



Exact confidence coefficients of simultaneous confidence intervals for multinomial proportions

Hsiuying Wang

Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan

Received 16 March 2006

Available online 18 May 2007

Abstract

Simultaneous confidence intervals for multinomial proportions are useful in many areas of science. Since 1964, approximate simultaneous $1 - \alpha$ confidence intervals have been proposed for multinomial proportions. Although at each point in the parameter space, these confidence sets have asymptotic $1 - \alpha$ coverage probability, the exact confidence coefficients of these simultaneous confidence intervals for a fixed sample size are unknown before.

In this paper, we propose a procedure for calculating exact confidence coefficients for simultaneous confidence intervals of multinomial proportions for any fixed sample size. With this methodology, exact confidence coefficients can be clearly derived, and the point at which the infimum of the coverage probability occurs can be clearly identified.

© 2007 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: 62F25

Keywords: Confidence coefficient; Coverage probability; Multinomial distribution; Simultaneous confidence intervals

1. Introduction

Let X_1, X_2, \dots, X_k be observed cell frequencies in a sample of size $N = \sum_{i=1}^k X_i$ from a multinomial distribution $M(N, p_1, \dots, p_{k-1})$ with cell probabilities p_1, \dots, p_k , and observations $X = (X_1, \dots, X_k)$, where $p_k = 1 - p_1 \cdots - p_{k-1}$. The problem of finding simultaneous confidence intervals for p_1, \dots, p_k was developed in the 1960s. Miller [7] gave a survey of this work, including the result of Goodman [4] and Quesenberry and Hurst [8]. Fitzpatrick and Scott [2] and Sison and Glaz [9] also proposed several simultaneous confidence intervals for p_1, p_2, \dots, p_k .

E-mail address: hywang@stat.sinica.edu.tw.

The confidence coefficients of these simultaneous confidence intervals are defined as the infimum of the coverage probabilities of these intervals in the parameter space $\Omega = \{(p_1, \dots, p_k) | 0 \leq p_i \leq 1, i = 1, \dots, k, \sum_{i=1}^k p_i = 1\}$. Usually, the exact confidence coefficients of these simultaneous confidence intervals are unknown since the infimum of the coverage probabilities may occur at any point in Ω . We do not know at which point in the parameter space the infimum of the coverage probability occurs. When N is large, a nominal coefficient is suggested as an approximation. However, for fixed N , the nominal coefficient may be much larger than its exact value. By applying other approaches to approximate the confidence coefficient, we can obtain an estimate, but not the exact value. Fitzpatrick and Scott [2] derived a lower bound for the asymptotic simultaneous confidence level of their confidence intervals. With this lower bound, the performance of the simultaneous confidence intervals is more apparent when the sample size is large. However, it is a lower bound for asymptotic coverage probability, and not the exact confidence coefficient for any sample size. Moreover, the confidence coefficients of other intervals are still unknown.

Thus, we propose a method for calculating the exact confidence coefficients of simultaneous confidence intervals for p_1, \dots, p_{k-1} for any fixed sample size in this paper. Note that since $p_k = 1 - p_1 \cdots - p_{k-1}$ for the multinomial distribution, it is natural to consider the simultaneous confidence intervals for p_1, \dots, p_{k-1} because there are only $k - 1$ variables indeed. Like in the binomial distribution ($k = 2$), we are interested in constructing a confidence interval for p_1 instead of simultaneous confidence intervals for p_1 and p_2 . With this proposed method, we only need to calculate the coverage probabilities at some finite points in the parameter space. The minimum of these coverage probabilities is the exact confidence coefficient, and not just an estimate.

Note that for the binomial distribution, it is a special case of this paper. A related result for calculating the confidence coefficients of confidence intervals for a binomial proportion is referred to Wang [10]. The techniques used for the binomial distribution cannot be directly applied to the multinomial distribution with $k > 2$ because there is only one variable and one unknown parameter p_1 involved. For the multinomial distribution case with $k > 2$, there are at least two dependent variables involved which causes that the derivation for the confidence coefficients of simultaneous confidence intervals for the multinomial distribution is much more difficult than that for the binomial distribution.

Methodology for calculating the exact confidence coefficient of the simultaneous confidence intervals is provided in Section 2. With this method, the exact confidence coefficient can be clearly derived, and the point at which the infimum of the coverage probabilities occurs can be identified. Section 3 provides a review of the simultaneous confidence intervals in the literature. In Section 4, the computation of the proposed method is illustrated using an example. Using the proposed procedure, the exact confidence coefficients of some simultaneous confidence intervals are presented. Simulation results to reinforce Theorem 1 are given in Section 5. The simulation results reveal that the minimum value of the coverage probabilities at many randomly chosen points in the parameter space is still larger than the value derived from the proposed method, which requires calculations on far fewer points.

2. The main result

Let $I(X) = (I^1(X), \dots, I^{k-1}(X))$, where $I^j(X) = (L_j(X), U_j(X))$, $j = 1, \dots, k - 1$ be simultaneous confidence intervals of p_1, \dots, p_{k-1} . Let $\hat{p}_i = X_i/N$ denote the maximum likelihood estimator of p_i . Assume that, for fixed N , $L_j(X)$ depends on \hat{p}_j , but does not depend on \hat{p}_i , $i \neq j$. We say that $L_j(X)$ has the same form as $L_i(X)$ if $L_j(X)$ is the same as $L_i(X)$ when the term \hat{p}_j in $L_j(X)$ is replaced by \hat{p}_i , and we say that $U_j(X)$ has the same form as

$U_i(X)$ if $U_j(X)$ is the same as $U_i(X)$ when the term \hat{p}_j in $U_j(X)$ is replaced by \hat{p}_i . In this paper, we assume that $L_j(X)$ and $U_j(X)$ have the same form as $L_i(X)$ and $U_i(X)$ for $i \neq j$, respectively. Assumption 1 covers some conditions for simultaneous confidence intervals required by Theorem 1.

