

Ivana Pobocikova \*

## BETTER CONFIDENCE INTERVALS FOR A BINOMIAL PROPORTION

*Interval estimation of a binomial proportion is one of the basic problems in statistics. In technical practice a binomial proportion is often used in statistical quality control. The standard Wald interval and the exact Clopper-Pearson interval are the most common and frequently used intervals. They are presented in the majority of statistical literature. It is known that the Wald interval performs poorly and this interval should not be used. In this paper we recommend the alternatives of confidence intervals that have a better performance and are appropriate for practical use. We compare the performance of six alternatives of confidence intervals for a binomial proportion: the Wald interval, the Clopper-Pearson interval, the Wilson score interval, the Wilson score interval with continuity correction, the Agresti-Coull interval and the Jeffreys interval in terms of the coverage probability, the interval length and the root mean square error.*

*Key words: binomial distribution, binomial proportion, confidence interval, coverage probability, interval length, root mean square error*

### 1. Introduction

Interval estimation of a binomial proportion is one of the basic problems in statistics. In technical practice the binomial proportion is often used in statistical quality control.

Let random variable  $X$  follow a binomial distribution with parameters  $n \in \mathbb{N}$  and  $\pi \in (0,1)$ , abbreviated  $X \sim Bi(n,\pi)$ . The probability that a random variable  $X$  is equal to the value  $x$  is given by

$$P(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x}, \quad x = 0, 1, \dots, n \quad (1)$$

The parameter  $\pi$  is also called binomial proportion. In practice the value of the parameter  $\pi$  is usually unknown and must be estimated from a sample. Let  $X$  be a number of successes in a random sample of size  $n$ . The maximum likelihood estimator for from the sample is  $p = X/n$ . This estimator is unbiased and consistent. The  $100 \cdot (1 - \alpha)\%$  two-sided confidence interval for parameter  $\pi$  is an interval  $\langle p_L, p_U \rangle$  such as  $P(p_L \leq \pi \leq p_U) \geq 1 - \alpha$ , where  $(1 - \alpha)$  is the desired confidence coefficient,  $\alpha \in (0,1)$ .

Due to the discrete nature of the binomial distribution the interval estimation of binomial proportion is a complicated problem. The standard Wald interval (Laplace, 1812) and the exact Clopper-Pearson interval (Clopper - Pearson, 1934) are the most common and most frequently used intervals. They are presented in the majority of statistical literature. The standard Wald interval (Wald) is based on the standard normal approximation to the binomial distribution. This interval is simple to compute, it is narrow, but the

interval has a poor performance. It is known that its coverage probability behaves irregularly even when  $\pi$  is not close to 0 and 1. The coverage probability is below a nominal level even for very large sample sizes. It is known that Wald interval has a problem with the zero width interval and overshoot (the lower bound can be below 0 and the upper bound can be above 1). Many authors have pointed out that this interval should not be used (Vollset, 1993, Newcombe, 1998, Brown, Cai, DasGupta, 2001, Pires, Amado, 2008).

The exact Clopper-Pearson interval is based on the exact binomial distribution. This interval eliminates overshoot and zero width intervals and it is known that this interval is strictly conservative and too wide (Newcombe, 1998, Brown, Cai, DasGupta, 2001, Pires, Amado, 2008). Its coverage probability is always equal to or above the nominal level.

In this paper we recommend the alternatives of confidence intervals for binomial proportion that have better performance and are often used in practice, but they are presented sporadically in the basic statistical literature. Here we consider the confidence intervals methods that are based on the standard normal approximation: Wilson score interval (Wilson), Wilson score interval with continuity correction (Wilson+CC), Agresti-Coull interval (Agresti-Coull), and finally the interval that is based on the Bayesian approach: Jeffreys interval (Jeffreys).

We demonstrate the performance of selected alternatives of confidence intervals. For comparison we use common criteria: coverage probability, average coverage probability, conservatism,

\* Ivana Pobocikova

Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Slovakia, E-mail: ivana.pobocikova@fstroj.uniza.sk

interval length, average expected length, root mean square error. We summarize the results for the coverage probability in terms of the observed minimum coverage probability and the average coverage probability and we classify the alternatives of confidence intervals into two classes of acceptable intervals- strictly conservative intervals and intervals that are not strictly conservative, but conservative on average.

Our recommendation of these selected alternatives of confidence intervals is based on our investigations of these intervals and on the existing comparative studies that were presented in recent statistical literature, see e. g. Newcombe (1998), Brown, Cai, DasGupta (2001) and Pires, Amado (2008).

