COMUNICAÇÃO

AVALIAÇÃO DE MÉTODOS DE ESTIMAÇÃO INTERVALAR PARA FUNÇÕES LINEARES BINOMIAIS VIA BOOTSTRAP INFINITO

Evaluation of intervalar estimation methods for binomial linear functions through infinity bootstrap

Marcelo Angelo Cirillo[1], Daniel Furtado Ferreira[2], Thelma Safádi[3]

RESUMO

Objetivou-se, neste trabalho, avaliar intervalos de confiança de diferentes funções lineares compostas por proporções binomiais construídas com base no método de Wald e Wald ajustado utilizando a técnica de *bootstrap* infinito. Considerados diversos tamanhos amostrais (n_{is}), parâmetros (p) e número de coeficientes das funções lineares. Concluiu-se, por meio das probabilidades de cobertura dos intervalos de confiança das populações binomiais, cujas proporções foram baixas (p=0,2) e considerando diferentes tamanhos de amostras, que a generalização *bootstrap* do método de Wald ajustado foi eficiente para funções lineares cujos coeficientes indicaram a comparação de uma proporção versus as demais.

Termos para indexação: Probabilidade de cobertura, intervalos de confiança, Wald.

ABSTRACT

This work aimed to evaluate confidence intervals of different linear functions of binomial proportions base on Wald and Wald's adjusted method using the infinity bootstrap technique. Several sample sizes (n_i), binomial parameters (p) and number of coefficients of the linear functions were considered. One concluded through the probabilities of binomial population confidence covered intervals that the bootstrap generalization of adjusted Wald's method was efficient for linear functions whose coefficients indicated a comparison of versus the others proportion (p=0.2) considering different sample sizes.

Index terms: Covered probabilities, confidence intervals, Wald.

(Recebido em 9 de março de 2006 e aprovado em 31 de março de 2008)

Os intervalos de confiança para diferença de duas proporções são usualmente estimados pelo método de Wald, cujo fundamento está na aproximação assintótica da distribuição normal. Contudo, para amostras pequenas, essas estimativas não são condizentes com o nível de confiança desejado. Como conseqüência desse problema, Agresti & Coull (1998) propuseram uma modificação no método, que consiste em acrescentar quatro pseudo-observações classificadas, empiricamente, como duas observações representando "sucessos" e duas representando "fracassos", para uma população binomial. Posteriormente, Agresti & Caffo (2000) realizaram essa modificação no intervalo de confiança para diferença de duas proporções. Com essas modificações, originou-se o teste de Wald ajustado, doravante representado nesse trabalho por Wald $_{(ajust)}$.

Price & Bonett (2004) apresentaram uma extensão desses métodos, considerando funções lineares de proporções binomiais. O termo função linear se dá por apenas assumir linearidade nos parâmetros p. Vale ressaltar que a praticidade dessas funções lineares dos parâmetros binomiais consiste em comparar tratamentos não estruturados, exemplificados pelas comparações múltiplas, ou estruturados, no sentido de grupos de tratamentos serem previamente determinados para análise. Tal estrutura poderá ser exemplificada pela comparação de um tratamento controle versus os demais. Portanto, nota-se que o uso de funções lineares, perfeitamente enquadra-se em aplicações da estatística experimental, diferindo apenas no fato de que as populações consideradas são binomiais e não normais.

Um exemplo prático da aplicação dessas funções pode ser encontrado em Cohen et al. (1991). O experimento

[1]Doutor em Estatística e Experimentação Agropecuária – Departamento de Ciências Exatas/DEX – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – Lavras, MG – marcufla@gmail.com, marcelocirillo@hotmail – Bolsista FAPEMIG
[2]PhD em Estatística e Experimentação Agropecuária, Professor Adjunto – Departamento de Ciências Exatas/DEX – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – Lavras, MG – danielff@ufla.br
[3]PhD em Estatística, Professora Adjunta – Departamento de Ciências Exatas/DEX – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – Lavras, MG – safadi@ufla.br

realizado por esses autores consistiu em aplicar as funções lineares binomiais para avaliar o efeito de diferentes dietas alimentares, em relação à formação ou não de tumores. Cada dieta foi formada a partir dos fatores: fibra (ausência/presença) e gordura, classificadas nos níveis: alto e baixo, sendo essas submetidas, aleatoriamente, a 120 ratos. As respostas de cada tratamento foram determinadas por meio de contagens de ratos que desenvolveram ou não tumores, de forma a propiciar o cálculo das proporções de cada tratamento, em relação ao número total de ratos.

