta} = E(U_\theta U_{\theta\theta\theta}).$$

Denota-se, também, as derivadas dos cumulantes com sobrescritos como

$$\kappa_{\theta\theta}^{(\theta)} = d\kappa_{\theta\theta}/d\theta, \quad \kappa_{\theta\theta}^{(\theta\theta)} = d^2\kappa_{\theta\theta}/d\theta^2, \text{ etc.}$$

Outras identidades de Bartlett usuais são definidas como:

$$\kappa_{\theta,\theta} = -\kappa_{\theta\theta},\ \kappa_{\theta,\theta,\theta} = 2\kappa_{\theta\theta\theta} - 3\kappa_{\theta\theta}^{(\theta)},\ \kappa_{\theta,\theta\theta} = \kappa_{\theta\theta}^{(\theta)} - \kappa_{\theta\theta\theta},$$
$$\kappa_{\theta,\theta,\theta,\theta} = -3\kappa_{\theta\theta\theta\theta} + 8\kappa_{\theta\theta\theta}^{(\theta)} - 6\kappa_{\theta\theta}^{(\theta\theta)} + 3\kappa_{\theta\theta,\theta\theta},$$
$$\kappa_{\theta,\theta,\theta\theta} = \kappa_{\theta\theta\theta\theta} - 2\kappa_{\theta\theta\theta}^{(\theta)} + \kappa_{\theta\theta}^{(\theta\theta)} - \kappa_{\theta\theta,\theta\theta},\ \kappa_{\theta,\theta\theta\theta} = \kappa_{\theta\theta\theta}^{(\theta)} - \kappa_{\theta\theta\theta\theta}.$$

A grande vantagem das identidades de Bartlett é facilitar a obtenção dos cumulantes κ's, pois determinada parametrização pode conduzir a um cálculo direto simples de alguns cumulantes, sendo os demais calculados de forma indireta através destas identidades (Cordeiro, 1999). Esses cumulantes têm grande aplicabilidade no cálculo de correções de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças e escore. Os detalhes estão descritos em Cribari-Neto e Cordeiro (1996).

4.1 Três estatísticas corrigidas

Os testes em grandes amostras, cujas distribuições de referência são qui-quadrado, mais conhecidos são: razão de verossimilhanças (w), escore (S) e Wald (W). Apresenta-se, aqui, três estatísticas corrigidas correspondentes aos testes em grandes amostras. Porém, na Seção 5, é apresentada a estatística proposta por Cordeiro (1998) cuja aplicabilidade será ilustrada para obtenção das probabilidades referentes às distribuições gama $G(\mu, \phi)$ e t-Student com ν graus de liberdade.

Para testar $H : \theta = \theta^{(0)}$ em qualquer distribuição regular uniparamétrica, em que $\theta^{(0)}$ é um valor especificado para θ, assume-se que $T \to \chi_1^2$ sob a hipótese nula H. Denote a log-verossimilhança total por $\ell_T(\theta)$ e a função escore total por $U_T(\theta) = d\ell_T(\theta)/d\theta$. Então, T pode tomar a forma de qualquer estatística a seguir

$$w = 2[\ell_T(\widehat{\theta}) - \ell_T(\theta^{(0)})], \quad S = U_T(\theta^{(0)})^2/(n\,\widetilde{\kappa}_{\theta,\theta}),$$

$$W = n(\widehat{\theta} - \theta^{(0)})^2\ \widehat{\kappa}_{\theta,\theta}, \quad MW = n(\widehat{\theta} - \theta^{(0)})^2\ \widetilde{\kappa}_{\theta,\theta},$$

em que $\widehat{\kappa}_{\theta,\theta}$ e $\widetilde{\kappa}_{\theta,\theta}$ são a informação para uma observação avaliada em $\widehat{\theta}$ e $\theta^{(0)}$, respectivamente.

Tem-se que

$$P(T \leq x) = P(\chi_1^2 \leq x) + O(n^{-1}).$$

e

$$P(T^* \leq x) = P(\chi_1^2 \leq x) + O(n^{-3/2}).$$

A seguir, apresentam-se três propostas para corrigir a estatística T:

- Cordeiro e Ferrari (1991):

$$T^* = T\left[1 - \frac{1}{n}\left(\alpha_1 + \alpha_2 T + \alpha_3 T^2\right)\right].$$

- Kakizawa (1996):

$$K(T) = T^* + \frac{1}{4}[\alpha_3^2 T + 2\alpha_2\alpha_3 T^2 + 2\alpha_1\alpha_3 + \frac{4}{3}\alpha_2^2 T^3 + 3\alpha_1\alpha_2 T^4 + \frac{9}{5}\alpha_1^2 T^5].$$

- Cordeiro, Ferrrari e Cysneiros (1998):

$$\widetilde{T} = \sqrt{\frac{\pi}{3\alpha_1}}\exp\left(\frac{\alpha_2^2}{3\alpha_1} - \alpha_3\right) \times \left\{\Phi\left(\sqrt{6\alpha_1}\,T + \sqrt{\frac{2}{3\alpha_1}}\alpha_2\right) - \Phi\left(\sqrt{\frac{2}{3\alpha_1}}\alpha_2\right)\right\},$$

se $\alpha_1 > 0$ (α_1 é sempre não negativo).

Quando $\alpha_1 = 0$ (e $\alpha_2 \neq 0$)

$$\widetilde{T} = \frac{1}{2\alpha_2}\exp(-\alpha_3)\{1 - \exp(-2\alpha_2 T)\}.$$

As expressões dos coeficientes α_1, α_2 e α_3 acima estão descritas para cada estatística nos papers correspondentes. Vide, também, as Seções 5.7 a 5.9 do livro de Cordeiro (1999). Diferentes α_i's são apresentados correspondendo às estatísticas de razão de verossimilhanças w, estatística escore S, estatística de Wald W e estatística de Wald modificada MW. Tem-se os seguintes α_i's para a estatística da razão de verossimilhanças w:

$$\alpha_1 = \frac{5\kappa_{\theta\theta\theta}^2 + 24\kappa_{\theta\theta}^{(\theta)}(\kappa_{\theta\theta}^{(\theta)} - \kappa_{\theta\theta\theta})}{12\,\kappa_{\theta\theta}^3} - \frac{\kappa_{\theta\theta\theta\theta} + 4(\kappa_{\theta\theta}^{(\theta\theta)} - \kappa_{\theta\theta\theta}^{(\theta)})}{4\kappa_{\theta\theta}^2},$$

$$\alpha_2 = \alpha_3 = 0.$$

Para a estatística escore S, tem-se que

$$\alpha_1 = \frac{-\kappa_{\theta,\theta,\theta}^2}{36\,\kappa_{\theta\theta}^3},$$

$$\alpha_2 = \frac{10\kappa_{\theta,\theta,\theta}^2 + 3\,\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta,\theta} - 9\kappa_{\theta\theta}^3}{36\,\kappa_{\theta\theta}^3},$$

$$\alpha_3 = \frac{-5\kappa_{\theta,\theta,\theta}^2 - 3\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta,\theta} + 9\kappa_{\theta\theta}^3}{12\,\kappa_{\theta\theta}^3}.$$