### 2. Alternatives of Confidence Intervals

**Clopper - Pearson interval.** The exact Clopper - Pearson interval (Clopper - Pearson, 1934) is based on inverting two-sided binomial tests on the null hypothesis  $H_0 : \pi = \pi_0$  against the alternative  $H_1 : \pi \neq \pi_0$ . If  $X = x$  is observed, the lower and upper bounds are the solutions of the equations

$$\sum_{k=0}^x \binom{n}{k} p_v^k (1 - p_v)^{n-k} = \frac{\alpha}{2} \tag{2}$$

$$\sum_{k=x}^n \binom{n}{k} p_L^k (1 - p_L)^{n-k} = \frac{\alpha}{2}$$

The lower and upper bounds of 100·(1-α)% Clopper - Pearson interval for  $0 < X < n$  are

$$P_L = \frac{x}{x + (n - x + 1)F_{1-\frac{\alpha}{2}}(2(n - x + 1), 2x)}, \tag{3}$$

$$P_U = \frac{(x + 1)F_{\frac{\alpha}{2}}(2(x + 1), 2(n - x))}{n - x + (x + 1)F_{\frac{\alpha}{2}}(2(x + 1), 2(n - x))}$$

where  $F_\alpha(k_1, k_2)$  is the  $\alpha$ -quantile of  $F$ -distribution with  $k_1$  and  $k_2$  degrees of freedom. For  $X = 0$  is  $p_L = 0$  and  $p_U = 1 - \left(\frac{\alpha}{2}\right)^{\frac{1}{n}}$ . For  $X = n$  is  $p_L = \left(\frac{\alpha}{2}\right)^{\frac{1}{n}}$  and  $p_U = 1$ .

**Wald interval.** Wald interval (Laplace, 1812) is based on inverting Wald test and is obtained by using the Central Limit Theorem

that the random variable  $U = (p - \pi) / \sqrt{\frac{p(1 - p)}{n}}$  is approximately standard normally distributed  $N(0,1)$ . Thus the lower and upper bounds of 100·(1-α)% Wald interval are

$$p_L = p - k_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1 - p)}{n}}, \tag{4}$$

$$p_U = p + k_{1-\frac{\alpha}{2}} \sqrt{\frac{p(1 - p)}{n}},$$

where  $k_\alpha$  is the  $\alpha$ -quantile of standard normal  $N(0,1)$ .

**Wilson score interval.** Wilson score interval (Wilson, 1927) is

obtained by noting that the random variable  $U = (p - \pi) / \sqrt{\frac{\pi(1 - \pi)}{n}}$  is approximately standard normally distributed  $N(0,1)$ . The lower and upper bounds are the solutions of the equation  $(p - \pi)^2 = k_\alpha^2 \frac{\pi(1 - \pi)}{n}$ . The lower and upper bounds of 100·(1-α)%

Wilson score interval for  $0 < X < n$  are

$$p_L = \frac{2np + k_{1-\frac{\alpha}{2}}^2 - k_\alpha \sqrt{4np(1 - p) + k_{1-\frac{\alpha}{2}}^2}}{2(n + k_{1-\frac{\alpha}{2}}^2)}, \tag{5}$$

$$p_U = \frac{2np + k_{1-\frac{\alpha}{2}}^2 + k_\alpha \sqrt{4np(1 - p) + k_{1-\frac{\alpha}{2}}^2}}{2(n + k_{1-\frac{\alpha}{2}}^2)},$$

where  $k_\alpha$  is the  $\alpha$ -quantile of standard normal  $N(0,1)$  distribution.

For  $X = 0$  is  $p_L = 0$ , for  $X = n$  is  $p_U = 1$ .

**Wilson score interval with continuity correction.** The continuity correction suggested by Blyth and Still (1983). The lower and upper bounds for  $0 < X < n$  are

$$p_L = \frac{np - \frac{1}{2} + \frac{1}{2}k_{1-\frac{\alpha}{2}}^2 - k_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{4}k_{1-\frac{\alpha}{2}}^2 + n\left(p - \frac{1}{2n}\right)\left(1 - p + \frac{1}{2n}\right)}}{n + k_{1-\frac{\alpha}{2}}^2}, \tag{6}$$

$$p_U = \frac{n\hat{p} + \frac{1}{2} + \frac{1}{2}k_{1-\frac{\alpha}{2}}^2 + k_{1-\frac{\alpha}{2}} \sqrt{\frac{1}{4}k_{1-\frac{\alpha}{2}}^2 + n\left(p + \frac{1}{2n}\right)\left(1 - p - \frac{1}{2n}\right)}}{n + k_{1-\frac{\alpha}{2}}^2},$$

where  $k_\alpha$  is the  $\alpha$ -quantile of standard normal  $N(0,1)$  distribution .

