Plugin Confidence Intervals in Discrete Distributions

T. Tony Cai*
Department of Statistics
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104

Abstract

The standard Wald interval is widely used in applications and textbooks because of its elementary motivation and simplicity of computation. It has been shown in several recent papers that the Wald interval suffers from a serious systematic negative bias in its coverage probability. In this article we propose confidence intervals in Binomial, Negative Binomial and Poisson distributions which have good coverage and parsimony properties while possess the same simple form as the Wald interval. We consider both one-sided and two-sided confidence intervals and give a unified treatment for all three distributions. The properties of the confidence intervals are studied through numerical and analytical calculations.

Keywords: Binomial distribution; Confidence intervals; Coverage probability; Edgeworth expansion; Expected length; Natural exponential family; Negative Binomial distribution; Normal approximation, Poisson distribution.

*Research supported in part by NSF Grant DMS-0306576.

1 Introduction

The Binomial, Negative Binomial, and Poisson distributions are three important discrete distributions which belong to the natural exponential family with quadratic variance functions (NEF-QVF). Interval estimation, both one-sided and two-sided, in these discrete distributions has a long history, an extensive literature and a wide range of applications. See, for example, Clevenson and Zidek (1975), Kaplan (1983), Duncan (1986), Santner and Duffy (1989), and Montgomery (2001).

The popular Wald interval has a very intuitive and simple form and is in nearly universal use. For binomial proportion, it has been generally known that the Wald interval is deficient in the coverage probability for p near 0 and 1. See, for example, Cressie (1980), Blyth and Still (1983), Vollset (1993), Agresti and Coull (1998), and Newcombe (1998). In recent articles, Brown, Cai and DasGupta (2001, 2002 and 2003) give a comprehensive treatment of two-sided confidence intervals in binomial and other distributions in the NEF-QVF. The Wald interval in the binomial case is shown to suffer from a systematic negative bias in its coverage probability far more persistent than is appreciated. It is also shown that the problems and the solutions in the binomial proportion case are common to all the distributions in the NEF-QVF. Alternative intervals with superior coverage properties are recommended. Among them, the Score interval, produced by inversion of Rao's Score test, was shown to always provide major improvements in coverage probability.

Cai (2005) considered one-sided confidence intervals for the discrete distributions in the NEF-QVF. It was shown that, although there are some common features, the one-sided interval estimation problem differs significantly from the two-sided problem. In particular, despite the good performance of the Score interval in the two-sided problem, the one-sided Score interval does not perform well for each of the three distributions. Both the one-sided Wald and Score intervals suffer a pronounced systematic bias in the coverage, although the severity and direction differ.

The Wald interval does not perform well, but its simple form is intuitive and appealing. The alternative confidence intervals proposed in the papers mentioned above have much more complicated forms than the Wald interval, with the exception of the Agresti-Coull interval in the binomial case. For applications as well as for teaching in introductory statistics courses, elementary motivation and simplicity are important. In the present paper we construct confidence intervals which have the same simple form as the Wald interval and have good coverage and parsimony properties. We call these confidence intervals *plugin intervals*. We consider both one-sided and two-sided confidence intervals and give a unified treatment for all three distributions. The properties of the confidence intervals are studied through numerical and analytical calculations. Edgeworth expansion is used for assessing the coverage property of the plugin interval. For confidence sets it is desirable to have the probability matching property. See Ghosh (1994) and Ghosh (2001). It is shown that all of our plugin intervals are first-order probability matching. In contrast, both the one-sided Wald and Score intervals are not first-order probability matching.

The paper is organized as follows. Section 2 considers the one-sided interval estimation problem. After basic notation and definitions are reviewed, the plugin interval is introduced. The coverage properties are studied through a two-term Edgeworth expansion and comparisons are made with the Wald and Score intervals. Parsimony property is considered through an asymptotic expansion of the expected distance from the mean. The analysis shows that the one-sided plugin interval, with its simple form and good coverage and parsimony properties, is much preferred over both the Wald and Score intervals in all three distributions. Section 3 gives an analogous treatment of the two-sided problem. The proofs are given in the Appendix.

2 One-Sided Interval Estimation

As in Brown, Cai and DasGupta (2003) and Cai (2005), the common setup in the present paper is that we have iid observations $X_1, X_2, ..., X_n \sim F$ where F is $\text{Bin}(1,p)$ in the binomial case, $\text{Pois}(\lambda)$ in the Poisson case, and $\text{NBin}(1,p)$, the number of successes before the first failure, in the Negative Binomial case. These three distributions form the discrete exponential family with a quadratic variance function where the variance

$$\sigma^2 \equiv V(\mu) = \mu + b_* \mu^2 \tag{1}$$

with $\mu = p$ and $b_* = -1$ in the binomial $\text{Bin}(1,p)$ case; $\mu = \lambda$ and $b_* = 0$ in the Poisson $\text{Pois}(\lambda)$ case; and $\mu = p/(1-p)$ and $b_* = 1$ in the Negative Binomial $\text{NBin}(1,p)$ case. See Morris (1982) and Brown (1986) for details on natural exponential family with quadratic variance functions. Set $X = \sum_{i=1}^n X_i$, $\hat{\mu} = \bar{X} = \sum_{i=1}^n X_i/n$ and denote by z_α the

$100(1-\alpha)$th percentile of the standard normal distribution. The objective is to construct confidence intervals for the mean μ based on the random sample $\{X_1, X_2, ..., X_n\}$.

2.1 The One-Sided Confidence Intervals

The Wald interval is the most commonly used confidence interval in practice and has a simple form. The $100(1-\alpha)\%$ upper limit Wald interval is based on the normal approximation

$$\frac{\sqrt{n}(\hat{\mu}-\mu)}{V^{\frac{1}{2}}(\hat{\mu})} = \frac{\sqrt{n}(\hat{\mu}-\mu)}{\sqrt{\hat{\mu}+b_*\hat{\mu}^2}} \xrightarrow{\mathcal{L}} N(0,1) \tag{2}$$

and has the form

$$CI_W^u = [0,\ \ \hat{\mu}+z_\alpha V^{\frac{1}{2}}(\hat{\mu})n^{-\frac{1}{2}}] = [0,\ \ \hat{\mu}+z_\alpha(\hat{\mu}+b_*\hat{\mu}^2)^{\frac{1}{2}}n^{-\frac{1}{2}}]. \tag{3}$$

The $100(1-\alpha)\%$ lower limit Wald interval is given by $CI_W^l = [\hat{\mu}-z_\alpha V^{\frac{1}{2}}(\hat{\mu})n^{-\frac{1}{2}},\ 1]$ in the Binomial case and $CI_W^l = [\hat{\mu}-z_\alpha V^{\frac{1}{2}}(\hat{\mu})n^{-\frac{1}{2}},\ \infty)$ in the Poisson and Negative Binomial cases. For simplicity we shall combine the three cases and write hereafter the upper limit of the lower limit intervals as ∞ for all three distributions with the understanding that it is actually 1 in the Binomial case. For one-sided confidence intervals our focus in this paper will be mainly on the upper limit intervals. The analysis for the lower limit intervals is analogous.

