DigitalCommons @WayneState

Journal of Modern Applied Statistical Methods

Volume 8 | Issue 1 Article 18

5-1-2009

Improved Confidence Intervals for the Difference between Two Proportions

James F. Reed III
Christiana Care Hospital System, Newark, Delaware, JaReed@ChristianaCare.org

Journal of Modern Applied Statistical Methods
May, 2009, Vol. 8, No. 1, 208-214

1538 – 9472/09/$95.00

Improved Confidence Intervals for the Difference between Two Proportions

James F. Reed III
Christiana Care Hospital System, Newark, Delaware

Wald-z asymptotic methods, with and without a continuity correction, have less than nominal coverage probability characteristics but continue to be used. Newcombe's hybrid method and the Agresti-Caffo methods have coverage probabilities that are near nominal for either equal or unequal samples. Newcombe's hybrid and Agresti-Caffo methods demonstrate superior coverage properties.

Key words: Wald-z asymptotic, Newcombe's hybrid, Agresti-Caffo.

Introduction

In reporting the results of medical studies the problem of comparing two binomial success probabilities p_1 and p_2, $p_1 > 0$ and $p_2 > 0$ is often encountered. Implicit in this comparison are the independent observations $X_1 \sim B(n_1, p_1)$ and $X_2 \sim B(n_2, p_2)$. The most common comparison is the hypothesis H_o: $p_1 = p_2$ versus H_a: $p_1 \neq p_2$. Accompanying the hypothesis test is the construction of a confidence interval for the difference between p_1 and p_2. Nearly all introductory statistics textbooks include a method for computing this confidence interval and issue a warning - usually in a footnote - when not to use the common method: this commonly described method is the Wald-z method. Occasionally, a continuity corrected version is given (Wald-c).

The problems associated with the confidence interval for the difference between two independent proportions are similar to the confidence interval of a single proportion. Despite these properties, the Wald-z and Wald-c methods continue to dominate. We review the coverage probability functions of the Wald methods and a set of alternative methods for computing a confidence interval for the difference between two independent proportions.

James F. Reed III, PhD, is a Senior Biostatistician. Email him at: JaReed@ChristianaCare.org.

Methodology

The Wald-z and Wald-c confidence interval lower upper bounds for the difference between two independent proportions are defined as (See Appendix A for a typical data structure):

Wald-z:

$$LB = (p_1 - p_2) - z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$$
$$UB = (p_1 - p_2) + z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$$

Wald-c:

$$LB=(p_1-p_2)-[z_{\alpha/2}\sqrt{\{ac/m^3+bd/n^3\}}+(1/m+1/n)/2]$$
$$UB=(p_1-p_2)+[z_{\alpha/2}\sqrt{\{ac/m^3+bd/n^3\}}+(1/m+1/n)/2]$$

