Stat Comput (2009) 19: 139–148
DOI 10.1007/s11222-008-9077-8

Exact average coverage probabilities and confidence coefficients of confidence intervals for discrete distributions

Hsiuying Wang

Received: 14 March 2006 / Accepted: 11 June 2008 / Published online: 4 July 2008

Abstract For a confidence interval $(L(X), U(X))$ of a parameter θ in one-parameter discrete distributions, the coverage probability is a variable function of θ. The confidence coefficient is the infimum of the coverage probabilities, $\inf_\theta P_\theta(\theta \in (L(X), U(X)))$. Since we do not know which point in the parameter space the infimum coverage probability occurs at, the exact confidence coefficients are unknown. Beside confidence coefficients, evaluation of a confidence intervals can be based on the average coverage probability. Usually, the exact average probability is also unknown and it was approximated by taking the mean of the coverage probabilities at some randomly chosen points in the parameter space. In this article, methodologies for computing the exact average coverage probabilities as well as the exact confidence coefficients of confidence intervals for one-parameter discrete distributions are proposed. With these methodologies, both exact values can be derived.

Keywords Confidence coefficient · Confidence interval · Coverage probability · Discrete distribution

1 Introduction

Let X be a one-dimensional discrete random variable with a probability mass function $f_\theta(x), x \in S = \{0, \dots, n\}$, where θ is an unknown parameter. Let Ω denote the parameter space of θ. For a confidence interval $(L(X), U(X))$ of the parameter θ, the coverage probability of $(L(X), U(X))$ is the probability that the random interval covers the true parameter θ. In this paper, we assume that Ω is a closed interval, that the confidence interval $(L(X), U(X))$ satisfies some conditions given in the next section and that the coverage probability of the discrete random variable is a continuous function of θ. The confidence coefficient of the confidence interval is defined as the infimum of the coverage probabilities, $\inf_{\theta\in\Omega} P_\theta(\theta \in (L(X), U(X)))$. For continuous distributions, the coverage probability may be the same for every point in the parameter space. However, for discrete distributions, the coverage probabilities are different when the true parameter varies in the parameter space. Since the infimum of the coverage probability may occur at any point in the parameter space, and we do not know at which point in the parameter space the infimum coverage probability occurs at, the exact confidence coefficient is unknown.

The asymptotic behavior of the coverage probability for the discrete distributions has been investigated by Brown et al. (2002, 2003), among others. However, there is a lack in the literature of investigation for the derivations of the exact minimun coverage probability for a fixed sample size. The previous derivations of the minimum coverage probability for one-parameter discrete distributions have been approximated by simulation. However, simulation only provides a rough estimation and is time consuming. Therefore, development of a methodology to calculate the exact value for the minimum coverage probability accurately and efficiently is needed.

Furthermore, from another aspect, since the confidence coefficient is the behavior of the confidence interval at one point in a discrete distribution, evaluation of a confidence

H. Wang (✉)
Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan
e-mail: hywang@stat.sinica.edu.tw

H. Wang
Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan

interval based only on the behavior of the infimum coverage probability is too conservative. Woodroofe (1986) and Stein (1985) suggested that evaluation can also be based on the average coverage probability. Let $\eta(\theta)$ be a density function on Ω. The average coverage probability under $\eta(\theta)$ is defined by

$$\int_{\Omega} P_\theta(\theta \in (L(X), U(X)))\eta(\theta)d\theta. \quad (1)$$

This value takes the average of the coverage probability with respect to the prior $\eta(\theta)$. Usually, the exact average coverage probability is also unknown as well as the confidence coefficient, and it has been approximated by taking the mean of coverage probabilities at some randomly chosen points in the parameter space. For evaluating a confidence interval, reporting both the confidence coefficient and average coverage probability can provide more objective information than only reporting the confidence coefficient.

In this paper, methodologies for calculating the exact confidence coefficient and exact average coverage probability for discrete families are proposed. We show that the confidence coefficient can be derived by computing the coverage probabilities at the endpoints of the confidence intervals corresponding to each observation, and that the minimum value of these coverage probabilities is the confidence coefficient. For the binomial distribution, some results of calculating the exact confidence coefficient are in Wang (2007). To derive the exact average coverage probability, we compute coverage probability functions in each interval separated by the endpoints of the confidence intervals corresponding to each observation in the parameter space, and compute the summation of the integrations of these coverage functions with respect to a prior measure. The methods proposed in this paper make the calculation of both exact values easy and plausible for any fixed sample size.

The paper is organized as follows. Section 2 gives some assumptions required for deriving the exact confidence coefficients for confidence intervals and distributions. A procedure for calculating the exact confidence coefficients for discrete families is proposed. In Sect. 3, a procedure for computing the average coverage probability of a confidence interval for discrete distributions is proposed. In Sect. 4, some examples of deriving the confidence coefficients and average coverage probabilities of confidence intervals for binomial and Poisson distributions are given. Finally, in Sect. 5, simulation studies are conducted to show that the proposed methods are efficient and valid.

2 Exact confidence coefficient

In this section, a methodology for deriving the confidence coefficient of a confidence interval for a discrete distribution is proposed. Let $(L(X), U(X))$ be a confidence interval of an unknown parameter θ in a discrete distribution with sample space $S = \{0, \dots, n\}$. Let l and u denote the lower boundary point and the upper boundary point of the parameter space Ω. The methodology proposed in this section applies to confidence intervals satisfying some of the following assumptions.