For  $X = 0$  is  $p_L = 0$ , for  $X = n$  is  $p_U = 1$ .

**Agresti - Coull interval.** Agresti and Coull (1998) introduced a slight modification of the Wald interval by adding two successes

and two failures. The point estimator of  $\pi$  is then  $p_w = \frac{X+2}{n+4}$ .

The lower and upper bounds of  $100 \cdot (1-\alpha)\%$  Agresti - Coull interval are

$$\begin{aligned} p_L &= p_w - k_{1-\frac{\alpha}{2}} \sqrt{\frac{p_w(1-p_w)}{n}}, \\ p_U &= p_w + k_{1-\frac{\alpha}{2}} \sqrt{\frac{p_w(1-p_w)}{n}}, \end{aligned} \quad (7)$$

where  $k_\alpha$  is the  $\alpha$ -quantile of standard normal distribution  $N(0,1)$ .

**Jeffreys interval.** This interval is based on the Bayesian approach. *Beta*-distribution is conjugate priors for binomial distribution. Let random variable  $X \sim Bi(n, \pi)$  and  $\pi \sim Beta(k_1, k_2)$ . Then the posterior distribution of  $\pi$  is  $Beta(x + k_1, n - x + k_2)$ . Thus  $100 \cdot (1-\alpha)\%$  Bayesian interval is

$$\begin{aligned} p_L &= Beta_{\frac{\alpha}{2}}(x + k_1, n - x + k_2), \\ p_U &= Beta_{1-\frac{\alpha}{2}}(x + k_1, n - x + k_2). \end{aligned}$$

It is known that non-informative Jeffreys prior is  $Beta\left(\frac{1}{2}, \frac{1}{2}\right)$ .

Then the lower and upper bounds of  $100 \cdot (1-\alpha)\%$  Jeffreys interval are

$$\begin{aligned} p_L &= Beta_{\frac{\alpha}{2}}\left(x + k_1, n - x + \frac{1}{2}\right), \\ p_U &= Beta_{1-\frac{\alpha}{2}}\left(x + \frac{1}{2}, n - x + \frac{1}{2}\right). \end{aligned} \quad (8)$$

where  $Beta(k_1, k_2)$  is the  $\alpha$ -quantile of *Beta*-distribution with  $k_1$  and  $k_2$  degrees of freedom.

For  $X = 0$  is  $p_L = 0$ , for  $X = n$  is  $p_U = 1$ .

### 3. Criteria for Comparing the Confidence Intervals

In this section we introduce the criteria that are used for comparing the confidence intervals.

**Coverage Probability.** For the fixed values  $n$  and  $\pi$  the coverage probability is the probability that the confidence interval  $CI(X, n)$  contains the parameter  $\pi$ . The coverage probability is defined for the given  $n$  and  $\pi$  as

$$C(n, \pi) = P_\pi(\pi \in CI(X, n)) = \sum_{x=0}^n I(x, \pi) \binom{n}{x} \pi^x (1-\pi)^{n-x} \quad (9)$$

where  $0 < \pi < 1$ ,  $I(x, \pi) = \begin{cases} 1 & \text{if } \pi \in CI(X, n) \\ 0 & \text{if } \pi \notin CI(X, n) \end{cases}$ .

The confidence interval is strictly conservative, if  $C(n, \pi) \geq 1 - \alpha$  for all  $\pi$ .

Due to the discrete nature of the binomial distribution the coverage probability can not be exactly equal to the nominal level  $(1 - \alpha)$  at all possible values. Therefore, our goal is to construct

such a confidence interval for the parameter  $\pi$  that the coverage probability is near the nominal level  $(1 - \alpha)$ .

**Average Coverage Probability.** The average coverage probability (*AVEC*) for the uniform averaging of the parameter values is defined as

$$AVEC(n) = \int_0^1 C(n, \pi) d\pi. \quad (10)$$

The confidence interval is conservative on average, if  $AVEC(n) \geq 1 - \alpha$ .