Assumption 1.

- (i) For fixed N , $L_j(X)$ and $U_j(X)$ depend on \hat{p}_j , but do not depend on $\hat{p}_i, i \neq j$.
- (ii) For fixed N , $L_j(X)$ and $U_j(X)$ are increasing functions of \hat{p}_j .
- (iii) When $\hat{p}_j = 0$, $L_j(X) \leq 0 \leq U_j(X)$.
- (iv) For any fixed $p^* = (p_1^*, \dots, p_{k-1}^*)$ in the parameter space, there exists an x_0 such that $p^* \in I(x_0)$, and $P_{p^*}(X = x_0) > 0$.

If simultaneous confidence intervals $I(X)$ do not satisfy the fourth condition, the confidence coefficient is zero.

Before proving the main result, concerning the derivation of the confidence coefficient of simultaneous confidence intervals for p_1, \dots, p_{k-1} , we define notations and give some lemmas. Let

$$\begin{aligned}
 & S_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-1}}(x_1, \dots, x_{k-2})(p_1, \dots, p_{k-1}) \\
 &= \sum_{x_1=a_1}^{b_1} \cdots \sum_{x_i=a_i}^{b_i(x_1, \dots, x_{i-1})} \cdots \sum_{x_{k-1}=a_{k-1}}^{b_{k-1}(x_1, \dots, x_{k-2})} \frac{N!}{x_1! \cdots x_{k-1}!(N - x_1 - \cdots - x_{k-1})!} p_1^{x_1} p_2^{x_2} \\
 & \quad \cdots \left(1 - \sum_{j=1}^{k-1} p_j \right)^{(N - x_1 - \cdots - x_{k-1})} \tag{1}
 \end{aligned}$$

denote the sum of the probabilities of the set

$$O = \{(x_1, \dots, x_{k-1}) : a_1 \leq x_1 \leq b_1, a_i \leq x_i \leq b_i(x_1, \dots, x_{i-1}), i = 2, \dots, k - 1\}, \tag{2}$$

with cell probabilities $(p_1 \dots p_{k-1})$, where $a_1 \geq 0, b_1 \geq a_1, a_i \geq 0$ and $b_i(x_1, \dots, x_{i-1})$ is a function of x_1, \dots, x_{i-1} not less than a_i , and $x_1 + \dots + x_{k-2} + b_{k-1}(x_1, \dots, x_{k-2}) \leq N$. Note that (1) is a function depending on $p_1, \dots, p_{k-1}, a_1, \dots, a_{k-1}$ and $b_1, \dots, b_{k-1}(x_1, \dots, x_{k-2})$. To simplify the notations, $S_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-1}}(x_1, \dots, x_{k-2})(p_1, \dots, p_{k-1})$ is replaced by

$$S(p_1, \dots, p_{k-1}) \tag{3}$$

in the rest of the paper.

Let

$$S(p_{k-1} | p_1, \dots, p_{k-2}) \tag{4}$$

denote function (3) for fixed p_1, \dots, p_{k-2} , which is a curve of p_{k-1} in the surface (3).

Lemma 1. Let $W^* = \{(p_1, \dots, p_{k-1}) : 0 < p_i < 1, i = 1, \dots, k - 1, p_1 + \dots + p_{k-1} < 1\}$. When $b_i(x_1, \dots, x_{i-1}), i = 1, \dots, k - 2$ are constants and satisfy $a_1 = b_1, \dots, a_{k-2} = b_{k-2}$ in (3), which implies that $b_{k-1}(x_1, \dots, x_{k-2})$ is a constant, say b_{k-1} , then for $p \in W^*$:

- (i) (4) is a unimodal function of p_{k-1} with one maximum if a_{k-1} is greater than zero and $b_1 + \dots + b_{k-1}$ is less than N ;

- (ii) (4) is a decreasing function of p_{k-1} if a_{k-1} is zero; and
- (iii) (4) is an increasing function of p_{k-1} if $b_1 + \dots + b_{k-1}$ is N .

In Lemma 1, we discussed the special case of $a_1 = b_1, \dots, a_{k-2} = b_{k-2}$. In Lemma 2, we discuss a general case.

Lemma 2. Assume that W^* is the same as in Lemma 1. If $b_i(x_1, \dots, x_{i-1}) > a_i$ for some $i \in 1, \dots, k - 2$, then for $p \in W^*$:

- (i) (4) is a unimodal function or an increasing function of p_{k-1} if a_{k-1} is greater than zero, and $x_1 + \dots + b_{k-1}(x_1, \dots, x_{k-2}) < N$;
- (ii) (4) is a decreasing function of p_{k-1} if a_{k-1} is zero; and
- (iii) (4) is an increasing function of p_{k-1} if $x_1 + \dots + x_{k-2} + b_{k-1}(x_1, \dots, x_{k-2}) = N$.

Lemma 3. For a subset in the parameter space separated by the endpoints in each p_i -axis of the simultaneous confidence intervals satisfying (i) and (ii) of Assumption 1, without other endpoints of the confidence intervals in the interior of this subset, there exist constants a_i and functions $b_i(x_1, \dots, x_{i-1}), i = 1, \dots, k - 1$ such that the coverage probabilities of the confidence intervals at the points in this subset are given by (1).