Assumption 1 A confidence interval $(L(X), U(X))$ satisfies the conditions:

(i) For any observations X_1 and X_2, $L(X_1) < L(X_2)$ and $U(X_1) < U(X_2)$ if $X_1 < X_2$.
(ii) $L(0) \le l \le U(0)$ and $L(n) \le u \le U(n)$.

Assumption 2 A confidence interval $(L(X), U(X))$ satisfies the conditions:

(i) For $X_1 > 0$ and $X_2 < n$, $L(X_1) < L(X_2)$ and $U(X_1) < U(X_2)$ if $X_1 < X_2$.
(ii) $L(0) = U(0) = l$ and $L(n) = U(n) = u$.

Assumption 3 A confidence interval $(L(X), U(X))$ satisfies the condition:

For any θ in the interior of the parameter space, there exists x_0 such that $\theta \in (L(x_0), U(x_0))$ and $P_\theta(X = x_0) > 0$.

Remark 1 In this paper, we need to assume that for each x, the intersection of $(L(X), U(X))$ and the parameter space is not empty. Since the assumption is satisfied for most confidence intervals, we do not list this condition in Assumptions 1–3.

Assumption 3 guarantees that there does not exist θ_0 in the parameter space, such that the coverage probability at θ_0 is zero. Most widely-used confidence intervals satisfy either Assumption 1 or Assumption 2. For example, the usual Wald interval for a proportion of a binomial distribution satisfies Assumption 2. The Clopper-Pearson interval (see, Clopper and Pearson 1934)

$$(l, u) = \left(\left(1 + \frac{(n-x+1)F_{(1-\alpha/2;2n-2x+2,2x)}}{x}\right)^{-1}, \left(1 + \frac{n-x}{(x+1)F_{(1-\alpha/2;2x+2,2n-2x)}}\right)^{-1}\right)$$

and the Agresti-Coull interval (see, Agresti and Coull 1998 and Brown et al. 2002)

$$CI_{AC}(X) = \left(\tilde{p} - z_{\alpha/2}\,(\tilde{p}\tilde{q})^{1/2}\,\tilde{n}^{-1/2}, \tilde{p} + z_{\alpha/2}\,(\tilde{p}\tilde{q})^{1/2}\,\tilde{n}^{-1/2}\right)$$

satisfy Assumption 1, where z_a denotes the upper $(1-\alpha)$th quantile of the standard normal distribution, $F_{(a;r_1,r_2)}$ denotes the ath quantile of the F distribution with r_1 and r_2 degrees of freedom, $\tilde{X} = X + z^2_{\alpha/2}/2$, $\tilde{n} = n + z^2_{\alpha/2}$, $\tilde{p} = \tilde{X}/\tilde{n}$ and $\tilde{q} = 1 - \tilde{p}$.

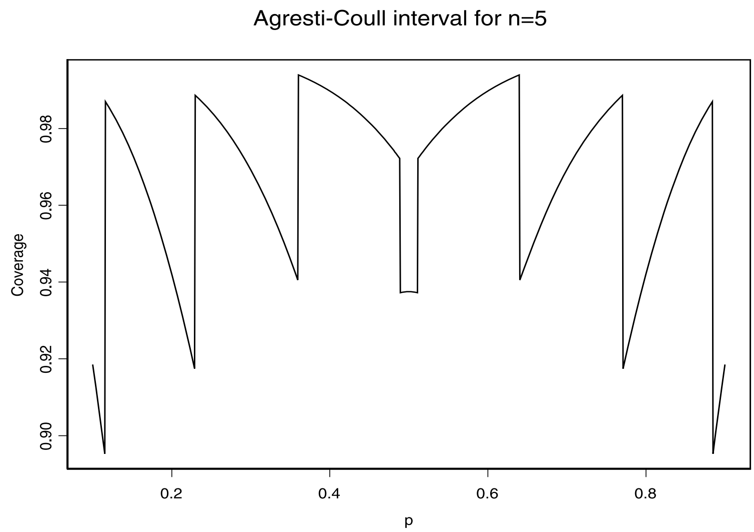

Fig. 1 Coverage probability of the Agresti-Coull interval for $n = 5$

We will use the Agresti-Coull interval for $n = 5$ as an example to illustrate the proposed methods throughout this paper. The 0.95 level intervals corresponding to $X = 0, \ldots, 5$ are $(-0.05457,\ 0.48906)$, $(0.02031,\ 0.64037)$, $(0.11598,\ 0.77091)$, $(0.22908, 0.88401)$, $(0.35962, 0.97968)$, $(0.51093,\ 1.05457)$. From these six intervals, it is clear that Assumption 1 and Assumption 3 are satisfied. Figure 1 shows the behavior of the coverage probability of this example.

For the general n and α of the Agresti-Coull interval, Wang (2006) demonstrated that Assumptions 1 and 3 are satisfied for $n \geq 2$ and $z_{\alpha/2} > 1$.

Assumptions 1–3 are conditions for confidence intervals, while Assumption 4 contains a condition for distributions. The methodology proposed in this paper can be employed for the distributions satisfying this condition.

Assumption 4 A probability mass function $f_\theta(x)$ satisfies that

$$\sum_{x=h_1}^{h_2} f_\theta(x) \tag{2}$$

is a unimodal function, a decreasing function or an increasing function of θ for each h_1 and h_2, where h_1 and h_2 satisfy $0 \leq h_1 \leq h_2 \leq n$.

The methodology proposed in this paper can be applied to exponential families.

