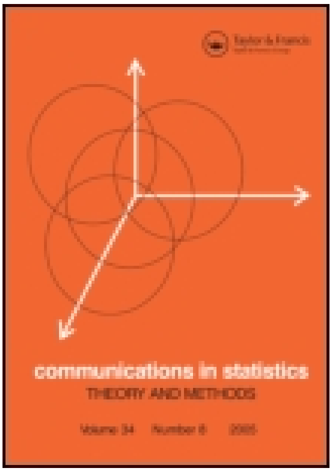

Communications in Statistics - Theory and Methods

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

Approximate and Fiducial Confidence Intervals for the Difference Between Two Binomial Proportions

K. Krishnamoorthy[a] & Dan Zhang[a]

[a] University of Louisiana at Lafayette, Lafayette, Louisiana, USA

Accepted author version posted online: 18 Jun 2013.

Communications in Statistics—Theory and Methods®, 44: 1745–1759, 2015

ISSN: 0361-0926 print / 1532-415X online
DOI: 10.1080/03610926.2013.765478

Approximate and Fiducial Confidence Intervals for the Difference Between Two Binomial Proportions

K. KRISHNAMOORTHY AND DAN ZHANG

University of Louisiana at Lafayette, Lafayette, Louisiana, USA

The problem of estimating the difference between two binomial proportions is considered. Closed-form approximate confidence intervals (CIs) and a fiducial CI for the difference between proportions are proposed. The approximate CIs are simple to compute, and they perform better than the classical Wald CI in terms of coverage probabilities and precision. Numerical studies indicate that these approximate CIs can be used safely for practical applications under a simple condition. The fiducial CI is more accurate than the approximate CIs in terms of coverage probabilities. The fiducial CIs, the Newcombe CIs, and the Miettinen–Nurminen CIs are comparable in terms of coverage probabilities and precision. The interval estimation procedures are illustrated using two examples.

Keywords Confidence limits; Exact coverage probability; Precision; Wilson's confidence interval.

Mathematics Subject Classification 62F25.

1. Introduction

There has been continuous interest in developing confidence intervals (CI) for the difference between two binomial proportions because of their common occurrence in medical and social sciences and clinical trials. The problem of testing the difference between two binomial proportions has been well addressed in the literature. Several tests, including exact methods, have been proposed in the literature for the testing problem; see Storer and Kim (1990) and the references therein. An exact test for discrete distributions means the use of exact probabilities in computing p-values associated with the test procedure, and it does not refer to the size of the test being exactly as the nominal level. It has been now well realized that the exact methods (e.g., Fisher's exact test) are too conservative, yielding tests that are less powerful than some approximate tests with satisfactory size properties. Furthermore, it is often difficult to invert exact tests (e.g., Suissa and Shuster, 1985) to obtain a CI for the difference between two proportions.

Comparison of two binomial probabilities can be made using the measure such as the relative risk, that is, ratio of the probabilities. Liu et al. (2006) noted that the choice

Received December 28, 2012; Accepted December 28, 2012.
Address correspondence to K. Krishnamoorthy, University of Louisiana at Lafayette, LA 70504, USA; E-mail: krishna@louisiana.edu

of relative risk or risk difference is somewhat arbitrary, but may provide quite different interpretations. Further, for rare adverse experiences, the ratio is not a meaningful measure if there are no events in a group. Regarding interval estimation of the difference between two binomial proportions, some exact methods are proposed in the literature; for example, see Peskun (1993) and Chan and Zhang (1999). As in the case of exact tests, these exact CIs are numerically involved and are not easy to calculate. Even though coverage probabilities of these exact CIs are at least the specified nominal level, their actual coverage properties were not studied because of computational complexity. In this article, we focus only on some conceptually simple approximate approaches to find CIs for the difference between two binomial proportions.

To describe the problem formally, let X_i be a binomial random variable with the number of trials n_i and the success probability p_i, $i = 1, 2$. We consider some approximate procedures to find CIs for $p_1 - p_2$ based on X_1 and X_2. Newcombe (1998) has compared 11 approximate CIs, including profile likelihood CI and the CI based on constrained maximum likelihood estimates (MLEs) (Miettinen and Nurminen, 1985), and concluded that his new CI based on Wilson's CIs for individual proportion is simple to use and comparable to the other CIs that are numerically complex. Specifically, the Newcombe CI for $p_1 - p_2$ can be calculated using the Wilson score CIs for p_1 and p_2. Recently, Li et al. (2011) proposed CIs for the difference between two Poisson means based on individual confidence limits. The article by Zou and Donner (2008) gives a justification of such CIs based on individual confidence limits.

In this article, we propose some simple CIs for the difference between two binomial probabilities that are easy to compute. The procedures are developed along the lines of methods for Poisson distributions in Krishnamoorthy and Lee (2012). Recently, Krishnamoorthy and Lee (2010) proposed a fiducial approach for finding CIs for a function of several binomial proportions. The fiducial approach is more general, and is useful to find CIs for any real-valued function of several binomial probabilities. So it is of interest to see the performance of the fiducial CIs for the difference between two binomial proportions. We also propose CIs based on the Wald statistic with the moment variance estimate. This approach is similar to the one by Miettinen and Nurminen (1985), who have used variance estimate based on the constrained MLEs. We also outline the Newcombe CI based on individual score CIs for p_1 and p_2. In Sec. 3, we evaluate exact coverage probabilities and expected widths of all CIs, and compare them. Based on our comparison studies, some recommendations are made as to the choice of CIs for practical applications. The interval estimation procedures are illustrated using two examples in Sec. 4, and some concluding remarks are given in Sec. 5.

2. Confidence Intervals

Let $X_1 \sim \text{binomial}(n_1, p_1)$ independently of $X_2 \sim \text{binomial}(n_2, p_2)$. Let $\widehat{p}_i = \frac{X_i}{n_i}$, $i = 1, 2$. In the following, we shall describe some methods of finding CIs for $p_1 - p_2$ based on X_1 and X_2.