The main result of this paper is briefly described as follows. For fixed N and specified simultaneous confidence intervals, there are $(N + 1)$ confidence intervals for p_i corresponding to $X_i = 0, \dots, N$. There are $2(N + 1)$ endpoints of these $(N + 1)$ confidence intervals. Assume that there are g endpoints between 0 and 1 out of these $2(N + 1)$ endpoints. Rank these g endpoints from the smallest value to the largest value, say v_1, \dots, v_g . The interval $(0, 1)$ can be separated into $(g + 1)$ intervals by these g endpoints. The set $\{p = (p_1, \dots, p_{k-1}) : 0 \leq p_i \leq 1, i = 1, \dots, k - 1, \sum_{i=1}^{k-1} p_i \leq 1\}$ can be separated into at most $(g + 1)^{k-1}$ subsets, see Fig. 1 for the $k = 3$ case.

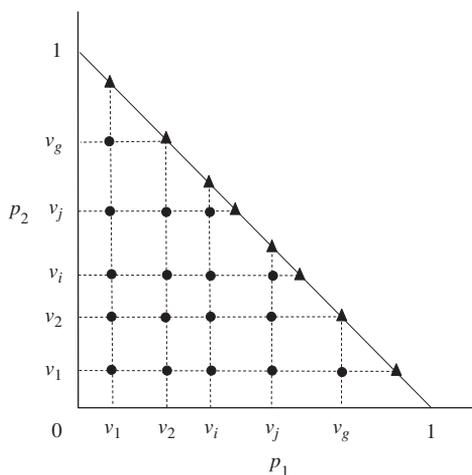


Fig. 1. For $k = 3$, the parameter space can be separated at most into $(g + 1)^2$ subsets. Each subset has four endpoints. The circles are the endpoints in E_1 of Theorem 1, and the triangles are the endpoints in E_2 of Theorem 1.

From Fig. 1, it is clear that the subset separated by the points v_i and v_j in the p_1 -axis and the point $v_{i'}$ and $v_{j'}$ in the p_2 -axis has four endpoints $(v_i, v_{i'})$, $(v_i, v_{j'})$, $(v_j, v_{i'})$, $(v_j, v_{j'})$. The endpoints of these subsets, which are in the interior of the parameter space, are the set $\{(m_1, m_2) : m_1 + m_2 < 1, m_i = v_j, 1 \leq j \leq g, i = 1, 2\}$, which are the circles in Fig. 1. The set $\{(m_1, m_2) : m_1 = v_j, 1 \leq j \leq g, m_2 = 1 - m_1\}$ is the set of triangles in Fig. 1. E_1 and E_2 in Theorem 1 are generations of the above two sets to the general k case. Theorem 1 shows that the minimum coverage probability is the minimum value of the coverage probabilities at the points in $E_1 \cup E_2$.

Theorem 1. *Let X have a multinomial distribution $M(N, p_1, \dots, p_{k-1})$ and $I(X)$ be simultaneous confidence intervals satisfying Assumption 1. Let $p = (p_1, \dots, p_{k-1})$ and $W = \{L_1(X_1 = 0), L_1(X_1 = 1), \dots, L_1(X_1 = N), U_1(X_1 = 0), U_1(X_1 = 1), \dots, U_1(X_1 = N)\}$. Assume that there are g points in W which are greater than 0 and less than 1. Then rank the g points from the smallest value to the largest value, say v_1, \dots, v_g . Let*

$$E_1 = \{p = (m_1, \dots, m_{k-1}) : m_i = v_j \text{ for } i = 1, \dots, k - 1, j = 1, \dots, g, m_1 + \dots + m_{k-1} < 1\},$$

and

$$E_2 = \{p = (m_1, \dots, m_{k-1}) : m_i = v_j \text{ for } i = 1, \dots, k - 2, j = 1, \dots, g, m_1 + \dots + m_{k-2} \leq 1, m_{k-1} = 1 - m_1 - \dots - m_{k-2}\},$$

then the infimum coverage probability, confidence coefficient, is the minimum value of the coverage probabilities at the points in $E_1 \cup E_2$.

Remark 1. $E_1 \cup E_2$ is equal to the set

$$E_1 = \{p = (m_1, \dots, m_{k-1}) : m_i = v_j \text{ for } i = 1, \dots, k - 1, j = 1, \dots, g, m_1 + \dots + m_{k-1} \leq 1\},$$

for some simultaneous confidence intervals, such as $I_{3,\alpha}$ and $I_{4,\alpha}$ in Section 4 for the $k = 3$ case because for each lower endpoint, there exists an upper endpoint such that the sum of these two points is one.

According to this result in Theorem 1, the procedure for establishing confidence coefficients is as follows:

Procedure for establishing confidence coefficients for simultaneous confidence intervals of multinomial proportions.

Step 1: Check if the simultaneous confidence intervals satisfy Assumption 1. If the fourth condition in Assumption 1 is not satisfied, then the exact confidence coefficient is zero.

Step 2: Calculate the coverage probabilities at the points of $(E_1 \cup E_2)$ in Theorem 1.

Step 3: Calculate the minimum value of the probabilities in Step 2. This value is the exact confidence coefficient.

Remark 2. In Step 2, when k is large, it may be not easy to calculate the coverage probabilities at the points of $(E_1 \cup E_2)$. In this case, these probabilities can be approximated by using simulation.