Com base nesses resultados, os autores construíram três funções lineares, de modo que os coeficientes utilizados para cada função permitiram estimar a significância dos efeitos principais e da interação.

Em se tratando do estudo dos métodos utilizados para estimar as funções lineares, Price & Bonett (2004) obtiveram as probabilidades de cobertura dos métodos Wald e $\text{Wald}_{(\text{ajust.})}$ por meio de simulação Monte Carlo, utilizando diferentes configurações, como número de populações, diferentes proporções e funções lineares. Assim, os mesmos autores concluíram que o intervalo de confiança de $\text{Wald}_{(\text{ajust.})}$ resultou em probabilidades de cobertura mais próximas ao nível de confiança desejado, em relação às probabilidades propiciadas pelo método de Wald original. Entretanto, convém salientar que, em algumas situações, Price & Bonett (2004) verificaram que os métodos em questão (Wald e $\text{Wald}_{(\text{ajust.})}$) revelaram resultados referentes à probabilidade de cobertura, inapropriados ao nível nominal de confiança fixado em 95%. Com base nessas evidências, julgou-se necessário avaliar o desempenho desses métodos na abordagem *bootstrap*, na expectativa de que resultados condizentes com o nível de confiança nominal fossem obtidos. Destaca-se, assim o objetivo deste trabalho de avaliar os métodos de Wald e $\text{Wald}_{(\text{ajust.})}$ por meio das probabilidades de cobertura e amplitude intervalar, submetidos à abordagem *bootstrap* proposta por Conlon & Thomas (1990), destacando seu mérito nas situações em que os estimadores intervalares convencionais apresentaram deficiências. Para isso, configurações não contempladas por Price & Bonett (2004) foram avaliadas, sendo essas diferenciadas pela estrutura dos tratamentos e diferentes valores paramétricos representados pelas proporções p.

Com esse propósito, consideraram-se k=4, 6 e 8 populações binomiais (n_i,p_i), em que os tamanhos amostrais foram fixados em n_i=10 e 60 (i=1,...,k) para funções lineares com coeficientes definidos pelos vetores c_1=(1,1,-1,-1); c_2=(1,1,1,-1,-1,-1) e c_3=(1,1,1,1,-1,-1,-1,-1) e p_i=0,2; 0,5 e 0,9. Tal situação se deu para amostras de tamanhos iguais para todas as populações. No caso de amostras diferentes, os tamanhos amostrais foram definidos pela relação $n_i=d_j\times 10$, sendo d_j a j-ésima componente do vetor d=(1,2,...,k) e os coeficientes das funções lineares foram c_1=(3,-1,-1,-1); c_2=(5,-1,-1,-1,-1,-1) e c_3=(7,-1,-1,-1,-1,-1,-1,-1). Com base nessas configurações, realizaram-se 2000 simulações Monte Carlo, nas quais foram geradas observações binomiais independentes, Y_i, (i=1, 2,..., k). Para cada simulação, foram estimados os intervalos de confiança de Wald (1) e $\text{Wald}_{(\text{ajust.})}$ (2), para funções de proporções binomiais representadas por $\Psi=\sum_{i=1}^{k}c_ip_i$:

$$\hat{\Psi}_1 \pm z_{\frac{\alpha}{2}}\sqrt{\sum_{i=1}^{k}c_i^2\,\frac{\hat{p}_i(1-\hat{p}_i)}{n_i}} \tag{1}$$

$$\hat{\Psi}_{(2)} \pm z_{\frac{\alpha}{2}}\sqrt{\sum_{i=1}^{k}c_i^2\,\frac{\tilde{p}_i(1-\tilde{p}_i)}{(n_i+4/k)}} \tag{2}$$

em que, $\hat{p}_i=\dfrac{Y_i}{n_i}$; $\hat{\Psi}_{(1)}=\sum_{i=1}^{k}c_i\hat{p}_i$; $\hat{\Psi}_{(2)}=\sum_{i=1}^{k}c_i\tilde{p}_i$; $\tilde{p}_i=\dfrac{(Y_i+2/k)}{(n_i+4/k)}$ e $Z_{\frac{\alpha}{2}}$ o quantil superior 100(a/2)% da distribuição normal padrão, cujo valor é 1,96, para a=5%.