2.1. Confidence Intervals Based on the Constrained MLEs

The well-known Wald statistic for testing $H_0 : p_1 - p_2 = d$ is given by

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2 - d}{\sqrt{\frac{\widehat{p}_1\widehat{q}_1}{n_1} + \frac{\widehat{p}_2\widehat{q}_2}{n_2}}}, \tag{1}$$

where $\widehat{q}_i = 1 - \widehat{p}_i$, $i = 1, 2$. The Wald CI is based on the asymptotic normality of Z, and is given by

$$\widehat{p}_1 - \widehat{p}_2 \pm z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{p}_1(1-\widehat{p}_1)}{n_1} + \frac{\widehat{p}_2(1-\widehat{p}_2)}{n_2}}. \tag{2}$$

Note that the above CI is formed by the roots of the quadratic equation $Z^2 = z^2_{1-\frac{\alpha}{2}}$ in d.

Instead of using the usual estimate of variance of $(\widehat{p}_1 - \widehat{p}_2)$ in (2), Miettinen and Nurminen (1985) have used the variance estimate based on the maximum likelihood estimate under the constraint that $p_1 = p_2 + d$. Specifically, they proposed the following test statistic:

$$T_M = \frac{\widehat{p}_1 - \widehat{p}_2 - d}{[(\widetilde{p}_2 + d)(1 - \widetilde{p}_2 - d)/n_1 + \widetilde{p}_2(1 - \widetilde{p}_2)/n_2 R_n]^{\frac{1}{2}}}, \tag{3}$$

where $R_n = (n_1 + n_2)/(n_1 + n_2 - 1)$, and $\widetilde{p}_2$ is the maximum likelihood estimator under the constraint that $p_1 - p_2 = d$. Even though the constrained likelihood equation (a function of $\widetilde{p}_2$) is a polynomial of order three, the likelihood equation has a unique closed-form solution; see Appendix I of Miettinen and Nurminen (1985). Since the constrained MLE is also a function of d, an approximate CI for $p_1 - p_2$ can be obtained by solving the equation $|T_M| = z_{1-\alpha/2}$ for d numerically. R package "gsDesign" can be used to compute the Miettinen–Nurminen (MN) CI for $p_1 - p_2$.

2.2. *Confidence Intervals Based on Constrained Moment Estimates*

The variance estimate based on the constrained MLE leads to a numerically complex procedure for finding CIs for $p_1 - p_2$. As noted by Krishnamoorthy and Thomson (2004) for the Poisson case, the test statistic based on constrained moment estimates (for method of moment estimates, see Section 7.2.1 of Casella and Berger, 2001) may lead to a simple interval estimation procedure. The constrained moment estimates can be obtained as follows. Let $a_1 = n_1/(n_1 + n_2)$ and $a_2 = 1 - a_1$ so that $\widehat{p} = (X_1 + X_2)/(n_1 + n_2) = a_1\widehat{p}_1 + a_2\widehat{p}_2$. Note that $E(\widehat{p}) = a_1 p_1 + a_2 p_2$. The moment estimate of p_2, under $H_0 : p_1 - p_2 = d$, is obtained by setting $\widehat{p} = a_1(d + p_2) + a_2 p_2$ and solving the equation for p_2. This procedure yields the moment estimates of p_1 and p_2 as

$$\widehat{p}_{20} = \widehat{p} - a_1 d \ \text{ and } \ \widehat{p}_{10} = \widehat{p} + a_2 d, \tag{4}$$

respectively. A $1 - \alpha$ CI for $p_1 - p_2$ is formed by the roots of the equation

$$\left| \frac{\widehat{p}_1 - \widehat{p}_2 - d}{\sqrt{\frac{\widehat{p}_{10}\widehat{q}_{10}}{n_1} + \frac{\widehat{p}_{20}\widehat{q}_{20}}{n_2}}} \right| = z_{1-\alpha/2},$$

where $\widehat{q}_{i0} = 1 - \widehat{p}_{i0}$, $i = 1, 2$. The above equation is quadratic in d, specifically, $ad^2 + bd + c = 0$ with

$$a = 1 + z^*_\alpha \left(\frac{1}{n_1} + \frac{1}{n_2} - \frac{3}{n_1 + n_2} \right), \ b = -\left(z^*_\alpha (\widehat{q} - \widehat{p}) \left(\frac{1}{n_1} - \frac{1}{n_2} \right) + 2(\widehat{p}_1 - \widehat{p}_2) \right),$$

and $c = -z_\alpha^* \widehat{p}\widehat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right) + (\widehat{p}_1 - \widehat{p}_2)^2$, where $\widehat{q} = 1 - \widehat{p}$ and $z_\alpha^* = z_{1-\alpha/2}^2$. The CI is given by

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{5}$$

For the case of $n_1 = n_2 = n$, the above CI can be expressed as

$$\frac{\widehat{p}_1 - \widehat{p}_2}{1 + .5z_\alpha^*/n} \pm \frac{\sqrt{\frac{z_\alpha^*}{n}\left[\widehat{p}\widehat{q}\left(\frac{z_\alpha^*}{n} + 2\right) - .5\left(\widehat{p}_1 - \widehat{p}_2\right)^2\right]}}{1 + .5z_\alpha^*/n}.$$

We shall refer to this CI as the moment-based CI or simply moment CI.

2.3. Fiducial Confidence Intervals

The fiducial CIs are based on the fiducial quantities of individual parameters given in Krishnamoorthy and Lee (2010). To describe this procedure, let $X \sim$ binomial(n, p) and let $B_{a,b}$ denote the beta random variable with shape parameters a and b. It is well known that, for an observed value k of X, $P(X \geq k|n, p) = P\left(B_{k,n-k+1} \leq p\right)$ and $P(X \leq k|n, p) = P\left(B_{k+1,n-k} \geq p\right)$. On the basis of this relation, we see that there is a pair of fiducial distributions for p, namely, $B_{k,n-k+1}$ for setting lower limit for p and $B_{k+1,n-k}$ for setting upper limit for p. Instead of having two fiducial variables, a random quantity that is "stochastically between" $B_{k,n-k+1}$ and $B_{k+1,n-k}$ can be used as a single approximate fiducial variable for p. On the basis of Cai's (2005) result, a simple choice is $B_{k+.5,n-k+.5}$. Hypothesis test or CI for p can be obtained from the distribution of $B_{k+.5,n-k+.5}$. For example, the $\alpha/2$ quantile and the $1-\alpha/2$ quantile of $B_{k+.5,n-k+.5}$ form a $1-\alpha$ CI for p.