Para a estatística de Wald W, tem-se que

$$\alpha_1 = \frac{-44\kappa_{\theta\theta\theta}^2 + 120\kappa_{\theta\theta\theta}\kappa_{\theta\theta}^{(\theta)} - 81(\kappa_{\theta\theta}^{(\theta)})^2 + 12\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta\theta}}{12\,\kappa_{\theta\theta}^3} - \frac{3\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta,\theta}}{12\,\kappa_{\theta\theta}^3}$$

$$\alpha_2 = \frac{-10\kappa_{\theta\theta\theta}{}^2 + 48(2\kappa_{\theta\theta\theta} - 3\kappa_{\theta\theta}^{(\theta)})^2}{72\kappa_{\theta\theta}^3} + \frac{6(\kappa_{\theta\theta}^{(\theta)} - \kappa_{\theta\theta\theta})(17\kappa_{\theta\theta\theta} - 45\kappa_{\theta\theta}^{(\theta)})}{72\kappa_{\theta\theta}^3} + \frac{3\kappa_{\theta\theta,\theta\theta} + 20\kappa_{\theta\theta\theta}^{(\theta)} - 11\kappa_{\theta\theta\theta\theta} - 12\kappa_{\theta\theta}^{(\theta\theta)}}{12\kappa_{\theta\theta}^2},$$

$$\alpha_3 = -\frac{\kappa_{\theta\theta\theta}^2}{36\,\kappa_{\theta\theta}^3}$$

E para a estatística de Wald modificada MW, tem-se que

$$\alpha_1 = \frac{-44\kappa_{\theta\theta\theta}^2 + 120\kappa_{\theta\theta\theta}\kappa_{\theta\theta}^{(\theta)} - 81(\kappa_{\theta\theta}^{(\theta)})^2 + 12\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta\theta}}{12\,\kappa_{\theta\theta}^3} - \frac{3\kappa_{\theta\theta}\kappa_{\theta,\theta,\theta,\theta}}{12\,\kappa_{\theta\theta}^3}$$

$$\alpha_2 = \frac{63\kappa_{\theta\theta\theta}\kappa_{\theta\theta}^{(\theta)} - 22\kappa_{\theta\theta\theta}^2 - 45(\kappa_{\theta\theta}^{(\theta)})^2}{18\,\kappa_{\theta\theta}^3} + \frac{4\kappa_{\theta\theta\theta\theta} - 4\kappa_{\theta\theta\theta}^{(\theta)} - 4\kappa_{\theta,\theta,\theta\theta} - 3\kappa_{\theta,\theta,\theta,\theta}}{12\,\kappa_{\theta\theta}^2}$$

$$\alpha_3 = -\frac{(3\kappa_{\theta\theta}^{(\theta)} - \kappa_{\theta\theta\theta})^2}{36\,\kappa_{\theta\theta}^3}$$

Tome-se a distribuição Poisson como forma de ilustração do uso dessas estatísticas (1, 2 e 3 são índices que se referem a α_1, α_2 e α_3). Desse modo, observa-se que

$$\begin{aligned}
&LR1 = -\frac{1}{6\theta}\,,\\
&S1 = \frac{1}{36\theta}\,,\quad S2 = -\frac{7}{36\theta}\,,\quad S3 = \frac{1}{6\theta},\\
&W1 = \frac{1}{6\theta}\,,\quad W2 = \frac{1}{18\theta}\,,\quad W3 = \frac{1}{9\theta},\\
&MW2 = -\frac{7}{36\theta}\,,\quad MW3 = \frac{1}{36\theta}\,.
\end{aligned}$$

Ao exemplificar para a distribuição em série logaritmica, tem-se que

$$\begin{aligned}
LR1 = {} & \frac{1}{12\theta\{\theta + \log(1-\theta)\}^3}\left\{2\theta^4 + 6\theta^3\log(1-\theta) + 8\theta^2(\theta+1)\log^2(1-\theta)\right.\\
& \left. -3\theta\left(\theta^2 - 2\theta - 2\right)\log^3(1-\theta) + 2\left(\theta^2 - \theta + 1\right)\log^4(1-\theta)\right\},
\end{aligned}$$

$$\begin{aligned}
S1 = {} & -\frac{\{2\theta^2 + 3\log(1-\theta)\theta + (\theta+1)\log^2(1-\theta)\}^2}{36\theta\{\theta + \log(1-\theta)\}^3},\\
S2 = {} & \frac{1}{36\theta\{\theta + \log(1-\theta)\}^3}\left\{22\theta^4 + 66\theta^3\log(1-\theta) + \theta^2(28\theta+73)\log^2(1-\theta)\right.\\
& \left. -3\theta\left(\theta^2 - 12\theta - 12\right)\log^3(1-\theta) + \left(7\theta^2 + 8\theta + 7\right)\log^4(1-\theta)\right\},\\
S3 = {} & \frac{1}{36\theta\{\theta + \log(1-\theta)\}^3}\left\{22\theta^4 + 66\theta^3\log(1-\theta) + \theta^2(28\theta+73)\log^2(1-\theta)\right.\\
& \left. -3\theta\left(\theta^2 - 12\theta - 12\right)\log^3(1-\theta) + \left(7\theta^2 + 8\theta + 7\right)\log^4(1-\theta)\right\},
\end{aligned}$$

$$\begin{aligned}
W1 = {} & -\frac{1}{12\theta\{\theta + \log(1-\theta)\}^3}\left\{2\theta^4 + 6\theta^3\log(1-\theta)\right.\\
& +8\theta^2(\theta+1)\log^2(1-\theta) - 3\theta\left(\theta^2 - 2\theta - 2\right)\log^3(1-\theta)\\
& \left. +2\left(\theta^2 - \theta + 1\right)\log^4(1-\theta)\right\},\\
W2 = {} & \frac{1}{36\theta\{\theta + \log(1-\theta)\}^3}\left\{22\theta^4 + 6\theta^3(12\theta - 1)\log(1-\theta)\right.\\
& +4\theta^2\left(6\theta^2 + 43\theta - 20\right)\log^2(1-\theta) + \left(70\theta^2 - 46\theta\right)\log^4(1-\theta)\\
& \left. -2\log^4(1-\theta) + 3\theta\left(25\theta^2 + 18\theta - 18\right)\log^3(1-\theta)\right\},\\
W3 = {} & -\frac{\left\{3\theta^2\log(1-\theta) + 2\theta^2 + (4\theta - 2)\log^2(1-\theta)\right\}^2}{36\theta\{\theta + \log(1-\theta)\}^3},
\end{aligned}$$

$$
\begin{aligned}
MW2 =& \frac{1}{36\theta\{\theta+\log(1-\theta)\}^3}\left\{22\theta^4+6\theta^3(6\theta+5)\log(1-\theta)\right.\\
&+3\theta^2(25\theta-6)\log^3(1-\theta)\\
&+\theta^2\left(24\theta^2+46\theta+1\right)\log^2(1-\theta)\\
&\left.+\left(43\theta^2-28\theta+7\right)\log^4(1-\theta)\right\},\\
MW3 =& -\frac{\{4\theta^2+3(\theta+1)\log(1-\theta)\theta+(5\theta-1)\log^2(1-\theta)\}^2}{36\theta\{\theta+\log(1-\theta)\}^3}.
\end{aligned}
$$

4.2 Resultados de simulação

Apresenta-se aqui alguns resultados de simulação com o intuito de avaliar a eficácia das correções de Bartlett e tipo-Bartlett para as estatísticas da razão de verossimilhanças, escore, Wald e Wald modificada no teste de hipótese sobre os graus de liberdade da distribuição t-Student. Comparam-se os desempenhos dos testes baseados nestas estatísticas com suas respectivas versões corrigidas em termos de tamanho. Estudos de poder serão realizados em pesquisa futura.

