Modified Clopper-Pearson Confidence Interval for Binomial Proportion

Desale Habtzghi, [1*] Chand K. Midha [2] and Ashish Das [3]

[1*]*The University of Akron, Akron, USA. E-mail: dh52@uakron.edu*
[2]*The University of Akron, Akron, USA. E-mail: chand@uakron.edu*
[3]*Indian Institute of Technology Bombay, Mumbai, India. Email: ashish@math.iitb.ac.in*

Abstract

We introduce expected coverage probability as a measure for constructing confidence intervals for the binomial proportion, π. We propose a model based confidence interval for π using the expected coverage probabilities of the Clopper-Pearson interval. The method provides intervals comparable or better than the alternative intervals, such as the Wilson, Agresti-Coull and Jeffreys intervals.

Keywords: Binomial proportion, Expected coverage probability, Coverage probability, Exact confidence interval, Logistic model.

AMS (2000) subject classification. Primary 62F30; Secondary 62F03, 62G10, 62N03.

*

1. Introduction

Recent years have seen considerable development in a very basic but important problem in statistical practice, namely interval estimation of the probability of success, π, from a binomial observation, $x \sim \text{Bin}(n, \pi)$. A point estimate of π is provided by the MLE $p = \frac{x}{n}$, where x is the number of successes in n Bernoulli trials with probability π succss on each trial. A considerable literature exists about different methods of forming confidence intervals for π (e.g. [1, 2, 3, 4]). These and other authors have used a number of alternative methods of forming confidence intervals for π. To name some, we have Wald, Wilson, Jeffreys, Agresti-Coull and Clopper-Pearson intervals among others. The two papers, [5, 6] provide a thorough study of the coverage characteristics of these intervals.

For most of these intervals their validity depends heavily on the large sample approximation. To improve the normal approximation, in particular for the Wald interval, some intervals based on transformation of sample proportion were proposed. These intervals include arcsine transformation, probit transformation and complementary log-log transformation. These approaches have better convergence rates than the Wald interval. To avoid normal approximation, exact intervals can be derived directly from the binomial distribution. Especially these days, with modern computational power, it is not necessary to rely on large sample approximations to

---

*Corresponding author. E-mail: dh52@uakron.edu.

obtain the distribution of test statistic and confidence intervals for the binomial proportion. Tests and confidence intervals can use the binomial distribution directly rather than its approximation. The best known interval of this type is Clopper-Pearson (CP) interval; it uses the tail method for forming confidence intervals. According to [7] and [8], there was a time when this method was treated as the "golden rule" for obtaining binomial proportion confidence intervals. In a traditional statistical sense, the appealing coverage probability of at least $100(1-\alpha)\%$ is accepted by many statisticians as well as practitioners. In other words, it is well known that the CP interval is conservative in the sense that $(1-\alpha)$ is not the greatest lower bound for the coverage probability.

To improve the conservativeness, in this paper we consider an alternative method to construct confidence intervals for π through introducing a measure of expected coverage probability (see Section 2 for the definition of the expected coverage). We then model these expected coverage probabilities as a function of the limits of the confidence interval. As mentioned before for a given value of α, the $100(1-\alpha)\%$ CP intervals are known to provide conservative intervals, wherein the true coverage probability is always equal to or above the nominal coverage probability. Furthermore, the expected coverage probabilities are bounded below by the nominal level (see Proposition 1). For instance, given x and n, a 95% CP interval has an expected coverage much greater than 95%.

Our approach enables one to determine comparable or better confidence interval than those of Wald, Wilson, Agresit-Coull and Jeffreys from observed proportion of success when these approaches are inadequate. For example, [5] and other authors have shown the inconsistency and poor performance of the Wald interval unless *npq* is large. The use of the Wilson interval or Jeffreys interval for a small *n* and the Agresti and Coull interval for a large *n* was recommended by [5]. Various evaluations indicate that the Clopper-Pearson interval tends to be extremely conservative for small to moderate *n*. We believe that our method is valuable for instructors as well, particularly when they conclude that the confidence intervals in many textbooks are erroneous. For example, most texts use the traditional (Wald) interval which has a long history. The methodology and tables of CP intervals can be found in many intermediate and most advanced statistics (e.g. [9]). But these two have been found to be somewhat inaccurate. Because of the accessibility of digital computers and programmable electronic devices such as calculators and smart phones, the new approach can be easily obtained.

In Section 2, we introduce the concept of expected coverage probabilities. In Section 3, we discuss the construction of the model based confidence interval for π and give the new confidence intervals *Mnew* for various values of *n* at a given level of confidence. Section 4 compares the performance of the *Mnew* and the most commonly used alternative confidence intervals using coverage probabilities, mean coverage probabilities, average expected length and mean absolute errors.

2. Expected Coverage Probability

The concept of expected coverage probability being introduced here is an extension of the true coverage which is usually provided for a given *n* and π. For each observed $x = 0, 1, \cdots, n$, let the $100(1-\alpha)\%$ confidence intervals be denoted by $CI_{x,n,1-\alpha}$. Furthermore, let $L_{x,n,1-\alpha}$ and $U_{x,n,1-\alpha}$ denote the lower and upper limits of the confidence interval, respectively. Now, for fixed n, π and α, the true coverage probability is

$$C(n, 1-\alpha, \pi) = \sum_{x=0}^{n} \delta(x) f(x|\pi) \tag{2.1}$$

where, $f(x|\pi) = \binom{n}{x} \pi^x (1-\pi)^{n-x}$, $x = 0, \ldots, n$ and $\delta(x) = 1$ if $CI_{x,n,1-\alpha}$ captures π and is 0, otherwise.

For a given x, n and α, we now define the expected coverage probability of $CI_{x,n,1-\alpha}$ as follows

$$E_C(x,n,1-\alpha) = \int_0^1 C(n,1-\alpha,\pi)w(\pi)d\pi. \tag{2.2}$$

Here $w(\pi) = \left\{ f(x|\pi)/\int_0^1 f(x|\pi)d\pi \right\}$. As the name suggests, the expected coverage probability is the expectation of the true coverage over π for a given x. It can be easily noticed that the expression of the expected coverage is the conditional distribution of π given x. Thus, the expected coverage can be considered as the posterior expected coverage given x. For the sake of simplicity and convenience, we place a prior distribution of $beta(1,1)$ on the parameter π of a binomial distribution. Just to illustrate, when $x=2$ and $n=15$, for the 95% confidence interval, the expected coverage for CP, Wilson and Jeffreys intervals are 98.28%, 95.69% and 96.72%, respectively. Also, when $x=8$ and $n=15$, the expected coverage for CP, Wilson and Jeffreys intervals are 97.42%, 95.11% and 94.78%, respectively.