Em ambos os métodos, as probabilidades de cobertura foram obtidas pelas porcentagens de intervalos, que contiveram o valor do parâmetro Ψ e o comprimento médio pela média das amplitudes intervalares no final das 2000 simulações. O *bootstrap* infinito, definido por Conlon & Thomas (1990), consiste em um método de computação intensiva utilizado para estimar diferença entre duas proporções binomiais, utilizando-se a distribuição conjunta do modelo binomial para cada população.

Neste trabalho, utilizou-se o algoritmo definido por esses autores, porém expandindo a metodologia proposta para k populações binomiais, referenciadas por meio de um único valor paramétrico definido por $\Psi=\sum_{i=1}^{k}c_ip_i$. Dessa forma, a distribuição de probabilidade conjunta a ser utilizada no processo de reamostragem das observações $(y_1,y_2,\ldots,y_k)$, uma vez que, os tamanhos amostrais n_i (i=1,...,k) foram previamente fixados, foi dada por:

$$\hat{P}\left(Y_1 = y_1, Y_2 = y_2, ..., Y_k = y_k\right) =$$

$$= \prod_{i=1}^{k} \binom{n_i}{y_i} p_{i(z)}^{y_i} \left(1 - p_{i(z)}^{y_i}\right)^{n_i - y_i} \tag{3}$$

em que z =1 e 2.

de modo que se z=1, o estimador de $p_i = \hat{p}_i$ e assumindo z=2 tem-se $p_i = \tilde{p}_i$. Assim, utilizou-se essa expressão para gerar amostras *bootstrap,* de tal forma que, cada combinação amostral, obtida pela realização das B=2000 reamostragens foi representada por $(y_1^*, y_2^*, ..., y_k^*)$. Cada valor gerado foi classificado como "sucesso", e utilizado para compor as estimativas $\hat{p}_i$ e $\tilde{p}_i$. Isso posto, para um melhor entendimento da obtenção da probabilidade conjunta (4), definiu-se o conjunto de pontos $\Omega = \left\{\{0,...,n_1\} \times \{0,...,n_2\} \times ... \times \{0,...,n_k\}\right\}$, no qual cada valor simulado estará contido.

Para cada estimativa de $\Psi_{b(z)}$ em que, z = 1 ou 2, estimou-se a probabilidade $D = \Psi_{b(z)}$, de acordo com a expressão

$$\hat{P}\left(D = \Psi_{b(z)}\right) = \sum_{\left\{(y_1^*, y_2^*, ..., y_k^*) \in \Omega \,/\, \Psi_{b(z)}\right\}} = \tag{4}$$

$$= \hat{P}(Y_1 = y_1^*, Y_2 = y_2^*, Y_3 = y_3^*, ..., Y_k = y_k^*)$$

Posteriormente, após a realização das 2000 reamostragens com reposição, obteve-se a distribuição empírica da variável aleatória D. Em seguida, esses valores foram ordenados e os limites inferiores (q_1) e superiores (q_2) foram estimados conforme as expressões (5) e (6).

$$q_1 = \min_{\Psi} \left\{ \sum_{\left\{\Psi_{b(z)} \geq \Psi\right\}} \hat{P}\left(D = \Psi_{b(z)}\right) \leq \frac{a}{2} \right\} \tag{5}$$

$$q_2 = \max_{\Psi} \left\{ \sum_{\left\{\Psi_{b(z)} < \Psi\right\}} \hat{P}\left(D = \Psi_{b(z)}\right) \leq \frac{a}{2} \right\} \tag{6}$$

sendo, z=1 ou 2.

Seguindo a metodologia, mencionada anteriormente, procedeu-se à apresentação e à discussão dos resultados encontrados na Tabela 1. Observou-se que, para o valor da proporção (p=0,2), quando consideradas amostras menores ($n_{i(s)}$=10), a probabilidade de cobertura foi superior a 95%, em todos os métodos avaliados.

