International Journal of Innovative Research and Advanced Studies (IJIRAS)
Volume 1 Issue1, December 2014

Comparison Of Different Interval Estimatorsfor The Binomial Proportion *P* Using Root Mean Squared Error (Rmse) Criterion

Prof. B. J. Guillen, MD
Professor, Iligan City

***Abstract*: This study proposed a quadratic-basedinterval estimator for binomial proportion*p*. This modified quadratic-based interval was compared to the different existing alternative interval estimators through empirical analysis using Root Mean Squared Error (RMSE) for various values of *n*, *p* and α = 0.05.**

Generated results showsthat the modified interval has a desirable RMSE property compared to the standard and non-modified intervals and the RMSE behavior of the modified and the alternative methods decreases as *n* increases. As observed and manifested, the study suggested that modified method is an enhancement of the standard method.

Thus, it is recommended to modify other existing alternative methods in such a way that there's an enhancement in performance in terms of RMSE and other evaluation criterion.

***Keywords*: Confidence Interval, Binomial Distribution, Standard Interval, Root Mean Squared Error**

I. INTRODUCTION

Interval estimation in the binomial proportion *p* is one of the classical problems in statistics offering many challenges. When someone constructing a confidence interval, one usually desires that the actual coverage probability is to be close to the nominal confidence level, that is, it closely approximates to $1-\alpha$. The unexpected difficulties inherent to the choice of a confidence interval estimate of the binomial parameter *p*, and the relative *inefficiency* (Marchand, et al., 2004) of the "standard" Wald confidence interval, has explored much in the previous literature (Brown, et. Al., 1999a, 1999b, 2001, and Agresti and Coull, 1998).

Some alternative intervals make use of a continuity correction while others guarantee a minimum $1-\alpha$ coverage probability for all values of the parameter *p*.

This study aims to develop an alternative method with some modifications of the continuity correction. This modification imposes a continuity correction factor.

A. OBJECTIVE OF THE STUDY

The objective of this study is to develop a non-randomized confidence interval $C(X)$ for *p*, such that the coverage probability $P_p(p \in C(x)) \approx 1-\alpha$, where α is some pre-specified value between 0 and 1 (Brown, et al., 2001). Specifically, the objective of this study is to compare numerically the performance of the standard and modified intervals and some alternative interval estimators basedon Root Mean Squared Error (RMSE).

II. BASIC CONCEPTS

A. CONFIDENCE INTERVAL

Definition 1 Let $X_1, X_2, \ldots X_n$ be a random sample from the density $f(x|\theta)$. Let $l(x) = l(x_1, x_2, \ldots x_n)$ and $u(x) = u(x_1, x_2, \ldots x_n)$ be two statistics satisfying $l(x) \le u(x)$ for which $P_\theta[l(x) < \theta < u(x)] = 1-\alpha$. Then the random interval $(l(x), u(x))$ is called a $100(1-\alpha)\%$ confidence interval for θ; $1-\alpha$ is called the confidence coefficient; and $l(x)$ and $u(x)$ are called the lower and upper confidence limits, respectively, for θ.

Definition 2 The root mean squared error (RMSE) is defined as:

$$RMSE_n = \left\{ \int_0^1 \left(P_\theta(\theta \in C(X)) - (1-\alpha)\right)^2 dp \right\}^{1/2},$$

which measures the fluctuation of the coverage probability around the nominal $1-\alpha$ level.

III. INTERVAL ESTIMATORS

A. STANDARD INTERVAL ESTIMATOR

A standard confidence interval for *p* based on normal approximation has gained universal recommendation in the introductory statistics textbooks and in statistical practice. The interval is known to guarantee that for any fixed *p*, the coverage probability $P(p \in C(x)) \to 1-\alpha \; as \; n \to \infty$.

Let $z = z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$, $\hat{p} = \frac{x}{n}$ and $\hat{q} = 1-\hat{p}$, where $\hat{p}+\hat{q} = 1$.The normal theory approximation of a confidence interval for binomial proportion is defined as:

$$C(X)_s = \hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

wherez is the $(1-\alpha/2)$thquantile of the standard normal distribution (Wald and Wolfowitz, 1939).

B. NON-MODIFIED INTERVAL ESTIMATOR

The approximate $1-\alpha$ confidence interval $[L(x,n), U(x,n)]$ is given by:

$$C(X) = \frac{2\hat{p} + \frac{z_{\alpha/2}^2}{n} \pm \left[\left(2\hat{p} + \frac{z_{\alpha/2}^2}{n}\right) - 4\hat{p}^2\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

, where

the lower limit and upper limit are

$$L(x,n) = \frac{2\hat{p} + \frac{z_{\alpha/2}^2}{n} - \left[\left(2\hat{p} + \frac{z_{\alpha/2}^2}{n}\right) - 4\hat{p}^2\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

$$U(x,n) = \frac{2\hat{p} + \frac{z_{\alpha/2}^2}{n} + \left[\left(2\hat{p} + \frac{z_{\alpha/2}^2}{n}\right) - 4\hat{p}^2\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

C. PROPOSED MODIFIED INTERVAL ESTIMATOR

Due to the discreteness of the binomial distribution and as suggested by Casella, et al., 1990), this proposed modified interval imposes a continuity correction, $c = \frac{1}{4n}$, over the modified interval. The factor is arbitrarily chosen.