Por conseguinte, há o interesse em testar o número de graus de liberdade θ em um modelo t_θ, em que $\theta = 2$ e 6. Os tamanhos amostrais considerados foram $n = 10, \ldots, 60$. Em cada simulação, computou-se o E.M.V. de θ e as quatro estatísticas não corrigidas juntamente com as versões corrigidas. Todas as entradas das tabelas correspondem a porcentagens.

A taxa de rejeição sob a hipótese nula, i.e., a percentagem de vezes que essas estatísticas excederam os limites superiores apropriados ($\alpha = 0,05$ e $\alpha = 0,10$) da distribuição χ^2_1 de referência são dados nas tabelas a seguir para $\theta = 2$ e 6, respectivamente.

Todas simulações foram obtidas usando a linguagem de programação R com uma maximização não-linear da log-verossimilhança com respeito a θ e todos resultados são baseados em 10.000 replicações. Os valores nessas tabelas mostram informação importante.

Em princípio, observa-se melhores resultados para as versões corrigidas das estatísticas tanto para $\alpha = 0,05$ quanto para $\alpha = 0,10$. Nota-se, ainda, que as taxas de rejeição se aproximam do valor nominal à medida que se aumenta n e θ.

Nos resultados da Tabela 1, para a distribuição t-Student com dois g.l. (t_2) e $\alpha = 0,05$, tem-se que as estatísticas corrigidas apresentam melhores resultados sendo a estatística S^* a que apresenta resultados próximos ao valor nominal para $n = 60$. Para $\alpha = 0,10$ correspondente à Tabela 2, observa-se um melhor desempenho da estatística corrigida S^* e, ainda, que esta apresenta resultados próximos ao valor nominal para $n > 30$. Ao analisar a Tabela 3 referente à distribuição t-Student com 6 g.l. (t_6) para $\alpha = 0,05$, tem-se que a estatística corrigida com melhor desempenho é S^* quando se tem valores de $n \geq 30$. Para a Tabela 4 em que $\alpha = 0,10$, observa-se um melhor desempenho da estatística corrigida S^* com relação às demais estatísticas a partir de $n \geq 20$ e, ainda, que seus resultados se aproximam do valor nominal para $n > 30$.

Tabela 1 - Taxa de Rejeição (%) obtida através das estatísticas não-corrigidas e corrigidas para a distribuição t-Student t_2

	$\alpha = 0,05$					
n	10	20	30	40	50	60
LR	54,5	25,7	18,3	14,6	11,9	8,3
LR^*	14,5	12,6	11,1	9,3	7,7	6,2
S	31,0	22,1	17,8	15,2	12,4	9,9
S^*	11,7	10,8	9,4	7,1	6,8	5,6
W	25,1	18,4	17,3	13,5	12,0	8,9
W^*	10,3	9,5	9,0	8,2	6,5	6,3
MW	27,2	23,9	19,2	16,7	14,1	9,4
MW^*	11,9	10,4	9,2	8,3	6,9	5,8

Tabela 2 - Taxa de Rejeição (%) obtida através das estatísticas não-corrigidas e corrigidas para a distribuição t-Student t_2

	$\alpha = 0,10$					
n	10	20	30	40	50	60
LR	70,3	33,6	24,7	19,4	16,2	13,6
LR^*	17,5	15,6	12,0	11,5	11,3	10,4
S	43,1	29,3	22,1	19,3	17,1	15,8
S^*	14,4	12,0	11,5	10,7	10,4	10,2
W	36,9	25,3	23,7	18,1	15,4	14,2
W^*	15,7	13,8	12,6	11,9	11,3	10,5
MW	41,3	29,0	27,4	22,3	18,0	15,6
MW^*	17,6	14,7	13,4	12,0	10,7	10,3

Tabela 3 - Taxa de Rejeição (%) obtida através das estatísticas não-corrigidas e corrigidas para a distribuição t-Student t_6

	$\alpha = 0,05$					
n	10	20	30	40	50	60
LR	49,1	23,6	16,8	13,7	9,2	7,6
LR^*	11,5	9,7	8,4	7,2	6,4	5,3
S	26,2	18,3	15,1	13,5	10,1	9,4
S^*	9,5	9,1	8,2	6,5	5,4	5,1
W	23,4	17,5	16,0	12,4	11,8	8,3
W^*	9,5	8,9	8,7	7,8	6,1	5,8
MW	25,7	20,4	17,9	15,4	13,6	8,7
MW^*	11,2	9,9	8,5	7,7	6,4	5,6

Tabela 4 - Taxa de Rejeição (%) obtida através das estatísticas não-corrigidas e corrigidas para a distribuição t-Student t_6

	$\alpha = 0,10$					
n	10	20	30	40	50	60
LR	63,1	32,3	22,1	17,3	14,1	13,3
LR^*	14,4	13,2	11,7	11,1	10,8	10,3
S	39,5	27,7	20,8	17,9	16,3	15,1
S^*	13,2	11,8	11,1	10,5	10,2	10,1
W	35,4	23,8	22,8	16,2	14,9	13,3
W^*	12,9	12,6	12,3	11,2	10,7	10,4
MW	39,3	27,2	25,1	19,7	16,1	14,0
MW^*	13,0	12,4	11,9	10,8	10,5	10,2

Nota-se, em todos os casos apresentados nas tabelas, que a estatística com melhor desempenho é a S^* e que as estatísticas não-corrigidas apresentam piores desempenhos dentre as demais quando comparadas aos valores nominais.

5 Correção de Bartlett generalizada

Seja S^* uma estatística unidimensional com f.d.a. $F_{S^*}(y)$ e função densidade $f_{S^*}(y)$. Considera-se que S^* é obtida de uma amostra Y de n observações independentes tendo funções densidades que dependem de um vetor de parâmetros desconhecidos θ.

A idéia por trás do procedimento de modificação de S^* é baseado na suposição de que a função distribuição $F_{S^*}(y)$ possa ser formalmente expandida como

$$F_{S^*}(y) = F_Z(y) + \sum_{i=1}^{m} (-1)^i \eta_i \frac{D^i F_Z(y)}{i!}, \tag{20}$$

em que $D^i F_Z(y) = d^i F_Z(y)/dy^i$, quando $m \to \infty$ (vide equação 5.6 em McCullagh, 1987).

A f.d.a. da estatística S^*, $F_{S^*}(y)$, pode ainda ser expandida como

$$F_{S^*}(y) = F_Z(y) + A_1(y) + A_2(y) + O(n^{-3/2}), \tag{21}$$

em que $F_Z(y)$ é a f.d.a. de uma variável aleatória escalar Z de referência (não necessariamente normal) usada para aproximar a distribuição de S^*, $A_1(y)$ e $A_2(y)$ são termos de ordens $O(n^{-1/2})$ e $O(n^{-1})$, respectivamente, que dependem de algumas diferenças dos cumulantes $(\kappa_i - \kappa_{i0})$ de S^* e Z, sendo κ_i o i-ésimo cumulante de S^* e κ_{i0} o i-ésimo cumulante de Z.

A forma (21) da f.d.a. de S^* sugere o uso de uma estatística modificada definida por

$$CF = S^* - b_1(S^*) - b_2(S^*),$$

em que $b_r(S^*) = O_p(n^{-r/2})$, para $r = 1$ e 2, são termos estocásticos aditivos como funções da estatística S^*. Cordeiro e Ferrari (1998) deduziram que $b_1(y) = -A_1(y)/f_Z(y)$ e $b_2(y) = -A_2(y)/f_Z(y) + A_1(y)^2 f'_Z(y)/\{2f_Z(y)^3\}$, supondo que $f_Z(y)$ é não-nula no suporte de S^*.