**Expected Length.** The expected length of the confidence interval is defined as

$$EL(n, \pi) = \sum_{x=0}^n [p_U(x, n) - p_L(x, n)] \binom{n}{x} \pi^x (1-\pi)^{n-x} \quad (11)$$

where  $p_L(x, n)$ ,  $p_U(x, n)$  are lower and upper bounds of a particular confidence interval.

This criterion measures the confidence interval length. In addition to the coverage probability the interval length is important for evaluating the confidence interval. The confidence interval is better if it has a shorter expected interval length with the similar performance of the other criteria.

**Average Expected Length.** The average expected length (*AVEL*) of the confidence interval is defined as

$$AVEL(n) = \int_0^1 EL(n, \pi) d\pi. \quad (12)$$

**Root Mean Square Error.** The root mean square error (*RMSE*) is defined as

$$RMSE(n) = \sqrt{\int_0^1 |C(n, \pi) - (1 - \alpha)|^2 d\pi} \quad (13)$$

This criterion is used to describe how far the coverage probability typically falls from the nominal level. This criterion measures the variability of coverage probability about the nominal level  $(1 - \alpha)$ .

### 4. Comparison of Confidence Intervals

In this section we demonstrate the performance of the confidence intervals which are compared in terms of the criteria mentioned above. The coverage probability, conservatism and interval length are important for evaluating the confidence intervals. To evaluate and compare the performance of confidence intervals the coverage probability was computed in 2001 values equally spaced in the interval  $(0,1)$  for  $n = 1$  to 1 000 and for  $\alpha = 0.05$ . The calculations were performed in Matlab. As it is impossible to analyze a large number of plots we summarize the results for the coverage probability in terms of the observed minimum coverage probability and *AVEC*. The confidence intervals were grouped into two classes of acceptable intervals:

1. **strictly conservative intervals** – intervals that have the minimum coverage probability at the least nominal level  $(1 - \alpha)$ , for all  $n$  and for all  $\pi : \min_{\pi} C(n, \pi) \geq 1 - \alpha$ .
2. **not strictly conservative intervals, but conservative on average** – intervals whose average coverage probability is at the least nominal level  $(1 - \alpha)$ , for all  $n : AVEC(n) \geq 1 - \alpha$ .

In the first class such a confidence interval is ideal whose minimum coverage probability is equal or a little above the nominal

level  $(1 - \alpha)$ . In the second class such a confidence interval is ideal whose AVEC is equal or a little above the nominal level  $(1 - \alpha)$  and minimum coverage probability is little below the nominal level  $(1 - \alpha)$ . A shorter expected length and a smaller average expected length is preferred.

**Coverage probability.** Fig. 1 shows the coverage probabilities of 95% confidence intervals for the case  $n = 50$ . The figures for other values of  $n$  are similar to this figure. It is evident why the Wald performs poorly and why the Clopper-Pearson is known as

