Original Article

MODERN SOLUTIONS TO CLASSICAL PROBLEMS: INNOVATIONS IN CONFIDENCE INTERVAL METHODS

Alicia Monique Barnes and Helena Grace Oduro

Department of Mathematics,
Florida Agricultural and Mechanical University, Florida, USA
DOI: https://doi.org/10.5281/zenodo.17152326

Abstract: Statistical inference frequently involves estimating confidence intervals for binomial parameters, especially the proportion ppp. Among the most commonly used techniques is the Wald interval, which utilizes the sample proportion p^\hat{p}p^, the sample size nnn, and the standard normal quantile zαz_\alphazα. Despite its simplicity, the Wald interval is known for its poor performance with small sample sizes and when ppp is near 0 or 1, often leading to inaccurate coverage probabilities.

To overcome these limitations, a range of alternative methods has been proposed. The Clopper-Pearson "exact" interval ensures a minimum coverage probability of 1−α1 - \alpha1−α for all values of ppp, though it tends to be conservative. The Score interval, introduced by Wilson and refined by subsequent researchers like Guan, offers improved accuracy and stability. Bayesian approaches, including those based on non-informative priors such as the Jeffreys prior, also provide flexible and effective solutions for constructing intervals. Additional techniques like the Arcsin and Logit transformations further expand the set of tools available for inference on binomial proportions.

This article reviews these key methods, comparing their theoretical properties, practical strengths and weaknesses, and applicability to real-world statistical problems. Emphasis is placed on understanding when and why certain methods are preferable, depending on factors like sample size, target coverage level, and the range of the proportion being estimated. Through this comparative analysis, the study highlights the intricacies involved in constructing reliable confidence intervals for binomial data and related linear functions.

Keywords: Confidence intervals, binomial parameter, Wald interval, Clopper-Pearson interval, Score method, Bayesian methods.

Original Article

1. Introduction:

A basic analysis in statistical inference is constructing a confidence interval for a binomial parameter p. The simplest interval which is almost universally used is

$$\hat{p} \pm z_{\frac{\alpha}{2}}[\frac{\hat{p}(1-\hat{p})}{n}]^{\frac{1}{2}} \quad (1)$$

where p is the sample proportion, n is the sample size and $z\alpha$ denotes the $1-\frac{\alpha}{2}$ quantile of the standard normal distribution. For instance, α = 0.05 for a 95 % confidence interval, α = 0.10for a 90% confidence interval, etc. This interval is derived from the Wald large sample confidence interval and is commonly referred to as the Wald interval.

So it seems at first glance that the problem is simple and has a clear solution. Actually, the problem is a difficult one with several complexities. It is widely recognized that Wald interval coverage probability is poor for p near 0 or 1. It is known that the Wald interval performs poorly unless n is large by Blyth, C. R. and Still, H. A.(1983). Most statistics books take this into account by requiring that this interval should be used only when min (n , n (1 $-p$)) is at least 5 or 10 by Brown, L. D., Cai, T., and Das Gupta, A.(2001).

A considerable literature exists about this and other less common methods for constructing a confidence interval for p. By Santner, T. J. (1998), and Vollset, S. E.(1993), reviewed a variety of methods. One of the methods is the Clopper-Pearson "exact" interval by Clopper, C. J. and Pearson, E. S.(1934). This method is widely used and has the advantage of a coverage probability of at least $1-\alpha$ for every possible value of p. The Score method by Wilson, E. B. (1927) discussed by Agresti, A., and Coull, B. A.(1998), is arguably the best procedure for constructing a confidence interval for a population proportion. Guan,Yu (2012) introduced the generalized score method which computes easily and reduces the spike fluctuations of the score method. Also, Bayesian methods are effective for constructing confidence intervals for a population proportion. In addition, other effective procedures such as the Arcsin, Logit and Jeffres prior intervals are discussed in Brown, L. D., Cai, T., and Das Gupta, A. (2001). The Jeffres prior interval is a special case of a Bayes procedure with a non-informative prior. Bayes procedures with a non- informative prior have a good track record in constructing confidence intervals for p; described by Wasserman, L. (1991). Wang, W. (2006) discusses methods for constructing the smallest exact confidence intervals. Zou, G. Y., Huang, W., and Zhang, X. (2009) use the Score interval to construct a confidence interval for a linear function of binomial proportions.