Assim, a estatística modificada CF cuja função de distribuição é $F_Z(y)$ até termos de ordem $O(n^{-1})$ é expressa como

$$CF = S^* \left[1 + \frac{A_1(S^*)}{f_Z(S^*)S^*} + \frac{1}{S^*} \left\{ \frac{A_2(S^*)}{f_Z(S^*)} - \frac{A_1(S^*)^2 f'_Z(S^*)}{2f_Z(S^*)^3} \right\} \right]. \quad (22)$$

O método que obtém (22) é formalmente válido e provado somente para que a distribuição de S^* inclua o termo de ordem $O(n^{-1})$ da expansão de Edgeworth. Tem-se, $P(CF \leq x) = P(Z \leq x) + O(n^{-3/2})$. O termo multiplicativo de S^* em (22) é um tipo de ajuste estocástico envolvendo as funções $A_1(S^*)$ e $A_2(S^*)$, de ordens $O(n^{-1/2})$ e $O(n^{-1})$, deduzidas da expansão (21), da densidade $f_Z(y)$ com sua primeira derivada $f'_Z(y)$ e da estatística S^*.

O fator multiplicativo estocástico em (22) pode ser escrito como $1 + b(S^*, \eta_i, F_Z^{(i)}(y))$, em que a notação enfatiza a dependência das derivadas da função de distribuição $F_Z(y)$, dos "momentos formais" η_i's e da estatística não modificada S^*. Esses momentos formais são definidos em termos de diferenças entre os cumulantes de S^* e Z. Cordeiro e Ferrari (1998) definem este fator de ajuste como *correção de Bartlett generalizada* tal que este é um resultado geral que pode ser utilizado em muitos testes importantes na estatística e econometria.

Suponha que existem $\mu = \mu(\theta)$ e $\sigma = \sigma(\theta)$ tais que a estatística padronizada $S^* = (S - \mu)/\sigma$, em que S é uma estatística que depende dos parâmetros n e θ. A estatística S^* tem média zero e variância unitária de forma a convergir em distribuição para uma variável aleatória com distribuição normal padrão.

Combinando a equação da expansão da f.d.a. em série de Edgeworth com (21), pode-se observar imediatamente que $A_1(y) = -\rho_3\phi(y)H_2(y)/(6\sqrt{n})$ e $A_2(y) = -\phi(y)\{\rho_4 H_3(y)/24 + \rho_3^2 H_5(y)/72\}/n$. Substituindo esses resultados na equação (22), obtém-se

$$CF = S^* - \frac{\rho_3}{6\sqrt{n}}(S^{*2} - 1) + \frac{1}{12n} \left\{ \frac{\rho_3^2(4S^{*2} - 7S^*)}{3} - \frac{\rho_4(S^{*3} - 3S^*)}{2} \right\}. \quad (23)$$

A equação (23) é a transformação clássica polinomial de Cornish-Fisher para obter normalidade quando os quantis estocásticos de ordem $O_p(n^{-3/2})$ e menores são omitidos, isto é, $CF \sim N(0,1) + O_p(n^{-3/2})$.

Exemplo 4. Mediante o uso da equação (20), apresenta-se a aproximação da distribuição gama $G(\mu, \phi)$ de média μ e parâmetro de dispersão ϕ considerando que é a f.d.a. da estatística S^*, em função da distribuição exponencial Z com média α (representando a f.d.a. da variável aleatória Z). Seja $F_Z(y) = 1 - e^{-y/\alpha}$, $y > 0$, $\alpha > 0$. Inicialmente, tem-se que

$$D^i F_Z(y) = (-1)^{i-1} \frac{e^{-y/\alpha}}{\alpha^i}, \; y > 0, \; \alpha > 0. \quad (24)$$

A relação dos momentos em função dos cumulantes para $r = 1,\ 2$ e 3 é $\mu_1 = \kappa_1$, $\mu_2 = \kappa_2$, $\mu_3 = \kappa_3$ e para $r \geq 4$ pode ser expressa como

$$\mu_r = \kappa_r + \sum_{j=2}^{r-2} \binom{r-1}{j-1} \kappa_j \mu_{r-j}.$$

Assim, para $r = 4,\ 5,\ 6$ e 7 tem-se as seguintes identidades: $\mu_4 = \kappa_4 + 3\kappa_2^2$, $\mu_5 = \kappa_5 + 10\kappa_2\kappa_3$, $\mu_6 = \kappa_6 + 15\kappa_2\kappa_4 + 10\kappa_3^2 + 15\kappa_2^3$ e $\mu_7 = \kappa_7 + 10\kappa_2\kappa_5 + 80\kappa_2^2\kappa_3 + 20\kappa_3\kappa_4$. Sejam ϵ's as diferenças entre os cumulantes de S^* e Z, tais que

$$\epsilon_i = (i-1)!\alpha^i \left[\frac{1}{\phi^{i-1}} \left(\frac{\mu}{\alpha} \right)^i - 1 \right].$$

Tem-se que os $\eta_i = O(\alpha^i) \times [\epsilon_i]$ representam momentos centrais em função dos cumulantes correspondentes, que são iguais às diferenças entre os cumulantes de S^* e Z. Seja $[\epsilon]_i$ um fator que representa uma combinação dos ϵ_i's.

A aproximação da distribuição gama $G(\mu, \phi)$ em função da distribuição exponencial de média α decorre substituindo na equação (20) a equação (24) e η_i. Logo,

$$F_{S^*}(y) = F_Z(y) - e^{-y/\alpha} \sum_{i=1}^{m} \frac{[\epsilon]_i}{i!}.$$

Supondo que $m = 7$, então a aproximação para $F_{S^*}(y)$, que representa a distribuição gama $G(\mu, \phi)$, será

$$\begin{aligned} F_{S^*}(y) = F_Z(y) - e^{-y/\alpha}\{&\epsilon_1 + \frac{1}{2!}\epsilon_2 + \frac{2}{3!}\epsilon_3 + \frac{3}{4!}(2\epsilon_4 + \epsilon_2^2) + \frac{4}{5!}(6\epsilon_5 + 5\epsilon_2\epsilon_3) \\ &+ \frac{5}{6!}[24\epsilon_6 + 9(2\epsilon_4 + \epsilon_2^2) + 3\epsilon_2^3 + 8\epsilon_3^2] + \frac{1}{7!}[720\epsilon_7 + 40\epsilon_2(6\epsilon_5 + 5\epsilon_2\epsilon_3) \\ &+ 160\epsilon_2^2\epsilon_3 + 120\epsilon_3(2\epsilon_4 + \epsilon_2^2)]\}, \end{aligned}$$

em que, $\alpha = \mu$ e $\epsilon_i = \phi^{-(i-1)} - 1$.

Na Tabela 5, apresentam-se os resultados exatos e aproximados para diferentes valores de μ e ϕ para a aproximação da distribuição gama $G(\mu, \phi)$ em função da distribuição exponencial de parâmetro α, sendo $\alpha = \mu = 1$. Observa-se que a aproximação (20) é adequada somente quando $\phi > 1$. Entretanto, pode ser usada quando $\phi < 1$ desde que $y \geq 10$. A Tabela 6 apresenta os resultados exato e aproximado para $y = 10$ e $\phi < 1$.