Lemma 1.

$$E_C(x,n,1-\alpha) = (n+1)\int_0^1 C(n,1-\alpha,\pi)f(x|\pi)d\pi. \tag{2.3}$$

Proof. The proof of this lemma is straightforward and trivial. First we write $f(x|\pi) = \binom{n}{x}\pi^x(1-\pi)^{n-x}$ and $\binom{n}{x} = \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n-x+1)}$ since $\Gamma(k) = (k-1)!$ when k is an integer.

Note that, $f(x|\pi)$ is a density function in π and a probability mass function in x. We can easily calculate the marginal density of X by integrating the joint over π.

Now, recall the beta density

$$g(u) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}u^{a-1}(1-u)^{b-1}, \ 0 \leqslant u \leqslant 1.$$

The fact that this density integrates to 1 implies

$$\int_0^1 u^{a-1}(1-u)^{b-1}du = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Thus, identifying u with π, $a-1$ with x, and $b-1$ with $n-x$,

$$\int_0^1 \binom{n}{x}\pi^x(1-\pi)^{n-x}d\pi = \frac{\Gamma(n+1)}{\Gamma(x+1)\Gamma(n-x+1)}\int_0^1 \pi^x(1-\pi)^{n-x}d\pi = \frac{1}{n+1}.$$

The rest follows using (2.2).

The integral on the right hand side of (2.3) does not have a closed form solution. Thus, we replace it by a numerical approximation involving summation over the range of π having a specified increment, say $\Delta\pi$. This $\Delta\pi$ could be a small fixed positive number. Practically, $\Delta\pi = 1/m$, where m is a finite large positive integer, such as 100, 1000 or 10000. Let $V = \{0, \Delta\pi, 2\Delta\pi, 3\Delta\pi, \ldots, m\Delta\pi\}$, then the numerical approximation of the expected coverage probability is

$$E_C(x,n,1-\alpha) \approx (n+1)\sum_{\pi \in V} C(n,1-\alpha,\pi)f(x|\pi)\Delta\pi. \tag{2.4}$$

In other words, the approximate expected coverage probability approach uses a weighted sum of the true coverage probabilities over distributions of π to derive its value at any x. In order to assign these weights,

we create a weight function that gives the largest weight to the true coverage corresponding to π such that, $(x-1)/n \leqslant \pi \leqslant (x+1)/(n+1)$. As $n\pi$ moves away from x, the weight will eventually decay to zero. We illustrate this using fixed values of $n = 12$, $x = 4$ and $0 \leqslant \pi \leqslant 1$ with an increment of 0.1. Table 1 provides computation of the expected coverage probabilities of CP, Wald, Wilson, Jeffreys, (which are 0.972, 0.883, 0.962 and 0.962, respectively) for $\alpha = 0.05$.

Table 1: A 95% coverage probabilities and expected coverage probabilities at $x = 4$ of CP, Wald, Wilson, and Jeffreys for $n = 12$, $0 \leqslant \pi \leqslant 1$ with increment of 0.1, and probabilities of success for $x = 4$ evaluated at the different values of π.

π	$C(12, 0.95, \pi)$	CP	Wald	Wilson	Jeffreys	$\frac{f(4\mid\pi)}{\sum_{\pi \in V} f(4\mid\pi)}$
0.000	$C(12, 0.95, 0.0)$	1.000	1.000	1.000	1.000	0.000
0.100	$C(12, 0.95, 0.1)$	0.996	0.713	0.974	0.974	0.027
0.200	$C(12, 0.95, 0.2)$	0.981	0.912	0.981	0.981	0.173
0.300	$C(12, 0.95, 0.3)$	0.977	0.876	0.948	0.948	0.300
0.400	$C(12, 0.95, 0.4)$	0.965	0.901	0.965	0.965	0.277
0.500	$C(12, 0.95, 0.5)$	0.961	0.854	0.961	0.961	0.157
0.600	$C(12, 0.95, 0.6)$	0.965	0.901	0.965	0.965	0.055
0.700	$C(12, 0.95, 0.7)$	0.977	0.876	0.948	0.948	0.010
0.800	$C(12, 0.95, 0.8)$	0.981	0.912	0.981	0.981	0.001
0.900	$C(12, 0.95, 0.9)$	0.996	0.713	0.974	0.974	0.000
1.000	$C(12, 0.95, 1.0)$	1.000	1.000	1.000	1.000	0.000
$E_C(4, 12, 0.95)$		0.972	0.883	0.962	0.962	
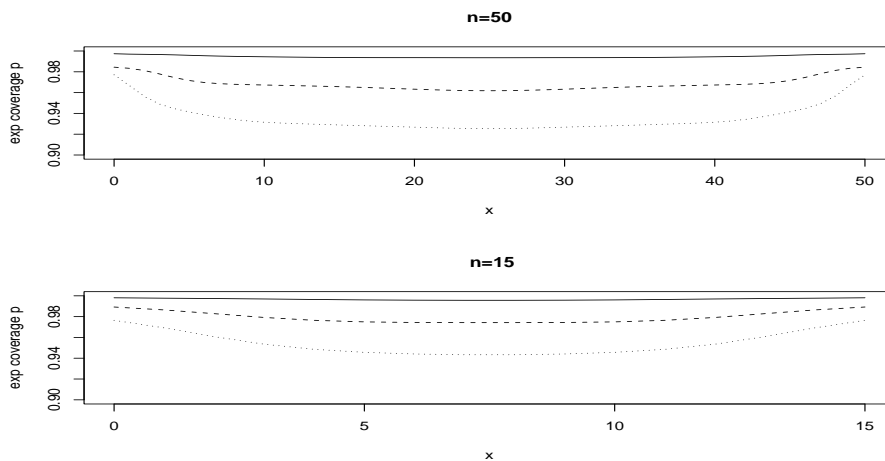
Figure 1: Display of the expected coverage probabilities of CP for $n = 50$ and $n = 15$, 99% (solid curve), 95% (dashed curve) and 90% (dotted curve).