Result: The approximate $1-\alpha$ confidence interval for *p* with $c = \frac{1}{4n}$ is given by

$$C(X) = \frac{\left(2\hat{p} \pm \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) \pm \left[\left(2\hat{p} \pm \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) - 4\left(\hat{p}^2 \pm \frac{1}{4n}\right)\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

where the lower limit is given by,

$$L(x,n) = \frac{\left(2\hat{p} - \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) - \left[\left(2\hat{p} - \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) - 4\left(\hat{p}^2 - \frac{1}{4n}\right)\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

and upper limit is given by

$$U(x,n) = \frac{\left(2\hat{p} + \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) + \left[\left(2\hat{p} + \frac{1}{2n} + \frac{z_{\alpha/2}^2}{n}\right) - 4\left(\hat{p}^2 + \frac{1}{4n}\right)\left(1+\frac{z_{\alpha/2}^2}{n}\right)\right]^{\frac{1}{2}}}{2\left(1+\frac{z_{\alpha/2}^2}{n}\right)}$$

D. ALTERNATIVE INTERVAL ESTIMATORS

It is very clear that the standard interval is just too problematic as partially mentioned in the preceding section. This brings us to consideration of some alternative intervals.

3.4.1 Arcsine Interval

This interval is based on a widely used variance stabilizing transformation for the binomial distribution (Bickel and Doksum, 1977: $T(\hat{p}) = \arcsin e(\hat{p}^{1/2})$). This variance stabilization is based on the delta method and is, of course only an asymptotic one.

Anscombe (1948) showed that replacing $\hat{p}$ by $\breve{p} = (X+3/8)/(n+3/4)$ gives better variance stabilization; furthermore

$$2n^{1/2}\left[\arcsin(\breve{p}^{1/2}) - \arcsin(p^{1/2})\right] \to N(0,1) \quad as \; n \to \infty$$

This leads to an approximate $100(1-\alpha)\%$ confidence interval for p, given as:

$$C(X)_{Arc} = \left[\sin^2\left(\arcsin(\breve{p}^{1/2}) - 1/2zn^{-1/2}\right), \sin^2\left(\arcsin((\breve{p}^{1/2}) + 1/2zn^{-1/2})\right)\right]$$

3.4.2Agresti – Coull Interval

Agresti and Coull (1998) proposed a method whose form is quite similar to the standard interval but with a new choice of $\hat{p}$. They consider $\tilde{x} = x + \frac{z^2}{2}$ and $\tilde{n} = n + z^2$. Let $\tilde{p} = \tilde{x}/\tilde{n}$ and $\tilde{q} = 1-\tilde{p}$. Then, define the confidence interval $C(X)_{AC}$ by $C(X)_{AC} = \tilde{p} \pm z(\tilde{p}\tilde{q})^{1/2}\tilde{n}^{-1/2}$

For the case when $\alpha = 0.05$ and replacing a value of 2 instead 1.96 for *z*, the interval becomes “add two successes and two failures” interval in Agresti and Coull (1998).

3.4.3 Wilson Score Interval

The Wilson interval is an improvement over the normal approximation interval and was first developed in Wilson (1927). This interval was based on inverting such that $\left|(\hat{\theta}-\theta)/se(\hat{\theta})\right| \le z$. This confidence interval has the form $C(X)_W = \frac{X + z^2/2}{n+z^2} \pm \frac{z\sqrt{n}}{n+z^2}\sqrt{\hat{p}\hat{q} + \frac{z^2}{4n}}$

wherez is the $z_{\alpha/2}$ quantile of the standard normal distribution.

3.4.4 Wilson Score Interval with Continuity Correction

The score (or test-based) interval (Wilson (1927)) has been recommended often over all alternative approximate intervals for the binomial parameter *p*. Blyth and Still (1983),Santner and Duffy (1989) and Vollset (1993) specifically recommend a modified score interval with continuity correction $1/2$.A modified score $100(1-\alpha)\%$ confidence interval with continuity correction c_c and Normal critical point *z* is:

$$SBS(c_c) = \frac{(x \pm c_c) + \frac{z^2}{2} \pm z\left[(x \pm c_c) - \frac{(x \pm c_c)^2}{n} + \frac{z^2}{4}\right]^{1/2}}{n+z^2},$$

except that for $x = 0$, the lower limit is zero; for $x = n$, the upper limit is one; for $x = 1$, the lower limit is $1-(1-\alpha)^{1/n}$ and for $x = n-1$, the upper limit is $(1-\alpha)^{1/n}$.