Proposition 1 *Exponential families with monotone likelihood ratio in x satisfy Assumption 4.*

Proof For exponential families with monotone likelihood ratio in x, it will be shown that (2) is a decreasing function if $h_1 = 0$, an increasing function if $h_2 = n$, and a unimodal function if $0 < h_1 < h_2 < n$, respectively. By Lemma 2 of Chap. 3 in Lehmann (1986), for $\theta < \theta_1$, the cumulative distribution functions of X under θ and θ_1 satisfy $\sum_{x=0}^{h_2} f_\theta(x) \geq \sum_{x=0}^{h_2} f_{\theta_1}(x)$ for all h_2, which implies that (2) is a decreasing function when $h_1 = 0$ and (2) is an increasing function when $h_2 = n$ because $\sum_{x=h_1}^{n} f_\theta(x) = 1 - \sum_{x=0}^{h_1-1} f_\theta(x)$. For $0 < h_1 < h_2 < n$, there exists a θ_1 and a θ_2 such that h_1 and h_2 can be viewed as C_1 and C_2 of $\phi(x)$ in Theorem 6 of Chap. 3 in Lehmann (1986). By (iii) in Theorem 6, the power function has a maximum at a point θ_0 between θ_1 and θ_2, and decreases strictly as θ tends away from θ_0 in either direction. Note that (2) is equal to the power function. Therefore, (2) can be shown to be a unimodal function of θ for $0 < h_1 < h_2 < n$. □

Before proving the main result, some notation is given. For a confidence interval $(L(X), U(X))$, there are $2(n+1)$ endpoints corresponding to $X = 0, \ldots, n$. Assume that there are g values of the $2(n+1)$ values between l and u. Rank the g values from the smallest value to the largest value, say $v_1, \ldots, v_g$. Then the parameter space can be separated into $g + 1$ intervals by these g points. Let Ω^o denote the interior of Ω, and $W = \{w : w = l, w = u, w = L(X)$ *or* $U(X), X = 0, \ldots, n,\ w \in \Omega^o\}$, where W is the set containing the lower and upper boundary points of the parameter space and the endpoints of the confidence intervals belonging to Ω^o. For the Agresti-Coull interval example, the set W is $\{0, 0.48906, 0.02031, 0.64037, 0.11598, 0.77091, 0.22908, 0.88401, 0.35962, 0.97968, 0.51093, 1\}$.

Theorem 1 *Let $f_\theta(x)$ satisfy Assumption 4. For interval $(L(X), U(X))$ of θ satisfying Assumption 3 and either Assumption 1 or Assumption 2, the confidence coefficient is the minimum value of the coverage probabilities at $\theta \in W$.*

The proof of Theorem 1 is in the Appendix, which is a generalization of Theorem 1 in Wang (2007).

Let $I(\cdot)$ denote the indicator function. For a fixed endpoint $w \in W$, the coverage probability at the endpoint w is

$$\sum_{X \in S} I(w \in (U(X), L(X)) f_w(X),$$

which can be calculated by straightforward calculation or derived from Theorem 2. Let $[x]$ indicate the Gaussian notation, the largest integer not exceeding the value x. Theorem 2 gives the exact value of the coverage probabilities at each endpoint.

Theorem 2 *Under the same assumptions as in Theorem 1,*

(i) *the coverage probability at a lower endpoint $L(x)$ is $\sum_{i=w_1(x)}^{x-1} f_\theta(i)$, where $w_1(x) = \max([U^{-1}(L(x))] + 1, 0)$;*
(ii) *the coverage probability at an upper endpoint $U(x)$ is $\sum_{i=x+1}^{w_2(x)} f_\theta(i)$, where $w_2(x) = \min([L^{-1}(U(x))], n)$;*
(iii) *the coverage probability at the lower boundary point l of the parameter space is $\sum_{i=0}^{[L^{-1}(l)]} f_\theta(i)$ or $\sum_{i=1}^{[L^{-1}(l)]} f_\theta(i)$ when the confidence interval satisfies Assumption 1 or Assumption 2, respectively;*
(iv) *the coverage probability at the upper boundary point u of the parameter space is $\sum_{i=[U^{-1}(u)]+1}^{n} f_\theta(i)$ or $\sum_{i=[U^{-1}(u)]+1}^{n-1} f_\theta(i)$ when the confidence interval satisfies Assumption 1 or Assumption 2, respectively.*

Proof For any point p in (v_i, v_{i+1}), where $v_i = L(x)$, the maximum value of X such that the confidence interval based on X covers p is x, because the left endpoint of the interval based on $x+1$ is $L(x+1)$ which is greater than $L(x)$ by Assumption 1 or Assumption 2. Therefore $L(x+1) \geq v_{i+1}$. However, the maximum value of x such that the confidence interval based on x covers v_i is $x-1$ because v_i does not belong to (v_i, v_{i+1}). By a similar argument, the minimum value of X, such that the confidence interval based on X covers p, can be derived by finding the smallest y such that $L(x) < U(y)$. This is because the interval based on y does not cover the point in (v_i, v_{i+1}) if $U(y) \leq L(x)$. The solution of y is $[U^{-1}(L(x))] + 1$. By the above argument and $x \geq 0$, $w_1(x)$ is $\max([U^{-1}(L(x))] + 1, 0)$. This completes the proof of (i). By a similar argument, (ii)–(iv) can be proved. □