3. Simultaneous confidence intervals

In this section, we will review some simultaneous confidence intervals for multinomial proportions that have been given in the literature. The maximum likelihood estimators of p_j are

$\hat{p}_j = X_j/N, j = 1, \dots, k$. The random vector $(\hat{p}_1, \dots, \hat{p}_k)$ has an asymptotic multivariate normal distribution with mean vector (p_1, \dots, p_k) and covariance matrix Σ/N , where the elements in Σ are

$$\sigma_{jj} = p_j(1 - p_j),$$

and

$$\sigma_{jj'} = -p_j p_{j'} \quad \text{for } j \neq j'.$$

The simultaneous confidence intervals of (p_1, \dots, p_k) proposed by Gold [3] are

$$I'_{1,\alpha} = \left(I_{1,\alpha}^1, I_{1,\alpha}^2, \dots, I_{1,\alpha}^k \right),$$

where

$$I_{1,\alpha}^j = \left(\hat{p}_j - \left(\chi_{k-1,\alpha}^2 \right)^{1/2} [\hat{p}_j(1 - \hat{p}_j)/N]^{1/2}, \hat{p}_j + \left(\chi_{k-1,\alpha}^2 \right)^{1/2} [\hat{p}_j(1 - \hat{p}_j)/N]^{1/2} \right).$$

Goodman [4] considers the Bonferroni intervals $I'_{2,\alpha} = \left(I_{2,\alpha}^1, \dots, I_{2,\alpha}^k \right)$, where

$$I_{2,\alpha}^j = \left(\hat{p}_j - z_{\alpha/2k} [\hat{p}_j(1 - \hat{p}_j)/N]^{1/2}, \hat{p}_j + z_{\alpha/2k} [\hat{p}_j(1 - \hat{p}_j)/N]^{1/2} \right),$$

and $z_{\alpha/2k}$ denotes the upper $\alpha/2k$ cutoff point of a standard normal distribution. Quesenberry and Hurst [8] proposed their intervals based on the χ^2 statistics

$$\sum_{j=1}^k \frac{(X_j - Np_j)^2}{Np_j}, \tag{5}$$

which is asymptotically distributed as a χ^2 distribution with $k - 1$ degrees of freedom. Note that (5) is equal to $N(\hat{p}_1 - p_1, \dots, \hat{p}_k - p_k)' \Sigma^{-1} (\hat{p}_1 - p_1, \dots, \hat{p}_k - p_k)$ (referred to by Miller [7]). The simultaneous $(1 - \alpha)$ confidence intervals considered by Quesenberry and Hurst [8] are

$$I'_{3,\alpha} = \left\{ p_j : \frac{|\hat{p}_j - p_j|}{\sqrt{p_j(1 - p_j)/N}} < (\chi_{k-1,\alpha}^2)^{1/2}, j = 1 \dots k \right\}.$$

Solving the equation $\hat{p}_j - \left(\chi_{k-1,\alpha}^2 \right)^{1/2} [p_j(1 - p_j)/N]^{1/2} = p_j, I_{3,\alpha}^j$ can be rewritten as

$$I_{3,\alpha}^j = \left(\frac{c + 2N\hat{p}_j - \sqrt{c^2 + 4Nc(1 - \hat{p}_j)\hat{p}_j}}{2(c + N)}, \frac{c + 2N\hat{p}_j + \sqrt{c^2 + 4Nc(1 - \hat{p}_j)\hat{p}_j}}{2(c + N)} \right),$$

where $c = \chi_{k-1,\alpha}^2$.

The other two approaches are proposed by Fitzpatrick and Scott [2] and Sison and Glaz [9]. Fitzpatrick and Scott [2] proposed the simultaneous confidence intervals $I'_{4,\alpha} = \left(I_{4,\alpha}^1, \dots, I_{4,\alpha}^k \right)$,

where

$$I_{4,\alpha}^j = \left(\hat{p}_j - \frac{z_{\alpha/2}}{2\sqrt{N}}, \hat{p}_j + \frac{z_{\alpha/2}}{2\sqrt{N}} \right). \tag{6}$$

A lower bound for the asymptotic simultaneous confidence level of (6) is also given in their paper.

Another simultaneous intervals based on the approximation for multinomial probabilities by Levin [6] is proposed by Sison and Glaz [9]. They considered the region $I'_{5,\alpha} = \left(I_{5,\alpha}^1, \dots, I_{5,\alpha}^k \right)$, where

$$I_{5,\alpha}^j = \left(\hat{p}_j - b/N, \hat{p}_j + (b + 2r)/N \right),$$

and $b \geq 1$ and r are two values depending on the observations given in Sison and Glaz [9]

4. Confidence coefficients

The intervals $I'_{i,\alpha}, i = 1, \dots, 5$ proposed in the literature are simultaneous confidence intervals of p_1, \dots, p_k constructed by an approximation approach, and are not exact $1 - \alpha$ confidence intervals. Since we consider the simultaneous confidence intervals of p_1, \dots, p_{k-1} in this paper, let $I_{i,\alpha}$ denote the simultaneous confidence interval $(I_{i,\alpha}^1, \dots, I_{i,\alpha}^{k-1})$, for $i = 1, \dots, 5$. In this section, we will employ the methodology in Section 2 to calculate the exact confidence coefficients of $I_{i,\alpha}$ for some i . First, note that $I_{1,\alpha}$ and $I_{2,\alpha}$ do not satisfy conditions (ii) in Assumption 1, thus the methodology cannot be applied to these two intervals. From Fitzpatrick and Scott [2] and Sison and Glaz [9], $I_{1,\alpha}$ and $I_{2,\alpha}$ have worse performance than the other three simultaneous confidence intervals. In fact, the confidence coefficients of these two intervals are zero, as shown in Theorem 2.