Nessa primeira análise, uma atenção especial deve ser dada aos resultados relativos ao método de Wald$_{(ajust.)}$, justamente por serem concordantes com os resultados obtidos por Price & Bonett (2004). Contudo, vale ressaltar que a metodologia utilizada por esses autores assumiu diferentes valores para os coeficientes c, tendo o(s) parâmetro(s) p sido gerados por uma distribuição uniforme. Isso posto, respalda-se em afirmar que, independente dos coeficientes assumidos, o método de Wald$_{(ajust.)}$ apresentou probabilidade de cobertura próxima do nível de confiança nominal especificado. Além do mais, notou-se similaridade com o método de Wald-Laplace, cujas probabilidades de cobertura foram calculadas de forma exata por Greenland (2001).

Aumentando-se o valor dessa proporção para p = 0,5, ainda considerando-se n_i=10 (i=1,...,k). As probabilidades de cobertura dos métodos de Wald e Wald$_{(ajust.)}$ foram reduzidas, apresentando valores levemente inferiores ao nível de confiança especificado (95%). Esse fato foi mais notável para proporções maiores, no caso p=0,9, em que apenas o método de Wald $_{(ajust.)}$ manteve probabilidades superiores ao nível de confiança nominal, corroborando os resultados obtidos por Price & Bonett (2004).

Em relação à eficiência da abordagem *bootstrap* dos métodos mencionados, observou-se que nenhum resultado agregou informação no sentido de corrigir os intervalos, de forma a proporcionar uma probabilidade de cobertura desejável em relação a um valor nominal do coeficiente de confiança. Diante dos resultados apresentados e discutidos, com base nos valores paramétricos avaliados, não se aconselha a generalização *bootstrap* desses métodos.

Analisando-se os resultados referentes às amostras $n_{i(s)}$=60, observou-se que, de modo geral, todos os métodos, inclusive a generalização *bootstrap*, mostraram probabilidades superiores ou aproximadas ao nível de confiança especificado em 95%, desprezando-se o erro de Monte Carlo. Portanto, pode-se inferir que os problemas de aproximação assintótica, observados por Agresti & Caffo (2000), no caso de diferença de duas proporções binomiais, mantêm-se para k proporções.

Considerando amostras de mesmo tamanho e funções lineares simétricas (ψ=0) para quaisquer valores de p fixado (ex. p=0,9), os resultados são iguais às funções em que os parâmetros são definidos como 1-p (ex. p=0,1).

Tabela 1 – Probabilidade de cobertura de 95% de confiança dos intervalos de Wald e $Wald_{(ajust.)}$, nas versões originais e *bootstrap*, para função de proporções binomiais para amostras de tamanhos iguais.

	Métodos	p=0,2			p=0,5			p=0,9		
		c_1	c_2	c_3	c_1	c_2	c_3	c_1	c_2	c_3
n=10	Wald	0,969	0,968	0,975	0,926	0,933	0,940	0,412	0,377	0,385
	Wald (Boot)	0,971	0,964	0,974	0,930	0,936	0,935	0,412	0,376	0,388
	$Wald_{(ajust.)}$	0,960	0,952	0,964	0,954	0,948	0,941	0,983	0,976	0,969
	$Wald_{(ajust.)}$ (Boot)	0,967	0,965	0,973	0,933	0,923	0,936	0,412	0,378	0,389
n=60	Wald	0,951	0,950	0,952	0,953	0,956	0,951	0,943	0,954	0,936
	Wald (Boot)	0,950	0,949	0,949	0,950	0,953	0,946	0,936	0,949	0,936
	$Wald_{(ajust.)}$.	0,954	0,953	0,954	0,959	0,956	0,951	0,956	0,964	0,956
	$Wald_{(ajust.)}$ (Boot)	0,948	0,950	0,949	0,950	0,953	0,946	0,936	0,949	0,937

c_1=(1,1-,1,-1); c_2=(1,1,1,-1,-1,-1); c_3=(1,1,1,1,-1,-1,-1,-1).

Facilmente isso é verificado, em virtude da propriedade de simetria da distribuição binomial.

Os resultados encontrados na Tabela 2 referem-se às probabilidades de cobertura para amostras de tamanhos diferentes, conforme especificado pela relação $n_i=d_j\times10$, sendo d_j a j-ésima componente do vetor d = (1,2,...,k) e exemplificado para k = 4 populações, como n_1 = 10; n_2 = 20; n_3 = 30 e n_4 = 40.