3.4.5 Wald Logit Interval

This interval is obtained by inverting a Wald type interval for the log odds $\lambda = \log\left(\frac{p}{1-p}\right)$ (Stone, 1995). The MLE of $\lambda(\text{for } 0 < X < n)$ is $\hat{\lambda} = \log\left(\frac{\hat{p}}{1-\hat{p}}\right) = \log\left(\frac{X}{n-X}\right)$,which is called the empirical logit transformation. The variance of $\hat{\lambda}$, by an application of the delta theorem, can be estimated by $\hat{V} = \frac{n}{X(n-X)}$ This leads to an approximate $100(1-\alpha)\%$ confidence interval for λ, $C(X)_\lambda = [\lambda_l, \lambda_u] = \left[\hat{\lambda} - z\hat{V}^{1/2}, \hat{\lambda} + z\hat{V}^{1/2}\right]$. The logit interval for *p* is obtained by inverting the preceding interval, $C(X)_{\text{logit}} = \left[\frac{e^{\lambda l}}{1+e^{\lambda l}}, \frac{e^{\lambda u}}{1+e^{\lambda u}}\right]$.

IV. RESULTS AND DISCUSSION

This section presents the comparative analysis graphicallyof the different interval estimators in terms of RMSE. In investigating the performance of the standard interval, non-modified, proposed modified and the alternative intervals, the usual α = 0.05 is utilized. The simulated data values are generated from a computer program using Maple software.

4.1 Comparison for Standard, Non-Modified and Modified Intervals in terms of Root Mean Squared Error (RMSE)

Figure 1 shows the RMSE behavior of the standard, the non-modified and the modified intervals for *n* = 5 to 100 with nominal 95% confidence level.

Results reveal that the modified interval has a significantly smaller RMSE for most values of *n*, while the standard and the non-modified have significantly larger RMSE especially for small *n*. This result gives evidences to the following conjecturethat the modified interval has a desirable RMSE property compared to the standard and the non-modified intervals.

4.2 Comparison for Modified and Alternative Intervals in terms of Root Mean Squared Error (RMSE)

Figure 2 displays the behavior of the six intervals in terms of RMSE for nominal 95%. The results show that as *n* increases, the RMSE performances of the six methods improve. It also shows that the RMSE of the alternative intervals are comparable except for small values of *n*, since the Modified and Wilson* intervals are significantly larger but the Agresti-Coull interval has the smaller RMSE for most values of *n* in the nominal 99% level. Thus, the preceding results verify the following finding that the RMSE behavior of the modified and the alternative methods decreases as n increases.

V. CONCLUSIONS

The existing and additional results would suggest rejection of the conditions made by several authors regarding the use of the standard interval, but instead utilize the alternative methods found in the literature which perform better in terms of RMSE. The performance of the alternative methods and the proposed method modified show some good RMSE behavior especially for the sample size *n* to become larger. Given the varied options, the best solution will no doubt be influenced by the user's personal preferences. A wise choice could be either one of the Wilson, Agresti-Coull, Wilson*, logit**, arcsine and modified intervals which show decisive improvement over the standard interval. Further study should consider the performance (like coverage properties, expected width and RMSE) of the most probable classical and Bayesian intervals, and explore the Poisson approximation to the binomial distribution in the context of the confidence interval for *p*.

.

REFERENCES

[1] Agresti, A., and Caffo, B. (2000). Simple and Effective Confidence Intervals for Proportions and Differences of Proportions Result from Adding Two Success and Two Failures. The American Statistician, 54, 280 – 288.

[2] Brown, L. D., Cai, T. T., and DasGupta, A. (1999a). Interval Estimation of a Binomial Proportion.Unpublished Technical Report.

[3] Brown, L. D., Cai, T. T., and DasGupta A. (1999b). Confidence Intervals for a Binomial Proportion and Edgeworth Expansion.Unpublished Technical Report.

[4] Brown, L. D., Cai, T. T., and DasGupta A. (2001). Interval Estimation for a Binomial Proportion (with discussion). Statistical Science, 16, 101 – 133.

[5] Casella, G., and Berger, R. (1990). Statistical Inference.Pacific Coast, CA. Woodsworth and Brooks/Cole.

[6] Edwardes, M. D. (1998). The Evaluation of Confidence Sets with Application to Binomial Intervals.StatisticaSinica, 8, 393 – 409.

[7] Marchand, E., Perron, F., Rokhaya, G., (2004). MinimaxEsimation of a Binomial Proportion p when [p – ½] is bounded.Universite' de Montreal.