Besides the Wald interval, the Score interval is also frequently used. The $100(1-\alpha)\%$ upper limit Score interval for μ is

$$CI_S^u = [0,\ \ \frac{X+z_\alpha^2/2}{n-b_*z_\alpha^2} + \frac{z_\alpha n^{\frac{1}{2}}}{n-b_*z_\alpha^2}(V(\hat{\mu})+\frac{z_\alpha^2}{4n})^{\frac{1}{2}}]. \tag{4}$$

This confidence interval is derived by inverting the Score test of the one-sided hypotheses $H_0: \mu \geq \mu_0$ against $H_a: \mu < \mu_0$ using the normal approximation

$$\frac{\sqrt{n}(\hat{\mu}-\mu)}{V^{\frac{1}{2}}(\mu)} = \frac{\sqrt{n}(\hat{\mu}-\mu)}{\sqrt{\mu+b_*\mu^2}} \xrightarrow{\mathcal{L}} N(0,1). \tag{5}$$

The $100(1-\alpha)\%$ lower limit Score interval is constructed similarly and has the form

$$CI_S^l = [\frac{X+z_\alpha^2/2}{n-b_*z_\alpha^2} - \frac{z_\alpha n^{\frac{1}{2}}}{n-b_*z_\alpha^2}(V(\hat{\mu})+\frac{z_\alpha^2}{4n})^{\frac{1}{2}},\ \ \infty).$$

As shown in Cai (2005) that both the one-sided Wald and Score intervals have unsatisfactory coverage properties. Figure 1 below plots the actual coverage probability of the

99% upper limit Wald and Score intervals for the mean of the binomial, Negative Binomial and Poisson distributions with $n = 25$. In addition to the unavoidable oscillations for the nonrandomized intervals, it is clear from the plots that both the Wald interval and Score interval contain significant systematic bias in the coverage probability. As alternatives to the Wald and Score intervals, the Jeffreys interval and the second-order corrected interval were introduced in Cai (2005). These two intervals were shown to possess better coverage and length properties. However these intervals are also of much more complicated forms and are difficult to motivate in elementary courses. As mentioned earlier, simplicity is important in many settings.

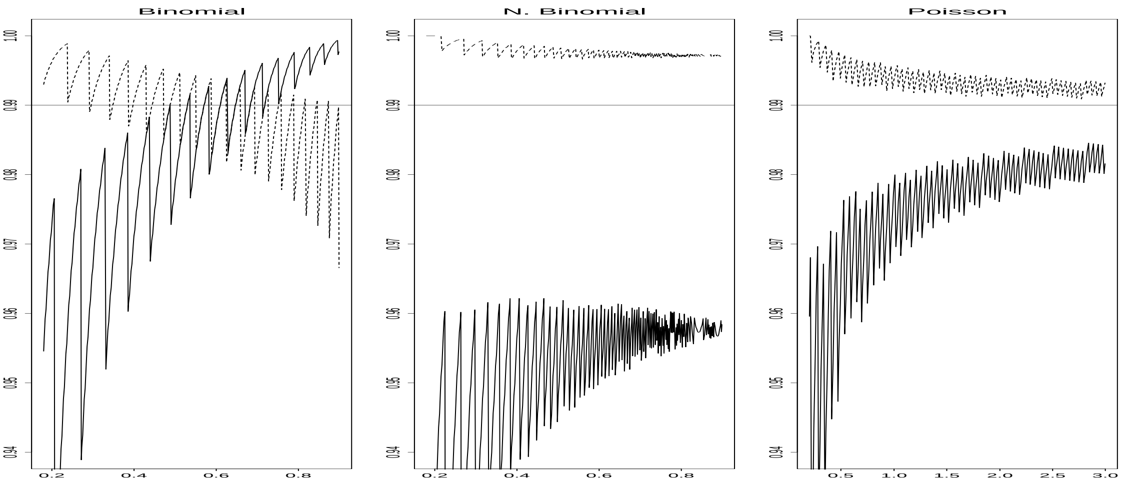

Figure 1: Coverage probability of the upper limit Wald interval (solid line) and upper limit Score interval (dashed line) for $n = 25$ and $\alpha = .01$. From left to right: Binomial, Negative Binomial and Poisson.

We now introduce a new confidence interval which has the same simple form as the Wald interval CI_W, but with a different $\hat{\mu}$ and a modified value for n. We shall call such an interval a *plugin interval*. Let $\eta = \frac{1}{3}z_\alpha^2 + \frac{1}{6}$ and set

$$\tilde{X} = X + \eta, \ \tilde{n} = n - 2b_*\eta \ \text{ and } \ \tilde{\mu} = \frac{\tilde{X}}{\tilde{n}}. \tag{6}$$

The *plugin interval* takes the same form as the Wald interval, by replacing $\hat{\mu}$ with $\tilde{\mu}$ and n with $\tilde{n}$. The upper limit and lower limit plugin intervals are defined, respectively, as

$$CI_P^u = [0, \ \ \tilde{\mu} + z_\alpha V^{\frac{1}{2}}(\tilde{\mu})\tilde{n}^{-\frac{1}{2}}] \ \text{ and } \ CI_P^l = [\tilde{\mu} - z_\alpha V^{\frac{1}{2}}(\tilde{\mu})\tilde{n}^{-\frac{1}{2}}, \ \ \infty). \tag{7}$$

Example: Consider the special Binomial case. The upper limit and lower limit plugin intervals for p are

$$CI_P^u = [0, \;\; \tilde{p} + z_\alpha \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}] \;\text{ and }\; CI_P^l = [\tilde{p} - \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}, \;\; \infty) \tag{8}$$

where $\tilde{n} = n + 2\eta$ and $\tilde{p} = \frac{X+\eta}{n+2\eta}$ with $\eta = \frac{1}{3}z_\alpha^2 + \frac{1}{6}$.

For the Poisson case, the plugin intervals for λ are

$$CI_P^u = [0, \;\; \tilde{\lambda} + z_\alpha \sqrt{\frac{\tilde{\lambda}}{\tilde{n}}}] \;\text{ and }\; CI_P^l = [\tilde{\lambda} - z_\alpha \sqrt{\frac{\tilde{\lambda}}{n}}, \;\; \infty) \tag{9}$$

where $\tilde{\lambda} = (X + \frac{1}{3}z_\alpha^2 + \frac{1}{6})/n$.

The plugin interval CI_P^u has better coverage property than both the Wald and Score intervals. This can be easily seen from Figure 2 which compares the coverage probabilities of the three one-sided intervals for $n = 35$ at the nominal level 99%. Analytical comparisons are given later.