Exemplo 5. Mediante o uso da equação (20), apresenta-se abaixo a aproximação da função de distribuição t-Student com ν g.l., considerando que esta é a f.d.a. da estatística S^*, em função da distribuição normal padrão Z. Tem-se,

$$P(T_\nu \leq t) = \Phi(t) - \phi(t) \left[\frac{t(t^2+1)}{4\nu} + \frac{t(3t^6 - 7t^4 - 5t^2 - 3)}{96\nu^2} \right] + O(\nu^{-3}). \qquad (25)$$

Tabela 5 - Resultados exato e aproximado da distribuição gama $G(\mu, \phi)$ pela distribuição exponencial de média um, para $\phi = 1, 5, 10, 20$ e 30 e diferentes valores de y

$y = 1$; $\phi =$	1	5	10	20	30
Exato	0.6321206	0.9932620	0.9999546	1	1
Aproximado	0.6321206	0.9915655	0.9960130	0.9983550	0.9991863
$y = 3$; $\phi =$	1	5	10	20	30
Exato	0.9502130	0.9999997	1	1	1
Aproximado	0.9502130	0.9988585	0.9994604	0.9997774	0.9998899
$y = 5$; $\phi =$	1	5	10	20	30
Exato	0.9932620	1	1	1	1
Aproximado	0.9932620	0.9998455	0.9999270	0.9999699	0.9999850
$y = 10$; $\phi =$	1	5	10	20	30
Exato	0.9999546	1	1	1	1
Aproximado	0.9999546	0.9999990	0.9999995	0.9999998	0.9999999

Tabela 6 - Resultados exato e aproximado da distribuição gama $G(\mu, \phi)$ pela distribuição exponencial de média um, para $\phi < 1$ e $y = 10$

$y = 10$; $\phi =$	0.2	0.4	0.6	0.8
Exato	0.8646647	0.9816844	0.9975212	0.9996645
Aproximado	0.8511665	0.9965158	0.9995440	0.9998807

A demonstração da equação (25) pode ser encontrada em Fisher (1925). Johnson *et al.* (1995), também, apresenta esta equação mas é importante a ressalva de que é preciso realizar uma correção na equação (28.15) contida na página 375 do livro, pois a segunda parte da fórmula está multiplicada pela função acumulada da distribuição normal, mas, a forma correta seria a função densidade da normal padrão.

Finner *et al.* (2008) investigaram o comportamento assintótico da f.d.p. e da f.d.a. da distribuição t-Student com $\nu > 0$ g.l. para ν tendendo a infinito quando o argumento $t = t_\nu$ da f.d.p. (f.d.a.) depende de ν e tende a $\pm\infty$. O respectivo artigo volta-se para a análise de algumas propriedades assintóticas da cauda da distribuição t-Student comparada com a distribuição normal padrão.

Objetiva-se, portanto, utilizar o método apresentado no artigo de Cordeiro e Ferrari (1998) para obter uma melhor aproximação da distribuição t-Student.

Em princípio, define-se $\nu = n^{1/2}$, $\nu^2 = n$ e $\nu^3 = n^{3/2}$, em que n é o tamanho amostral. Tem-se, então,

$$P(T_\nu \leq t) = \Phi(t) - \phi(t)\left[\frac{t(t^2+1)}{4n^{1/2}} + \frac{t(3t^6 - 7t^4 - 5t^2 - 3)}{96n}\right] + O(n^{-3/2}). \quad (26)$$

Seja agora $F_S(t) = \Phi(t) + A_1(t) + A_2(t) + O(n^{-3/2})$. Observa-se a partir de (26) que

$$A_1(t) = -\phi(t)\Big[\frac{t(t^2+1)}{4n^{1/2}}\Big]$$

e

$$A_2(t) = -\phi(t)\Big[\frac{t(3t^6 - 7t^4 - 5t^2 - 3)}{96n}\Big].$$

Considerando como base a distribuição normal, através da equação (6) de Cordeiro e Ferrari (1998), segue-se que

$$t^* = t\left[1 - \frac{A_1(t)}{\phi(t)t} - \frac{1}{t}\Big\{ - \frac{A_1(t)A_1'(t)}{\phi(t)^2} + \frac{A_2(t)}{\phi(t)} + \frac{A_1^2(t)\phi'(t)}{2\phi^3(t)}\Big\}\right]$$

Observa-se, então, a necessidade da derivada de $A_1(t)$, sendo esta

$$A_1'(t) = -\phi(t)\Big[\frac{3t^2+1}{4\sqrt{n}}\Big] - \phi'(t)\Big[\frac{t(t^2+1)}{4\sqrt{n}}\Big].$$

Tem-se que $\frac{\partial^n \phi(t)}{\partial t^n} = \phi(t)(-1)^n H_n(t)$, sendo assim, $\phi'(t) = -\phi(t)t$, pois $H_1(t) = t$. Logo, obtém-se $A_1'(t) = \frac{\phi(t)}{4\sqrt{n}}\Big(t^4 - 2t^2 - 1\Big)$. Portanto,

$$t^* = t + \frac{t(t^2+1)}{4n^{1/2}} - \frac{(t^3+t)(t^4 - 2t^2 - 1)}{16n} + \frac{t^3(t^2+1)^2}{32n} + \frac{t(3t^6 - 7t^4 - 5t^2 - 3)}{96n}$$

ou

$$t^* = t + \frac{t(t^2+1)}{4n^{1/2}} + \frac{t(5t^2+1)(t^2+3)}{96n}. \tag{27}$$

Pode-se, assim, expressar a equação (27) como

$$t^* = t + \frac{Q_1(t)}{\sqrt{n}} + \frac{Q_2(t)}{n},$$

em que $Q_1(t) = \frac{t(t^2+1)}{4}$ e $Q_2(t) = \frac{t(5t^2+1)(t^2+3)}{96}$.

Seguindo Cordeiro e Ferrari (1998, p. 515), deseja-se obter $F_{T_\nu}(t^*) = \Phi(t)$, em que t^* é expresso pela equação (27). Dessa forma, é necessário obtermos uma equação que expresse t em termos de t^*. Esta equação é deduzida a partir da equação de inversão (6.74) em Kendall (1945, p. 167), sendo esta expressa por

$$\begin{aligned} t - t^* &= g(t) = g(t^* + t - t^*) \\ &= g(t^*) + g(t^*)g'(t^*) + g(t^*)g'(t^*)^2 + \frac{1}{2}g(t^*)^2 g''(t^*) + g(t^*)g'(t^*)^3 \\ &\quad + \frac{3}{2}g(t^*)^2 g'(t^*)g''(t^*) + \frac{1}{6}g(t^*)^3 g'''(t^*) + \cdots \end{aligned}$$

Da equação (27), tem-se, então, $t - t^* = g(t)$, em que $g(t) = -\frac{Q_1(t)}{\sqrt{n}} - \frac{Q_2(t)}{n}$. Para utilizar a equação dada por Kendall (1945), torna-se necessário obter as derivadas de $g(t)$. Portanto,

$$g'(t) = -\frac{3t^2+1}{4\sqrt{n}} - \frac{25t^4+48t^2+3}{96n},$$

$$g''(t) = -\frac{3t}{2\sqrt{n}} - \frac{t(25t^2+24)}{24n} \quad \text{e}$$

$$g'''(t) = -\frac{3}{2\sqrt{n}} - \frac{25t^2+8}{8n}.$$

Após determinar as derivadas de $g(t)$, obtém-se a expansão de t em função de t^*, ou seja,

$$\begin{aligned}
t \approx t^*\Big\{ & 1 - \frac{(t^{*2}-1)}{4\sqrt{n}} + \frac{(13t^{*4}+8t^{*2}+3)}{96n} + \frac{t^{*2}(2t^{*4}-1)}{24n^{3/2}} \\
& + \frac{455t^{*8} - 656t^{*6} - 1158t^{*4} - 408t^{*2} - 9}{9216n^2} \\
& + \frac{8970t^{*10} + 30745t^{*8} + 31464t^{*6} + 12810t^{*4} + 1950t^{*2} + 45}{36864n^{5/2}} \\
& + \frac{656760t^{*10} + 1174855t^{*8} + 827808t^{*6} + 228375t^{*4} + 19080t^{*2} + 297}{884736n^3} \\
& + \frac{25500t^{*14} + 185500t^{*12} + 479115t^{*10} + 528055t^{*8} + 239250t^{*6} + 40194t^{*4} + 2295t^{*2} + 27}{884736n^{7/2}} \\
& + \frac{178125t^{*16} + 1632000t^{*14} + 5586000t^{*12} + 8765120t^{*10} + 6213130t^{*8} + 1733760t^{*6}}{84934656n^4} \\
& + \frac{201096t^{*4} + 8640t^{*2} + 81}{84934656n^4}\Big\}.
\end{aligned}$$