Figure 1 shows the trend in expected coverage probabilities of CP for $n = 15$ and $n = 50$ at 90%, 95% and 99% nominal coverage probabilities. One can see that the discrepancies of the expected coverage probabilities from nominal levels increases for x near 0 or n and decreases when x is close to $n/2$. In addition, one can

observe that the conservativeness of the CP intervals increases as the confidence level decreases and decreases as the sample size increases as expected.

Proposition 1 *Consider any fixed n and $0 < \alpha < 1$. Let $E_C(x, n, 1 - \alpha)$ be the expected coverage probabilities of Clopper-Pearson for any fixed $1 \leqslant x \leqslant n$ then $E_C(x, n, 1 - \alpha) \geqslant 1 - \alpha$.*

Proof. Suppose $0 < \alpha < 1$ and $1 \leqslant x \leqslant n$ with $E_C(x, n, 1 - \alpha)$ as in (2.2).

$$E_C(x, n, 1 - \alpha) = \int_0^1 C(n, 1 - \alpha, \pi) \left\{ \frac{f(x|\pi)}{\int_0^1 f(x|\pi)d\pi} \right\} d\pi$$

$$\geqslant \int_0^1 (1 - \alpha) \left\{ \frac{f(x|\pi)}{\int_0^1 f(x|\pi)d\pi} \right\} d\pi$$

since $C(n, 1 - \alpha, \pi)$ is bounded below by $(1 - \alpha)$ (see, [10] and [11], i.e. $C(n, 1 - \alpha, \pi) > (1 - \alpha)$

$$= (1 - \alpha) \int_0^1 \left\{ \frac{f(x|\pi)}{\int_0^1 f(x|\pi)d\pi} \right\} d\pi = (1 - \alpha).$$

3. Construction of Modified Clopper-Pearson Intervals

In this section, before we introduce our method, we review some fundamental concepts of the Clopper-Pearson interval that can help us to lay groundwork for the construction of our interval. The methodology and results along with more details about the properties of Clopper-Pearson method can be found in [11] and [9].

The "exact" $100(1 - \alpha)\%$ confidence interval of Clopper and Pearson is obtained by inverting the binomial test of $H_0 : \pi = \pi_0$. For $0 \leqslant x \leqslant n$, to obtain the lower and upper confidence limits, we solve

$$P_{\pi_0}(X \geqslant x) = \sum_{i=x}^n \binom{n}{i} \pi_0^i (1 - \pi_0)^{n-i} = \frac{\alpha}{2}, \tag{3.1}$$

and

$$P_{\pi_0}(X \leqslant x) = \sum_{i=0}^x \binom{n}{i} \pi_0^i (1 - \pi_0)^{n-i} = \frac{\alpha}{2}. \tag{3.2}$$

The explicit form of the solutions for the confidence limits using the F-dist is

$$\left(\frac{x F_{2x, 2(n-x+1)}(1 - \alpha/2)}{n - x + 1 + x F_{2x, 2(n-x+1)}(1 - \alpha/2)}, \frac{(x+1) F_{2(x+1), 2(n-x)}(\alpha/2)}{n - x + (x+1) F_{2(x+1), 2(n-x)}(\alpha/2)} \right), \tag{3.3}$$

for more details about (3.3) see [11].

Here $F_{a,b}(c)$ denotes the $1 - c$ quantile of the F-distribution with degrees of freedom a and b. The following endpoint adjustments are needed: the lower limit is 0 if $x = 0$ and the upper limit is 1 when $x = n$. As mentioned before, the interval has true coverage probabilities of at least $1 - \alpha$ for all n and π (see, for example, [5]). The limits can also be obtained using root finding routines such as bisection method or Newton-Raphson method for a given x and n using (3.1) and (3.2).

In order to construct the model based confidence interval *Mnew,* the following algorithms are proposed to compute the confidence limits for given x, n and α^*.

Step 1. For a given n and all outcomes, $x = 0, 1, \ldots, n$, obtain the $100(1-\alpha)\%$ CP confidence intervals. Denoting the confidence interval by $CI_{(CP)(x,n,1-\alpha)}$, let $L_{(CP)(x,n,1-\alpha)}$ and $U_{(CP)(x,n,1-\alpha)}$ be the corresponding lower and upper limits of this confidence interval, respectively.

Step 2. For fixed π and α, obtain the true coverage probability for the CP confidence intervals using (2.1).

Step 3. Taking $\Delta\pi = 0.001$, find the expected coverage probability of $CI_{(CP)(x,n,1-\alpha)}$ for fixed x, n and α using (2.4).

Step 4. For given x and n, in order to find the desired $100(1-\alpha^*)\%$ confidence interval, model the expected coverage probabilities for various values of the lower limits, $L_{(CP)(x,n,1-\alpha)}$ by varying $1-\alpha$. The range of $1-\alpha$, for building the model, is chosen such that the corresponding expected coverage probabilities contain $1-\alpha^*$.

Similarly, model the expected coverage probabilities for various values of the upper limits, $U_{(CP)(x,n,1-\alpha)}$.

Finally, fit the two logistic regression models

$$\text{Logit}(E_C(x,n,1-\alpha)) = \beta_{L0} + \beta_{L1}L_{(CP)(x,n,1-\alpha)}$$

and

$$\text{Logit}(E_C(x,n,1-\alpha)) = \beta_{U0} + \beta_{U1}U_{(CP)(x,n,1-\alpha)}.$$

Step 5. For $i = 0, 1$, let the estimate of the coefficients, β_{Li} and β_{Ui} be b_{Li} and b_{Ui}, respectively. Then the modified upper and lower limits for the $100(1-\alpha^*)\%$ confidence interval are obtained by inverting the equations given in step 4, as follows

$$\text{Lower Limit}: L_{(new)(x,n,1-\alpha^*)} = \left(\log\left[\frac{1-\alpha^*}{\alpha^*}\right] - b_{L0}\right)/b_{L1},$$

$$\text{Upper Limit}: U_{(new)(x,n,1-\alpha^*)} = \left(\log(\frac{1-\alpha^*}{\alpha^*}) - b_{U0}\right)/b_{U1}.$$