Theorem 2. *The confidence coefficients of simultaneous confidence intervals $I_{1,\alpha}$ and $I_{2,\alpha}$ are zero.*

Proof. Using the result for the binomial case in Lehmann and Loh [5] and Blyth and Still [1], for $k = 2$, we have

$$\inf_{p_1 \in (0,1)} P_{p_1} \left(p_1 \in I_{i,\alpha}^1 \right) = 0 \quad \text{for } i = 1, 2 \tag{7}$$

for all α .

For $k \geq 2$, we have

$$\inf_{p \in \Omega} P_p(p \in I_{i,\alpha}) \leq \inf_{p_1} P_{p_1}(p_1 \in I_{1,\alpha}^1) = 0, \tag{8}$$

because $P_p(p \in I_{1,\alpha}) \leq P_{p_1}(p_1 \in I_{1,\alpha}^1)$. Thus, the confidence coefficients of $I_{i,\alpha}, i = 1, 2$ are zero. \square

We can provide an intuitive explanation of the result of Theorem 2 by observing the behavior of the intervals when $X_1 = 0$. For the intervals $I_{1,\alpha}$ and $I_{2,\alpha}, I_{1,\alpha}^1$ and $I_{1,\alpha}^2$ are $(0, 0)$ when $X_1 = 0$. When p_1 is close to zero, the probability of any observation greater than zero is close to zero because $\binom{n}{x} p_1^x (1 - p_1)^{n-x}$ goes to zero as p_1 goes to zero. Since $(0, 0)$ does not contain any points $p_1 \in (0, 1)$, therefore, the coverage probabilities of $I_{1,\alpha}$ and $I_{2,\alpha}$ go to zero as p_1 goes to zero. Consequently, the infimum coverage probabilities of $I_{1,\alpha}$ and $I_{2,\alpha}$ are zero.

The lower and upper endpoints of $I_{3,\alpha}^j$ and $I_{4,\alpha}^j$ satisfy conditions (i)–(iii) in Assumption 1. Differentiating the lower bound of $I_{3,\alpha}^j$ with respect to \hat{p}_j , gives

$$\begin{aligned} \partial L_j(X)/\partial \hat{p}_j &= \frac{N(c^2 + 4Nc(1 - \hat{p}_j)\hat{p}_j)^{-1/2}}{c + N} \\ &\quad \times [(c^2 + 4Nc(1 - \hat{p}_j)\hat{p}_j)^{1/2} - (c - 2c\hat{p}_j)] \\ &\geq 0 \end{aligned}$$

for all N . This implies that the lower bound of $I_{3,\alpha}^j$ is an increasing function of \hat{p}_j for all N . By a straightforward calculation, the upper bound of $I_{3,\alpha}^j$, and the lower and upper bounds of $I_{4,\alpha}^j$ are also increasing functions of \hat{p}_j for all N . The procedure to derive the exact confidence coefficient of $I_{3,\alpha}$ is illustrated by an example.

Example 1. Let $X = (X_1, X_2, X_3)$ have a multinomial distribution $M(N, p_1, p_2)$, where $N = 5$. Consider the simultaneous confidence intervals $I_{3,\alpha}$ of the probabilities, p_1, p_2 , where $\alpha = 0.05$. Applying the procedure in Section 2 to compute the confidence coefficient, we list the confidence intervals of p_1 , corresponding to $X_1 = 0, 1, \dots, 5$, which are (0, 0.5451016), (0.02595312, 0.7011078), (0.08872917, 0.8202911), (0.17970885, 0.9112708), (0.29889218, 0.9740469) and (0.45489843, 1). It is clear that the intervals satisfy Assumption 1. Although Assumption 1 is for simultaneous confidence intervals of (p_1, p_2) , we only need to check the confidence intervals for p_1 because the forms of confidence intervals are the same for p_1 and p_2 . The ranked endpoints of these six intervals from smallest to largest value are 0, 0.02595312, \dots , 0.9740469, 1. Let $v_1 = 0.02595312, v_2 = 0.08872917, \dots, v_9 = 0.9112708, v_{10} = 0.9740469$, which are the endpoints between 0 and 1. Then by step 2 of the procedure, we need to compute the coverage probabilities at the points in the sets $E_1 \cup E_2 = \{(p_1, p_2) = (m_1, m_2) : m_1 = v_i, m_2 = v_j, 1 \leq i \leq 10, 1 \leq j \leq 10, m_1 + m_2 \leq 1\}$. In this case, there are 47 points in $E_1 \cup E_2$. Therefore, there are 47 coverage probabilities that need to be calculated. The minimum value of these 47 coverage probabilities is 0.766. However, less than 47 coverage probabilities actually need to be calculated because the coverage probabilities at the points (v_i, v_j) and (v_j, v_i) are the same.

Note that the confidence coefficient for $I'_{i,\alpha}, i = 1, \dots, 5$ should be less than that for $I_{i,\alpha}$. The proposed procedure cannot be directly applied to compute the confidence coefficient for $I'_{i,\alpha}$. The derivation of computing the exact confidence coefficient for $I'_{i,\alpha}$ needs further investigations.