For the Agresti-Coull interval example, we have $w_1(x) = a_1(x)$ and $w_2(x) = a_2(x)$, where

$$a_1(x) = \max([U^{-1}(L(X))] + 1, 0)$$
$$= \max\left(\left[\frac{\tilde{n}((1-\frac{\tilde{x}}{\tilde{n}})z_{\frac{\alpha}{2}}^2 + n(\frac{\tilde{x}}{\tilde{n}} - 2z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{x}(\tilde{n}-\tilde{x})}{\tilde{n}}}/\tilde{n}))}{\tilde{n} + z_{\frac{\alpha}{2}}^2} - k^2/2\right] + 1, 0\right),$$

and

$$a_2(x) = \min([L^{-1}(U(X))], n)$$
$$= \min\left(\left[\frac{\tilde{n}((1-\frac{\tilde{x}}{\tilde{n}})z_{\frac{\alpha}{2}}^2 + n(\frac{\tilde{x}}{\tilde{n}} + 2z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{x}(\tilde{n}-\tilde{x})}{\tilde{n}}}/\tilde{n}))}{\tilde{n} + z_{\frac{\alpha}{2}}^2} - k^2/2\right], n\right).$$

Note that $U^{-1}(L(X))$ and $L^{-1}(U(X))$ can be derived by solving the largest integer y satisfying

$$\tilde{x}/\tilde{n} - z_{\alpha/2}\sqrt{\tilde{x}(\tilde{n}-\tilde{x})/\tilde{n}}/\tilde{n} < y/\tilde{n} + z_{\alpha/2}\sqrt{y(\tilde{n}-y)/\tilde{n}}/\tilde{n}$$

and solving the smallest integer y satisfying

$$\tilde{x}/\tilde{n} + z_{\alpha/2}\sqrt{\tilde{x}(\tilde{n}-\tilde{x})/\tilde{n}}/\tilde{n} > y/\tilde{n} - z_{\alpha/2}\sqrt{y(\tilde{n}-y)/\tilde{n}}/\tilde{n},$$

respectively.

By Theorem 2, the coverage probabilities of a lower endpoint $L(x)$ and an upper endpoint $U(x)$ are $\sum_{i=a_1(x)}^{x-1} \binom{n}{i} p^i \times (1-p)^{(n-i)}$ and $\sum_{i=x+1}^{a_2(x)} \binom{n}{i} p^i (1-p)^{(n-i)}$, respectively. The coverage probability at the lower boundary point 0 and the upper boundary point 1 are $\sum_{i=0}^{[nz_{\alpha/2}^2/(2n+4z_{\alpha/2}^2)]} \binom{n}{i} p^i \times (1-p)^{(n-i)}$ and

$$\sum_{i=[n(2n+3z_{\alpha/2}^2)/(2(n+2z_{\alpha/2}^2))]+1}^{n} \binom{n}{i} p^i (1-p)^{(n-i)},$$

respectively.

Therefore, by Theorem 2, the coverage probabilities corresponding to the endpoints between 0 and 1 are listed in Table 1.

The coverage probabilities at both the lower and upper boundary points are 1. The minimum value of these coverage probabilities is 0.89405, which is the coverage probability of the Agresti-Coull interval at $p = 0.11598$ and

Table 1 Coverage probabilities of the 95% Agresti-Coull confidence intervals at the endpoints between 0 and 1 for $n = 5$

Endpoint	0.020	0.116	0.229	0.359	0.489	0.511	0.640	0.771	0.884	0.979
Probability	0.902	0.894	0.917	0.940	0.937	0.937	0.940	0.917	0.894	0.902

$p = 0.88405$. Thus, the exact confidence coefficient for this example is 0.89405.

In this case, we only need to consider the coverage probabilities at the endpoints between 0 and 1 because Wang (2007) showed that the minimum coverage probability does not occur at the lower boundary point and the upper boundary point. The confidence coefficients of the interval corresponding to different n are provided in Wang (2007).

Note that the confidence intervals considered in this paper are open intervals. For a closed confidence interval, the coverage probability behavior is the same as for the corresponding open interval except at the endpoints of the confidence intervals. The coverage probability at each endpoint for closed confidence intervals needs to include the probability of the observation corresponding to the endpoint. Since for an open confidence interval, the minimum coverage probability occurs at one of the endpoints, for a closed confidence interval, there exists an infimum coverage probability, which is the same as the minimum coverage probability for the corresponding open confidence interval. The infimum coverage probability does not occur at any point in the parameter space for a closed confidence interval because the coverage probabilities at the endpoints are higher than those for the corresponding open interval.

Procedure 1 is proposed to obtain the exact confidence coefficient of a confidence interval $(L(X), U(X))$ according to Theorem 1.

Procedure 1 (Computing the exact confidence coefficient)

Step 1. For a confidence interval, check if Assumption 3 is satisfied and if either Assumption 1 or Assumption 2 is satisfied. If Assumption 3 is not satisfied, the confidence coefficient is zero, and we do not need to go to step 2.

Step 2. If Assumption 3 and either Assumption 1 or Assumption 2 are satisfied, list the endpoints belonging to the parameter space.

Step 3. Calculate the coverage probabilities at each endpoint of Step 2 and at the lower boundary point and the upper boundary point of the parameter space. Then calculate the minimum value of these coverage probabilities, which is the exact confidence coefficient.

Remark 2 If the parameter space is known to be a restricted parameter space, Procedure 1 still holds by replacing the set of endpoints in Step 2 by the subset of endpoints belonging to the restricted parameter space. The minimum coverage probability is the minimum value of the coverage probabilities at θ in the subset and at the lower boundary point and upper boundary point of the subset.

3 Average coverage probability

In this section, we propose a procedure to compute the exact average coverage probability of a confidence interval. To derive (1), we need to calculate the coverage probability $P_\theta(\theta \in (L(X), U(X)))$ in (1). Since the coverage probability is a function of θ, we call it a coverage probability function.