Following the above steps, one can obtain the $100(1-\alpha^*)\%$ confidence intervals for various values of n and x. In general, for building the logistic model, the values of $1-\alpha$ can range from 0.50 to 0.999 with increments of 0.001. However, for more commonly used values of $1-\alpha^*$, such as 0.99, 0.95 or 0.90, we prescribe more precise ranges for $1-\alpha$. For example, for the $100(1-\alpha^*)\% = 95\%$ confidence interval, in order to fit the logistic model, the values of $1-\alpha$ may range from 0.7 to 0.95 with increments of 0.001. Similarly, for fitting the 99% and 90% confidence intervals, the values of $1-\alpha$ range from 0.90 to 0.999 (with increment of 0.001) and 0.55 to 0.90 (with increment of 0.001), respectively. For building the model, the above ranges of $1-\alpha$ ensure that the corresponding expected coverage probabilities contains 0.95, 0.99 and 0.90, respectively. We obtain the 95% confidence intervals for various values of n ($= 5, 6, \ldots, 40$) and x ($= 0, 1, \ldots, n$). The results are displayed in Tables 2 and 3. For each n, we give the intervals for $x = 0, 1, \ldots, [n/2]$, where $[z]$ represents the largest integer less than or equal to z. The intervals for the other values of x follows from the relation

$$L_{(new)(x,n,1-\alpha^*)} = 1 - U_{(new)(n-x,n,1-\alpha^*)}.$$

Some select confidence limits using the new approach are also provided for 90% and 99% confidence intervals. The results are given in Table 4. The results in this article can be reproduced using our **R** code. The authors have provided a confidence interval function in **R** on the Web page at http: //gozips.uakron.edu/˜dh52/confnew.html. The function allows the user to request any of the confidence intervals for choices *n*, *x* and α^*.

4.   Comparison of the Confidence Intervals

The comparison of the Wald, Clopper-Pearson (CP), Wilson, Agresit-Coull (AC), Jeffreys and our *Mnew* intervals are carried out in this section. In Figure 2 we provide the coverage probabilities for the six intervals for a fixed $n = 50$ and variable π while taking 95% confidence level. From Figure 2, the coverage probability of the Wald is strictly smaller than the nominal coverage probability of 0.95. The coverage probabilities of Wilson, AC, Jeffreys and *Mnew* fluctuate around the target confidence level. The oscillation in the coverage probabilities is caused by the discreteness of the binomial distribution. See [5], [12] and [13] for a more detailed discussion. The plots in Figure 2 and our evaluation show that when π nears 0.5, the performance of *Mnew* is better than the other alternative intervals. Similarly, in Figure 3, we compare the six intervals using fixed $\pi = 0.25$ and variable *n*. In this figure the coverage probabilities of *Mnew* shows fewer oscillations and the coverage probabilities do not deviate significantly from the nominal coverage probability.

Instead of comparing the performance of our method using coverage probabilities, in some cases, the coverage probabilities vary according to the parameter value due to the discreteness of the binomial distribution. So we used mean coverage probability, expected length along with the average of expected length of the intervals and mean absolute errors for $n = 5, 6, \ldots, 55$. The results are given in Figures 4-7. The mean coverage probability, expected length, average expected length and mean absolute errors for a given *n* and $1 - \alpha^*$ are computed as follows, respectively:
Mean coverage probability

$$\overline{C}(n, 1-\alpha) = \int_0^1 C(n, 1-\alpha^*, \pi)g(\pi)d\pi,$$

expected length

$$E_{n,\pi}(length(CI)) = \sum_{x=0}^n (U_{x,n,1-\alpha} - L_{x,n,1-\alpha}) \binom{n}{x} \pi^x (1-\pi)^{n-x}$$

and the average expected length is given by the following integral

$$\int_0^1 E_{n,\pi}(length(CI))d\pi,$$

mean absolute errors,

$$\int_0^1 |C(n, 1-\alpha) - (1-\alpha)|d\pi,$$

where $g(\pi) = 1$ for $0 \leqslant \pi \leqslant 1$. From Figures 4-7, we see that *Mnew* provides comparable estimates vis-a-vis those of Wilson, Jeffreys and Agresit-Coull. The performance of *Mnew* is consistently comparable or better than those of Wilson, AC and Jeffreys. This supports our observations made in Figures 2 and 3. In these examples,

Table 2: The 95% confidence limits based on "*Mnew*" approach for $5 \leqslant n \leqslant 28$.

x	L_{new} 5	U_{new}	L_{new} 6	U_{new}	L_{new} 7	U_{new}	L_{new} 8	U_{new}	L_{new} 9	U_{new}	L_{new} 10	U_{new}
0	0.000	0.360	0.000	0.314	0.000	0.279	0.000	0.250	0.000	0.228	0.000	0.208
1	0.020	0.597	0.016	0.528	0.013	0.474	0.011	0.428	0.010	0.391	0.009	0.359
2	0.100	0.774	0.078	0.696	0.065	0.631	0.055	0.575	0.048	0.528	0.043	0.487
3			0.174	0.826	0.141	0.758	0.120	0.697	0.104	0.644	0.092	0.598
4							0.202	0.798	0.173	0.743	0.152	0.693
5											0.224	0.776

x	11		12		13		14		15		16	
0	0.000	0.193	0.000	0.179	0.000	0.166	0.000	0.155	0.000	0.145	0.000	0.136
1	0.008	0.332	0.007	0.309	0.006	0.289	0.006	0.271	0.006	0.255	0.005	0.241
2	0.039	0.452	0.035	0.422	0.032	0.395	0.030	0.372	0.028	0.351	0.026	0.332
3	0.082	0.557	0.075	0.521	0.068	0.489	0.063	0.461	0.058	0.436	0.054	0.413
4	0.136	0.648	0.123	0.609	0.112	0.573	0.103	0.541	0.096	0.512	0.089	0.486
5	0.199	0.729	0.179	0.687	0.163	0.648	0.150	0.614	0.138	0.582	0.129	0.554
6			0.243	0.757	0.220	0.717	0.201	0.681	0.186	0.647	0.172	0.616
7							0.258	0.742	0.237	0.707	0.220	0.674
8											0.271	0.729