To develop a fiducial quantity for $p_1 - p_2$, let (k_1, k_2) be an observed value of (X_1, X_2). The fiducial quantity for p_i is given by $B_{k_i+.5,n_i-k_i+.5}$, $i = 1, 2$. The fiducial quantity for the difference $\delta = p_1 - p_2$ is given by $Q_\delta = B_{k_1+.5,n_1-k_1+.5} - B_{k_2+.5,n_2-k_2+.5}$, and appropriate percentiles form a CI for $p_1 - p_2$. Specifically, the $1-\alpha$ fiducial CI is given by

$$\left(Q_{\delta;\frac{\alpha}{2}}, Q_{\delta;1-\frac{\alpha}{2}}\right), \tag{6}$$

where $Q_{\delta;\alpha}$ is the α quantile of Q_δ. Note that, for a given (k_1, k_2), the distribution of Q_δ does not depend on any unknown parameters, and the percentiles of Q_δ can be estimated by Monte Carlo simulation as described in the following algorithm.

Algorithm 1

For a given (k_1, n_1, k_2, n_2) and a confidence level $1-\alpha$:

(1) Generate a $B_{k_1+.5,n_1-k_1+.5}$ variate, and a $B_{k_2+.5,n_2-k_2+.5}$ variate.
(2) Set $Q_\delta = B_{k_1+.5,n_1-k_1+.5} - B_{k_2+.5,n_2-k_2+.5}$.
(3) Repeat steps 1 and 2 for a large number of times, say, 10,000.

The $100\alpha/2$ percentile, and the $100(1-\alpha/2)$ percentile of 10,000 Q_δ's generated above is a $1-\alpha$ fiducial CI for $p_1 - p_2$.

The percentiles of Q_δ can also be obtained numerically. Note that the cumulative distribution of Q_δ can be expressed as

$$P(Q_\delta \leq t) = \frac{1}{\text{beta}(k_2 + .5, n_2 - k_2 + .5)} \int_0^1 F(x + t; k_1 + .5, n_1$$

$$-k_1 + .5)x^{k_2-.5}(1-x)^{n_2-k_2-.5}dx, \quad (7)$$

where $F(x; a, b)$ denotes the beta distribution function with shape parameters a and b. The α quantile of Q_δ is the value of t for which $P(Q_\delta \le t) = \alpha$, and the value of t can be obtained numerically.

An approximation to Q_δ can be obtained using the standard result that for large a and b,

$$\frac{B_{a,b} - E(B_{a,b})}{\sqrt{\text{var}(B_{a,b})}} \sim N(0, 1), \quad \text{approximately.}$$

Since the beta random variables in Q_δ are independent, the distribution of Q_δ is approximately normal with mean $\widehat{\mu} = E(Q_\delta)$ and variance $\widehat{\sigma}^2 = \text{var}(Q_\delta)$, which are given by

$$\widehat{\mu} = \frac{k_1 + .5}{n_1 + 1} - \frac{k_2 + .5}{n_2 + 1} \text{ and } \widehat{\sigma}^2 = \frac{(k_1 + .5)(n_1 - k_1 + .5)}{(n_1 + 1)^2(n_1 + 2)} + \frac{(k_2 + .5)(n_2 - k_2 + .5)}{(n_2 + 1)^2(n_2 + 2)}.$$

Thus, the approximate fiducial CI can be expressed as

$$\left(Q_{\delta;\frac{\alpha}{2}}, Q_{\delta;\frac{1-\alpha}{2}}\right) \simeq \widehat{\mu} \pm z_{1-\frac{\alpha}{2}}\widehat{\sigma}. \quad (8)$$

It is easy to see that, $\widehat{\mu} \simeq \widehat{p}_1 - \widehat{p}_2$ and $\widehat{\sigma}^2 \simeq \widehat{p}_1\widehat{q}_1/n_1 + \widehat{p}_2\widehat{q}_2/n_2$, for large sample sizes. So the Wald CI and the above approximate CI are asymptotically similar, but their finite sample properties are quite different as will be seen in the sequel.

2.4. The Newcombe Confidence Interval

The Newcombe CI for the difference $\delta = p_1 - p_2$ is calculated using individual $1-\alpha$ CIs for p_1 and p_2. Newcombe (1998) proposed the one based on the Wilson score intervals for individual proportions, and is given by

$$(l_i, u_i) = \left(\frac{\hat{p}_i + \frac{z^2_{\alpha/2}}{2n}}{1 + \frac{z^2_{\alpha/2}}{n_i}}\right) \pm \frac{\frac{z_{\alpha/2}}{\sqrt{n_i}}\sqrt{\hat{p}_i(1-\hat{p}_i) + z^2_{\alpha/2}/(4n_i)}}{1 + \frac{z^2_{\alpha/2}}{n_i}}, \ i = 1, 2, \quad (9)$$

where z_α is the α quantile of the standard normal distribution. Newcombe's CI for $p_1 - p_2$ is based on (l_i, u_i)'s, and is given by (L, U), with

$$(L, U) = \left(\widehat{\delta} - z_{1-\frac{\alpha}{2}}\sqrt{\frac{l_1(1-l_1)}{n_1} + \frac{u_2(1-u_2)}{n_2}},\ \widehat{\delta} + z_{1-\frac{\alpha}{2}}\sqrt{\frac{u_1(1-u_1)}{n_1} + \frac{l_2(1-l_2)}{n_2}}\right), \quad (10)$$

where $\widehat{\delta} = \widehat{p}_1 - \widehat{p}_2$. In the above formula, l_i should be taken as 0 if $l_i < 0$ and u_i should be taken as 1 if $u_i > 1$.