E, considerando termos até ordem $O(n^{-1})$, tem-se

$$\begin{aligned}
t &\approx t^* - \frac{t^*(t^{*2}+1)}{4\sqrt{n}} - \frac{t^*(5t^{*2}+1)(t^{*2}+3)}{96n} + \frac{t^*(t^{*2}+1)(3t^{*2}+1)}{16n} \\
&= t^* - \frac{t^*(t^{*2}+1)}{4\sqrt{n}} + \frac{t^*(13t^{*4}+8t^{*2}+3)}{96n}. \qquad (28)
\end{aligned}$$

Para diferentes valores de ν, apresenta-se na Tabela 7 a probabilidade exata para a distribuição t-Student, a probabilidade aproximada baseada na distribuição normal padrão (primeiro termo do lado direito de (26)), a probabilidade aproximada baseada na equação (26) e a probabilidade da normal padrão utlizando o valor da estatística t definida pela equação (28), isto é, a f.d.a. $\Phi(\cdot)$ no valor de t correspondente à equação (28).

Na Tabela 7 observa-se ainda que, à medida que ν aumenta os valores obtidos para (26), pela aproximação de t obtida através das equações (27) e (28) sob a distribuição normal padrão, se aproximam dos valores exatos da distribuição

Tabela 7 - Valores aproximados da estatística t para diferentes valores de $\nu = \sqrt{n}$

ν	t	*Exato*	$\Phi(t)$	(26)	(27)	(28)
	0.0	0.0	0.5	0.5	0.5	0.5
2	0.5	0.6666667	0.691462	0.666085	0.659960	0.666608
	1.5	0.8638034	0.933193	0.862118	0.738720	1.000000
	2.0	0.9082483	0.977250	0.893733	0.493767	1.000000
	0.0	0.0	0.5	0.5	0.5	0.5
3	0.5	0.674276	0.691462	0.674071	0.671365	0.674171
	1.5	0.884708	0.933193	0.884065	0.836974	0.961253
	2.0	0.930337	0.977250	0.925134	0.795708	1.000000
	0.0	0.0	0.5	0.5	0.5	0.5
4	0.5	0.678170	0.691462	0.678242	0.676724	0.678279
	1.5	0.896404	0.933193	0.895693	0.871289	0.915385
	2.0	0.940904	0.977250	0.939498	0.881713	0.999900
	0.0	0.0	0.5	0.5	0.5	0.5
5	0.5	0.6808506	0.691462	0.680801	0.679831	0.680819
	1.5	0.9030482	0.933193	0.902879	0.888018	0.908915
	2.0	0.9490303	0.977250	0.947690	0.915821	0.987824
	0.0	0.0	0.5	0.5	0.5	0.5
6	0.5	0.68256	0.691462	0.682530	0.681858	0.682541
	1.5	0.9078596	0.933193	0.907757	0.897773	0.910083
	2.0	0.9537868	0.977250	0.952973	0.932968	0.971901
	0.0	0.0	0.5	0.5	0.5	0.5
7	0.5	0.6837964	0.691462	0.683778	0.683284	0.683784
	1.5	0.9113508	0.933193	0.911284	0.904119	0.912351
	2.0	0.9571903	0.977250	0.956659	0.942987	0.965457
	0.0	0.0	0.5	0.5	0.5	0.5
8	0.5	0.684732	0.691462	0.684719	0.684341	0.684724
	1.5	0.9139984	0.933193	0.913952	0.908563	0.914511
	2.0	0.9597419	0.977250	0.959376	0.949461	0.963790
	0.0	0.0	0.5	0.5	0.5	0.5
9	0.5	0.6854644	0.691462	0.685455	0.685157	0.685458
	1.5	0.9160747	0.933193	0.916042	0.911841	0.916365
	2.0	0.9617236	0.977250	0.961461	0.953949	0.963869
	0.0	0.0	0.5	0.5	0.5	0.5
10	0.5	0.6860532	0.691462	0.686047	0.685805	0.686049
	1.5	0.9177463	0.933193	0.917722	0.914356	0.917924
	2.0	0.9633060	0.977250	0.963111	0.957227	0.964526
	0.0	0.0	0.5	0.5	0.5	0.5
20	0.5	0.6887341	0.691462	0.688733	0.688673	0.688734
	1.5	0.9253821	0.933193	0.925379	0.924577	0.925393
	2.0	0.9703672	0.977250	0.970341	0.969073	0.970408
	0.0	0.0	0.5	0.5	0.5	0.5
30	0.5	0.6896385	0.691462	0.689638	0.689611	0.689638
	1.5	0.9279670	0.933193	0.927966	0.927615	0.927970
	2.0	0.9726875	0.977250	0.972679	0.972143	0.972695
	0.0	0.0	0.5	0.5	0.5	0.5
100	0.5	0.6909132	0.691462	0.690913	0.690911	0.690913
	1.5	0.9316175	0.933193	0.931617	0.931587	0.931618
	2.0	0.9758940	0.977250	0.975894	0.975849	0.975894
	0.0	0.0	0.5	0.5	0.5	0.5
150	0.5	0.6910961	0.691462	0.691096	0.691095	0.691096
	1.5	0.9321419	0.933193	0.932142	0.932128	0.932142
	2.0	0.9763472	0.977250	0.976347	0.976327	0.976347

t-Student. Para o caso referente a f.d.a. $\Phi(\cdot)$ no valor de t correspondente à equação (27), observa-se que os valores das probabilidades aproximam-se do valor exato apenas à medida que se aumentam os graus de liberdade. Para o caso referente a f.d.a. $\Phi(\cdot)$ no valor de t correspondente à equação (28), nota-se que os correspondentes valores das probabilidades se aproximam do valor exato da distribuição t-Student para valores pequenos de g.l. e para valores de t pequenos (exceto quando $\nu < 5$) e, ainda, à medida que o g.l. é aumentado esse valor de probabilidade se torna ainda mais próximo do valor exato para qualquer valor de t apresentado.

Dessa maneira, baseado em Cordeiro e Ferrari (1998), vê-se a possibilidade de uso da propriedade $F_{T_\nu}(t^*) = \Phi(t)$, em que t é agora expresso pela equação (28), ou seja, fazer uso da equação (28) em termos da distribuição normal padrão para obter valores de probabilidade de uma distribuição t-Student com ν graus de liberdade.