x	17		18		19		20		21		22	
0	0.000	0.128	0.000	0.122	0.000	0.117	0.000	0.112	0.000	0.107	0.000	0.103
1	0.005	0.229	0.005	0.217	0.004	0.207	0.004	0.198	0.004	0.189	0.004	0.181
2	0.024	0.315	0.023	0.300	0.021	0.286	0.020	0.273	0.019	0.261	0.018	0.251
3	0.051	0.392	0.048	0.374	0.045	0.357	0.043	0.341	0.041	0.327	0.039	0.314
4	0.083	0.463	0.078	0.441	0.074	0.422	0.070	0.404	0.066	0.387	0.063	0.372
5	0.120	0.528	0.113	0.504	0.106	0.482	0.100	0.462	0.095	0.443	0.091	0.426
6	0.161	0.588	0.151	0.562	0.142	0.538	0.134	0.516	0.127	0.496	0.121	0.477
7	0.205	0.644	0.192	0.617	0.180	0.591	0.170	0.568	0.161	0.546	0.153	0.525
8	0.252	0.698	0.236	0.668	0.221	0.642	0.209	0.617	0.197	0.593	0.187	0.571
9			0.282	0.718	0.265	0.689	0.249	0.663	0.236	0.639	0.223	0.616
10							0.292	0.708	0.276	0.683	0.261	0.658
11											0.301	0.699

x	23		24		25		26		27		28	
0	0.000	0.098	0.000	0.095	0.000	0.092	0.000	0.089	0.000	0.086	0.000	0.083
1	0.003	0.174	0.003	0.168	0.003	0.162	0.003	0.156	0.003	0.150	0.003	0.146
2	0.017	0.241	0.017	0.232	0.016	0.224	0.015	0.216	0.015	0.208	0.014	0.202
3	0.037	0.302	0.035	0.291	0.034	0.280	0.032	0.271	0.031	0.262	0.030	0.253
4	0.060	0.358	0.057	0.345	0.055	0.333	0.053	0.322	0.051	0.311	0.049	0.301
5	0.086	0.410	0.082	0.395	0.079	0.382	0.076	0.369	0.073	0.357	0.070	0.346
6	0.115	0.459	0.110	0.443	0.105	0.428	0.101	0.414	0.097	0.400	0.093	0.388
7	0.146	0.506	0.139	0.489	0.133	0.472	0.127	0.457	0.122	0.442	0.118	0.428
8	0.178	0.551	0.170	0.532	0.162	0.514	0.156	0.498	0.149	0.482	0.143	0.467
9	0.212	0.594	0.202	0.574	0.193	0.555	0.185	0.538	0.178	0.521	0.171	0.505
10	0.248	0.636	0.236	0.615	0.226	0.595	0.216	0.576	0.207	0.558	0.199	0.542
11	0.285	0.676	0.272	0.654	0.259	0.633	0.248	0.613	0.238	0.595	0.228	0.578
12			0.308	0.692	0.294	0.670	0.281	0.650	0.269	0.630	0.258	0.612
13							0.315	0.685	0.302	0.665	0.289	0.646
14											0.321	0.679

Table 3: The 95% confidence limits based on " *Mnew*" approach for $29 \leqslant n \leqslant 40$.

x	L_{new} 29	U_{new}	L_{new} 30	U_{new}	L_{new} 31	U_{new}	L_{new} 32	U_{new}	L_{new} 33	U_{new}	L_{new} 34	U_{new}
0	0.000	0.081	0.000	0.078	0.000	0.076	0.000	0.074	0.000	0.071	0.000	0.069
1	0.003	0.141	0.003	0.136	0.003	0.132	0.002	0.128	0.002	0.125	0.002	0.121
2	0.014	0.195	0.013	0.189	0.013	0.183	0.012	0.178	0.012	0.173	0.012	0.169
3	0.029	0.245	0.028	0.237	0.027	0.231	0.026	0.224	0.025	0.218	0.025	0.212
4	0.047	0.291	0.045	0.282	0.044	0.274	0.042	0.267	0.041	0.259	0.040	0.252
5	0.067	0.335	0.065	0.324	0.063	0.316	0.061	0.307	0.059	0.298	0.057	0.290
6	0.090	0.376	0.086	0.364	0.083	0.355	0.080	0.345	0.078	0.335	0.076	0.326
7	0.113	0.415	0.109	0.403	0.105	0.392	0.102	0.381	0.098	0.371	0.095	0.361
8	0.138	0.453	0.133	0.440	0.128	0.428	0.124	0.417	0.120	0.405	0.116	0.395
9	0.164	0.490	0.158	0.476	0.153	0.463	0.147	0.451	0.143	0.439	0.138	0.427
10	0.191	0.526	0.184	0.511	0.178	0.497	0.172	0.484	0.166	0.471	0.161	0.459
11	0.219	0.561	0.211	0.545	0.203	0.531	0.196	0.516	0.190	0.503	0.184	0.490
12	0.248	0.595	0.239	0.578	0.230	0.563	0.222	0.548	0.215	0.534	0.208	0.521
13	0.278	0.628	0.267	0.611	0.258	0.595	0.249	0.579	0.240	0.565	0.232	0.550
14	0.308	0.660	0.297	0.642	0.286	0.626	0.276	0.609	0.266	0.594	0.258	0.579
15			0.327	0.673	0.315	0.656	0.303	0.639	0.293	0.623	0.283	0.608
16							0.332	0.668	0.320	0.652	0.310	0.636
17											0.337	0.663