From Tables 1 and 2, the confidence coefficients for $I_{4,\alpha}$ are higher than $I_{3,\alpha}$. Confidence coefficients for $I_{3,\alpha}$ and $I_{4,\alpha}$ do not vary a lot with changing sample sizes. The simultaneous confidence intervals $I_{5,\alpha}$ do not satisfy condition (ii) in Assumption 1 just as $I_{1,\alpha}$ and $I_{2,\alpha}$. Since the result of Theorem 2.1 in Sison and Glaz [9] is for $I'_{5,\alpha}$, we calculate the coverage probabilities of $I'_{5,\alpha}$, instead of $I_{5,\alpha}$, at some randomly chosen points in the parameter space, and have a minimum coverage probability of 0.3091 at the point $(p_1, p_2) = (0.0031, 0.9824)$ for $k = 3$ and $N = 25$. This implies that the exact confidence coefficient is smaller than 0.309. By checking cases for some k and N , we conjecture that the confidence coefficient of $I_{5,\alpha}$ is greater than zero because it satisfies condition (iv) in Assumption 1. By calculating the minimum coverage probability of $I'_{3,\alpha}$ and $I'_{4,\alpha}$ at randomly chosen points in the parameter space, the minimum coverage probabilities of $I'_{3,\alpha}$ and $I'_{4,\alpha}$ are about 0.788 and 0.869, for $N = 25$ and $k = 3$, which are higher than $I'_{5,\alpha}$. Note that, in this case, the confidence coefficient of $I'_{4,\alpha}$ is higher than the exact confidence coefficient of $I_{4,\alpha}$ is due to simulation error.

Table 1

The table gives the exact confidence coefficients for $I_{3,0.05}$ corresponding to different N for $k = 3$

N	Confidence coefficient	(p_1, p_2)
5	0.7660488	(0.02595312, 0.02595312)
10	0.7708778	(0.01284343, 0.01284343)
15	0.772429	(0.00853351, 0.00853351)
20	0.773194	(0.00638945, 0.00638945)
25	0.7736498	(0.00510646, 0.00510646)
30	0.7739523	(0.00425256, 0.00425256)
50	0.774554	(0.00254816, 0.00254816)
100	0.7750025	(0.00127282, 0.00127282)

The points (p_1, p_2) are the points at which the minimum coverage probability occurs.

Table 2

The table gives the exact confidence coefficients for $I_{4,0.05}$ corresponding to different N for $k = 3$

N	Confidence coefficient	(p_1, p_2)
5	0.8783224	(0.3617386, 0.3617386)
10	0.8971484	(0.3901024, 0.3901024)
15	0.8914067	(0.4136361, 0.4136361)
20	0.8898067	(0.3808693, 0.4308693)
25	0.8917298	(0.4040035, 0.4040035)
30	0.8970429	(0.4210805, 0.3877472)
50	0.899548	(0.4214095, 0.4014095)
100	0.9110785	(0.4120017, 0.4120017)

The points (p_1, p_2) are the points at which the minimum coverage probability occurs.

The lower bound for the asymptotic coverage probability of $I'_{4,\alpha}$ in Fitzpatrick and Scott [2] is $1 - 2\alpha$ for $\alpha \leq 0.016$ and $6\Phi(3z_{\alpha/2}/\sqrt{8}) - 5$ for $0.016 \leq \alpha \leq 0.15$. They conjecture that $1 - 2\alpha$ is actually a lower bound for all values of α . Although the confidence coefficient result for $I_{4,\alpha}$ in this paper is not directly related to the confidence coefficient for $I'_{4,\alpha}$, from Table 2, their conjecture may be correct because the exact confidence coefficient is greater than $1 - 2\alpha = 0.9$ for $I_{4,\alpha}$ when $N = 100$.

Note that the programs for computing the exact confidence coefficients are available from the author upon request.

5. Simulations

The result of Theorem 1 is also examined by conducting simulations to calculate the minimum coverage probability of coverage probabilities at randomly chosen points in the parameter space Ω . For $k = 3$ and $N = 20$, the minimum coverage probability of $I_{3,0.05}$ is 0.77824 at one million randomly chosen points in the parameter space. The exact confidence coefficient derived by the procedure proposed in this paper is 0.773194031, which is the minimum value of 192 coverage probabilities at $E_1 \cup E_2$. For $k = 3$ and $N = 20$, the minimum coverage probability of $I_{4,0.05}$ is 0.8898634 at one million randomly chosen points in the parameter space. The exact confidence

coefficient derived by the procedure proposed in this paper is 0.8898067, which is the minimum value of 178 coverage probabilities at points in $E_1 \cup E_2$. The simulation results show that even considering 10^6 randomly chosen points in the parameter space, the minimum value of these coverage probabilities is near the exact confidence coefficient, but still cannot reach the exact value of the confidence coefficient. The simulation calculation is much more time-consuming than the proposed method, which calculates at far fewer points.

6. Conclusion

In this paper, a procedure for calculating the exact confidence coefficient is proposed, and we apply this procedure to derive the exact confidence coefficients of some simultaneous confidence intervals, which were previously unknown. Compared with other approaches that only provide an estimate of confidence coefficients, this method provides an efficient and accurate way to obtain confidence coefficients.

Acknowledgment

The author would like to thank the referee for valuable comments, which help to improve the paper a lot.

Appendix A.