6 Quadratura de Gauss-Hermite

A quadratura Gaussiana escolhe os pontos para se calcular a aproximação da integral de maneira otimizada, em vez de considerar apenas os pontos igualmente espaçados. Os nós $y_1, y_2, \ldots, y_n$ no intervalo $[a, b]$ e os coeficientes $c_1, c_2, \ldots, c_n$ são escolhidos de modo a minimizar o erro esperado no cálculo da aproximação

$$\int_a^b f(y)dy \approx \sum_{i=1}^{m} c_i f(y_i).$$

Para medição da precisão é assumido que a melhor escolha desses valores seja a que produza o resultado exato para a maior classe de polinômios, isto é, a escolha que apresente o maior grau de precisão (Burden e Faires, 2003).

Na quadratura de Gauss-Hermite, considera-se um método sistemático para transformar a variável de integração tal que o integrando seja amostrado em uma região apropriada. A quadratura de Gauss-Hermite é usada muitas vezes em estatística para integração numérica por causa da sua relação com a função densidade Gaussiana, mas parece que há uma freqüente idéia inadequada dada a sua implementação. A quadradura é definida em termos de integrais da forma

$$\int_{-\infty}^{\infty} f(y) \exp(-y^2) dy. \tag{29}$$

Na quadratura de Gauss-Hermite, a integral (29) é aproximada por

$$\int_{-\infty}^{\infty} f(y) \exp(-y^2) dy \approx \sum_{i=1}^{m} w_i f(w_i), \tag{30}$$

em que os y_i são zeros de m-ésima ordem do polinômio de Hermite e os w_i são pesos adequados correspondentes (Liu e Pierce, 1994). Os y_i são simétricos em relação a zero, de forma que a função $f(y)$ é amostrada independente do campo em que se está interessado.

Para bons resultados, é necessário realizar algumas transformações ordinariamente, de forma que o integrando original seja amostrado num campo razoável.

Liu e Pierce (1994) consideram uma forma sistemática para aplicar a quadratura de Gauss-Hermite para integrais da forma

$$\int_{-\infty}^{\infty} g(t)dt, \tag{31}$$

em que $g(t) > 0$. Como se pode observar, o requisito para resultados efetivos é que a razão de $g(t)$ com alguma curva Gaussiana seja uma função moderadamente suave. Isto resulta muitas vezes, por exemplo, quando $g(t)$ é uma função de verossimilhança, no produto de uma função de verossimilhança e uma função densidade Gaussiana, e no produto de várias funções de verossimilhança, etc. Liu e Pierce (1994) assumem que $g(t)$ apresentam tais características e, em particular, que é unimodal.

O objetivo da quadratura de Gauss-Hermite é realizar uma transformação em t tal que o integrando $g(t)$ seja amostrado num campo aceitável. Liu e Pierce (1994) definiram uma forma clara para alcançar esse objetivo, tal que fosse re-expressada a quadratura de Gauss-Hermite (29) como a integral da forma

$$\int_{-\infty}^{\infty} f(t)\phi(t;\mu,\sigma)dt, \tag{32}$$

em que $\phi(t;\mu,\sigma)$ é uma densidade Gaussiana arbitrária (Naylor e Smith, 1982). Os nós amostrais são, então, $t_i = \mu + 2^{\frac{1}{2}}\sigma y_i$, e os pesos são modificados por $w_i/\sqrt{\pi}$. Escolhe-se, então, μ e σ tal que $g(t)$ seja amostrada na região apropriada. Em particular, toma-se $\hat{\mu}$ para ter a forma de $g(t)$, e $\hat{\sigma} = 1/\sqrt{\hat{j}}$, em que

$$\hat{j} = -\left.\frac{\partial^2}{\partial t^2}\log g(t)\right|_{t=\hat{\mu}}.$$

Isto resulta em uma densidade Gaussiana $\phi(t;\hat{\mu},\hat{\sigma})$ que tem as mesmas derivadas logaritmas de segunda ordem como na forma do integrando $g(t)$. Define-se

$$h(t) = \frac{g(t)}{\phi(t;\mu,\sigma)} \tag{33}$$

tal que se possa escrever

$$\int_{-\infty}^{\infty} g(t)dt = \int_{-\infty}^{\infty} h(t)\phi(t;\hat{\mu},\hat{\sigma})dt. \tag{34}$$

Ao aplicar a quadratura de Gauss-Hermite em (32) usando $\phi(t;\hat{\mu},\hat{\sigma})$, a função $h(t)$ e, por conseguinte $g(t)$, serão amostradas no campo relevante, resultando

$$\begin{aligned}\int_{-\infty}^{\infty} g(t)dt \approx& \sum_{i=1}^{m} \frac{w_i}{\sqrt{\pi}} h(\hat{\mu} + 2^{\frac{1}{2}}\hat{\sigma}y_i)\\ =& 2^{\frac{1}{2}}\hat{\sigma}\sum_{i=1}^{m} w_i^* g(\hat{\mu} + 2^{\frac{1}{2}}\hat{\sigma}y_i),\end{aligned} \tag{35}$$

em que $w_i^* = w_i \exp(y_i^2)$.

Liu e Pierce (1994) afirmam que a m-ésima ordem da quadratura de Gauss-Hermite, conforme implementada em (35), será altamente eficiente se a razão de $g(t)$ com uma função densidade Gaussiana $\phi(t; \hat{\mu}, \hat{\sigma})$ possa ser bem aproximada por um polinômio de ordem $2m + 1$ na região onde $g(t)$ é substancial.

Quando (35) é aplicada com apenas uma observação, o resultado é

$$\int_{-\infty}^{\infty} g(t)dt \approx h(\hat{\mu}) = (2\pi)^{\frac{1}{2}} \hat{\sigma} g(\hat{\mu}), \tag{36}$$

que é a aproximação de Laplace para a integral. Assim, a m-ésima ordem da quadratura de Gauss-Hermite, como implementada em (35), pode ser imaginada como uma forma alternativa da "m-ésima ordem da aproximação de Laplace".

A precisão assintótica da aproximação de Laplace para inferência Bayesiana foi estudada por Tierney e Kadane (1986). Wong e Li (1992) consideram que, ao incorporar um termo de correção de ordem superior, ocorre melhora das aproximações de Laplace para aplicações na Estatística. Esse método envolve cálculos de terceira e quarta derivadas do integrando $g(t)$. A quadratura de Gauss-Hermite (35) envolve derivadas de $g(t)$ apenas até segunda ordem, substituindo em essência o uso de derivadas de ordem superior pela amostragem da função $g(t)$.

As aplicações estatísticas muitas vezes envolvem uma família paramétrica de integrais

$$I(\beta) = \int_{-\infty}^{\infty} g(t; \beta)dt,$$

em que a função $I(\beta)$ necessita apenas ser aproximada por uma constante de proporcionalidade. Nesses casos, tudo que interessa sobre a precisão da quadratura ou aproximação de Laplace é que o erro relativo varie lentamente com β (Liu e Pierce, 1994).