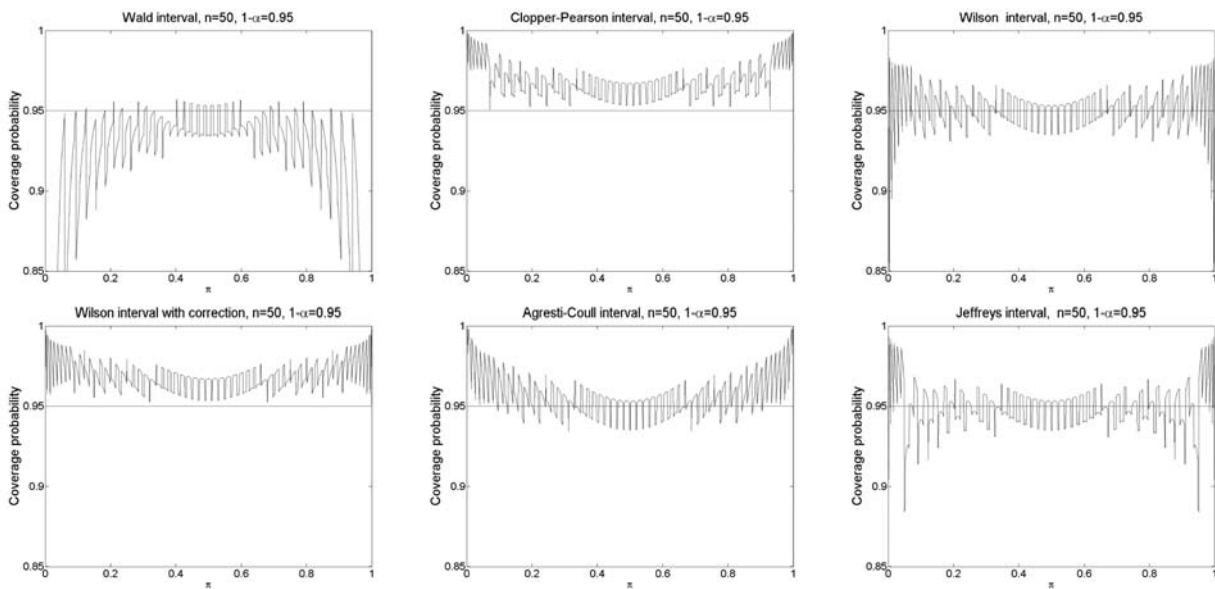


Fig. 1 Coverage probability of 95% confidence intervals for  $n = 50$

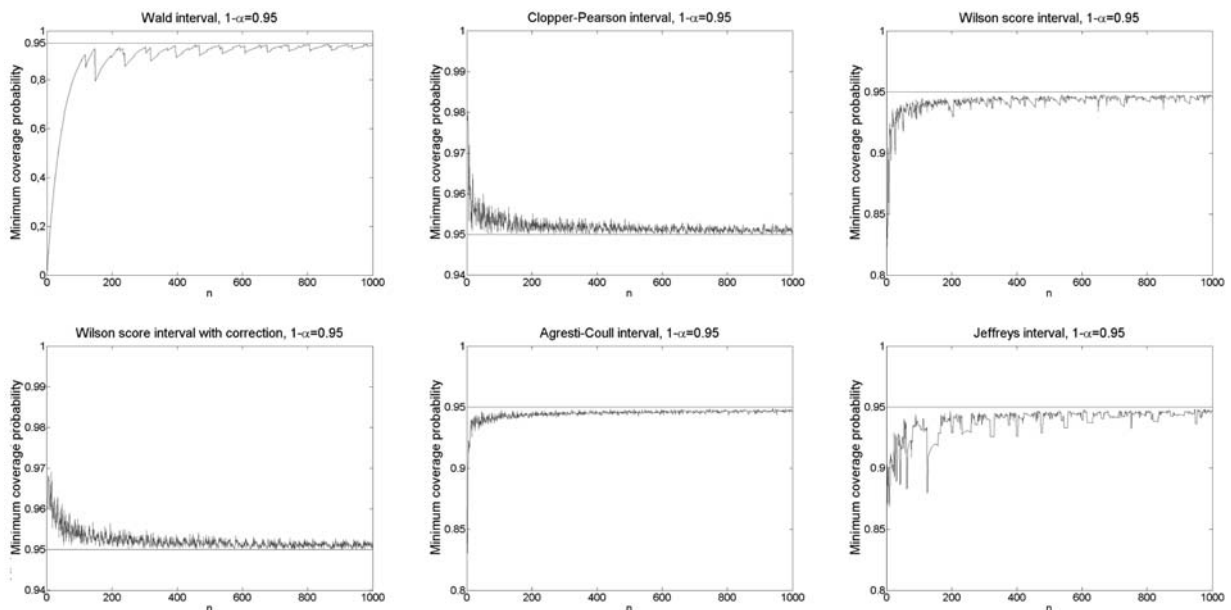


Fig. 2 Minimum coverage probability of 95% confidence intervals for  $n = 1$  to  $1000$

an overly conservative interval. The Clopper-Pearson guarantees that the coverage probability is always equal to or above the nominal level  $(1 - \alpha)$ . The coverage probability of the Wald is very poor for  $\pi$  near boundaries 0 and 1. The problems with coverage probability exist even for  $n$  large. This interval has more chaotic properties and can not be used (Brown, Cai, DasGupta, 2001). The Wilson has coverage probability which fluctuates near the nominal level  $(1 - \alpha)$ . As  $n$  gets larger it comes to the significant improvement. The coverage probability near to boundaries 0 and 1 is problematic. The Wilson+CC falls into conservative intervals, with performance similar to the Clopper-Pearson. In comparison to the Clopper-Pearson, the Clopper-Pearson is more conservative for  $\pi$  near 0 and 1. The Agresti-Coull is even more conservative especially for  $n$  small. In comparison to the Wilson, the coverage probability is as good as the Wilson, but the Agresti-Coull is quite conservative for  $\pi$  near the boundaries. The Jeffreys has a coverage probability qualitatively similar to the Wilson. Its coverage probability is reasonable, except for the very deep spikes near 0 and 1. As  $n$  gets larger it comes to the improvement. Fig. 2 shows the minimum coverage probabilities of 95% confidence intervals for  $n = 1$  to 1 000.