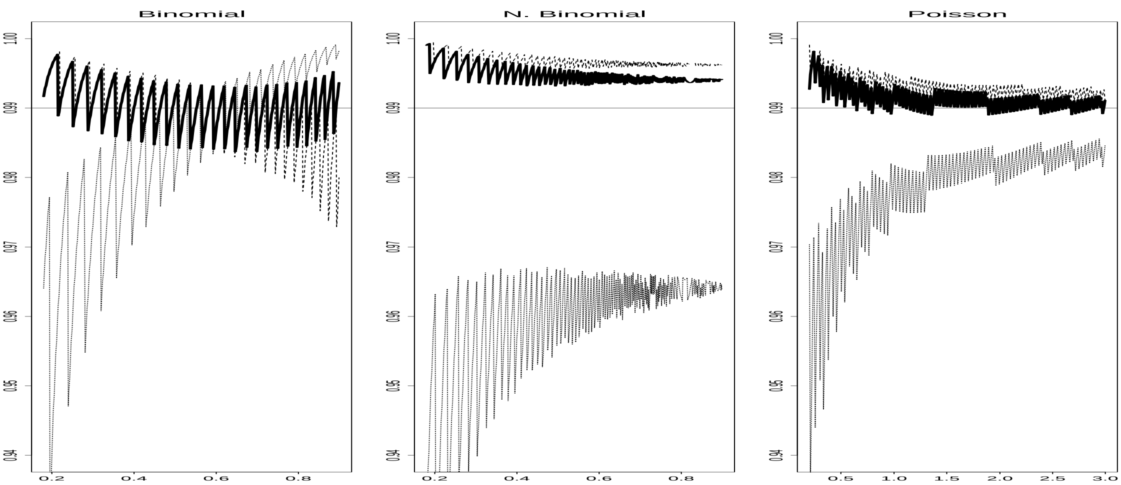

Figure 2: Coverage probability of the upper limit Plugin interval (dark solid line) for $n = 35$ and $\alpha = .01$. From left to right: Binomial, Negative Binomial and Poisson. The dotted line is the coverage probability of the Wald interval and the dashed line is that of the Score interval.

Remark: The plugin intervals can be further simplified for easy remembrance and for teaching in elementary courses. Take the commonly used 95% and 99% confidence intervals as examples. In the first case $\alpha = .05$, $z_{.05} = 1.645$ and $\eta = \frac{1}{3}z_{.05}^2 + \frac{1}{6} = 1.068 \approx 1$.

- Binomial: $\tilde{n} = n + 2$ and $\tilde{\mu} \approx \frac{X+1}{n+2}$, which can be translated into "adding 1 success and 1 failure".

- Negative Binomial: Recall that X is the number of successes before the nth failure. In this case $\tilde{n} = n - 2$ and $\tilde{\mu} \approx \frac{X+1}{n-2}$ which means "add 1 success and remove 2 failures".

- Poisson: $\tilde{n} = n$ and $\tilde{\mu} \approx \frac{X+1}{n}$ which simply means "add 1 arrival".

For the nominal 99% intervals, $\alpha = .01$ and $\eta = 1.97 \approx 2$. To construct the plugin interval, one can apply the simple form of the Wald interval after "adding 2 successes and 2 failures" in the binomial case, "adding 2 successes and removing 4 failures" in the Negative Binomial case, and "adding 2 arrivals" in the Poisson case.

2.2 Comparisons of Coverage Probability

The Edgeworth expansions provide an accurate and useful tool in analyzing the coverage properties of confidence intervals. See Brown, Cai and DasGupta (2002 and 2003) and Cai (2005). For example, the Edgeworth expansion is particularly useful in understanding analytically why the Score interval performs better in the two-sided problem than in the one-sided problem.

A two-term Edgeworth expansion of the coverage probability has a general form of

$$P(\mu \in CI) = 1 - \alpha + S_1 \cdot n^{-\frac{1}{2}} + Osc_1 \cdot n^{-\frac{1}{2}} + S_2 \cdot n^{-1} + Osc_2 \cdot n^{-1} + O(n^{-\frac{3}{2}}) \qquad (10)$$

where the first $O(n^{-\frac{1}{2}})$ term, $S_1 n^{-\frac{1}{2}}$, and the first $O(n^{-1})$ term, $S_2 n^{-1}$, are respectively the first and second order smooth terms, and $Osc_1 n^{-\frac{1}{2}}$ and $Osc_2 n^{-1}$ are the oscillatory terms. See Bhattacharya and Rao (1976) and Hall (1992) for details on Edgeworth expansions. The smooth terms in (10) capture the systematic bias in the coverage probability as seen in Figures 1 and 2. A confidence interval is called *first-order probability matching* if the first order smooth term $S_1 n^{-\frac{1}{2}}$ is vanishing. See Ghosh (1994) and Ghosh (2001) for general discussions on probability matching and confidence sets. Hall (1982) used the Edgeworth expansion to construct first-order probability matching one-sided intervals for a binomial proportion and Poisson mean. However his intervals are not of the simple plugin form.

Denote by $(x)_-$ the largest integer less than or equal to x and set

$$g(\mu, z) = g(\mu, z, n) = n\mu + n^{\frac{1}{2}}\sigma z - (n\mu + n^{\frac{1}{2}}\sigma z)_-. \tag{11}$$

So $g(\mu, z)$ is the fractional part of $n\mu + n^{\frac{1}{2}}\sigma z$. Let

$$Q_1(\mu,\ z) = g(\mu,\ z) - \frac{1}{2} \ \text{ and } \ Q_2(\mu,\ z) = -\frac{1}{2}g^2(\mu,\ z) + \frac{1}{2}g(\mu, z) - \frac{1}{12}. \tag{12}$$

Note that $Q_1(\mu,\ z)$ and $Q_2(\mu,\ z)$ are oscillatory functions. They appear in the Edgeworth expansions to accurately capture the oscillation in the coverage probability. Denote by ϕ and Φ respectively the density and cumulative distribution function of the standard normal distribution.

We now give the two-term Edgeworth expansion for the upper limit plugin interval. Let $0 < \alpha < 1$ and assume that μ is a fixed point in the interior of the parameter spaces.

Theorem 1 *Let z_P be defined as in (23) in the appendix. Suppose $n\mu + n^{\frac{1}{2}}\sigma z_P$ is not an integer. Then the coverage probability of the confidence interval CI_P^u defined in (7) satisfies*

$$\begin{aligned} P(\mu \in CI_P^u) \ &= \ (1-\alpha) + Q_1(\mu,\ z_P)\sigma^{-1}\phi(z_\alpha)n^{-\frac{1}{2}} \\ &\quad + \{\frac{b_*}{36}(23z_\alpha^3 + z_\alpha) + \frac{1}{72\sigma^2}(10z_\alpha^3 - z_\alpha)\}\phi(z_\alpha)n^{-1} \\ &\quad + \{\frac{1}{3}(1 + 2b_*\mu)Q_1(\mu,\ z_P) + Q_2(\mu,\ z_P)\}\sigma^{-2}z_\alpha\phi(z_\alpha)n^{-1} + O(n^{-\frac{3}{2}}) \end{aligned} \tag{13}$$

Remark: If $n\mu + n^{\frac{1}{2}}\sigma z_P$ is an integer, then an additional term $P_p(X = n\mu + n^{\frac{1}{2}}\sigma z_P) = \phi(z_\alpha)n^{-1/2}\sigma^{-1} + O(n^{-1})$ should be added to the left-side of (13). The expansions for the lower limit intervals can be obtained by first replacing α by $1-\alpha$ and z_α by $-z_\alpha$ in (13) and then subtracting it from 1.