Theorem 3 *Let X be a random variable from a discrete distribution with a probability mass function $f_\theta(x)$. For a confidence interval satisfying Assumption 3 and either Assumption 1 or Assumption 2, assume there are g endpoints which belong to Ω^o out of $2(n+1)$ endpoints. The g points, $v_1, \ldots, v_g$, separate the parameter space into $(g+1)$ intervals. Besides the first interval (l, v_1), the left endpoint of each g interval is a lower endpoint or an upper endpoint of the confidence interval. The left endpoint of the first interval is lower boundary point l of the parameter space.*

For θ belonging to the first interval (l, v_1), the coverage probability function is $\sum_{i=0}^{[L^{-1}(l)]} f_\theta(i)$ or $\sum_{i=1}^{[L^{-1}(l)]} f_\theta(i)$ when the confidence interval satisfies Assumption 1 or Assumption 2, respectively.

For θ belonging to (v_i, v_{i+1}) of the g intervals,

(i) *if the left endpoint v_i is a lower endpoint $L(x)$, the coverage probability function for θ belonging to the interval is*

$$\sum_{i=w_1(x)}^{x} f_\theta(i) \quad \text{for } \theta \in (v_i, v_{i+1}); \tag{3}$$

(ii) *if the left endpoint v_i is an upper endpoint $U(x)$, the coverage probability function for θ belonging to the interval is*

$$\sum_{i=x+1}^{w_2(x)} f_\theta(i) \quad \text{for } \theta \in (v_i, v_{i+1}),$$

where $w_1(x)$ and $w_2(x)$ are the same as in Theorem 2.

Proof For θ belonging to the first interval (l, v_1), since $L(0) \le l \le U(0)$, the minimum value of x such that the confidence interval based on x covers the points belonging to (l, v_1) is 0 or 1 when the confidence interval satisfies Assumption 1 or Assumption 2, respectively. The maximum value of x such that the confidence interval based on x covers the points belonging to (l, v_1) is the largest y such that $L(y) < l$. Thus, the coverage probability function is $\sum_{i=0}^{[L^{-1}(l)]} f_\theta(i)$ or $\sum_{i=1}^{[L^{-1}(l)]} f_\theta(i)$ when the confidence interval satisfies Assumption 1 or Assumption 2.

When $v_i = L(x)$ is a lower endpoint, the confidence intervals covering the points $p \in (v_i, v_{i+1})$ are the set of the confidence intervals covering the point v_i and the confidence interval based on x. Thus, by (i) of Theorem 2, the proof of (i) is complete. When $v_i = U(x)$ is an upper point, the confidence interval based on x does not cover the points $p \in (v_i, v_{i+1})$ and the point v_i. Therefore, the confidence intervals covering the points $p \in (v_i, v_{i+1})$ are exactly the same as those which cover v_i. Thus, by a similar argument (ii) in Theorem 2, the proof of (ii) is complete. □

Procedure 2 (Computing the average coverage probability (ACP) under the prior $\eta(\theta)$)

Step 1. For a confidence interval, check if Assumption 1 or Assumption 2 is satisfied.

Step 2. If Assumption 1 or Assumption 2 is satisfied and assuming that there are g endpoints in Ω^o, rank the g endpoints from the smallest to the largest, say $v_1, \dots, v_g$.

Step 3. Calculate the coverage probability function of each interval (v_i, v_{i+1}) by Theorem 3, say e_i for $i = 1, \dots, g$ and the coverage probability function of the interval (l, v_1), say e_0.

Step 4. The exact average coverage probability under $\eta(\theta)$ is

$$\int_l^{v_1} e_0(\theta)\eta(\theta)d\theta + \int_{v_1}^{v_2} e_1(\theta)\eta(\theta)d\theta + \cdots + \int_{v_i}^{v_{i+1}} e_i(\theta)\eta(\theta)d\theta + \cdots + \int_{v_g}^{u} e_g(\theta)\eta(\theta)d\theta.$$

Remark 3 If the parameter space is known to be restricted, Procedure 2 still holds by replacing the set of the endpoints in Step 2 with the subset of the endpoints belonging to the restricted parameter space. The average coverage probability is

$$S_1 = \int_a^{v_{o_1}} e_{o_1-1}(\theta)\eta(\theta)d\theta + \sum_{i=o_1}^{o_2} \int_{v_i}^{v_{i+1}} e_i(\theta)\eta(\theta)d\theta + \int_{v_{o_2}}^{b} e_{o_2+1}(\theta)\eta(\theta)d\theta,$$

where a and b are the lower endpoint and the upper endpoint of the restricted parameter space, and $v_{o_1}, v_{o_1+1}, \dots, v_{o_2}$ are the endpoints in Step 2 belonging to the interior of the restricted parameter space.

The computation of the integrations in Step 4 can be numerically calculated by software such as Mathematica or Matlab. The numerical integration can provide a good approximation of the exact value of the integration.