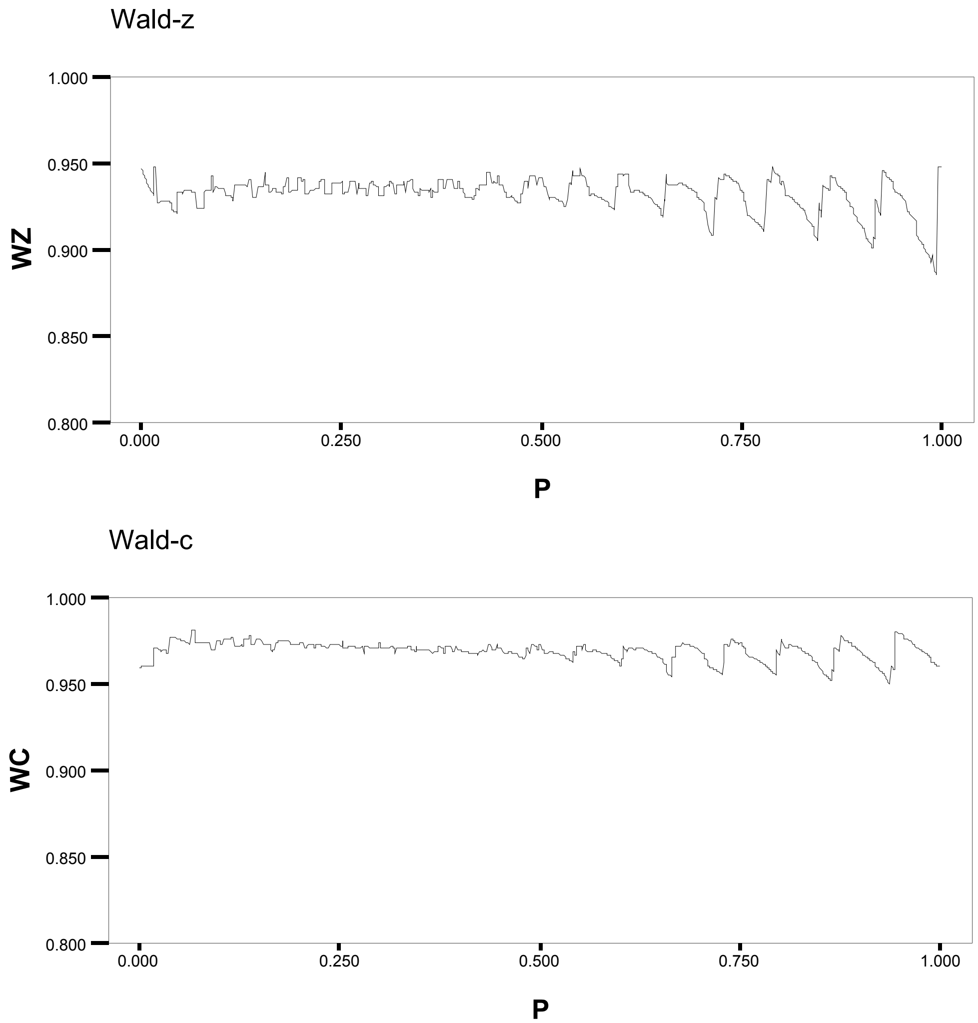
The primary criteria for evaluating a confidence interval method is the coverage probability function. This coverage probability for the difference between two independent proportions, $C(\pi_1,\pi_2|n_1,n_2,\alpha)$, is found by fixing n_1, n_2, π_1, and π_2, then computing the confidence interval for each $x_i = 0, \ldots, n_i$ for $i= 1, 2$. The coverage probability is then defined by:

$$C(\pi_1,\pi_2|n_1,n_2,\alpha) = \Sigma Pr(X_1 = x_1|n_1,\pi_1)Pr(X_2 = x_2|n_2,\pi_2) \delta(\pi_1,\pi_2|x_1,x_2,n_1,n_2,\alpha).$$

If $(\pi_1-\pi_2)\in[LB(x_1,x_2,n_1,n_2,\alpha), UB(x_1,x_2,n_1,n_2,\alpha)]$, $\delta(\pi_1,\pi_2|x_1,x_2,n_1,n_2,\alpha) = 1$, and 0 otherwise.

Figure 1 shows the 95% confidence interval coverage probability function for the Wald-z and Wald-c methods as a function of π_1, $\pi_1 \in [0,1]$ for $n_1 = n_2 = 20$ and $p_2 = 0.3$. The sawtooth appearance of the coverage functions

Figure 1: Coverage probabilities for nominal 95% Wald-z and Wald-c as a function of p_1 when p_2=0.3 with n_1=n_2=20