x	35		36		37		38		39		40	
0	0.000	0.067	0.000	0.066	0.000	0.065	0.000	0.063	0.000	0.062	0.000	0.061
1	0.002	0.118	0.002	0.115	0.002	0.112	0.002	0.110	0.002	0.107	0.002	0.104
2	0.011	0.164	0.011	0.160	0.011	0.156	0.010	0.152	0.010	0.148	0.010	0.145
3	0.024	0.206	0.023	0.201	0.022	0.196	0.022	0.191	0.021	0.186	0.021	0.182
4	0.039	0.246	0.037	0.240	0.036	0.234	0.036	0.228	0.035	0.222	0.034	0.217
5	0.055	0.283	0.054	0.276	0.052	0.269	0.051	0.262	0.049	0.256	0.048	0.250
6	0.073	0.318	0.071	0.310	0.069	0.303	0.067	0.295	0.065	0.288	0.064	0.282
7	0.092	0.352	0.090	0.343	0.087	0.335	0.085	0.327	0.082	0.319	0.080	0.312
8	0.113	0.385	0.109	0.375	0.106	0.366	0.103	0.358	0.100	0.349	0.098	0.341
9	0.134	0.417	0.130	0.406	0.126	0.397	0.122	0.388	0.119	0.379	0.116	0.370
10	0.156	0.448	0.151	0.437	0.146	0.427	0.142	0.417	0.138	0.407	0.135	0.398
11	0.178	0.478	0.173	0.467	0.168	0.456	0.163	0.445	0.158	0.435	0.154	0.426
12	0.201	0.508	0.195	0.496	0.189	0.484	0.184	0.473	0.179	0.463	0.174	0.452
13	0.225	0.537	0.218	0.524	0.212	0.512	0.205	0.501	0.200	0.489	0.194	0.479
14	0.249	0.566	0.242	0.552	0.234	0.540	0.228	0.528	0.221	0.516	0.215	0.505
15	0.274	0.594	0.266	0.580	0.258	0.567	0.250	0.554	0.243	0.542	0.236	0.530
16	0.300	0.621	0.290	0.607	0.281	0.593	0.273	0.580	0.265	0.567	0.258	0.555
17	0.326	0.648	0.315	0.633	0.306	0.619	0.297	0.605	0.288	0.592	0.280	0.580
18			0.341	0.659	0.330	0.645	0.320	0.631	0.311	0.617	0.302	0.604
19							0.345	0.655	0.335	0.641	0.325	0.628
20											0.348	0.652

Table 4: The 90% and 99% confidence limits based on "*Mnew*" approach for $n = 10$, 25 and 50.

x	L_{new} 90%	U_{new}	L_{new}	U_{new}	L_{new}	U_{new}	L_{new} 99%	U_{new}	L_{new}	U_{new}	L_{new}	U_{new}
	10		25		50		10		25		50	
0	0.000	0.163	0.000	0.069	0.000	0.038	0.000	0.308	0.000	0.142	0.000	0.075
1	0.015	0.312	0.006	0.137	0.003	0.071	0.004	0.463	0.002	0.214	0.001	0.114
2	0.060	0.439	0.023	0.197	0.011	0.103	0.007	0.591	0.002	0.281	0.000	0.149
3	0.119	0.550	0.045	0.252	0.022	0.132	0.036	0.699	0.011	0.342	0.005	0.183
4	0.188	0.647	0.069	0.302	0.034	0.159	0.079	0.789	0.025	0.398	0.012	0.214
5	0.266	0.734	0.096	0.350	0.047	0.185	0.138	0.862	0.042	0.450	0.020	0.244
6			0.125	0.395	0.061	0.210			0.063	0.498	0.029	0.271
7			0.155	0.439	0.075	0.234			0.086	0.542	0.040	0.298
8			0.187	0.481	0.090	0.258			0.111	0.585	0.051	0.324
9			0.220	0.522	0.105	0.281			0.138	0.625	0.063	0.349
10			0.254	0.562	0.121	0.303			0.167	0.664	0.075	0.373
11			0.288	0.601	0.137	0.326			0.198	0.701	0.089	0.396
12			0.324	0.639	0.153	0.348			0.230	0.736	0.102	0.419
13					0.170	0.370					0.117	0.442
14					0.187	0.391					0.131	0.464
15					0.204	0.412					0.146	0.486
16					0.221	0.433					0.162	0.507
17					0.239	0.454					0.178	0.528
18					0.257	0.475					0.194	0.549
19					0.275	0.495					0.210	0.569
20					0.293	0.515					0.227	0.589
21					0.311	0.535					0.244	0.609
22					0.330	0.555					0.261	0.629
23					0.349	0.575					0.278	0.648
24					0.367	0.594					0.296	0.667
25					0.387	0.613					0.315	0.685

the mean coverage probability of *Mnew* are closer to 95% than those of Wald, CP, AC and Jeffreys in most cases. When $22 \leqslant n \leqslant 55$ *Mnew* performs better than Wilson. As expected Wald has smaller mean coverage probabilities for smaller sample sizes. As *n* increases the mean coverage of *Mnew*, Wilson and Jeffreys becomes closer and closer to the nominal level.

We also computed the average expected lengths of the six intervals for $n = 5, \cdots, 55$ and $\alpha = 0.05$; the results are given in Figure 5. The average expected lengths of CP, Wilson, Jeffreys and AC are larger than *Mnew* by 0.0165 to 0.158, 0.0001 to 0.0218, 0 to 0.1021 and 0.005 to 0.0702, respectively. In this case, the average expected length of *Mnew* is the smallest among the six intervals especially when the sample size is small ($n \leqslant 32$). The *Mnew* intervals have considerably higher coverage probabilities than Wald intervals but they are not wider. Unless the *p* is very close to 0 or 1, the *Mnew* intervals are shorter.

Figure 6 shows the percentage increase of the average expected lengths of five intervals compare to *Mnew*. For $n > 40$ the difference of the average expected lengths between *Mnew* and Wilson as well as *Mnew* and Jeffreys diminishes gradually. These differences can have a practical relevance. Our interval can be a plausible choice for all sample sizes because of its good coverage and smaller expected length.