**Average coverage probability.** Fig. 3 shows the AVEC for  $n = 1$  to 100. As it is showed in the figure the AVEC of the Wald tends to be under other intervals and under the nominal level  $(1 - \alpha)$  for all values of  $n$ . The Clopper-Pearson and the Wilson+CC are comparable intervals and are too conservative on average. The Clopper-Pearson tends to be higher than other intervals and above the nominal level  $(1 - \alpha)$  for all values of  $n$ . The Agresti-Coull is slightly conservative on average. The Wilson and the Jeffreys have the AVEC quite close the nominal level  $(1 - \alpha)$  for most of values  $n$ , the Wilson even for  $n$  small.

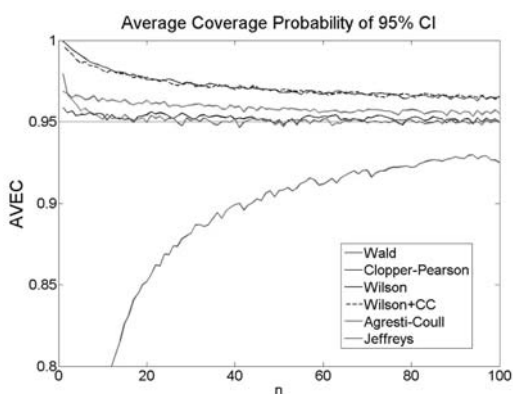


Fig. 3 Average coverage probability of 95% confidence intervals for  $n = 1$  to 100

**Expected length.** Fig. 4 shows the expected lengths of 95% confidence intervals for case  $n = 50$ . The figures for other values of  $n$  are similar to this figure. It is evident that the Clopper-Pearson and the Wilson+CC are wider than other intervals. For  $\pi$  moderate the Wilson+CC is shorter. For  $\pi$  small or large the Wald is the shortest. The Jeffreys and the Wilson are comparable intervals

with a relatively small width. The Jeffreys performs better. The Agresti-Coull is quite wide. For  $\pi$  moderate the Agresti-Coull, the Wilson and the Jeffreys are narrower than other intervals.

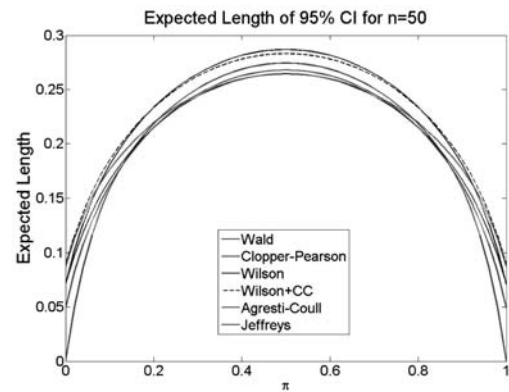


Fig. 4 Expected length of 95% confidence intervals for  $n = 50$

**Average expected length.** Fig. 5 shows the AVEL of 95% confidence intervals for  $n = 1$  to 100. As it is showed in the figure the Clopper-Pearson and the Wilson+CC are comparable intervals and their AVEL is the biggest of all the intervals. The Wilson and the Jeffreys are comparable intervals. In comparison to them the Agresti-Coull is larger for  $n$  small. From the given figure it is evident that as  $n$  gets larger the difference between intervals starts to wear off.

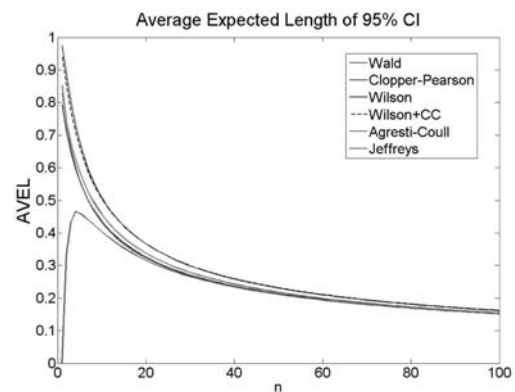


Fig. 5 Average expected length of 95% confidence intervals for  $n = 1$  to 100

**Root mean absolute square error.** Fig. 6 shows the RMSE of 95% confidence intervals for  $n = 1$  to 100. It is evident that the RMSE of the Wald is much larger than the other intervals. The Clopper-Pearson and the Wilson+CC are comparable intervals, the RMSE of the Clopper-Pearson is slightly larger than the Wilson+CC. The RMSE of the Jeffreys and the Agresti-Coull are comparable. The Wilson has the smallest RMSE.