Note that for $h \in (2, \dots, k - 1)$, set (2) can be represented as

$$\{(x_1, \dots, x_{k-1}) : a'_1 \leq x_1 \leq b'_1, a'_i \leq x_i \leq b'_i(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{i-1}), i = 1, \dots, h - 1, h + 1, \dots, k - 1, a'_h \leq x_h \leq b'_h(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-1})\},$$

where $a'_1 \geq 0, b'_1 \geq a'_1, a'_i \geq 0$ and $b'_h(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-1})$ is a function of $x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-1}$ not less than a'_h , and $x_1 + \dots + x_{k-1} + b'_h(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-1}) \leq N$. Thus, (3) can be represented as

$$\begin{aligned} & S(p_1, \dots, p_{k-1}) \\ &= \sum_{x_1=a'_1}^{b'_1} \dots \sum_{x_{k-1}=a'_{k-1}}^{b'_{k-1}(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-2})} \sum_{x_h=a'_h}^{b'_h(x_1, \dots, x_{h-1}, x_{h+1}, \dots, x_{k-1})} \\ & \times \frac{N!}{x_1! \dots x_{k-1}!(N - x_1 - \dots - x_{k-1})!} \\ & \times p_1^{x_1} \dots p_{h-1}^{x_{h-1}} p_{h+1}^{x_{h+1}} \dots p_{k-1}^{x_{k-1}} p_h^{x_h} \left(1 - \sum_{j=1}^{k-1} p_j\right)^{(N-x_1-\dots-x_{k-1})}. \end{aligned} \tag{9}$$

Then

$$S(p_h | p_1, \dots, p_{h-1}, p_{h+1}, \dots, p_{k-1}) \tag{10}$$

denotes function (9) for fixed $p_1, \dots, p_{h-1}, p_{h+1}, \dots, p_{k-1}$. If $h = 1$ in (9), (9) and (10) are equal to

$$\begin{aligned}
 & S(p_1, \dots, p_{k-1}) \\
 &= \sum_{x_2=a'_2}^{b'_2} \cdots \sum_{x_{k-1}=a'_{k-1}}^{b'_{k-1}(x_2, \dots, x_{k-2})} b'_1(x_2, \dots, x_{k-1}) \frac{N!}{x_1! \cdots x_{k-1}!(N - x_1 - \cdots - x_{k-1})!} \\
 & \quad \times p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} p_1^{x_1} \left(1 - \sum_{j=1}^{k-1} p_j \right)^{(N-x_1-\cdots-x_{k-1})}, \tag{11}
 \end{aligned}$$

and $S(p_1|p_2, \dots, p_{k-1})$, which denotes function (11) for fixed p_2, \dots, p_{k-1} , respectively.

The following three lemmas are for (3) and (4), the case of $h = k - 1$. The results of Lemmas 1–3 can be applied to (9) and (10) for general h .

Proof of Lemma 1. (i) We first consider the case of $a_{k-1} > 0$ and $b_1 + \cdots + b_{k-1} < N$. Since $a_1 = b_1, \dots, a_{k-2} = b_{k-2}$, (3) equals

$$\begin{aligned}
 & \sum_{x_{k-1}=a_{k-1}}^{b_{k-1}} \frac{N!}{a_1! \cdots x_{k-1}!(N - a_1 \cdots - a_{k-2} - x_{k-1})!} p_1^{a_1} \\
 & \quad \cdots p_{k-1}^{x_{k-1}} \left(1 - \sum_{j=1}^{k-1} p_j \right)^{(N-a_1 \cdots - x_{k-1})}.
 \end{aligned}$$

For any fixed a_1, \dots, a_{k-2} and p_1, \dots, p_{k-2} , differentiating $S(p_{k-1}|p_1, \dots, p_{k-2})$ with respect to p_{k-1} , gives

$$\begin{aligned}
 & \frac{\partial S(p_{k-1}|p_1, \dots, p_{k-2})}{\partial p_{k-1}} \\
 &= \frac{a_{k-1}N!}{a_1! \cdots a_{k-1}!(N - a_1 \cdots - a_{k-1})!} p_1^{a_1} \cdots p_{k-1}^{a_{k-1}-1} \left(1 - \sum_{j=1}^{k-1} p_j \right)^{(N-a_1 \cdots - a_{k-1})} \\
 & \quad - \frac{(N - a_1 \cdots - b_{k-1})N!}{a_1! \cdots b_{k-1}!(N - a_1 \cdots - b_{k-1})!} p_1^{a_1} \cdots p_{k-1}^{b_{k-1}} \left(1 - \sum_{j=1}^{k-1} p_j \right)^{(N-a_1 \cdots - b_{k-1})-1} \tag{12}
 \end{aligned}$$

$$\begin{aligned}
 &= N! p_1^{a_1} \cdots p_{k-1}^{a_{k-1}-1} \left(1 - \sum_{j=1}^{k-1} p_j \right)^{N-a_1 \cdots - a_{k-1}} \\
 & \quad \times \left(\frac{a_{k-1}}{a_1! \cdots a_{k-1}!(N - a_1 \cdots - a_{k-1})!} \right. \\
 & \quad \left. - \frac{(N - a_1 \cdots - b_{k-1})(p_{k-1}/(1 - \sum_{j=1}^{k-1} p_j))^{b_{k-1}-a_{k-1}+1}}{a_1! \cdots b_{k-1}!(N - a_1 \cdots - b_{k-1})!} \right). \tag{13}
 \end{aligned}$$

Eq. (12) holds because middle terms are canceled, leaving only the first and the last terms. For deriving the unimodal property, it is necessary to solve Eq. (13) set to 0, which is equivalent to the equation

$$A = p_{k-1} / \left(1 - \sum_{j=1}^{k-1} p_j \right), \tag{14}$$

where

$$A = \left(\frac{a_{k-1}}{(N - a_1 \cdots - b_{k-1})} \frac{b_{k-1}!(N - a_1 \cdots - b_{k-1})!}{a_{k-1}!(N - a_1 \cdots - a_{k-1})!} \right)^{1/b_{k-1} - a_{k-1} + 1}.$$