We will now use the two term Edgeworth expansions to compare the coverage properties of the standard Wald interval CI_W^u and the Score interval CI_S^u with the new plugin interval. The comparisons are consistent for all three distributions and the conclusions therefore carry a unifying character.

The Edgeworth expansion for the one-sided Wald and Score intervals are given in Cai (2005). Denote the nonoscillating terms in the two term expansion of the coverage probability of CI_S^u, CI_P^u, and CI_W^u by B_S^u, B_P^u, and B_W^u, respectively. Then we have:

$$B_W^u \ = \ -\frac{1}{6}(2z_\alpha^2 + 1)(1 + 2b_*\mu)\sigma^{-1}\phi(z_\alpha)n^{-\frac{1}{2}}$$

$$- \{\frac{b_*}{36}(8z_\alpha^5 - 11z_\alpha^3 + 3z_\alpha) + \frac{1}{36\sigma^2}(2z_\alpha^5 + z_\alpha^3 + 3z_\alpha)\}\phi(z_\alpha)n^{-1} \quad (14)$$

$$B_S^u = \frac{1}{6}(z_\alpha^2 - 1)(1 + 2b_*\mu)\sigma^{-1}\phi(z_\alpha)n^{-\frac{1}{2}}$$

$$- \{\frac{b_*}{36}(2z_\alpha^5 - 11z_\alpha^3 + 3z_\alpha) + \frac{1}{72\sigma^2}(z_\alpha^5 - 7z_\alpha^3 + 6z_\alpha)\}\phi(z_\alpha)n^{-1} \quad (15)$$

$$B_P^u = \{\frac{b_*}{36}(23z_\alpha^3 + z_\alpha) + \frac{1}{72\sigma^2}(10z_\alpha^3 - z_\alpha)\}\phi(z_\alpha)n^{-1} \quad (16)$$

Note that the plugin interval CI_P^u is first-order probability matching whereas both the Wald and Score intervals are not probability matching. The first order smooth term is the main contributor of the systematic bias of the Wald and Score intervals seen in Figures 1 and 2.

Figure 3 plots the smooth terms in the Edgeworth expansions for the three intervals in the binomial case. The systematic bias in the coverage of the plugin interval is much less severe compare to that of Wald and Score intervals. In addition, the Plugin interval CI_P^u is more balanced and is conservative near both boundaries. It does not have the problem of serious under-coverage near one end and serious over-coverage near another end which is shared by both the Wald and Score intervals. See also Figure 2.

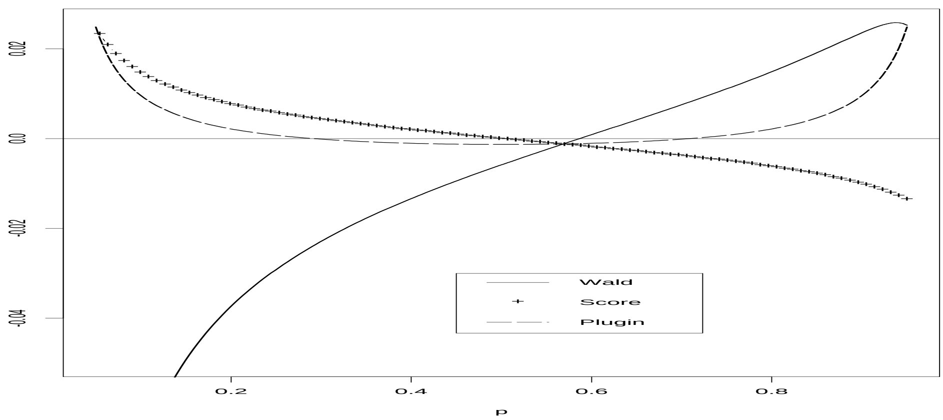

Figure 3: Comparison of the nonoscillating terms in the binomial case with $n = 40$ and $\alpha = .05$.

We now turn to the Negative Binomial and Poisson cases. Figure 4 displays the systematic bias for these two cases with $n = 40$ and $\alpha = .05$. In both cases there is a consistent significant negative bias in the coverage of the Wald interval while the Score interval has a non-negligible positive systematic bias. The coverage of the plugin interval is positively biased and the magnitude of the bias is much small than that of both the Wald and Score intervals, especially in the Poisson case.

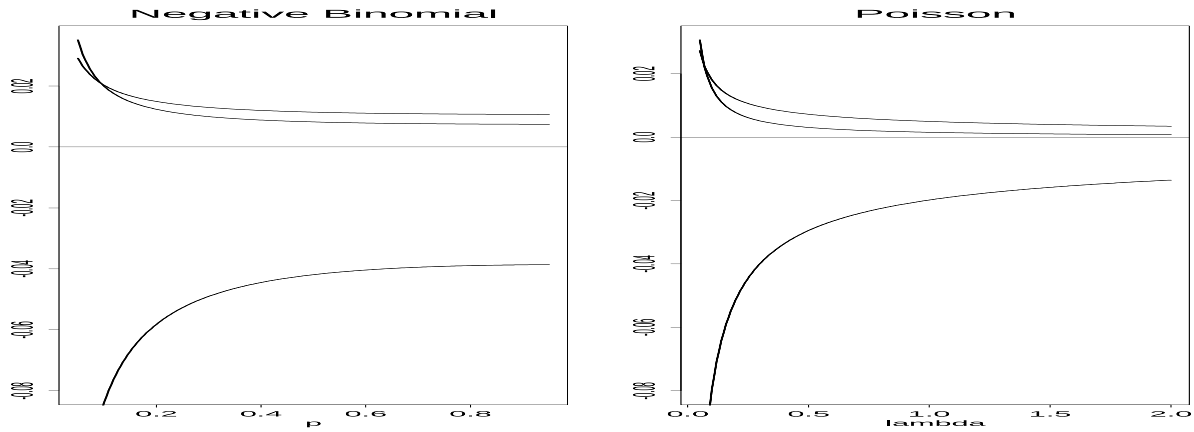

Figure 4: Comparison of the nonoscillating terms in the Negative Binomial and Poisson cases with $n = 40$ and $\alpha = .05$. From top to bottom: B_S^u, B_P^u and B_W^u.

2.3 Expected Distance from the Mean

Besides the coverage probability, parsimony, naturally measured by expected distance from the mean for one-sided intervals, is another important criterion. The expansion for the expected distance includes terms of the order $n^{-\frac{1}{2}}$, n^{-1} and $n^{-\frac{3}{2}}$. The coefficient of the $n^{-\frac{1}{2}}$ term is the same for all the intervals, but the coefficients for the n^{-1} and $n^{-\frac{3}{2}}$ terms differ. So, naturally, the coefficients of the n^{-1} and $n^{-\frac{3}{2}}$ terms will be used as a basis for comparison of their expected length.