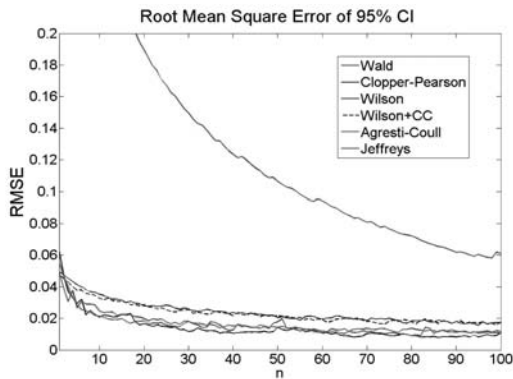


Fig. 6 Root mean square error of 95% confidence intervals for  $n = 1$  to  $100$

The comparison of 95% confidence intervals in terms of minimum coverage probability (MCP), average coverage probability (AVEC), root mean square error (RMSE) and average expected length (AVEL) for  $n = 10, 30, 50, 100, 500, 1000$ .

Table 1.

	Methods	MCP	AVEC	RMSE	AVEL
$n=10$	Clopper-Pearson	0.9610	0.9837	0.0349	0.5085
	Wald	0.0050	0.7696	0.2858	0.4035
	Wilson	0.8382	0.9540	0.0218	0.4354
	Wilson+CC	0.9511	0.9819	0.0330	0.5043
	Agresti-Coull	0.9168	0.9638	0.0217	0.4565
	Jeffreys	0.8681	0.9531	0.0244	0.4326
$n=30$	Clopper-Pearson	0.9506	0.9734	0.0256	0.2990
	Wald	0.0149	0.8753	0.1674	0.2663
	Wilson	0.8475	0.9525	0.0141	0.2708
	Wilson+CC	0.9528	0.9730	0.0245	0.2998
	Agresti-Coull	0.9339	0.9599	0.0163	0.2789
	Jeffreys	0.8887	0.9505	0.0172	0.2687
$n=50$	Clopper-Pearson	0.9509	0.9692	0.0215	0.2306
	Wald	0.0247	0.9009	0.1299	0.2113
	Wilson	0.8392	0.9516	0.0118	0.2129
	Wilson+CC	0.9512	0.9693	0.2008	0.2313
	Agresti-Coull	0.9345	0.9578	0.0141	0.2177
	Jeffreys	0.8842	0.9502	0.0135	0.2117
$n=100$	Clopper-Pearson	0.9504	0.9646	0.0169	0.1614
	Wald	0.0488	0.9226	0.0915	0.1518
	Wilson	0.8606	0.9510	0.0087	0.1523
	Wilson+CC	0.9512	0.9647	0.0164	0.1619
	Agresti-Coull	0.9390	0.9664	0.0113	0.1543
	Jeffreys	0.8806	0.9499	0.0098	0.1517
$n=500$	Clopper-Pearson	0.9503	0.9573	0.0091	0.0706
	Wald	0.2212	0.9434	0.0382	0.0687
	Wilson	0.9099	0.9504	0.0041	0.0687
	Wilson+CC	0.9502	0.9573	0.0087	0.0707
	Agresti-Coull	0.9448	0.9519	0.0059	0.0689
	Jeffreys	0.9145	0.9499	0.0046	0.0686
$n=1000$	Clopper-Pearson	0.9501	0.9551	0.0068	0.0496
	Wald	0.3934	0.9464	0.0243	0.0486
	Wilson	0.9098	0.9501	0.0041	0.0486
	Wilson+CC	0.9501	0.9552	0.0070	0.0496
	Agresti-Coull	0.9453	0.9511	0.0045	0.0487
	Jeffreys	0.9172	0.9499	0.0040	0.0486

### 5. Concluding Remarks

In this section we summarize the classification and performance of the alternatives of confidence intervals.

The Wald interval should not be used. It performs poorly in terms of the coverage probability and the RMSE, though the expected length is short. In comparison to the Wald all the intervals mentioned above outperform the Wald.