3. Exact Coverage Probabilities and Expected Widths

Exact coverage probabilities or expected widths can be evaluated using the binomial probabilities for each pair of observed samples. Specifically, for a given (n_1, p_1, n_2, p_2), the exact coverage probability of a $1-\alpha$ CI for $(L(X_1, X_2), U(X_1, X_2))$ is given by

$$\sum_{x_1=0}^{n_1}\sum_{x_2=0}^{n_2}\binom{n_1}{x_1}p_1^{x_1}(1-p_1)^{n_1-x_1}\binom{n_2}{x_2}p_2^{x_2}(1-p_2)^{n_2-x_2}I_{[(L(x_1,x_2),U(x_1,x_2))]}(p_1-p_2), \quad (11)$$

where $I_{[A]}(x)$ is the indicator function. For an accurate CI, the above coverage probability should be close to the nominal level $1-\alpha$. Exact expected width of a CI can be calculated using (11) with the indicator function replaced by the width $U(x_1, x_2) - L(x_1, x_2)$.

As the approximate fiducial CI in (8) is similar to the Wald CI, we first compare them with respect to coverage probabilities and expected widths. Toward this, we note that the Wald type CIs could be too anticonservative when the sample sizes are small and/or the parameter values are at the boundary. Therefore, we study their properties over the parameter space where $\min\{n_1p_1, n_1q_1, n_2p_2, n_2q_2\} \geq 2$. To obtain such a constrained set of parameter values, for a given (n_1, n_2), we generated 10,000 points (p_1, p_2) from uniform(0, 1) distributions, and computed the coverage probabilities and expected widths only for those p_1 and p_2 that satisfy the above inequality. Five-number statistics, along with the 5th and 95th percentiles, of the coverage probabilities and expected widths are given in Table 1 for some selected values of (n_1, n_2) that are not too small. We included the 5th and the 95th percentiles, to understand the performance of the CIs in most cases. We observe from Table 1 that the approximate fiducial CIs are superior to the Wald CIs in terms of coverage probabilities and expected widths; specifically, we note that the expected widths are shorter than those of the Wald CIs with better coverage probabilities. We also observe that the Wald CI could be anticonservative even for large samples; see the results for $(n_1, n_2) = (30, 30)$ and (50, 100) in Table 1. Thus, the approximate fiducial CI is certainly preferable to the Wald CI with respect to coverage properties and precision.

We next compare the approximate fiducial CIs, moment-based CIs, the Newcombe CIs, and the MN CIs for some small, moderate, and large sample sizes. We once again consider the constrained parameter space, that is, $\{(p_1, p_2) : \min\{n_1p_1, n_1q_1, n_2p_2, n_2q_2\} \geq 2\}$. The reported summary statistics of the coverage probabilities and expected widths in Table 2 indicate that all four CIs are in general comparable, and the approximate fiducial CIs could be anticonservative for some parameter values when sample sizes are small. The expected widths of moment CIs are shorter than those of the MN CIs for almost all cases. In terms of simplicity and accuracy, the moment CI and the Newcombe CI have an edge over the MN CI.

To judge the performance of the CIs over the entire parameter space, we evaluated coverage probabilities and expected widths of the moment CIs, the Newcombe CIs and the MN CIs over the parameter space $\{(p_1, p_2) : 0 < p_1 < 1, 0 < p_2 < 1\}$. Toward this, we generated 10,000 points (p_1, p_2) from uniform(0, 1) distributions, and computed exact coverage probabilities and expected widths of all three CIs. The summary statistics of these quantities are given in Table 3 for some small to moderate sample sizes. The fiducial CIs are based on 10,000 simulation runs. We see from these summary statistics that the Fiducial, Newcombe and MN CIs are comparable with respect to coverage probabilities and expected widths for all sample sizes considered. The moment CI also performs satisfactorily for most cases except that it could be too liberal for extreme parameter values; see the minimum and the 5th percentile of coverage probabilities of the moment CIs in Table 3.

Finally, it may be of interest to examine the large sample properties of the closed-form CIs, namely, the Wald, approximate fiducial, Newcombe, and moment CIs. We evaluated the merits of these CIs for sample sizes (80, 80) and (80, 100) and presented the coverage probabilities and expected widths in Table 4. The table values clearly indicate that fiducial

Table 1
Summary statistics of coverage probabilities (CPs) and expected widths (EWs) of confidence intervals, based on points (p_1, p_2) generated from uniform(0, 1) distributions with the constraint that $\min\{n_1p_1, n_1q_1, n_2p_2, n_2q_2\} \geq 2$