Theorem 2 *Let U be the upper limit of the plugin interval CI_P^u and let $L_P^u = U - \mu$ be the distance of U from the mean μ. Then*

$$\begin{aligned} E(L_P^u) &= z_\alpha(\mu + b_*\mu^2)^{\frac{1}{2}} n^{-\frac{1}{2}} + (\frac{1}{3}z_\alpha^2 + \frac{1}{6}) \cdot (1 + 2b_*\mu) n^{-1} \\ &\quad + [(\frac{1}{6}z_\alpha^3 - \frac{1}{24}z_\alpha)(\mu + b_*\mu^2)^{-\frac{1}{2}} + (z_\alpha^3 + \frac{1}{2}z_\alpha) b_*(\mu + b_*\mu^2)^{\frac{1}{2}}] n^{-\frac{3}{2}} + O(n^{-2}). \end{aligned} \tag{17}$$

We first consider the Poisson and Negative Binomial cases. In these two cases, up to an error of order $O(n^{-2})$, there is a uniform ranking of the intervals in expected distance from the mean pointwise for every value of the parameter.

Corollary 1 *Consider the special Poisson case. Then the expected lengths of CI_W^u, CI_P^u, and CI_S^u admit the expansions*

$$
\begin{aligned}
E(L_W^u) &= z_\alpha \lambda^{\frac{1}{2}} n^{-\frac{1}{2}} - \frac{1}{8} z_\alpha \lambda^{-\frac{1}{2}} n^{-\frac{3}{2}} + O(n^{-2}) \\
E(L_P^u) &= z_\alpha \lambda^{\frac{1}{2}} n^{-\frac{1}{2}} + (\frac{1}{3} z_\alpha^2 + \frac{1}{6}) n^{-1} + (\frac{1}{6} z_\alpha^3 - \frac{1}{24} z_\alpha) \lambda^{-\frac{1}{2}} n^{-\frac{3}{2}} + O(n^{-2}) \\
E(L_S^u) &= z_\alpha \lambda^{\frac{1}{2}} n^{-\frac{1}{2}} + \frac{1}{2} z_\alpha^2 n^{-1} + \frac{1}{8} (z_\alpha^3 - z_\alpha) \lambda^{-\frac{1}{2}} n^{-\frac{3}{2}} + O(n^{-2}).
\end{aligned}
$$

Hence, up to the error n^{-2}, for every $\lambda > 0$, the ranking of the intervals is CI_W^u, CI_P^u and CI_S^u, from the shortest to the longest, as long as $z_\alpha \geq 1$. In practice, z_α will certainly be larger than 1 and so, we have a uniform ranking of the intervals. The exactly same ranking holds in the Negative Binomial case.

Unlike the Poisson and the Negative Binomial cases, a uniform ranking in length pointwise for all p is not valid in the Binomial case. Assume $z_\alpha > 1$ and note that $b_* = -1$ in the binomial case. For $p < \frac{1}{2}$ the Wald interval is too short and the Score interval too long while for $p > \frac{1}{2}$ the Wald interval is too long and the Score interval too short. Considering together with the coverage properties (see Equations (14) and (15) and Figure 1), it is clear that this is not desirable in either case. The expected distance of CI_P^u is always between those of CI_W^u and CI_S^u.

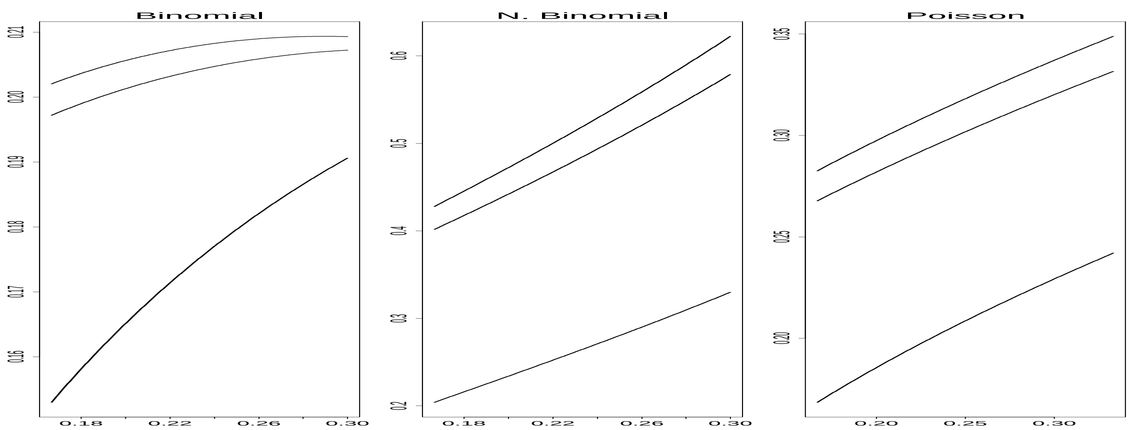

Figure 5: Expected distance of the upper limit of the three confidence intervals from the mean μ for $n = 30$ and $\alpha = .01$. For all three distributions, from top to bottom, the expected distance of the upper limit of CI_S^u, CI_P^u and CI_W^u.

3 Two Sided Confidence Intervals

The two sided Wald interval for μ is constructed based on the normal approximation (2) and has the simple form of

$$CI_W = \hat{\mu} \pm z_{\alpha/2} V^{\frac{1}{2}}(\hat{\mu}) n^{-\frac{1}{2}} = \hat{\mu} \pm z_{\alpha/2}(\hat{\mu} + b_* \hat{\mu}^2)^{\frac{1}{2}} n^{-\frac{1}{2}}. \tag{18}$$

The two-sided Score interval is formed by inverting the normal approximation (5) to the family of equal-tailed tests of $H_0 : \mu = \mu_0$ versus $H_a : \mu \neq \mu_0$ and has the form of

$$CI_S = \frac{X + z_{\alpha/2}^2/2}{n - b_* z_{\alpha/2}^2} \pm \frac{z_{\alpha/2} n^{\frac{1}{2}}}{n - b_* z_{\alpha/2}^2} (V(\hat{\mu}) + \frac{z_{\alpha/2}^2}{4n})^{\frac{1}{2}}. \tag{19}$$

The twos-sided plugin confidence interval has the same simple form as the Wald interval CI_W, but with a different $\hat{\mu}$ and a modified value for n. Let $\tilde{n}$ and $\tilde{\mu}$ be defined as in (6) with $\eta = z_{\alpha/2}^2/2$. Then the two-sided *plugin interval* is defined as

$$CI_P = \tilde{\mu} \pm z_{\alpha/2} V^{\frac{1}{2}}(\tilde{\mu}) \tilde{n}^{-\frac{1}{2}} = \tilde{\mu} \pm z_{\alpha/2}(\tilde{\mu} + b_* \tilde{\mu}^2)^{\frac{1}{2}} \tilde{n}^{-\frac{1}{2}}. \tag{20}$$

Remark: As in the one-sided case, the two-sided plugin intervals can also be further simplified for easy remembrance and for teaching in elementary courses. For the commonly used two-sided 95% confidence intervals, $\alpha = .05$, $z_{.025} = 1.96$ and $\eta = z_{.025}^2/2 = 1.921 \approx 2$.