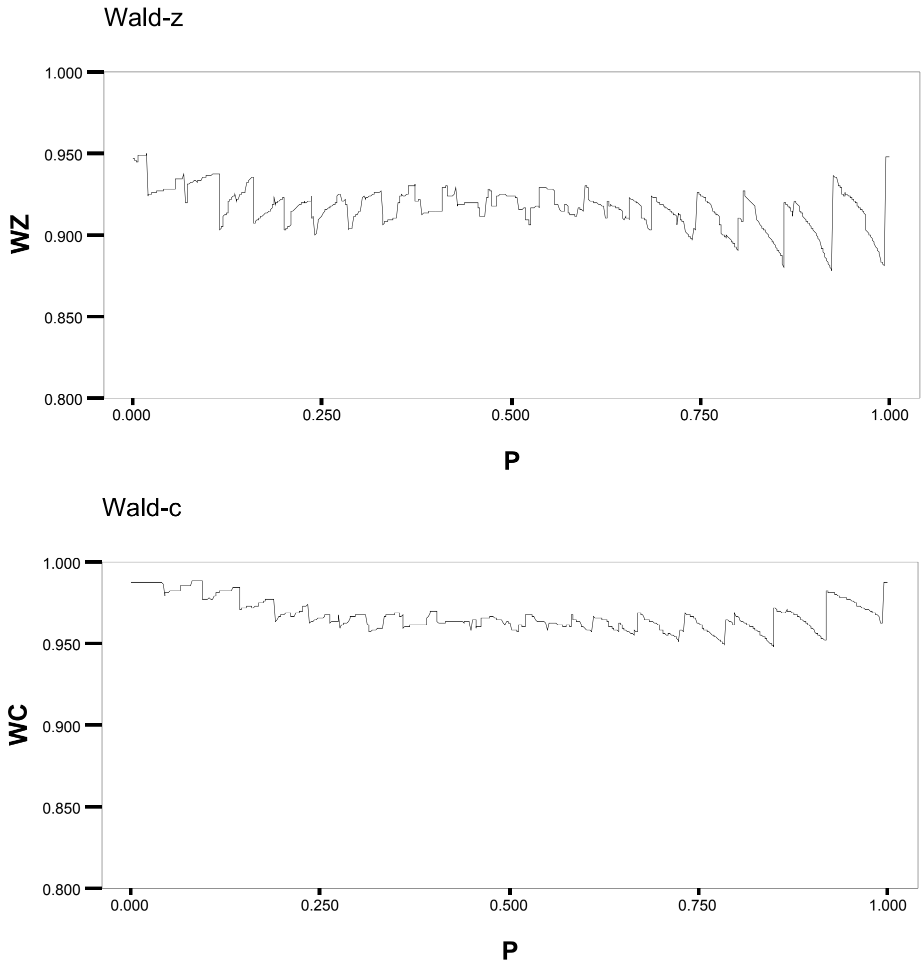


is due to the discontinuities for values of p_1 corresponding to any lower or upper limits in the set of confidence intervals. Like its one sample cousin, the Wald-z coverage probability curve is subnominal and less than 0.95 overall. The Wald-c coverage probability always exceeds 0.95 overall with interval widths larger than Wald-z.

Figure 2 shows the 95% confidence interval coverage probability function for the Wald-z and Wald-c methods as a function of π_1, $\pi_1 \in [0,1]$ for $n_1 = 20$, $n_2 = 10$ and $p_2 = 0.3$. The Wald-z coverage probability curve is subnominal for differences in proportions near 0 and 1 and less than 0.95 overall.

Beal evaluated several asymptotic methods for computing a confidence interval between the differences of two independent proportions. All involved identifying the interval within which $(\theta - \theta')^2 \leq z^2 V(\psi, \theta')$, where $\theta' = p_1 - p_2$, and $V(\psi, \theta') = u\{4\psi(1 - \psi)\theta = \pi_1(1 - \pi_1)/m + \pi_2(1 - \pi_2)/n$ (Beal, 1987). Beal examined two methods, labeled the Haldane (H) and Jeffreys-Perks (JP) methods. The JP method provides non-degenerative confidence intervals for all values of p_1 and p_2 unlike Wald-z or Wald-c. H and JP generally performed better than the Wald-z and Wald-c and of the two, JP was preferred (Beal, 1987; Radhakrishna, et. al., 1992).