However, most effective procedures are too complicated to use in an introductory statistics course. Therefore Agresti, A., and Coull, B. A. (1998) introduced the Adjusted Wald (AC) procedure. The AC method consists of adding two successes and two failures to the data and then proceeding as in the Wald interval. This method is simple, easy to use and accurate. The accuracy of the AC procedure is due to its midpoint and width being almost the same as those of the Score procedure. Actually, the Adjusted Wald (AC) interval is a simplified version of the Score interval.

At the present time the Wald interval is almost exclusively used in everyday practical statistics. Some reasons for its popularity are that it is easy to motivate and easy to use. Under the right conditions such as np (1 −) ≥ 1 0, it is reasonably accurate.

Original Article

We agree with Brown, L. D., Cai, T., and Das Gupta, A.(2001) (it is generally true that only those methods that are easy to describe, use and compute will be widely used). The purpose of this article is to present an easy to describe, use and compute alternative to the Wald procedure.

2. The T- Wald:

As mentioned earlier the Wald confidence interval procedure gives satisfactory results only under right conditions such as $np(1-\ ) \geq 1\ 0$. Under these conditions, the sample proportion, p is approximately normally distributed. So there is a reason to use the procedure for constructing a confidence interval for the mean of a normal population with unknown variance. This inverval is

$$\bar{x} \pm t\frac{S}{\sqrt{n}} \qquad (2)$$

3. An Improved Confidence Interval:

In equation (2), x is the sample mean and S is the sample standard deviation and n is the sample size. In binomial notation, equation (2) becomes

$$\hat{p} \pm t\sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}} \qquad (3)$$

where p is the proportion of successes in the sample.

To see the conversion of equation (2) to (3), consider a sample consisting of zeros and ones from a binomial population. Note that a zero is a failure and one is a success. Let x be an observation and y be the number of success in the sample. Then S^2, the sample variance is

$$S^2 = \frac{\sum x^2 - (\sum x)^2}{n}\left(\frac{1}{n-1}\right)$$

$$= n\left[\frac{y}{n} - \left(\frac{y}{n}\right)^2\right]\left(\frac{1}{n-1}\right)$$

$$= \frac{n[\hat{p}(1-\hat{p})]}{n-1}$$

Thus

$$(S^2)^{\frac{1}{2}} = S = \sqrt{\frac{\hat{p}(1-\hat{p})n}{n-1}} \quad (4) \text{ and}$$

$$\left(\frac{S^2}{n}\right)^{\frac{1}{2}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}} \qquad (5)$$

So equation (3) is a new formula which we refer to as the T-Wald. It could be used as an alternative to the Wald interval.

One can think of the T-Wald as a simple modification of the Wald procedure with “n" replaced by “n-1" and “z" replaced by the “t" with $n-1$ degrees of freedom.

Both the Wald and T-Wald procedures are reasonably accurate for $min\,[p\ ,\ n(1-p)] \geq 1\ 0$. But the TWald is more accurate.

4. Average Coverage:

Original Article

All confidence intervals for a population proportion are dependent on p, the true population proportion. It would be nice if there was a 95% confidence interval procedure for p that would cover each value of p with a probability of exactly 0.95. However, no such procedure exists. The most one can hope for is that the average coverage of all p to be close to 0.95.

Thus, one way (but not the only way) of comparing different confidence interval procedures for p is to look at average coverage. This is shown in Table 1. In Table 1, p is restricted to values between 0.3 and 0.7. This insures that the sample proportion p will be approximately normally distributed.

Table 1 shows comparison of the average coverage of the Wa ld (CI_W), T − Wa ld (CI_T) for 95% confidence intervals with $0.3 < p < 0.7$ for different values of n.