- Binomial: $\tilde{n} = n + 4$ and $\tilde{\mu} \approx \frac{X+2}{n+4}$ which means "add 2 success and 2 failure". This is the Agresti-Coull interval. See Agresti and Coull (1998).

- Negative Binomial: $\tilde{n} = n - 4$ and $\tilde{\mu} \approx \frac{X+2}{n-4}$ which translates to "add 2 success and subtract 4 failures".

- Poisson: $\tilde{n} = n$ and $\tilde{\mu} \approx \frac{X+2}{n}$ which means "add 2 arrivals".

For two-sided 90% and 99% confidence intervals, the values of η are 1.35 and 3.32 respectively.

In the binomial case the plugin interval is the Agresti-Coull interval discussed in detail in BCD (2001 and 2002). See also Agresti-Coull (1998). Here we shall focus on the Negative Binomial and Poisson cases. Figure 6 displays the coverage probability of the plugin interval together with those of the Wald and Score intervals. It can be seen from

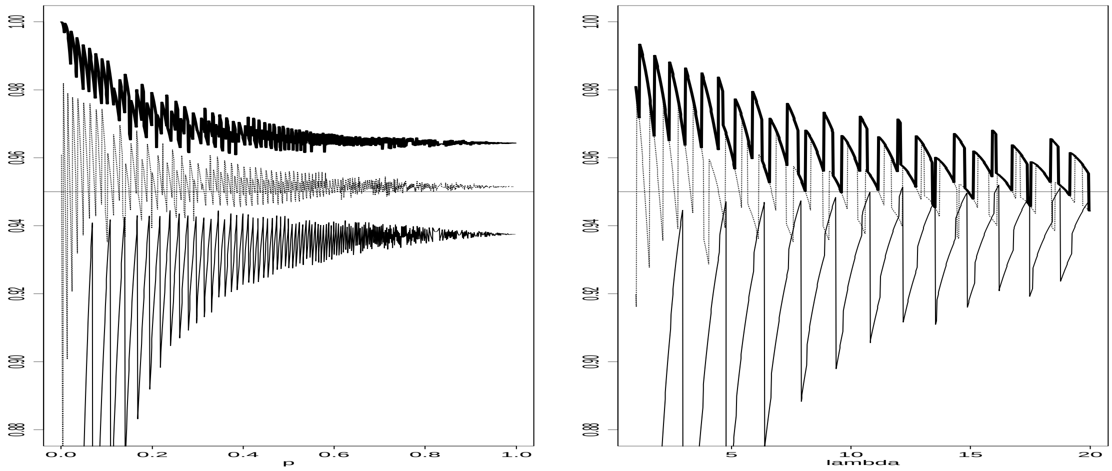

Figure 6: Comparison of the coverage probability of the three 95% confidence intervals for Negative Binomial with $n = 40$ (left panel) and for Poisson(λ) with $n = 1$ (right panel). The top dark line is the coverage probability of the plugin interval, the middle dotted line is that of the Score interval, and the bottom line is that of the Wald interval.

the plots that, as in the binomial case, the plugin interval is more conservative than the other two intervals in both Negative Binomial and Poisson distributions.

Similar to the one-sided interval case, an Edgeworth expansion for the coverage probability of the two-sided plugin interval can be derived.

Theorem 3 *Let $0 < \alpha < 1$. Suppose $n\mu + n^{1/2}\sigma\ell_P$ is not an integer. Then the coverage probability of the confidence interval CI_P defined in (20) satisfies*

$$\begin{aligned} P_\mu(\mu \in CI_P) &= (1-\alpha) + \sigma^{-1}\{g(\mu,\ell_P) - g(\mu,u_P)\} \cdot \phi(z_{\alpha/2})n^{-1/2} \\ &+ \{-\frac{b_*}{18}(2z_{\alpha/2}^5 - 29z_{\alpha/2}^3 + 3z_{\alpha/2}) - \frac{1}{36\sigma^2}(z_{\alpha/2}^5 - 16z_{\alpha/2}^3 + 6z_{\alpha/2})\} \cdot \phi(z_{\alpha/2})n^{-1} \\ &+ \{(1+2b_*\mu)(\frac{1}{6}z_{\alpha/2}^2 - \frac{1}{2})Q_{21}(\ell_P,u_P) + Q_{22}(-z_{\alpha/2},z_{\alpha/2})\}\sigma^{-2}z_{\alpha/2}\phi(z_{\alpha/2})n^{-1} \\ &+ O(n^{-3/2}) \end{aligned} \quad (21)$$

where the quantities ℓ_P and u_P are defined in (24) in the appendix.

Two term Edgeworth expansions for the Wald and Score intervals are given in BCD (2003). As in the one-sided problem, the smooth terms in the Edgeworth expansions can be used as the basis for the comparison of the coverage properties.

Denote the nonoscillating terms in the two term expansion of the coverage probability of CI_S, CI_P, and CI_W by B_S, B_P, and B_W, respectively. Then

$$
\begin{aligned}
B_W &= \{-\frac{b_*}{18}(8z_{\alpha/2}^5 - 11z_{\alpha/2}^3 + 3z_{\alpha/2}) - \frac{1}{18\sigma^2}(2z_{\alpha/2}^5 + z_{\alpha/2}^3 + 3z_{\alpha/2})\} \cdot \phi(z_{\alpha/2})n^{-1}, \\
B_S &= \{-\frac{b_*}{18}(2z_{\alpha/2}^5 - 11z_{\alpha/2}^3 + 3z_{\alpha/2}) - \frac{1}{36\sigma^2}(z_{\alpha/2}^5 - 7z_{\alpha/2}^3 + 6z_{\alpha/2})\} \cdot \phi(z_{\alpha/2})n^{-1}, \\
B_P &= \{-\frac{b_*}{18}(2z_{\alpha/2}^5 - 29z_{\alpha/2}^3 + 3z_{\alpha/2}) - \frac{1}{36\sigma^2}(z_{\alpha/2}^5 - 16z_{\alpha/2}^3 + 6z_{\alpha/2})\} \cdot \phi(z_{\alpha/2})n^{-1}.
\end{aligned}
$$

It can be verified directly that $B_P > B_S > B_W$ for all three distributions. Hence the plugin interval has the highest coverage probability in general. Figure 7 compares the smooth terms in the Edgeworth expansions for the three intervals.