For the Agresti-Coull interval example, by Theorem 3, when v_i is a lower endpoint $L(x) = x/n - z_{\alpha/2}\sqrt{x(n-x)/n}/n$, the coverage probability function of $p \in (v_i, v_{i+1})$ is

$$\sum_{i=a_1(x)}^{x} \binom{n}{i} p^i (1-p)^{n-i}, \tag{4}$$

and when v_i is an upper endpoint $U(x) = x/n + z_{\alpha/2}\sqrt{x(n-x)/n}/n$, the coverage probability function is

$$\sum_{i=x+1}^{a_2(x)} \binom{n}{i} p^i (1-p)^{n-i}. \tag{5}$$

The average coverage probability for the Agresti-Coull example is

$$\begin{aligned}
&\int_0^{0.0203} \sum_{i=0}^{0} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.0203}^{0.116} \sum_{i=0}^{1} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.116}^{0.2291} \sum_{i=0}^{2} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.2291}^{0.3596} \sum_{i=0}^{3} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.3596}^{0.4891} \sum_{i=0}^{4} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.4891}^{0.5109} \sum_{i=1}^{4} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.5109}^{0.6404} \sum_{i=1}^{5} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.6404}^{0.7709} \sum_{i=2}^{5} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.7709}^{0.8840} \sum_{i=3}^{5} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.8840}^{0.9797} \sum_{i=4}^{5} \binom{n}{i} p^i (1-p)^{(n-i)} dp \\
&+ \int_{0.9797}^{1} \sum_{i=5}^{5} \binom{n}{i} p^i (1-p)^{(n-i)} dp = 0.9666.
\end{aligned} \tag{6}$$

Table 2 lists the coverage probabilities with respect to the prior $\eta(p) = 1$ for different n.

Note that the confidence intervals discussed are open intervals. If confidence intervals are changed to closed inter-

Table 2 Average coverage probabilities of the 95% Agresti-Coull confidence intervals corresponding to different n

n	Average coverage probability
5	0.9666
10	0.9645
15	0.9630
20	0.9618
25	0.9609
30	0.9601
50	0.9580
100	0.9555

vals, the average coverage probabilities are the same as the open intervals except in the cases where the prior $\eta(\theta)$ is not continuous and has mass on the endpoints.

4 Examples

In this section, two more examples are provided to illustrate Procedures 1 and 2.

Example 1 Let X have a binomial distribution $B(n,p)$. The usual Wald confidence interval for p is

$$\left(\frac{X}{n}-\frac{1}{n}z_{\alpha/2}\sqrt{\frac{X(n-X)}{n}},\ \frac{X}{n}+\frac{1}{n}z_{\alpha/2}\sqrt{\frac{X(n-X)}{n}}\right). \quad (7)$$

By Blyth and Still (1983) and Lehmann and Loh (1990), the confidence coefficient of (7) is zero because it decreases to zero as p goes to zero. By applying Procedure 2 to compute the ACP, it is necessary to know the coverage probability function of p for each interval (v_i, v_{i+1}). By Theorem 3, when v_i is a lower endpoint $L(x) = x/n - z_{\alpha/2}\sqrt{x(n-x)/n}/n$, the coverage probability function of $p \in (v_i, v_{i+1})$ is

$$\sum_{i=b_1(x)}^{x}\binom{n}{i}p^i(1-p)^{n-i}, \quad (8)$$

where

$$b_1(x) = \max([U^{-1}(L(X))]+1, 0) = \max\left(\left[\frac{n((1-\frac{x}{n})z^2_{\frac{\alpha}{2}} + n(\frac{x}{n} - 2z_{\frac{\alpha}{2}}\sqrt{\frac{x(n-x)}{n}}/n))}{n+z^2_{\frac{\alpha}{2}}}\right] + 1, 0\right).$$

When v_i is an upper endpoint $U(x) = x/n + z_{\alpha/2}\sqrt{x(n-x)/n}/n$, the coverage probability function is

$$\sum_{i=x+1}^{b_2(x)}\binom{n}{i}p^i(1-p)^{n-i}, \quad (9)$$

where

$$b_2(x) = \min([L^{-1}(U(X))], n) = \min\left(\left[\frac{n((1-\frac{x}{n})z^2_{\frac{\alpha}{2}} + n(\frac{x}{n} + 2z_{\frac{\alpha}{2}}\sqrt{\frac{x(n-x)}{n}}/n))}{n+z^2_{\frac{\alpha}{2}}}\right], n\right).$$

Since the Wald interval satisfies Assumption 2, by Theorem 3, the coverage probability function for p belonging to the first interval $(0, v_1)$ is

$$\sum_{i=1}^{b_3(x)}\binom{n}{i}p^i(1-p)^{n-i}, \quad (10)$$

where

$$b_3(x) = [L^{-1}(0)] = [nz^2_{\alpha/2}/(n+z^2_{\alpha/2})].$$

Procedure 2 is illustrated by calculating an example for the case of $n = 5$. Assume the prior on the parameter space is $\eta(p) = 1$. For $n = 5$ and $z_{\alpha/2} = 1.96$, the six confidence intervals are E, (−0.1506, 0.5506), (−0.0294, 0.8294), (0.1706, 1.0294), (0.4494, 1.1506), E corresponding to $x = 0, \ldots, 5$, where E denotes the empty set. The parameter space (0, 1) can be separated into these intervals: (0, 0.1706), (0.1706, 0.4494), (0.4494, 0.5506), (0.5506, 0.8294), (0.8294, 1). By (8), (9) and (10), the exact average coverage probability of the interval under $\eta(p)$ is

$$\begin{aligned}&\int_0^{0.1706}\sum_{i=1}^{2}\binom{n}{i}p^i(1-p)^{(n-i)}dp \\ &+\int_{0.1706}^{0.4494}\sum_{i=1}^{3}\binom{n}{i}p^i(1-p)^{(n-i)}dp \\ &+\int_{0.4494}^{0.5506}\sum_{i=1}^{4}\binom{n}{i}p^i(1-p)^{(n-i)}dp \\ &+\int_{0.5506}^{0.8294}\sum_{i=2}^{4}\binom{n}{i}p^i(1-p)^{(n-i)}dp \\ &+\int_{0.8294}^{1}\sum_{i=3}^{4}\binom{n}{i}p^i(1-p)^{(n-i)}dp = 0.6406. \quad (11)\end{aligned}$$

Table 3 lists the average coverage probabilities of the 95% Wald intervals under the prior $\eta(p) = 1$ corresponding to different n.