Figure 2: Coverage probabilities for nominal 95% Wald-z and Wald-c as a function of p_1 when p_2=0.3 with n_1=20, n_2=10



The Haldane and Jeffreys-Perks lower and upper limits are defined by:

H

$$LB=\theta^* - w,$$

and

$$UB=\theta^* + w,$$

where

$$\theta^*=(\theta'+z^2v(1-2\psi'))/(1+z^2u),$$

$$w=[z/(1+z^2u)]\sqrt{[u\{4\psi'(1-\psi')-\theta'^2\}+2v(1-2\psi')\theta'+4z^2u^2(1-\psi')\psi'+z^2v^2(1-2\psi')^2]}$$

$$\psi'=(a/m+b/n)/2,$$

$$u=(1/m+1/n)/4,$$

and

$$v=(1/m-1/n)/4.$$

JP

$$LB=\theta^* - w,$$

and

$$UB=\theta^* + w,$$

where ψ' from the Haldane method is:

$$\psi'=[(a+0.5)/(m+1)+(b+0.5)/(n+1)]/2.$$

Newcombe (1998) compared eleven methods for estimating the difference between independent proportion. Similar to the single proportion, the virtues of Wald-z and Wald-c methods are in their simplicity, but overshoot and inappropriate intervals are still common. The Haldane and Jeffreys-Perks methods attempt to overcome the overshoot and inappropriate intervals while maintaining closed-form tractability. Newcombe concluded that both H and JP were improvements over the Wald-z and Wald-c methods, but both were still inadequate. Newcombe recommended a hybrid method based on Wilson's score method for a single proportion without continuity correction (NS). The LB and UB for the NS method are:

NS

$$LB=(p_1-p_2)-\delta,$$

where

$$\delta=\sqrt{(a/m-l_1)^2+(u_2-b/n)^2}$$
$$=z_{\alpha/2}\sqrt{l_1(1-l_1)/m+u_2(1-u_2)/n}.$$

$$UB = (p_1 - p_2) + \varepsilon,$$

where

$$\varepsilon=\sqrt{(u_1-a/m)^2+(b/n-l_2)^2}$$
$$=z_{\alpha/2}\sqrt{u_1(1-u_1)/m+l_2(1-l_2)/n},$$

and l_1, l_2, u_1, u_2 are the lower and upper bounds for the two proportions p_1 and p_2 using Wilson's score method.

Agresti & Coull's (1998) adjustment to the Wald method for a single proportion adds t/2 successes and t/2 failures. Agresti & Caffo (2000) later suggested that by adding two successes and two failures (total) to the two-sample method would improve the simple Wald method. This is an adjustment that adds a pseudo observation of each type to each sample. For instance, for sample *i*, $p_i = (r_i+1)/(n_i+2)$.

Results

Figure 3 shows the 95% confidence interval coverage probability function for the Newcombe NS, Haldane, Jeffreys-Perks, and Agresti-Caffo methods as a function of π_1, $\pi_1 \in [0,1]$ for $n_1 = n_2 = 20$ and $p_2 = 0.3$. The NS and Agresti-Caffo methods demonstrate coverage probabilities that are near nominal over $\pi_1 \in [0, 1]$.

Figure 4 shows the 95% confidence interval coverage probability function for the Newcombe NS, Haldane, Jeffreys-Perks, and Agresti-Caffo methods as a function of π_1, $\pi_1 \in [0,1]$ for $n_1 = 20$, $n_2 = 10$ and $p_2 = 0.3$. In the unequal sample size situation, Newcombe NS and Agresti-Caffo coverage probability functions are near nominal over $\pi_1 \in [0, 1]$.