		90% CIs Percentiles							95% CIs Percentiles						
Method		min	.05	.25	.5	.75	.95	max	min	.05	.25	.5	.75	.95	max
		$n_1 = 20, n_2 = 20$													
Wald	CP	.797	.858	.875	.882	.888	.896	.912	.862	.916	.929	.934	.939	.943	.949
	EW	0.30	0.37	0.42	0.45	0.48	0.50	0.51	0.36	0.44	0.50	0.54	0.57	0.60	0.60
Appr. Fid.	CP	.852	.872	.884	.890	.896	.907	.934	.916	.928	.937	.941	.945	.951	.974
	EW	0.31	0.37	0.41	0.44	0.46	0.48	0.48	0.37	0.44	0.49	0.52	0.55	0.57	0.58
		$n_1 = 20, n_2 = 30$													
Wald	CP	.827	.867	.878	.883	.887	.892	.905	.881	.919	.931	.935	.938	.941	.950
	EW	0.26	0.33	0.38	0.41	0.44	0.46	0.46	0.30	0.39	0.45	0.49	0.52	0.55	0.55
Appr. Fid.	CP	.869	.881	.886	.890	.894	.903	.927	.925	.935	.938	.941	.944	.951	.970
	EW	0.27	0.33	0.37	0.40	0.42	0.44	0.45	0.32	0.39	0.44	0.48	0.50	0.53	0.53
		$n_1 = 30, n_2 = 120$													
Wald	CP	.819	.869	.881	.884	.887	.890	.895	.864	.917	.931	.935	.938	.940	.942
	EW	0.15	0.21	0.25	0.29	0.31	0.33	0.33	0.17	0.24	0.30	0.35	0.37	0.39	0.39
Appr. Fid.	CP	.862	.885	.889	.891	.894	.904	.918	.909	.937	.939	.941	.943	.951	.962
	EW	0.16	0.21	0.25	0.29	0.30	0.32	0.32	0.18	0.25	0.30	0.34	0.36	0.38	0.38
		$n_1 = 30, n_2 = 30$													
Wald	CP	.789	.869	.883	.888	.892	.900	.909	.855	.925	.936	.940	.942	.946	.950
	EW	0.21	0.28	0.33	0.36	0.39	0.41	0.42	0.26	0.33	0.40	0.43	0.47	0.49	0.50
Appr. Fid.	CP	.862	.881	.889	.893	.898	.907	.936	.926	.935	.941	.944	.947	.954	.971
	EW	0.22	0.28	0.33	0.36	0.38	0.40	0.40	0.27	0.34	0.39	0.43	0.46	0.48	0.48
		$n_1 = 50, n_2 = 100$													
Wald	CP	.839	.883	.890	.893	.894	.897	.904	.867	.934	.941	.943	.945	.946	.954
	EW	0.10	0.17	0.21	0.24	0.26	0.28	0.28	0.12	0.20	0.25	0.29	0.31	0.33	0.34
Appr. Fid.	CP	.881	.890	.894	.896	.899	.904	.931	.936	.942	.945	.946	.948	.952	.976
	EW	0.11	0.17	0.21	0.24	0.26	0.27	0.28	0.13	0.20	0.25	0.28	0.31	0.33	0.33

Table 2

Summary statistics of coverage probabilities (CPs) and expected widths (EWs) of confidence intervals, based on points (p_1, p_2) generated from uniform(0, 1) distributions with the constraint that $\min\{n_1p_1, n_1q_1, n_2p_2, n_2q_2\} \geq 2$; MN – Miettinen–Nurminen CIs

		90% CIs Percentiles							95% CIs Percentiles						
Method		min	.05	.25	.5	.75	.95	max	min	.05	.25	.5	.75	.95	max
		$n_1 = 12, n_2 = 15$													
Appr. Fid.	CP	.834	.867	.876	.882	.891	.901	.924	.911	.924	.931	.935	.940	.947	.963
	EW	0.43	0.47	0.51	0.53	0.55	0.57	0.57	0.52	0.57	0.61	0.64	0.66	0.68	0.68
Moment	CP	.865	.876	.886	.894	.900	.907	.919	.927	.936	.943	.947	.950	.955	.964
	EW	0.41	0.47	0.51	0.54	0.54	0.55	0.56	0.49	0.56	0.60	0.63	0.66	0.68	0.68
Newcombe	CP	.867	.877	.887	.895	.902	.911	.935	.927	.936	.942	.946	.951	.958	.974
	EW	0.44	0.49	0.53	0.55	0.57	0.59	0.59	0.52	0.57	0.60	0.62	0.64	0.65	0.65
MN	CP	.863	.880	.891	.898	.905	.913	.923	.930	.939	.946	.951	.955	.960	.973
	EW	0.44	0.48	0.51	0.53	0.55	0.56	0.56	0.52	0.59	0.63	0.65	0.67	0.69	0.69
		$n_1 = 15, n_2 = 20$													
Appr. Fid.	CP	.855	.875	.881	.886	.892	.903	.930	.923	.930	.935	.938	.942	.950	.969
	EW	0.36	0.41	0.45	0.47	0.49	0.51	0.52	0.42	0.48	0.53	0.56	0.59	0.61	0.62
Moment	CP	.871	.880	.889	.895	.900	.906	.920	.927	.938	.944	.947	.950	.954	.961
	EW	0.34	0.40	0.45	0.47	0.50	0.52	0.52	0.40	0.47	0.52	0.56	0.59	0.61	0.61
Newcombe	CP	.869	.884	.891	.896	.901	.911	.931	.927	.940	.944	.947	.951	.960	.975
	EW	0.37	0.42	0.45	0.47	0.49	0.51	0.51	0.43	0.49	0.53	0.55	0.57	0.59	0.59
MN	CP	.868	.884	.893	.899	.903	.910	.922	.931	.942	.947	.951	.954	.958	.971
	EW	0.36	0.42	0.46	0.48	0.51	0.52	0.53	0.43	0.51	0.55	0.57	0.60	0.62	0.62
		$n_1 = 20, n_2 = 20$													
Appr. Fid.	CP	.852	.872	.884	.890	.896	.907	.933	.916	.929	.937	.941	.945	.952	.974
	EW	0.32	0.37	0.41	0.44	0.46	0.48	0.48	0.37	0.44	0.49	0.52	0.55	0.57	0.58