From the alternatives of confidence intervals mentioned above only the Clopper-Pearson is strictly conservative. The Clopper-Pearson guarantees the minimum coverage probability which is equal to or is above the nominal level. This interval is too conservative on average, too wide and has the larger RMSE.

The other confidence intervals the Wilson, the Wilson+CC, the Agresti-Coull and the Jeffreys are not strictly conservative, but are conservative on average.

The Wilson and the Jeffreys have similar properties such as a relatively small length, comparable AVEL and RMSE. The Wilson has excellent properties, the coverage probability near the nominal level, except for the problems with the coverage probability for values near 0 and 1 that makes a very low minimum coverage probability. Similar problems with the minimum coverage probability exist for the Jeffreys due to unlucky deep spikes near boundaries 0 and 1. But otherwise the Jeffreys has also good properties.

The Agresti-Coull has the minimum coverage probability better than others. To compare it to the intervals in this class, except for the Wilson+CC, the Agresti-Coull is slightly conservative and wider on average, but its advantages are easy calculation and presentation.

The Wilson+CC is similar to the Clopper-Pearson. It is too conservative on average, wide and has the larger RMSE. This interval is almost strictly conservative. The coverage probability for some values that are near boundaries 0 and 1 is slightly below the nominal level. (For example, for  $n = 15, \pi = 0.003485, \alpha = 0.05$  is  $C(n, \pi) = 0.9490$ .)

Which method should be used in practical applications? The choice from the alternatives depends on the situation where they should be used and on preferences of users. The strictly conservative Clopper-Pearson is a choice for a situation when the coverage probability must be guaranteed to be equal to or above the nominal level. Otherwise if strict conservativeness is not a major criterion, the preference is to use the confidence intervals which are conservative on average, and their coverage probability is quite close the nominal level and are narrower. The almost strict Wilson+CC is also a valid choice. The Jeffreys is also an appropriate choice for practice but it is more complicated to compute. Considering properties of alternatives of the confidence intervals the Wilson and the Agresti-Coull are the best choice in this class. They perform very well and are simple to compute.

These recommended confidence intervals are much better to guarantee the estimation of a binomial proportion when compared with the standard and frequently used Wald interval.

#### Acknowledgement

This paper was supported by the Grant VEGA No. 1/0249/09.

#### References

- [1] AGRESTI, A., COULL, D. A.: *Approximate is Better than "Exact" for Interval Estimation of Binomial Proportion*. American Statistician 52, p. 119-126, 1998
- [2] BLYTH, C. R., STILL, H. A.: *Binomial Confidence Intervals*. Journal of the American Statistical Association 78, pp. 108-116, 1983
- [3] BROWN, D. L., CAI, T. T., DASGUPTA, A.: *Interval Estimation for a Binomial Proportion*. Statistical Science 16, pp. 101-133, 2001
- [4] CLOPPER, C. AND PEARSON, S.: *The use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial*. Biometrika 26, pp. 404-413, 1934.
- [5] FENG, X.: *Confidence Intervals for Proportions with Focus on the US National Health and Nutrition Examination Survey*. Master thesis, Simon Fraser University, 2006.
- [6] LAPLACE, P. S.: *Theorie Analytique des Probabilities*. Paris, France, Courier, 1812
- [7] NEWCOMBE, R. G.: *Two-sided Confidence Intervals for the Single Proportion; Comparison of Several Methods*. Statistics in Medicine 17, pp. 857-872, 1998
- [8] PIRES, M. A., AMADO, C.: *Interval Estimators for a Binomial Proportion: Comparison of Twenty Methods*. REVSTAT - Statistical Journal, Volume 6, Number 2, p. 165-197, 2008
- [9] POBOČÍKOVÁ, I.: *Confidence Intervals for a Binomial Proportion*. Proc. 8th International Conference APLIMAT, Bratislava, February 3.-6. , pp. 791-800, 2009
- [10] REICZIGEL, J.: *Confidence Intervals for a Binomial Proportion: Some New Considerations*. Statistics in Medicine 22, pp. 611-621, 2003
- [11] VOLLSET, S. E.: *Confidence Intervals for a Binomial Proportion*. Statistics in Medicine 12, pp. 809-824, 1993
- [12] WILSON, E. B.: *Probable Inference, the Law of Succession, and Statistical Inference*. Journal of the American Statistical Association 22, pp. 209-212, 1927