Conclusion

In the case of differences between two independent proportions the Wald-z confidence interval behaves poorly with coverage probabilities below nominal values. Considering the coverage probability criterion, two alternative methods demonstrate superior coverage properties and both are easily programmable. Based on these results, the recommendation is to use either the NS or the Agresti-Caffo methods.

References

Agresti, A., & Coull, B.A. (1998). Approximate is better than 'exact' for interval estimation of binomial proportions. *The American Statistician, 52*, 119-126.

Agresti, A., & Caffo, B. (2000). Simple and effective confidence intervals for proportions and differences of proportions result from adding two successes and two failures. *The American Statistician, 54*, 280-288.

Beal, S. L. (1987). Asymptotic confidence intervals for the difference between two binomial parameters for use with small samples. *Biometrics, 43*, 941-950.

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: comparison of eleven methods. *Stat Med*, *17*, 873-890.

Radhakrishna, S., Murthy, B. N., Nair, N. G. K., Jayabal, P., & Jayasri, R. (1991). Confidence intervals in medical research. *Indian J Med Res [B], 96*, 199-205.

Figure 3: Coverage probabilities for nominal 95% Newcombe NS, Haldane, Jeffreys-Perks, and Agresti-Caffo as a function of p_1 when p_2=0.3 with n_1=n_2=20