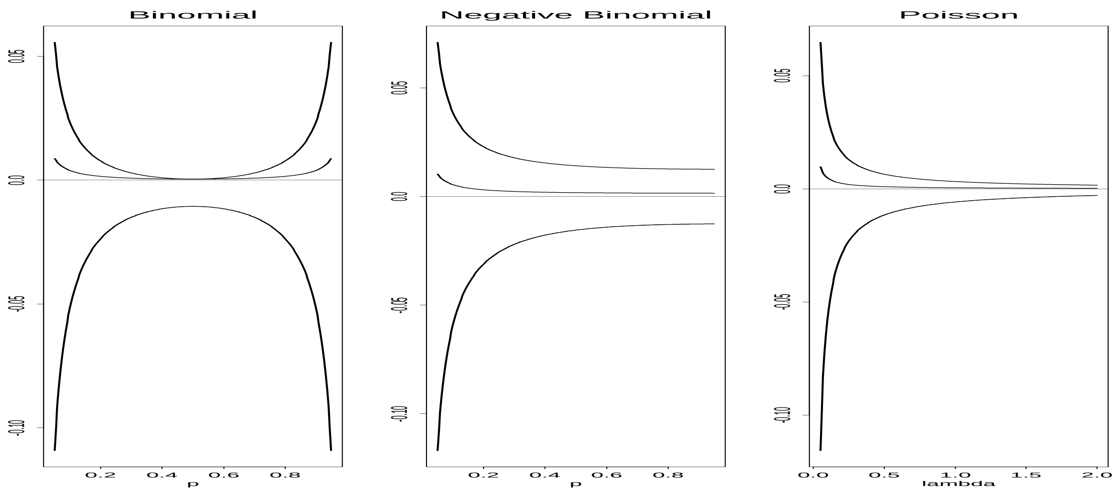

Figure 7: Comparison of the nonoscillating terms in the coverage probability of the Wald, Score and Plugin intervals (from bottom to top) for $n = 40$ and $\alpha = .05$.

Similarly, an expansion of the expected length can be given for the plugin interval.

Theorem 4 *Let L_P be the length of the two-sided plugin interval CI_P. Then*

$$
\begin{aligned}
E(L_P) &= 2z_{\alpha/2}(\mu + b_*\mu^2)^{\frac{1}{2}}n^{-\frac{1}{2}} + [3z_{\alpha/2}^3 b_*(\mu + b_*\mu^2)^{\frac{1}{2}} + \frac{1}{4}(2z_{\alpha/2}^3 - z_{\alpha/2})(\mu + b_*\mu^2)^{-\frac{1}{2}}]n^{-\frac{3}{2}} \\
&\quad + O(n^{-2}). \qquad (22)
\end{aligned}
$$

We now compare the expansion of $E(L_P)$ given in (22) with the those of the expected length for the Wald and Score intervals which are given in BCD (2003).

Corollary 2 *Consider the special Poisson case. Then the expected lengths of CI_W, CI_S and CI_P admit the expansions*

$$\begin{aligned}
E(L_W) &= 2z_{\alpha/2}\lambda^{\frac{1}{2}}n^{-\frac{1}{2}} - \frac{1}{4}z_{\alpha/2}\lambda^{-\frac{1}{2}}n^{-\frac{3}{2}} + O(n^{-2}) \\
E(L_S) &= 2z_{\alpha/2}\lambda^{\frac{1}{2}}n^{-\frac{1}{2}} + \frac{1}{4}(z_{\alpha/2}^3 - z_{\alpha/2})\lambda^{-\frac{1}{2}}n^{-\frac{3}{2}} + O(n^{-2}) \\
E(L_P) &= 2z_{\alpha/2}\lambda^{\frac{1}{2}}n^{-\frac{1}{2}} + \frac{1}{4}(2z_{\alpha/2}^3 - z_{\alpha/2})\lambda^{-\frac{1}{2}}n^{-\frac{3}{2}} + O(n^{-2})
\end{aligned}$$

Therefore, pointwise at every $\lambda > 0$, there is a uniform ranking of the intervals CI_W, CI_S and CI_P, from the shortest to the longest, provided $\alpha > 0.5$. The exact same ranking holds in the Negative Binomial case.

4 Concluding Remarks

The numerical and analytical results given in the present paper show that the plugin interval not only has the appealing simple form but also has good coverage and parsimony properties for all three distributions and in both one-sided and two-sided interval estimation problems. These desirable features make the plugin interval an excellent alternative to the standard Wald interval. The simple form of the plugin interval makes it easy for computation and for teaching in elementary statistics course.

5 Appendix: Proofs

5.1 Edgeworth Expansions

Proof of Theorem 1: The Edgeworth expansion for the coverage probability of the one-sided plugin interval CI_P^u can be derived by using Proposition 1 of BCD (2003). Set

$$\begin{aligned}
A &= n - b_*(z_\alpha^2 + 2\eta) \\
B &= 2(n - 2b_*\eta)(n\mu - (1 + 2b_*\mu)\eta) + z_\alpha^2 n \\
C &= (n - 2b_*\eta)(n\mu - (1 + 2b_*\mu)\eta)^2 - z_\alpha^2\eta n + b_* z_\alpha^2 \eta^2.
\end{aligned}$$

We have, after some straightforward algebra,

$$P_\mu(\mu \in CI_P^u) = P(n^{\frac{1}{2}}(\hat{\mu} - \mu)/\sigma \geq z_P^u)$$

where

$$z_P^u = \left(\frac{B - \sqrt{B^2 - 4AC}}{2A} - \mu \right) \sigma^{-1} n^{\frac{1}{2}} \tag{23}$$

Expanding z_P, one has

$$z_P^u = -z_\alpha + \frac{1}{6}(z_\alpha^2 - 1)(1 + 2b_*\mu)\sigma^{-1}n^{-\frac{1}{2}} - \{\frac{1}{6}(4z_\alpha^3 - z_\alpha)b_* + \frac{1}{8}z_\alpha^3\sigma^{-2}\}n^{-1} + O(n^{-\frac{3}{2}}).$$

Now (13) follows from Proposition 1 in BCD (2003) on some algebra. ∎

Proof of Theorem 3: The Edgeworth expansion for the two-sided plugin interval CI_P can be derived in a similar way. Denote

$$\begin{aligned} A &= (n - 2z_{\alpha/2}^2 b_*)n^2 \\ B &= 2\mu(n - z_{\alpha/2}^2 b_*)^2 n + z_{\alpha/2}^4 b_* n \\ C &= \mu^2(n - z_{\alpha/2}^2 b_*)^3 - z_{\alpha/2}^2 \mu(n - z_{\alpha/2}^2 b_*)^2 - \frac{1}{4} z_{\alpha/2}^4 n \end{aligned}$$

We have, after some straightforward algebra,

$$P_\mu(\mu \in CI_P) = P(\ell_P \leq n^{\frac{1}{2}}(\hat{\mu} - \mu)/\sigma \leq u_P)$$

where

$$(\ell_P,\ u_P) = \left(\frac{B \pm \sqrt{B^2 - 4AC}}{2A} - \mu \right) \sigma^{-1} n^{1/2} \tag{24}$$