Para o caso dos modelos lineares generalizados com efeitos aleatórios, Liu e Pierce (1993), descobriram que a aproximação de Laplace é muitas vezes praticamente exata. Essa mesma questão recái, também, na inferência Bayesiana quando a integral de interesse define uma função de densidade a posteriori. Por exemplo, se $g(t)$ é somente proporcional à densidade a posteriori, então o significado da posteriori para uma função $a(t)$ seria numericamente computada como

$$\int_{-\infty}^{\infty} a(t)g(t)dt \bigg/ \int_{-\infty}^{\infty} g(t)dt.$$

7 Considerações finais

Neste artigo são apresentadas algumas expansões assintóticas utilizadas para obter expansões do tipo ponto de sela. O artigo faz uma revisão detalhada das expansões de Edgeworth, Lugannani-Rice e Daniels com exemplos. Foi feito um estudo detalhado da estatística proposta por Cordeiro e Ferrari (1998), que aproxima, até a ordem de $O(n^{-1})$, a soma padronizada de variáveis aleatórias independentes identicamente distribuídas. Através de avaliações numéricas compara-se a peformance da estatística de Cordeiro e Ferrrari (1998) com as expansões de Edgeworth, Lugannani e Rice(1980) e Daniels (1987). Mostra-se como utilizar os resultados de Cordeiro e Ferrari (1998) na aproximação da distribuição t-Student para a distribuição normal padrão. As expansões ponto de sela apresentam diversas aplicações na estatística, a exemplo da grande precisão em aproximar funções densidade e de distribuição da soma e da média de variáveis aleatórias i.i.d.. Ainda, mediante aproximação da função de distribuição proposta por Cordeiro (1998), foi possível obter uma aproximação da distribuição gama $G(\mu, \phi)$ em função da distribuição exponencial de média α. E, ainda, ao utilizar a estatística, também, proposta por Cordeiro (1998), foi obtida uma estatística que ao ser usada em função da f.d.a. da distribuição normal $\Phi(t)$, resulta no valor aproximado da probabilidade da distribuição t de Student com ν graus de liberdade. Isso, portanto, corresponde à propriedade $F_{T_\nu}(t^*) = \Phi(t)$. É importante observar que melhores resultados dessa relação entre a estatística (28) e a f.d.a. da distribuição normal são obtidos quando $\nu > 9$.

Agradecimentos.

Os autores agradecem a dois revisores e ao editor por várias sugestões que melhoraram o artigo em várias seções.

CORDEIRO, G. M.; BRITO, R. S. Comparing asymptotic expansions from Edgeworth, Lugannani-Rice, Daniels and Cordeiro-Ferrari with applications in statistics. *Rev. Bras. Biom.,* São Paulo, v.27, n.2, p.225-254, 2009.

- *ABSTRACT: In this paper we explain the saddlepoint approximation and present examples to show the applicability of these asymptotic expansions in statistics. Using the Laplace approximation for integrals and Edgeworth expansions we show how to derive the saddlepoint approximation to the density of a single random variable and then to the density of the sample mean of independent and identically distributed random variables. Cordeiro and Ferrari (1998) proposed a statistic which approximates to order $O(n^{-1})$ the standardized sum of independent and identically distributed random variables, where n is the samplesize. We show the performance of this statistic when compared to the expansions due to Edgeworth, Lugannani and Rice (1980) and Daniels (1987).*

- *KEYWORDS: Edgeworth approximation; Laplace approximation; Saddlepoint approximation; generalized Bartlett correction.*

Referências

BAHADUR, R. R.; RANGA RAO, R. On deviations of the sample mean. *Ann. Math. Stat.*, Providence, v.31, p.1015-1027, 1960.

BARTLETT, M. S. Properties of sufficiency and statistical tests. *Proc. R. Soc. A*, London, v.160, p.268-282, 1937.

BARNDORFF-NIELSEN, O.; COX, D. R. Edgeworth and saddlepoint approximations with satistical applications. *J. R. Statist. Soc. B*, London, v.41, p.279-312, 1979.

BLACKWELL, D.; HODGES, J. L. The probability in the extreme tail of a convolution. *Ann. Math. Stat.*, Ann Habor, v.31, p.1113-1120, 1959.

BURDEN, R. L.; FAIRES, J. D. *Análise numérica.* São Paulo: Pioneira Thomson Learning, 2003.

CORDEIRO, G. M. *Introdução à teoria assintótica.* livro texto do 22º Colóquio Brasileiro de Matemática. Rio de Janeiro: IMPA, 1999.

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

CORDEIRO, G. M.; FERRARI, S. L. P. Generalized Bartlett Correction. *Commun. Stat. - Theor. Methods*, Philadelphiav.27, p.509-527, 1998.

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

CRAMÉR, H. *Sur un nouveu théorèm-limite de la théorie des probabilités.* Paris: Hermann et Cie, 1938. (Actualites Scientifiques et Industrielles, 736).

CRIBARI-NETO, F.; CORDEIRO, G. M. On Bartlett and Bartlett-type corrections. *Econ. Rev.*, New York, v.15, p.339-367, 1996.

DANIELS, H. E. Saddlepoint approximations in statistics. *Ann. Math. Stat.*, Ann Harbor, v.25, p.631-650, 1954.

DANIELS, H. E. Exact saddlepoint approximations. *Biometrika*, London, v.67, p.59-63, 1980.

DANIELS, H. E. Tail probability approximations. *Int. Stat. Rev.*, Edinburg, v.55, 1, p.37-48, 1987.

DAVISON, A. C. Biometrika centenary: theory and general methodology. *Biometrika*, London, v.88, p.13-52, 2001.

ESSCHER, F. The probability function in the collective theory of risk. *Skand. Akt. Tidsskr. (Scand. Actuarial J.)*, Stockholm, v.15, p.175-195, 1932.

FINNER, H; DICKHAUS, T.; ROTERS, M. Asymptotic tail properties of "Student's" t-distribution. *Commun. Stat. - Theor. Methods*, New York, v.37, p.175-179, 2008.

FISHER, R. A. Expansion of "Student's" integral in power of n^{-1}. *Metron*, Roma, v.5, p.109-120, 1925.

GOUTIS, C.; CASELLA, G. Explaining the saddlepoint approximation. *Am. Stat.*, Shashington, v.53, p.216-224, 1999.

HINKLEY, D. V.; REID, N.; SNELL, E. J. *Statistical theory and modelling.* Boca Raton: Chapman and Hall, 1990. 376p.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. *Continuous univariate distribution.* A Wiley-Interscience Publication, 1995. 719p.

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

KENDALL, M. G. *The advanced theory of statistics.* London, 1945. v.1, cap.6, p.158.

KOLASSA, J. E. *Series approximation methods in statistics.* 2nd. ed., Springer, 1997. (Lecture Notes in Statistics, 88).

LIU, Q.; PIERCE, D. A. Heterogeneity in Mantel-Haenszel-type models. *Biometrika*, London, v.80, p.543-556, 1993.

LIU, Q.; PIERCE, D. A. A note on Gauss-Hermite quadrature. *Biometrika*, London, v.81, p.624-629, 1994.

LUGANNANI, R.; RICE, S. Saddlepoint approximation for the distribution of the sum of independent random variables. *Adv. Appl. Prob.*, Sheffield, v.12, p.475-490, 1980.

McCULLAGH, P. *Tensor methods in statistics.* London: Chapman & Hall, 1987. 285p.

NAYLOR, J. C.; SMITH, A. F. M. Applications of a method for the efficient computation of posterior distributions. *Appl. Stat.*, London, v.31, p.214-225, 1982.

REID, N. Saddlepoint methods and statistical inference (with discussion). *Stat. Sci.*, Bethesda, v.3, p.213-238, 1988.

ROBINSON, J. Saddlepoint approximations for permutation tests and confidence intervals. *J. R. Stat. Soc B*, London, v.44, p.91-101, 1982.

TIERNEY, L.; KARDANE, J. B. Accurate approximations for posterior moments and marginal densities. *J. Am. Stat. Assoc.*, New York, v.81, p.82-86, 1986.

WONG, H. W.; Li, B. Laplace expansion for posterior densities of nonlinear functions of parameters. *Biometrika*, London, v.79, p.393-398, 1992.

Recebido em 26.01.2009.

Aprovado após revisão em 28.08.2009.