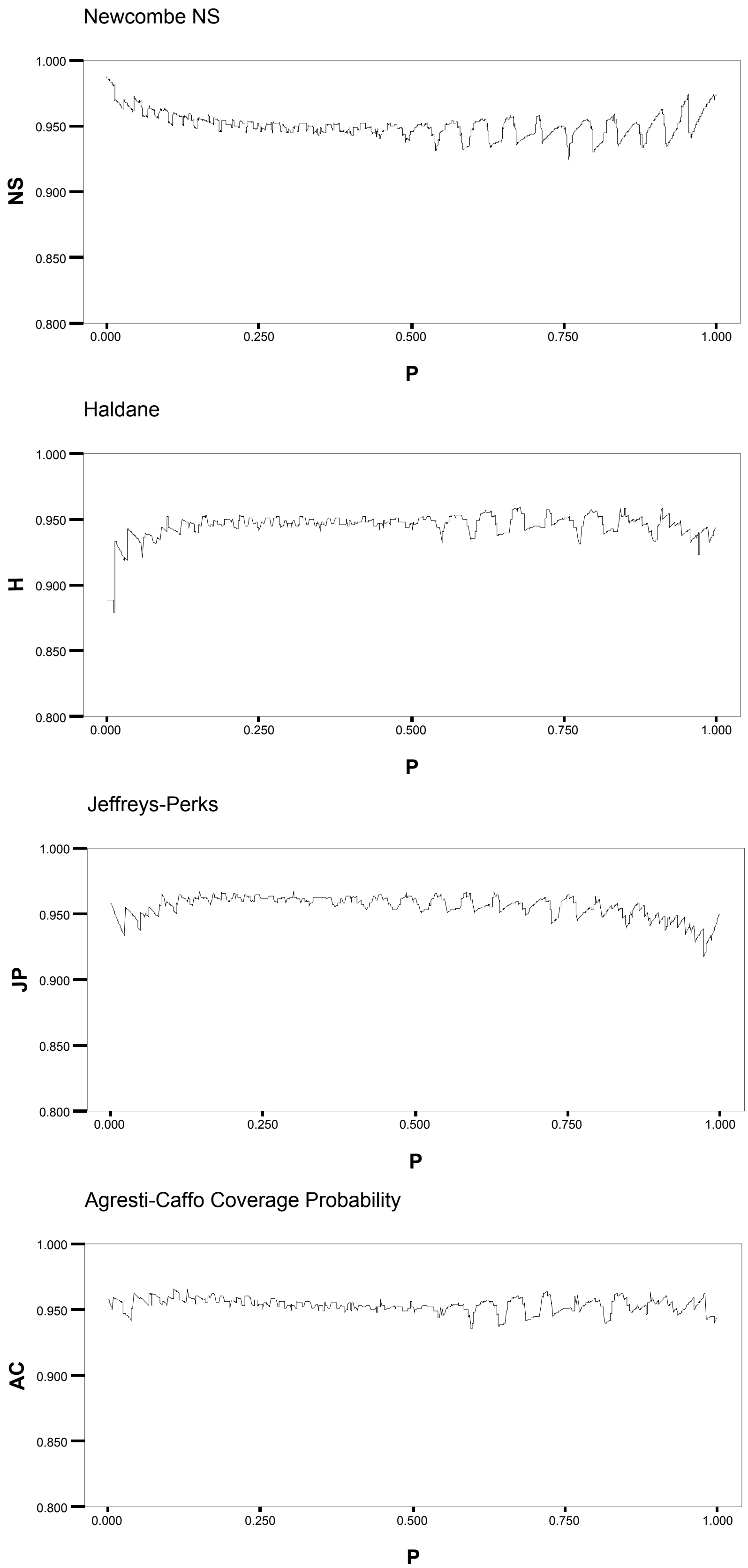
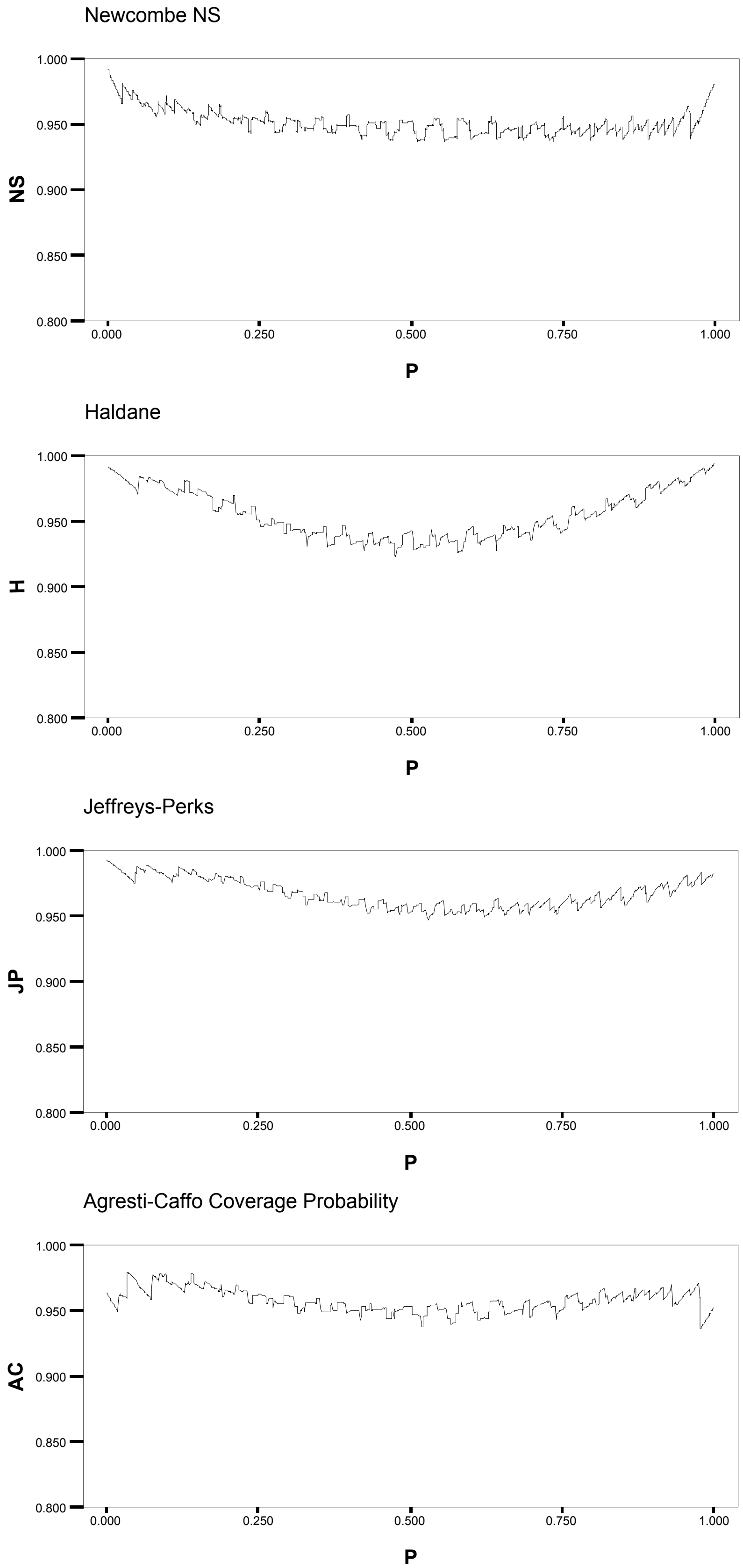