Moment	CP	.858	.882	.890	.895	.900	.911	.932	.928	.939	.944	.947	.950	.955	.968
	EW	0.29	0.36	0.41	0.44	0.47	0.48	0.48	0.34	0.43	0.48	0.52	0.55	0.57	0.58
Newcombe	CP	.854	.880	.891	.897	.903	.912	.931	.913	.936	.945	.949	.953	.959	.975
	EW	0.32	0.38	0.41	0.44	0.46	0.47	0.48	0.38	0.45	0.49	0.51	0.53	0.55	0.56
MN	CP	.859	.885	.894	.898	.903	.914	.927	.928	.941	.947	.950	.953	.958	.977
	EW	0.31	0.38	0.42	0.45	0.48	0.49	0.50	0.38	0.46	0.50	0.53	0.56	0.58	0.59
		$n_1 = 20, n_2 = 50$													
Appr. Fid.	CP	.869	.880	.885	.888	.893	.903	.919	.911	.933	.936	.939	.943	.950	.961
	EW	0.24	0.29	0.34	0.37	0.39	0.41	0.41	0.29	0.35	0.40	0.44	0.46	0.48	0.49
Moment	CP	.879	.889	.895	.898	.901	.905	.916	.931	.945	.947	.949	.951	.954	.969
	EW	0.23	0.29	0.34	0.37	0.39	0.40	0.41	0.28	0.35	0.40	0.43	0.46	0.48	0.48
Newcombe	CP	.881	.891	.896	.898	.902	.908	.933	.935	.945	.947	.949	.952	.957	.976
	EW	0.25	0.30	0.34	0.36	0.39	0.40	0.43	0.30	0.36	0.40	0.43	0.45	0.47	0.47
MN	CP	.863	.888	.896	.899	.903	.906	.928	.939	.947	.949	.951	.953	.956	.970
	EW	0.24	0.30	0.34	0.37	0.39	0.41	0.42	0.29	0.37	0.41	0.43	0.46	0.48	0.49
		$n_1 = 30, n_2 = 30$													
Appr. Fid.	CP	.864	.881	.888	.893	.898	.907	.936	.926	.935	.941	.944	.947	.953	.977
	EW	0.22	0.29	0.33	0.36	0.38	0.40	0.40	0.26	0.34	0.39	0.42	0.45	0.48	0.48
Moment	CP	.865	.884	.892	.896	.901	.909	.929	.930	.941	.945	.948	.950	.954	.966
	EW	0.20	0.28	0.33	0.36	0.38	0.40	0.41	0.24	0.33	0.39	0.42	0.45	0.48	0.48
Newcombe	CP	.860	.884	.894	.899	.904	.911	.933	.922	.939	.947	.950	.953	.958	.985
	EW	0.23	0.29	0.33	0.36	0.38	0.40	0.40	0.28	0.35	0.40	0.42	0.45	0.46	0.47
MN	CP	.866	.886	.895	.899	.904	.911	.934	.933	.944	.948	.950	.952	.956	.977
	EW	0.22	0.29	0.34	0.36	0.39	0.41	0.41	0.26	0.35	0.40	0.43	0.46	0.48	0.49

Table 3
Summary statistics of coverage probabilities (CPs) and expected widths (EWs) of fiducial confidence intervals, based on 10,000 points (p_1, p_2) generated from uniform(0, 1) distributions; MN – Miettinen–Nurminen CIs

		90% CIs Percentiles							95% CIs Percentiles						
Method		min	.05	.25	.5	.75	.95	max	min	.05	.25	.5	.75	.95	max
		$n_1 = 12, n_2 = 15$													
Moment	CP	.417	.859	.881	.891	.900	.913	.961	.540	.905	.936	.944	.949	.957	.999
	EW	0.09	0.32	0.41	0.46	0.51	0.55	0.56	0.12	0.39	0.51	0.57	0.63	0.67	0.68
Fiducial	CP	.823	.869	.884	.895	.904	.924	.997	.881	.931	.940	.945	.951	.967	.999
	EW	0.23	0.35	0.43	0.48	0.53	0.56	0.57	0.31	0.42	0.51	0.57	0.61	0.64	0.65
Newcombe	CP	.821	.879	.892	.902	.915	.950	.999	.879	.935	.944	.950	.961	.982	.999
	EW	0.24	0.39	0.46	0.50	0.53	0.56	0.58	0.32	0.48	0.55	0.58	0.62	0.64	0.65
MN	CP	.826	.876	.892	.903	.913	.947	.999	.889	.939	.948	.954	.962	.979	.999
	EW	0.27	0.38	0.45	0.51	0.55	0.58	0.59	0.34	0.46	0.54	0.60	0.66	0.69	0.69
		$n_1 = 15, n_2 = 20$													
Moment	CP	.434	.865	.885	.893	.900	.910	.991	.563	.914	.940	.945	.950	.955	.999
	EW	0.11	0.30	0.39	0.44	0.48	0.52	0.52	0.14	0.35	0.46	0.52	0.57	0.61	0.61
Fiducial	CP	.803	.876	.889	.896	.903	.918	.997	.879	.937	.946	.951	.952	.958	.979
	EW	0.20	0.32	0.39	0.44	0.48	0.51	0.51	0.28	0.41	0.49	0.52	0.55	0.56	0.58
Newcombe	CP	.806	.885	.893	.900	.912	.941	.999	.881	.939	.945	.950	.958	.978	.999
	EW	0.20	0.34	0.41	0.44	0.48	0.50	0.51	0.28	0.41	0.49	0.52	0.56	0.58	0.59
MN	CP	.805	.880	.893	.900	.908	.937	.999	.905	.942	.948	.952	.958	.973	.999
	EW	0.22	0.33	0.40	0.45	0.49	0.52	0.53	0.28	0.40	0.48	0.53	0.58	0.61	0.62
		$n_1 = 20, n_2 = 25$													
Moment	CP	.577	.874	.888	.895	.900	.907	.991	.540	.925	.944	.947	.950	.955	.999
	EW	0.10	0.27	0.35	0.39	0.43	0.46	0.47	0.11	0.30	0.40	0.45	0.49	0.52	0.53