Figure 7 displays the mean absolute error, *Mnew* does not have significant small mean absolute errors in comparison to the Jeffreys, Wilson and AC intervals. This indicates that the variability about the nominal level is similar for these four approaches.
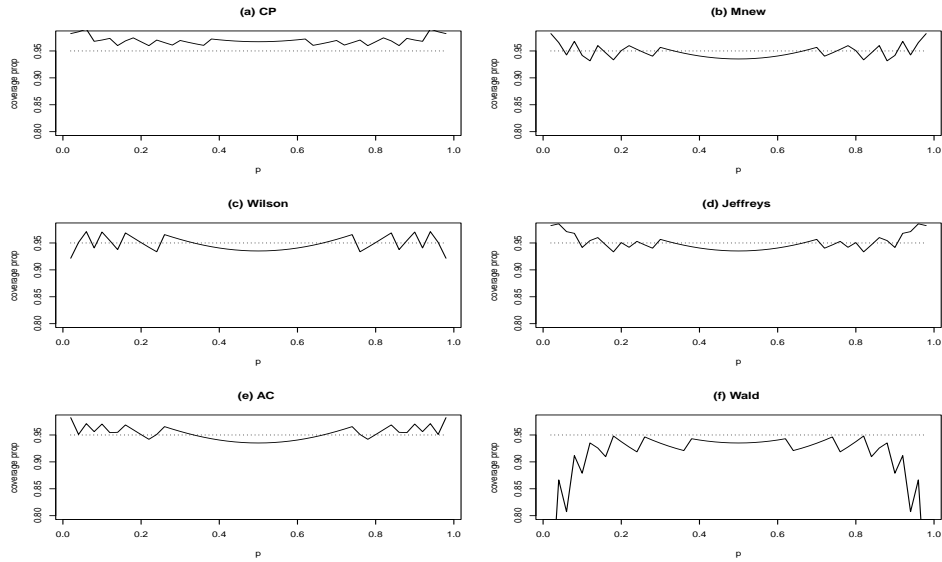
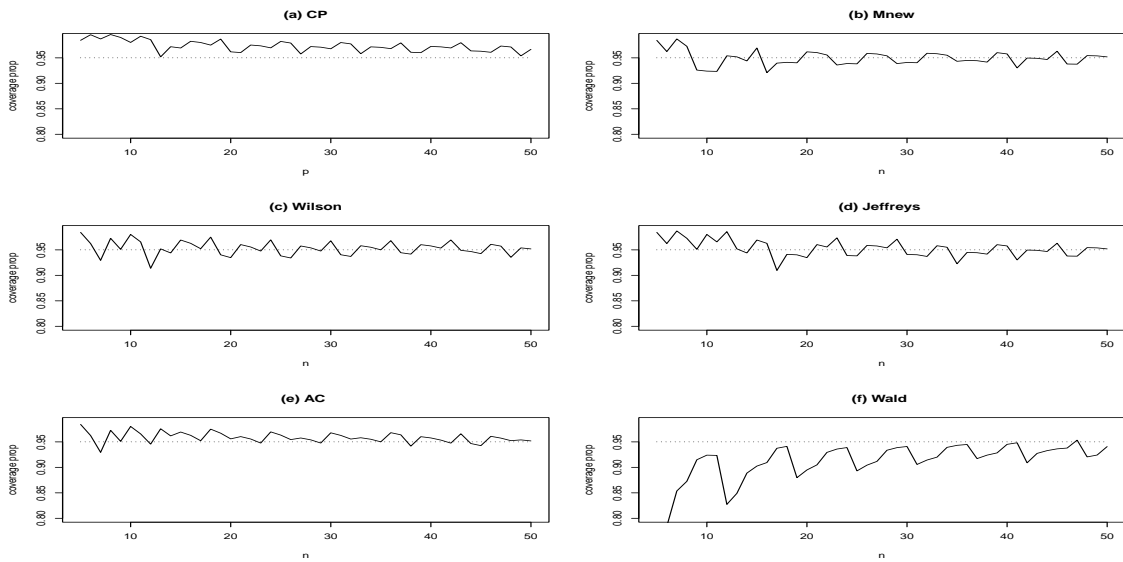Figure 2: Coverage probability for $n = 50$ and variable $p \in (0, 1)$.

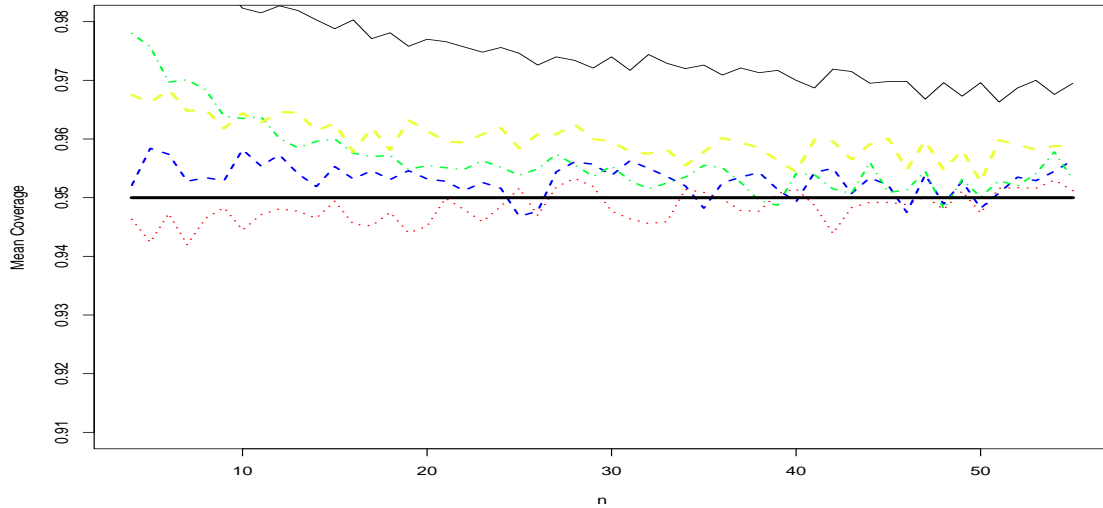Figure 3: Coverage probability for $p = 0.25$ and $5 \leqslant n \leqslant 50$.