Table 3 Average coverage probability of the 95% Wald confidence intervals corresponding to different n

n	Average coverage probability
5	0.6406
10	0.7692
15	0.8188
20	0.8458
25	0.8629
30	0.8749
50	0.9006
100	0.92225

Based on these results, we do not advocate use of the Wald interval for very small samples.

Example 2 Let Y follow a Poisson distribution with mean λ. The natural parameter space for λ is $(0,\infty)$. However, in practical applications, it is possible to assume that the parameter space is bounded from the data information. For applying the procedures, we assume the parameter space is $(0,5)$ in this example. The score confidence interval for λ is

$$\left(Y+\frac{z_{\alpha/2}^2}{2}-\frac{z_{\alpha/2}}{2}\sqrt{4Y+z_{\alpha/2}^2}, Y+\frac{z_{\alpha/2}^2}{2}+\frac{z_{\alpha/2}}{2}\sqrt{4Y+z_{\alpha/2}^2}\right).$$

By applying Theorem 2, the coverage probabilities at a left endpoint $v=y+z_{\alpha/2}^2/2-z_{\alpha/2}/2\sqrt{4y+z_{\alpha/2}^2}$, and a right endpoint $v'=y+z_{\alpha/2}^2/2+z_{\alpha/2}/2\sqrt{4y+z_{\alpha/2}^2}$ are $\sum_{i=c_1(y)}^{y}\frac{e^{-v}v^i}{i!}$ and $\sum_{i=y+1}^{c_2(y)}\frac{e^{-v}v'^i}{i!}$, respectively, where

$$c_1(y)=\max\left([y+z_{\alpha/2}^2-z_{\alpha/2}\sqrt{4y+z_{\alpha/2}^2}]+1,0\right)$$

and

$$c_2(y)=\left[y+z_{\alpha/2}^2+z_{\alpha/2}\sqrt{4y+z_{\alpha/2}^2}\right].$$

According to Procedure 1, we list all the endpoints in the parameter space. There are nine endpoints in the restricted parameter space out of the endpoints of the confidence intervals (0, 3.8416), (0.1765, 5.6651), ..., (4.7350, 17.1066) corresponding to $y=0,\dots,9$. Following Procedure 1, we calculate the coverage probabilities at the 9 endpoints, which are $0.8382, 0.89478, \dots, 0.93906$ corresponding to endpoints $0.1765, 0.5484, \dots, 4.73502$, respectively. The minimum value of these coverage probabilities is 0.8382.

For deriving the ACP, let the prior of the parameter space be a Gamma distribution with density function $\lambda^{\alpha-1}e^{-\lambda/\beta}/(\Gamma(\alpha)\beta^\alpha)$. Applying Procedure 2, we list all the endpoints in the parameter space. The coverage probability function of $\lambda\in(v_i,v_{i+1})$ in the parameter space is

$$\sum_{i=c_1(y)}^{y}\frac{e^{-\lambda}\lambda^i}{i!}, \tag{12}$$

when v_i is a lower endpoint $L(y)=y+z_{\alpha/2}^2/2-z_{\alpha/2}/2\sqrt{4y+z_{\alpha/2}^2}$.

When v_i is an upper endpoint $y+z_{\alpha/2}^2/2+z_{\alpha/2}/2\sqrt{4y+z_{\alpha/2}^2}$, the coverage probability function is

$$\sum_{i=y+1}^{c_2(y)}\frac{e^{-\lambda}\lambda^i}{i!}. \tag{13}$$

The exact ACP with respect to different α and β can be calculated by Procedure 2. Table 4 lists the average coverage probabilities corresponding to the Gamma prior with different α and β.

Table 4 Average coverage probabilities of the 95% score confidence intervals based on $x=0,\dots,10$ corresponding to different (α,β)

(α,β)	Average coverage probability
(1, 2)	0.9526
(1, 1)	0.9494
(2, 0.25)	0.9446
(2, 2)	0.9567
(3, 0.25)	0.9499
(3, 2)	0.9571

5 Simulation study

In this section, we conduct a simulation study to reinforce the theoretical results. The minimum values of the coverage probabilities of the Agresti-Coull interval at 10,000 randomly chosen points in the parameter space are listed in Table 5 for different n. The results reveal that the minimum values are greater than the confidence coefficients derived from the proposed method, which both supports the valid of the methodology and shows its efficiency, since we need to calculate the coverage probabilities at much fewer points.

We also calculate the mean of the coverage probabilities of the Agresti-Coull interval and Wald interval at m randomly chosen points in the parameter space. As shown in Table 6, the mean is approximating the exact average coverage probability as m goes to infinity.