Fiducial	CP	.853	.880	.890	.895	.902	.916	.999		.894	.941	.945	.948	.954	.968	.972
	EW	0.15	0.26	0.35	0.39	0.43	0.46	0.46		0.22	0.35	0.42	0.45	0.51	0.53	0.53
Newcombe	CP	.849	.887	.894	.900	.909	.933	.999		.892	.941	.946	.950	.957	.974	.999
	EW	0.16	0.29	0.36	0.39	0.43	0.45	0.46		0.22	0.36	0.43	0.46	0.50	0.53	0.53
MN	CP	.854	.884	.893	.899	.905	.925	.999		.914	.942	.947	.951	.954	.968	.999
	EW	0.17	0.28	0.35	0.39	0.43	0.46	0.47		0.22	0.34	0.42	0.47	0.52	0.55	0.55
		$n_1 = 30, n_2 = 35$														
Moment	CP	.646	.881	.891	.897	.901	.906	.991		.714	.934	.944	.948	.950	.953	.997
	EW	0.07	0.22	0.29	0.33	0.36	0.39	0.39		0.10	0.26	0.35	0.39	0.43	0.46	0.47
Fiducial	CP	.871	.885	.892	.897	.901	.910	.988		.890	.943	.946	.947	.953	.956	.971
	EW	0.13	0.22	0.29	0.33	0.36	0.39	0.39		0.16	0.29	0.35	0.39	0.42	0.44	0.44
Newcombe	CP	.876	.889	.896	.901	.907	.926	.999		.893	.942	.947	.950	.955	.969	.999
	EW	0.12	0.23	0.30	0.33	0.36	0.38	0.39		0.16	0.29	0.35	0.39	0.42	0.45	0.45
MN	CP	.865	.887	.895	.900	.905	.913	.999		.910	.943	.948	.951	.953	.959	.999
	EW	0.11	0.23	0.29	0.33	0.37	0.39	0.40		0.15	0.28	0.35	0.39	0.43	0.46	0.47
		$n_1 = 30, n_2 = 50$														
Moment	CP	.575	.883	.894	.898	.900	.905	.968		.871	.938	.942	.944	.948	.958	.999
	EW	0.08	0.21	0.27	0.31	0.34	0.36	0.37		0.09	0.24	0.32	0.37	0.40	0.43	0.43
Fiducial	CP	.848	.887	.899	.902	.905	.913	.990		.880	.941	.947	.949	.951	.957	.971
	EW	0.11	0.21	0.27	0.31	0.34	0.37	0.37		0.15	0.26	0.33	0.36	0.40	0.42	0.42
Newcombe	CP	.852	.893	.897	.900	.905	.923	.999		.884	.945	.948	.950	.954	.967	.999
	EW	0.11	0.21	0.27	0.31	0.34	0.37	0.37		0.15	0.26	0.33	0.36	0.41	0.42	0.42
MN	CP	.865	.889	.897	.900	.903	.911	.999		.668	.933	.946	.948	.950	.953	.993
	EW	0.10	0.21	0.27	0.31	0.34	0.36	0.37		0.10	0.24	0.32	0.37	0.40	0.43	0.43

Table 4
Coverage probabilities (CPs) and expected widths (EWs) of CIs for large samples, based on 10,000 points (p_1, p_2) generated from uniform(0, 1) distributions

		90% CIs Percentiles							95% CIs Percentiles						
Method		min	.05	.25	.5	.75	.95	max	min	.05	.25	.5	.75	.95	max
		$n_1 = 80, n_2 = 80$													
Wald	CP	.840	.883	.892	.895	.897	.902	.909	.861	.937	.944	.946	.947	.949	.953
	EW	0.07	0.14	0.19	0.22	0.24	0.26	0.26	0.09	0.18	0.23	0.26	0.28	0.30	0.31
Appr. Fid.	CP	.881	.889	.895	.898	.901	.907	.961	.939	.943	.946	.948	.949	.954	.993
	EW	0.09	0.15	0.19	0.21	0.24	0.25	0.26	0.08	0.18	0.23	0.25	0.28	0.30	0.30
Newcombe	CP	.870	.891	.898	.900	.903	.910	.944	.931	.945	.949	.951	.952	.958	.990
	EW	0.09	0.15	0.19	0.21	0.24	0.25	0.25	0.10	0.18	0.23	0.25	0.28	0.30	0.30
Moment	CP	.883	.890	.896	.899	.901	.907	.950	.936	.944	.949	.950	.952	.958	.984
	EW	0.08	0.15	0.19	0.21	0.24	0.25	0.26	0.13	0.18	0.23	0.25	0.28	0.30	0.30
		$n_1 = 80, n_2 = 100$													
Wald	CP	.784	.887	.894	.895	.897	.899	.905	.828	.938	.944	.946	.947	.949	.952
	EW	0.06	0.13	0.18	0.20	0.23	0.24	0.25	0.06	0.16	0.21	0.24	0.27	0.29	0.29
Appr. Fid.	CP	.889	.895	.896	.898	.900	.906	.958	.936	.945	.947	.948	.949	.952	.984
	EW	0.08	0.14	0.18	0.20	0.22	0.24	0.24	0.08	0.17	0.21	0.24	0.27	0.28	0.29
Newcombe	CP	.885	.896	.898	.899	.901	.908	.967	.935	.948	.949	.950	.951	.956	.981
	EW	0.07	0.14	0.18	0.20	0.22	0.24	0.24	0.09	0.17	0.21	0.24	0.26	0.28	0.29
Moment	CP	.881	.894	.898	.899	.900	.903	.948	.943	.948	.949	.950	.951	.957	.969
	EW	0.07	0.14	0.18	0.20	0.22	0.24	0.24	0.11	0.18	0.22	0.24	0.27	0.28	0.29

approximate, Newcombe, and moment CIs perform very similar in terms of coverage properties and precision while the Wald CI is not quite accurate.

On an overall basis, our numerical studies indicate that the fiducial, Newcombe and MN CIs are satisfactory in terms of coverage probabilities and precision and they are preferable to others in practical applications.

4. Examples

Example 1. This example is adapted from Example 3.7.1 of Krishnamoorthy (2006), and is concerned if the long-term exposure to a chemical causes a particular disease. Data were obtained from independent samples from exposed and unexposed groups as shown in the following table:

Group	Symptoms present	Symptoms absent	Totals
Exposed	13	19	32
Unexposed	4	21	25
Totals	17	40	57

Let p_e and p_u denote the true proportions of adults with symptoms in the exposed and unexposed groups, respectively. We calculated one- and two-sided 95% CIs for $p_e - p_u$ and reported them in Table 5. The fiducial CIs are calculated using the integral equation (7), and the MN CI is calculated using R package "GsDesign."