The + sign goes with u_P and the − sign with ℓ_P. Expanding ℓ_P and u_P, one has

$$(\ell_P,\ u_P) = \pm\{z_{\alpha/2} + (\frac{1}{2}b_* + \frac{1}{8}\sigma^{-2})z_{\alpha/2}^3 n^{-1}\} + O(n^{-3/2}) \tag{25}$$

with the + sign going with u_P and the − sign with ℓ_P. Now (21) follows from Proposition 1 of BCD (2003). ∎

5.2 Expansions for Expected Length

Proof of Theorem 2: For the one-sided plugin interval CI_P^u set $Z_n = (\bar{X} - \mu)(\mu + b_*\mu^2)^{-\frac{1}{2}} n^{\frac{1}{2}}$ and write $\tilde{\mu} = (\mu + Z_n(\mu + b_*\mu^2)^{\frac{1}{2}} n^{-\frac{1}{2}} + \eta n^{-1})(1 - 2b_*\eta n^{-1})^{-1}$. Then

$$\begin{aligned} L_P^u &= (1 + 2b_*\mu)\eta n^{-1} + Z_n(1 + 2b_*\eta n^{-1})(\mu + b_*\mu^2)^{\frac{1}{2}} n^{-\frac{1}{2}} \\ &\quad + z_\alpha(\mu + b_*\mu^2)^{\frac{1}{2}}(n - 2b_*\eta)^{-\frac{1}{2}} \cdot \{1 + [2b_* + \frac{1}{2}(\mu + b_*\mu^2)^{-1}]\eta n^{-1}\} \\ &\quad - \frac{1}{8} z_\alpha (n - 2b_*\eta)^{-\frac{1}{2}}(\mu + b_*\mu^2)^{-\frac{1}{2}} Z_n^2 n^{-1} + R_P(Z_n) \end{aligned}$$

where $E(|R_P(Z_n)|) = O(n^{-2})$. Hence

$$\begin{aligned} E(L_P^u) &= (1+2b_*\mu)\eta n^{-1} + z_\alpha(\mu+b_*\mu^2)^{\frac{1}{2}}(1+b_*\eta n^{-1})n^{-\frac{1}{2}} \cdot \\ &\quad \left\{1+2b_*\eta n^{-1} + \frac{1}{8}(4\eta-1)(\mu+b_*\mu^2)^{-1}n^{-1}\right\} + O(n^{-2}) \end{aligned}$$

and equation (17) follows on a few steps of algebra. ∎

Proof of Theorem 4: For the interval CI_P, the length is $L_P = 2z_{\alpha/2}(\tilde{\mu}+b_*\tilde{\mu}^2)^{\frac{1}{2}}\tilde{n}^{-1/2}$, where $\tilde{n} = n - b_* z_{\alpha/2}^2$ and $\tilde{\mu} = (X + z_{\alpha/2}^2/2)/\tilde{n}$. Set $Z_n = (\bar{X}-\mu)(\mu+b_*\mu^2)^{-\frac{1}{2}}n^{\frac{1}{2}}$. Then $E(Z_n)=0$ and $E(Z_n^2)=1$. Write

$$\tilde{\mu} = (\mu + Z_n(\mu+b_*\mu^2)^{1/2}n^{-1/2} + \frac{1}{2}z_{\alpha/2}^2 n^{-1})(1-b_* z_{\alpha/2}^2 n^{-1})^{-1},$$

and hence, after some algebra,

$$\begin{aligned} L_P &= 2z_{\alpha/2}(\mu+b_*\mu^2)^{1/2}(n-b_*z_\alpha^2)^{-1/2}\left\{1+[b_*+\frac{1}{4}(\mu+b_*\mu^2)^{-1}]z_{\alpha/2}^2 n^{-1}\right. \\ &+ \left.[\frac{1}{2}b_* - \frac{1}{8}(1+2b_*\mu)^2(\mu+b_*\mu^2)^{-1}]Z_n^2 n^{-1} + R_P(Z_n)\right\}, \end{aligned}$$

where $E(|R_P(Z_n)|) = O(n^{-3/2})$. Thus,

$$\begin{aligned} E(L_P) &= 2z_{\alpha/2}(\mu+b_*\mu^2)^{1/2}(1+\frac{1}{2}b_*z_{\alpha/2}^2 n^{-1})n^{-\frac{1}{2}}\left\{1+b_*z_{\alpha/2}^2 n^{-1} + \frac{1}{2}b_* n^{-1}\right. \\ &+ \left.\frac{1}{4}z_{\alpha/2}^2(\mu+b_*\mu^2)^{-1}n^{-1} - \frac{1}{8}(1+2b_*\mu)^2(\mu+b_*\mu^2)^{-1}n^{-1}\right\} + O(n^{-2}) \end{aligned}$$

which simplifies to equation (22) stated in Theorem 4. ∎

References

[1] Agresti, A. and Coull, B. A. (1998). Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician* **52**, 119-126.

[2] Bhattacharya, R. N. and Rao, R. R. (1976). *Normal Approximation And Asymptotic Expansions.* Wiley, New York.

[3] Blyth, C.R. and Still, H. A. (1983). Binomial confidence intervals. *J. Amer. Statist. Assoc.* **78**, 108-116.

[4] Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory.* Lecture Notes-Monograph Series, Institute of Mathematical Statistics, Hayward.

[5] Brown, L. D., Cai, T. and DasGupta, A. (2001). Interval estimation for a binomial proportion (with discussion). *Statist. Sci.* **16**, 101-133.

[6] Brown, L. D., Cai, T. and DasGupta, A. (2002). Confidence intervals for a binomial proportion and Edgeworth expansions. *Ann. Statist.* **30**, 160-201.

[7] Brown, L. D., Cai, T. and DasGupta, A. (2003). Interval estimation in exponential families. *Statistica Sinica* **13**, 19-49.

[8] Cai, T. (2005). One-sided confidence intervals in discrete distributions. *J. Statist. Planning and Inf.* **131**, 63-88.

[9] Clevenson, M. and Zidek, J. (1975). Simultaneous estimation of the means of independent Poisson laws. *J. Amer. Stat. Assoc.* **70**, 698-705.

[10] Cressie, N. (1980). A finely tuned continuity correction. *Ann. Inst. Statist. Math.* **30**, 435-42.

[11] Duncan, A. J. (1986). *Quality Control and Industrial Statistics.* 5th ed., Irwin, Homewood, IL.

[12] Ghosh, J. K. (1994). *Higher Order Asymptotics.* NSF-CBMS Regional Conference Series, Institute of Mathematical Statistics, Hayward.

[13] Ghosh, M. (2001). Comment on “Interval estimation for a binomial proportion” by L. Brown, T. Cai and A. DasGupta. *Statist. Sci.* **16**, 124-125.

[14] Hall, P. (1982). Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters. *Biometrika* **69**, 647-52.

[15] Hall, P. (1992). *The Bootstrap And Edgeworth Expansion.* Springer-Verlag, New York.

[16] Montgomery, D. C. (2001). *Introduction to statistical quality control.* 4th ed., John Wiley, New York.

[17] Morris, C.N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10**, 65-80.

[18] Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion; comparison of several methods. *Statistics in Medicine* **17**, 857-872.

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