Table 1: Several properties of confidence intervals for population proportion

n	12	24	48	96	192
CI_W	0.9111	0.9298	0.9400	0.9449	0.9475
CI_T	0.9405	0.9466	0.9484	0.9490	0.9495
ReEr	0.2442	0.1765	0.2072	0.1850	0.1863

Table 1 illustrates several properties of average coverage of confidence intervals for p, a population proportion. One is that the Wald interval (CI_W) is satisfactory as far as average coverage is concerned. But the T-Wald is much better. The row ReEr (relative error) is the relative error of the T-Wald and Wald. For instance the relative error for $n = 48$ is

$$E = z_{\frac{\alpha}{2}}[\frac{\hat{p}(1-\hat{p})}{n}]^{\frac{1}{2}}$$

$$\frac{0.95 - 0.9484}{0.95 - 0.9400} = 0.2072 \qquad n \qquad p$$

$$\bar{x} \text{ and } \frac{\hat{p}(1-\hat{p})}{n-1} \qquad \frac{s}{n}$$

Table 1 indicates that the relative error of the T-Wald relative to the Wald is roughly $\frac{1}{5}$ the relative error of the Wald. The reader must remember that both the T-Wald and the Wald should be used only for certain and combinations such as $np(1-) \geq 5$.

One might ask how does the average coverage of the T-Wald compare with certain other confidence interval procedures such as the Wilson(Score) or Adjusted Wald by Agresti, A., and Coull, B. A.(1998). Both of these procedures can be used for most combinations of n and p The Score procedure has probably the best average coverage of any confidence interval procedure. For instance the average coverage of the Score procedure for a 95% confidence interval with $n \geq 96$ and $0.3 < p < 0.7$ is 0.9501.

The average coverage of the adjusted Wald (AC) procedure is about the same as that of the T-Wald for a 95% confidence interval.

5. Discussion

A common analysis in statistical inference is forming a confidence interval for a binomial parameter p. The Wald interval is almost universally used because of its simplicity. We think a simple, easy to use procedure such as the T-Wald has a better chance of partially replacing the Wald. We think the T-Wald serves this purpose. The TWald also has the form $p \pm E$ so that it's center is , where, . Users seems to prefer this type of estimator. Finally, using the T-Wald procedure is the same as constructing a confidence interval for a population mean with the notation changed. That is, p replaces replaces .

6. Conclusion:

This new method is a good compromise between simplicity and accuracy.It is also a “natural” method to use since one used the same steps as those used in constructing a confidence interval as those used in constructing a confidence interval for a population mean with unknown variance.

References

Agresti, A.(1996),An introduction to Categorical Data Analysis, New York: Wiley.

Agresti, A., and Coull, B. A.(1998),Approximate is Better than Exact for Interval Estimation of Binomial Proportions, The American Statistician, 52, 119-126.

Blyth, C. R. and Still, H. A.(1983), Binomial Confidence Intervals, Journal of the American Statistical Association, 78, 108-116.

Brown, L. D., Cai, T., and Das Gupta, A.(2001), Interval Estimation for a Binomial Proportion, Statistical Science, 16, 101-133.

Cai, T.T.(2005), One-sided Confidence Intervals in Discrete Distributions, Journal of Statistical Planning and Inference, vol 131, 63-88.

Clopper, C. J. and Pearson, E. S.(1934), The Use of Confidence or Fiducial Limits Illustrated in the Cas of the Binomial, Biometrika, 26, 404-413.

Guan,Yu (2012), A generalized score CI for a binomial proportion, Journal of Statistical Planning and Inference, Vol 142, 785- 793.

Newcombe, R. G.(1998), Two-sided Confidence Intervals for the Single Proportion: Comparison of Seven Methods, Statistics in Medicine, 17, 857-873.

Santner, T. J.(1998), A Note on Teaching Binomial Confidence Intervals , Teaching Statistics, 20, 20-23.

Triola, M. F.(2006), Elementary Statistics, 10th edition.

Vollset, S. E.(1993), Confidence intervals for a binomial proportion, Statistics in Medicine, 12, 809-824.

Wei, Yu, Xu ,Guo, and Wangli (2012), An Improved Score Interval with a Modified Mid-Point for a Binomial Proportion, Journal of Statistical Planning and Inference, vol 142, 785-793.

Wang, W.(2006), Smallest confidence intervals for one binomial proportion, Journal of Statistical Planning and Inference, vol 136, 4293-4306.

Wasserman, L.(1991), An inferential interpretation of default priors, Technical report, Carnegie-Mellon University.

Wilson, E. B.(1927), Probable Inference the Law of Succession and Statistical Inference, Journal of the American Statistical Association, 22, 209-212.

Original Article

Zou, G. Y., Huang, W., and Zhang, X.(2009), A note on confidence interval estimation for a linear function of binomial proportions, Computational Statistics and Data Analysis, 53, 1080-1085.