Bonnet & Woodward (1987) destacaram que as funções lineares de proporções binomiais correspondem a casos mais complexos de comparações, pois um número maior de populações poderá ser exigido, podendo envolver não apenas os efeitos principais, mas também as interações. Em particular, para duas populações, a análise dessas funções coincide com a comparação de duas proporções binomiais, comumente realizada por meio de intervalos de confiança ou testes de hipóteses. Nesse caso, podem-se citar importantes resultados obtidos por Agresti & Caffo (2000), os quais, em síntese, afirmam que o intervalo de confiança de Wald, construído por meio dos estimadores de máxima verossimilhança apresentou deficiências em sua cobertura.

Posteriormente, Carari (2004) avaliou, por meio de simulação Monte Carlo, este método e concluiu que a probabilidade de cobertura apresentou probabilidades inferiores aos valores nominais do coeficiente de confiança, comprometendo sua aplicação prática para pequenas amostras. Assim sendo, verifica-se a inviabilidade da aplicação do método de Wald, seja na forma convencional ou na sua generalização *bootstrap*, inferida nesse trabalho para amostras provenientes de populações binomiais com o valor da proporção p=0,9, dado que, nessa situação, as probabilidades de cobertura foram bem inferiores ao nível de confiança especificado em 95%, A deficiência no método de Wald verificada nas funções lineares torna-se plausível de ser aceita. Entretanto, a principal vantagem da generalização *bootstrap* se deu para o método de $Wald_{(ajust.)}$, no qual, notoriamente, observou-se que, para populações com baixo valor (p = 0,2), esse método revelou probabilidades inferiores ao nível de confiança desejado. Porém, comparando-se os resultados desse método na abordagem *bootstrap*, verificou-se que essa deficiência foi corrigida, revelando em algumas funções lineares ganhos acima de 7 pontos porcentuais, como pode ser percebido nos resultados dos contrastes c_4, c_5 e c_6 (Tabela 2). Porém, não se pode expandir essa vantagem para p=0,9, de modo que, nesse caso, a generalização desse método apresentou resultados bastante insatisfatórios, em relação ao nível de confiança desejado.

Outro fator relevante é que a eficiência do método $Wald_{(ajust.)}$, mencionada anteriormente, se deu em funções lineares que retratam o teste de uma proporção em relação às demais, de maneira semelhante ao caso de contraste, envolvendo um tratamento controle com os demais. Além do mais, tais deficiências apontadas por esse método ratificam os resultados obtidos por Price & Bonett (2004), considerando pequeno número de população ($K \leq 4$) e diferentes combinações dos tamanhos amostrais (n) fixados em 10, 20 e 30. No entanto, todos os resultados obtidos por esses autores e citados em toda discussão desse trabalho, limitam-se apenas à simulação Monte Carlo. Evidentemente, são resultados empíricos e avaliando algumas situações escolhidas ao acaso por esses autores, o que, de certa forma, restringe a obtenção de conclusões mais amplas.

O método Wald ajustado foi proposto por Agresti & Coull (1998) com a finalidade de melhorar a cobertura do intervalo de Wald, para duas populações binomiais. Considerando esse fato, os resultados encontrados na

Tabela 2 – Probabilidade de cobertura de 95% de confiança dos intervalos de Wald e $Wald_{(ajust.)}$, nas versões originais e *bootstrap*, considerando as funções c_4, c_5 e c_6 de proporções binomiais para amostras de tamanhos diferentes.

Métodos	p=0,2			p=0,5			p=0,9		
	c_4	c_5	c_6	c_4	c_5	c_6	c_4	c_5	c_6
Wald	0,984	0,989	0,985	0,920	0,891	0,906	0,573	0,641	0,655
Wald (Boot)	0,984	0,988	0,980	0,935	0,906	0,919	0,574	0,652	0,654
$Wald_{(ajust.)}$.	0,913	0,880	0,883	0,937	0,901	0,910	0,989	0,981	0,948
$Wald_{(ajust.)}$(Boot)	0,980	0,986	0,973	0,937	0,907	0,919	0,572	0,649	0,65

c_4=(3,1-,1,-1); c_5=(5,1,1,-1,-1,-1); c_6=(7,1,1,1,-1,-1,-1,-1).

Tabela 2 evidenciam que a expansão desses métodos para um número maior de populações, representadas por meio das funções lineares, resultou em probabilidades questionáveis.