Table 5 Exact confidence coefficients and minimum coverage probabilities at 10000 randomly chosen points in the parameter space of the 95% Agresti-Coull confidence intervals corresponding to different n

n	The minimum coverage probability at 10^4 randomly chosen points	Exact confidence coefficient
10	0.923956	0.923944
30	0.9338114	0.9338105
50	0.934517	0.934515834
100	0.9379707	0.9379661

Table 6 Means of m coverage probabilities at m randomly chosen points in the parameter space of the 95% Wald confidence intervals, where the exact values for the average coverage probabilities are 0.6406 and 0.8458 corresponding to $n = 5$ and $n = 20$

m	$n = 5$, means of coverage probabilities of Wald intervals	$n = 20$, means of coverage probabilities of Wald intervals
500	0.63397	0.84118
2000	0.63041	0.84650
5000	0.63559	0.84791

Table 7 Means of m coverage probabilities at m randomly chosen points in the parameter space of the 95% Agresti-Coull confidence intervals, where the exact values for the average coverage probabilities are 0.9645 and 0.9601 corresponding to $n = 10$ and $n = 30$

m	$n = 10$, means of coverage probabilities of Agresti-Coull intervals	$n = 30$, means of coverage probabilities of Agresti-Coull intervals
500	0.96402	0.95995
2000	0.96415	0.95995
5000	0.96448	0.96038

From the simulation study, we conclude that the proposed procedures provide a valid and efficient way to calculate the exact confidence coefficient and exact average coverage probability.

Appendix

Proof of Theorem 1 If the interval $(L(X), U(X))$ does not satisfy Assumption 3, then the confidence coefficient is zero. We do not need to calculate the confidence coefficient. If Assumption 3 is satisfied and there are g endpoints, $v_1, \ldots, v_g$, of the confidence intervals belonging to Ω^o, where $v_1 < \cdots < v_g$, then the parameter space can be separated into $g + 1$ intervals by these endpoints. The infimum coverage probability of each interval $\theta \in (v_i, v_{i+1})$ is the minimum coverage probability at $\theta = v_i$ or $\theta = v_{i+1}$ if the following equations can be proved:

$$\begin{aligned}
&\inf_{v_i<\theta<v_{i+1}} P(\theta \in (L(X), U(X))) \\
&\quad = \min_{\theta=v_i \text{ or } v_{i+1}} P(\theta \in (L(X), U(X))), \\
&\inf_{l<\theta<v_i} P(\theta \in (L(X), U(X))) \\
&\quad = \min_{\theta=l \text{ or } v_i} P(\theta \in (L(X), U(X)))
\end{aligned} \tag{14}$$

and

$$\begin{aligned}
&\inf_{v_g<\theta<u} P(\theta \in (L(X), U(X))) \\
&\quad = \min_{\theta=v_g \text{ or } u} P(\theta \in (L(X), U(X))).
\end{aligned}$$

Note that v_i and v_{i+1} are the lower endpoints of the confidence intervals, the upper endpoints of the confidence intervals, the lower boundary point of the parameter space or the upper boundary point of the parameter space.

Since

$$\begin{aligned}
&\inf_{\theta\in\Omega} P(\theta \in (L(X), U(X))) \\
&\quad = \min_{i=1,\ldots,n}\left(\inf_{v_i<\theta<v_{i+1}} P(\theta \in (L(X), U(X)))\right),
\end{aligned}$$

the infimum coverage probability of $\theta \in (0, 1)$ is the minimum coverage probability of θ, at one of the points in W or the two boundary points of the parameter space.

Now we have to prove (14). For a $\theta_0 \in (v_i, v_{i+1})$, by either Assumption 1 or Assumption 2, the coverage probability of the interval at θ_0 is

$$P(\theta_0 \in (L(X), U(X))) = \sum_{i=k_0(\theta_0)}^{k_1(\theta_0)} f_{\theta_0}(i), \tag{15}$$

where $k_0(\theta_0)$ is the smallest x, such that $\theta_0 < U(X)$, and $k_1(\theta_0)$ is the largest X, such that $L(X) < \theta_0$. Note that for a fixed i, $k_0(\theta_0)$ and $k_1(\theta_0)$ are the same for all $\theta \in (w_i, w_{i+1})$ because there are no endpoints between v_i and v_{i+1}. According to Assumption 4, (15) is an increasing function, a decreasing function or a unimodal function. Thus, by the properties of these functions, the minimum value of $\inf_{v_i<\theta<v_{i+1}} P(\theta \in (L(X), U(X)))$ happens at $\theta_0 = v_i$ or v_{i+1}. Thus, the proof is completed. □

Acknowledgements The author thanks an associate editor and referees for helpful comments.

References

Agresti, A., Coull, B.A.: Approximate is better than "exact" for interval estimation of binomial proportions. Am. Stat. **52**, 119–126 (1998)

Blyth, C.R., Still, H.A.: Binomial confidence intervals. J. Am. Stat. Assoc. **78**, 108–116 (1983)

Brown, L.D., Cai, T., DasGupta, A.: Confidence intervals for a binomial and asymptotic expansions. Ann. Stat. **30**, 160–201 (2002)
Brown, L.D., Cai, T., DasGupta, A.: Interval estimation in exponential families. Stat. Sin. **13**, 19–49 (2003)
Clopper, C.J., Pearson, E.S.: The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika **26**, 404–413 (1934)
Lehmann, E.L.: Testing Statistical Hypotheses, 2nd edn. Wiley, New York (1986)
Lehmann, E.L., Loh, W.Y.: Pointwise versus uniform robustness of some large-sample tests and confidence intervals. Scand. J. Stat. **17**, 177–187 (1990)
Stein, C.: On the coverage probability of confidence sets based on a prior distribution. In: Sequential Methods in Statistics. Banach Center Publ., vol. 16, pp. 485–514. PWN, Warsaw (1985)
Wang, H.: The monotone bound property and the full coverage property of confidence intervals for a binomial proportion. Technical report (2006)
Wang, H.: Exact confidence coefficients of confidence intervals for a binomial proportion. Stat. Sin. **17**, 361–368 (2007)
Woodroofe, M.: Very weak expansions for sequential confidence levels. Ann. Stat. **14**, 1049–1067 (1986)