Figure 4: Mean coverage probabilities for $n = 5, \cdots, 55$ and variable $p \in (0, 1)$, CP (black solid curve), Wilson (blue dashed curve), Jefferys (green dashed-dotted curve), *Mnew* (red dotted), AC (yellow dashed curve) and nominal mean coverage (solid straight line).
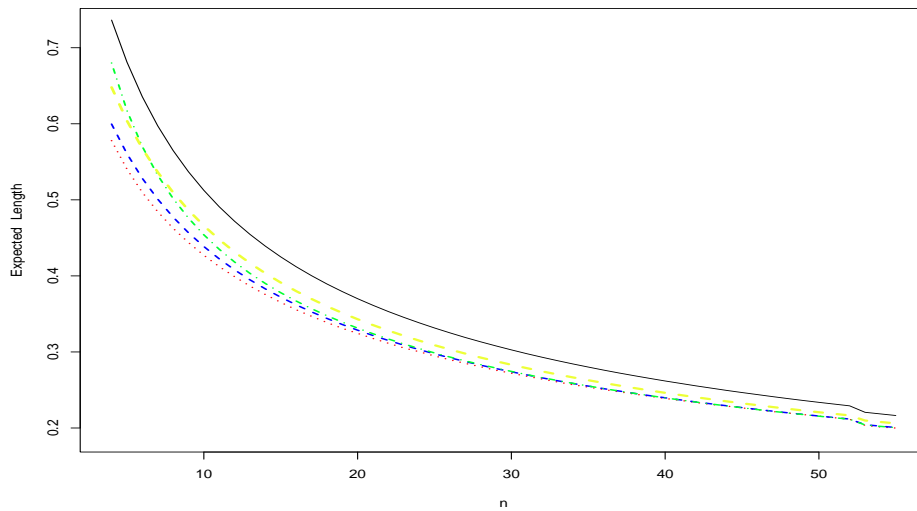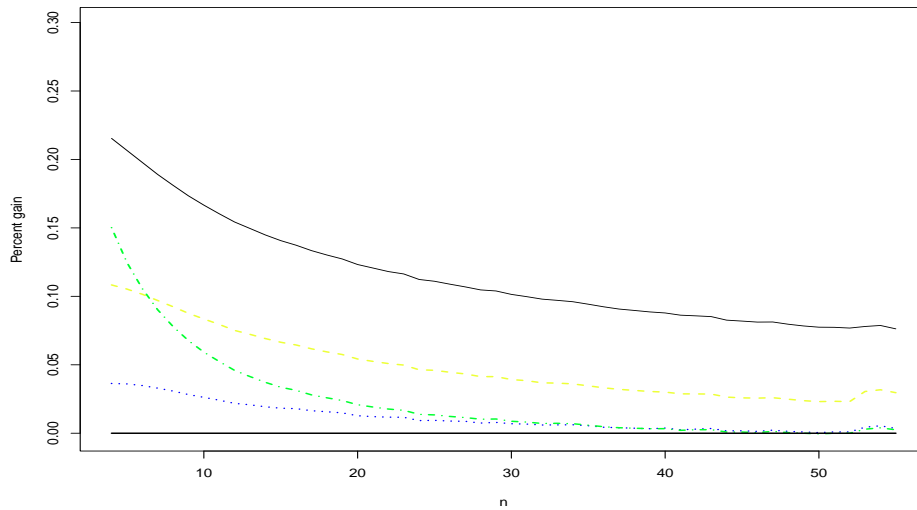

Figure 5: Average expected lengths of CP (black solid curve), Wilson (blue dashed curve), Jefferys (green dashed-dotted curve), *Mnew* (red dotted) and AC (yellow dashed curve) for $n = 5, \cdots, 55$.

Figure 6: Percentage increase of the average expected lengths for CP (black solid curve), Wilson (blue dashed curve), Jefferys (green dashed-dotted curve), and AC (yellow dashed curve) for $n = 5, \cdots, 55$ in comparison to *Mnew* .
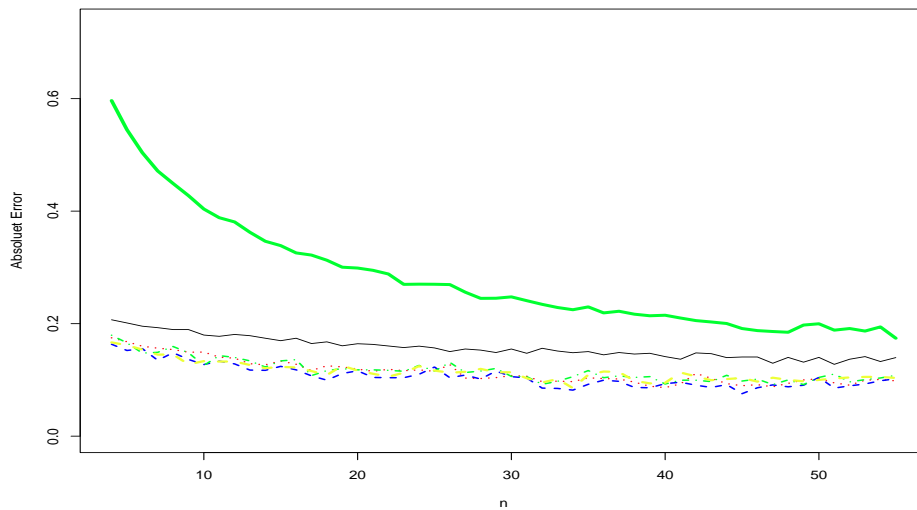

Figure 7: For $n = 5, \cdots, 55$ display of the mean absolute errors of CP (black solid curve), Wilson (blue dashed curve), Jefferys (green dashed-dotted curve), *Mnew* (red dotted), AC (yellow dashed curve) and Wald (green solid curve).

5.   Acknowledgements

References

1. A. Agresti and B. A. Couli, B. A., Approximate is better than "exact" for interval estimation of binomial proportions, *The American Statistician*, **52** (1998), 119-126.
2. C. R. Blyth and H. A. Still, Binomial confidence intervals, *Journal of the American Statistical Association*, **78** (1983), 108-116.
3. C. J. Clopper and E. S. Pearson, The use of confidence or fiducial limits illustrated in the case of the binomial, *Biometrika*, **26** (1934), 404-413.
4. E. B. Wilson, Probable inference, the law of succession, and statistical inference, *Journal of the American Statistical Association*, **22** (1927), 209-212.
5. L. D. Brown, T. T. Cai and A. DasGupta, Interval estimation for a binomial proportion, *Statistical Science*, **16** (2001), 101-117.
6. L. D. Brown, T. T. Cai and A. DasGupta, Confidence intervals for a binomial proportion and asymptotic expansions, *The Annals of Statistics*, **30** (2002), 160-201.
7. L. M. Leemis and K. S. Trivedi, A Comparison of Approximate Interval Estimators for the Bernoulli Parameter, *The American Statistician*, **50** (1996), 63-68.
8. B. D. Jovanovic and P. S. Levy, A Look at the Rule of Three, *The American Statistician*, **51** (1997), 137-139.
9. W. J. Conover, Practical Nonparametric Statistics, 3rd edn. John Wiley and Sons, New York, 1999.
10. J. Neyman, On the problem of confidence limits, *The Annals of Mathematical Statistics*, **6** (1935), 111-116.
11. A. Agresti, Categorical data analysis, 2nd edn. John Wiley and Sons, New York, 2002.
12. A. Agresti, Dealing with discreteness: making 'exact' confidence intervals for proportions, differences of proportions, and odds ratios more exact, *Statistical Methods in Medical Research*, **12** (2003), 3-21.
13. A. Agresti, Exact inference for categorical data: Recent advances and continuing controversies, *Statistics in Medicine*, **20** (2001) 2709-2722.