Esse questionamento é coerente, ao avaliar o método de Wald ajustado, em que as populações foram simuladas considerando-se o valor paramétrico em p=0,2. Nesse caso, suas probabilidades de cobertura foram inferiores às probabilidades do método de Wald, além de terem sido inferiores ao nível de confiança nominal. Contudo, enfatizando o estudo da abordagem *bootstrap* nos métodos em questão, pode-se observar que, os resultados propiciados pela generalização *bootstrap* do método de Wald ajustado, referidos na Tabela 2 por $Wald_{(ajust.)}$(Boot), corrigiu essa deficiência, apresentando probabilidades próximas ao método de Wald original.

Diante desses resultados, para julgar a viabilidade de aplicar a generalização *bootstrap* no método de Wald ajustado, uma vez que tal generalização envolve um esforço computacional relativamente custoso, justifica-se realizar um estudo de precisão. Procedeu-se, assim, a uma análise que contemplasse medidas das amplitudes médias intervalares (Tabela 3), considerando apenas os resultados mais expressivos discutidos anteriormente.

Tabela 3 – Comprimento médio dos métodos de Wald e $Wald_{(ajust.)}$ na versão *bootstrap*, fixado p=0,2 e funções lineares c_4, c_5 e c_6.

Métodos	p=0,2		
	c_4	c_5	c_6
Wald	1,546	2,468	3,475
$Wald_{(ajust.)}$(Boot)	1,631	2,693	3,276

c_4=(3,1-,1,-1); c_5=(5,1,1,-1,-1,-1); c_6=(7,1,1,1,-1,-1,-1,-1).

De modo geral, os resultados encontrados (Tabela 3) revelaram que os métodos de Wald e $Wald_{(ajust.)}$(Boot) mostraram praticamente a mesma precisão, com leves oscilações ocasionadas, supostamente devido ao erro Monte Carlo. Portanto, no que tange à precisão dos métodos, a generalização *bootstrap* não apresentou resultados que possam justificar seu uso.

Dadas as configurações entre os tamanhos amostrais e os valores paramétricos e a formalização dos contrastes avaliados nesse trabalho, concluiu-se que a generalização do método *bootstrap* infinito apresentou desempenho similar ao dos métodos convencionais. A exceção foi apenas para o método de $Wald_{(ajust.)}$, cujos resultados referentes à probabilidade de cobertura indicaram que a generalização *bootstrap* foi compensável, no sentido de justificar seu esforço computacional, nas situações (p=0,2), em que os coeficientes da função linear foram definidos de forma a comparar uma proporção versus as demais.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRESTI, A.; CAFFO, B. Simple and effective confidence intervals for proportions and differences of proportions result from adding two failures. **American Statistician**, Alexandria, v. 54, n. 4, p. 280-288, 2000.

AGRESTI, A.; COULL, B. A. Aproximate is better than “exact” for interval estimation of binomial proportions. **American Statistician**, Alexandria, v. 52, p. 119-126, 1998.

BONETT, D. G.; WOODWARD, J. A. Application of the kronecker product and Wald test in log-linear models. **Comput Statistic Quarterly**, [S.l.], v. 3, p. 235-243, 1987.

CARARI, M. L. **Intervalo de confiança para a diferença de duas proporções binomiais utilizando bootstrap infinito**. 2004. 66 f. Dissertação (Mestrado em Agronomia) - Universidade Federal de Lavras, Lavras, 2004.

COHEN, L. A.; KENDALL, M. E.; ZANG, E.; MESCHTER, C.; ROSE, D. P. Modulation of N-nitrosomethylurea-induced mammary tumor promotion by dietary fiber and fat. **Journal Natural Cancer Institute**, [S.l.], v. 83, p. 496-501, 1991.

CONLON, M.; THOMAS, R. G. A new confidence interval for the difference of two binomial proportions. **Computational Statistics & Data Analysis**, Amsterdam, v. 9, n. 2, p. 237-241, 1990.

GREENLAND, S. Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. **American Statistician**, Washington, v. 55, p. 172, 2001.

PRICE, M. R.; BONETT, D. G. An improved confidence interval for the difference of two binomial proportions. **Computational Statistics & Data Analysis**, Amsterdam, v. 45, p. 449-456, 2004.