Figure 4: Coverage probabilities for nominal 95% Newcombe NS, Haldane, Jeffreys-Perks, and Agresti-Caffo as a function of p_1 when p_2=0.3 with n_1=20, n_2=10



Appendix A: Methods for calculation of confidence intervals for the difference between independent proportions

	Sample 1	Sample 2	
+	a	b	$p_1 = a/m$
−	c	d	$p_2 = b/n$
Total	m	m	$\theta = \pi_1 - \pi_2$ $\theta' = p_1 - p_2$

Method	Formula
Wald-z	$LB=(p_1 - p_2) - z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$ $UB=(p_1 - p_2) + z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$
Wald-c	$LB=(p_1 - p_2) - [z_{\alpha/2}\sqrt{\{ac/m^3 + bd/n^3\}} + (1/m + 1/n)/2]$ $UB=(p_1 - p_2) + [z_{\alpha/2}\sqrt{\{ac/m^3 + bd/n^3\}} + (1/m + 1/n)/2]$
Haldane-H	$LB=\theta^* - w$ $UB=\theta^*+w$, where $\theta^*=(\theta'+z^2v(1-2\psi'))/1+z^2u)$, $w=[z/(1+z^2u)]\sqrt{[u\{4\psi'(1-\psi')-\theta'^2\}+2v(1-2\psi')\theta'+4z^2u^2(1-\psi')\psi'+z^2v^2(1-2\psi')^2]}$ $\psi'=(a/m+b/n)/2$, $u=(1/m+1/n)/4$, and $v=(1/m - 1/n)/4$
Jeffreys-Perks-JP	$LB=\theta^* - w$ $UB=\theta^*+w$, where ψ' (from Haldane method) is: $\psi'=[(a+0.5)/(m+1)+(b+0.5)/(n+1)]/2$
Newcombe-NS	$LB=(p_1-p_2) - \delta$, where $\delta=\sqrt{\{(a/m-l_1)^2+(u_2-b/n)^2\}}=z_{\alpha/2}\sqrt{\{l_1(1-l_1)/m+u_2(1-u_2)/n\}}$ $UB=(p_1-p_2) + \varepsilon$, where $\varepsilon=\sqrt{\{(u_1-a/m)^2+(b/n-l_2)^2\}}=z_{\alpha/2}\sqrt{\{u_1(1-u_1)/m+l_2(1-l_2)/n\}}$ l_1, l_2, u_1, u_2 are the LB and UB for p_1 and p_2 using Wilson's score method
Agresti & Caffo	$LB = (p_1-p_2) - z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$ $UB = (p_1-p_2) + z_{\alpha/2}\sqrt{(ac/m^3 + bd/n^3)}$