It is clear from Table 5 that all CIs indicate that the proportion of adults with symptoms in the exposed group is significantly larger than the one in the unexposed group. We also note that the moment, fiducial, Newcombe, and the MN CIs are in close agreement while the approximate fiducial CIs are in the proximity of these CIs. The Wald CI stands out among all CIs. Among all CIs the fiducial CI is the shortest followed by the moment CI.

Example 2. The data for this example is taken from Example 22.4 of Zar (1996). In this example, it is of interest to find if the frequency of occurrence of a specific intestinal parasite is the same in two animal populations. In a random sample of 24 animals from population 1, 18 are infected by the parasite, and in a random sample of 25 animals from population 2, 10 are infected. We computed 95% one-sided as well as two-sided CIs, and reported them in Table 6.

Table 5
95% confidence intervals for the difference between proportions of people with symptoms in the exposed and unexposed groups

Method	CI	One-sided lower	One-sided upper
Wald	(.0235, .4690)	.0593	.4332
App. Fid.	(.0177, .4544)	.0528	.4193
Moment	(.0071, .4399)	.0468	.4130
Fiducial	(.0110, .4437)	.0489	.4130
Newcombe	(.0062, .4425)	.0459	.4146
MN	(.0051, .4554)	.0459	.4240

Table 6
95% confidence intervals for the difference between infection rates

Method	CI	One-sided lower	One-sided upper
Wald	(.0914, .6086)	.1330	.5670
App. Fid.	(.0858, .5865)	.1260	.5463
Moment	(.0737, .5745)	.1192	.5435
Fiducial	(.0754, .5738)	.1199	.5386
Newcombe	(.0731, .5608)	.1176	.5339
MN	(.0710, .5783)	.1171	.5465

All CIs indicate that the proportion of infected animals in population 1 significantly larger than the proportion in population 2. As in the previous example, there is a close agreement among the moment, fiducial, Newcombe, and MN CIs. The approximate fiducial CI is in the proximity of the Wald CI. Among all two-sided CIs, the Newcombe CI is the shortest, followed by the fiducial CI.

5. Conclusions

In this article, we considered some simple approximate CIs for the difference between two binomial proportions. The usual condition for practical application of the Wald CIs is that the expected counts in each cell should be at least five; however, even under this condition, the Wald CI could be anticonservative. The proposed approximate CIs work very satisfactorily even under somewhat relaxed condition that $\min\{n_1p_1, n_1q_1, n_2p_2, n_2q_2\} \geq 2$. Our exact coverage studies show that the usual Wald CIs are not satisfactory even for large samples. The approximate fiducial CI, which is as simple as the Wald CI, is accurate for moderate to large sample sizes. Following Miettinen and Nurminen (1985) and Krishnamoorthy and Thomson (2004), we proposed a CI using the variance estimate based on moment estimates. The moment CIs and other approximate CIs are conceptually simple and they can be calculated in the absence of a computer. The fiducial approach is more general, and it can be used to find CIs for any real-valued function of several binomial probabilities such as the product of binomial probabilities. For the present problem, the fiducial CI is comparable with the Newcombe CI and the MN CI, and the latter two CIs do not involve simulation. These three CIs are quite comparable and safe to use in practical applications.

Acknowledgment

The authors are grateful to two reviewers for providing valuable comments and suggestions.

References

Casella, G., Berger, R. L. (2001). *Statistical Inference*. Pacific Grove CA: Duxbury Press.

Chan, I. S. F., Zhang, Z. (1999). Test based exact confidence intervals for the difference of two binomial proportions. *Biometrics* 55:1202–1209.

Cai, T. (2005). One-sided confidence intervals in discrete distributions. *J. Stat. Plann. Inference* 131:63–88.

Krishnamoorthy, K. (2006). *Handbook of Statistical Distributions with Applications*. Boca Raton, FL: Chapman & Hall/CRC.

Krishnamoorthy, K., Lee, M. (2010). Inference for functions of parameters in discrete distributions based on fiducial approach: Binomial and Poisson cases. *J. Stat. Plann. Inference* 140:1182–1192.

Krishnamoorthy, K., Lee, M. (2012). Score and fiducial confidence intervals for the difference between two Poisson means. *J. Stat. Comput. Simulat.* doi:10.1080/00949655.2012.686616.

Krishnamoorthy, K., Thomson, J. (2004). A more powerful test for comparing two Poisson means. *J. Stat. Plann. Inference* 119:23–35.

Li, H.-Q., Tang, M.-L., Poon, W.-Y., Tang, N.-S. (2011). Confidence intervals for difference between two Poisson rates. *Commun. Stat. Simulat. Comput.* 40:1478–1491.

Liu, G. F., Wang, J., Liu, K., Snavely, D. B. (2006). Confidence intervals for an exposure adjusted incidence rate difference with applications to clinical trials. *Stat. Med.* 25:1275–1286.

Peskun, P. H. (1993). A new confidence interval method based on the normal approximation for the difference of two binomial probabilities. *J. Amer. Stat. Assoc.* 88:656–661.

Miettinen, O. S., Nurminen, M. (1985). Comparative analysis of two rates. *Stat. Med.* 4:213–226.

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

Storer, B. E., Kim, C. (1990). Exact properties of some exact test statistics for comparing two binomial proportions. *J. Amer. Stat. Assoc.* 85:146–155.

Suissa, S., Shuster, J. J. (1985). Exact unconditional sample sizes for the 2×2 binomial trial. *J. Royal Stat. Soc.* Ser. A 148:317–327.

Zar, J. H. (1996). *Biostatistical Analysis*. Upper Saddle River,NJ: Prentice Hall.

Zou, G. Y., Donner, A. (2008). Construction of confidence limit about effect measures: A general approach. *Stat. Med.* 27:1693–1702.