Solving (14) with respect to p_{k-1} , we have a maximum of p_{k-1} at

$$u = \frac{A(1 - \sum_{j=1}^{k-2} p_j)}{1 + A}. \tag{15}$$

Since $a_{k-1} > 0, a_1 + \cdots + b_{k-1} < N$, and $p_{k-1}/(1 - \sum_{j=1}^{k-1} p_j)$ is an increasing function of p_{k-1} for fixed p_1, \dots, p_{k-2} , (13) is less than zero if $p_{k-1} > u$, and (13) is greater than zero if $p_{k-1} < u$. Thus, for fixed p_1, \dots, p_{k-2} , (4) is a unimodal function with a maximum at (15) when $a_1 = b_1, \dots, a_{k-2} = b_{k-2}$.

(ii) When $a_{k-1} = 0$, (13) is not greater than zero. Consequently, (4) is decreasing in p_{k-1} .

(iii) When $a_1 + \cdots + a_{k-2} + b_{k-1} = b_1 + \cdots + b_{k-2} + b_{k-1} = N$, (13) is not less than zero. Consequently, (4) is an increasing function. \square

Proof of Lemma 2. First, consider case (i). For fixed p_i, \dots, p_{k-2} , differentiating (4) with respect to p_{k-1} , gives

$$\begin{aligned} & \frac{\partial S(p_{k-1} | p_1, \dots, p_{k-2})}{\partial p_{k-1}} \\ &= N! p_{k-1}^{a_{k-1}-1} \sum_{x_1=a_1}^{b_1} \cdots \sum_{x_{k-2}=a_{k-2}}^{b_{k-2}(x_1, \dots, x_{k-3})} p_1^{x_1} \cdots p_{k-2}^{x_{k-2}} \\ & \times \left(1 - \sum_{j=1}^{k-1} p_j \right)^{N-x_1 \cdots -a_{k-1}} \left(\frac{a_{k-1}}{x_1! \cdots a_{k-1}!(N - x_1 \cdots - a_{k-1})!} \right. \\ & \left. - \frac{(N - x_1 \cdots - b_{k-1}(x_1, \dots, x_{k-2})) (p_{k-1}/(1 - \sum_{j=1}^{k-1} p_j))^{b_{k-1}(x_1, \dots, x_{k-2}) - a_{k-1} + 1}}{x_1! \cdots b_{k-1}(x_1, \dots, x_{k-2})!(N - x_1 \cdots - b_{k-1}(x_1, \dots, x_{k-2}))!} \right), \end{aligned} \tag{16}$$

which is a summation of higher order polynomials in p_{k-1} .

Unlike (13), which is only one higher order polynomial and is easy to solve, it is difficult to use an argument similar to that used in Lemma 1 to solve Eq. (16) set to 0 since it does not have a closed

form. Although it is difficult to know the exact value of the root, for proving the unimodal property of (4), it is only necessary to show that there is only one root of p_{k-1} of Eq. (16) set to 0. To show this, let

$$B_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-1}}(x_1, \dots, x_{k-2})(p_{k-1}) = \sum_{x_1=a_1}^{b_1} \cdots \sum_{x_{k-2}=a_{k-2}}^{b_{k-2}(x_1, \dots, x_{k-3})} \frac{p_1^{x_1} \cdots p_{k-2}^{x_{k-2}} (1 - \sum_{j=1}^{k-1} p_j)^{-x_1 \cdots - a_{k-1}} a_{k-1}}{x_1! \cdots a_{k-1}! (N - x_1 \cdots - a_{k-1})!},$$

which are the positive terms divided by $N! p_{k-1}^{a_{k-1}-1} (1 - \sum_{j=1}^{k-1} p_j)^N$ in (16), and let

$$C_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-1}}(x_1, \dots, x_{k-2})(p_{k-1}) = \sum_{x_1=a_1}^{b_1} \cdots \sum_{x_{k-2}=a_{k-2}}^{b_{k-2}(x_1, \dots, x_{k-3})} \frac{(N - x_1 \cdots - b_{k-1}(x_1, \dots, x_{k-2})) p_1^{x_1} \cdots p_{k-2}^{x_{k-2}} (1 - \sum_{j=1}^{k-1} p_j)^{-x_1 \cdots - b_{k-1}(x_1, \dots, x_{k-2})}}{x_1! \cdots b_{k-1}(x_1, \dots, x_{k-2})! (N - x_1 \cdots - b_{k-1}(x_1, \dots, x_{k-2}))!} \times p_{k-1}^{b_{k-1}(x_1, \dots, x_{k-2}) - a_{k-1} + 1},$$

which are the negative terms divided by $N! p_{k-1}^{a_{k-1}-1} (1 - \sum_{j=1}^{k-1} p_j)^N$ in (16). To simplify the notations, we use $B(p_{k-1})$ and $C(p_{k-1})$ to replace

$$B_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-2}}(x_1, \dots, x_{k-3})(p_{k-1}),$$

and $C_{a_1, \dots, a_{k-1}, b_1, \dots, b_{k-2}}(x_1, \dots, x_{k-3})(p_{k-1})$ in the rest of the proof. Note that for fixed p_1, \dots, p_{k-2} , $C(p_{k-1})$ is an increasing function of p_{k-1} because p_{k-1} and $1/(1 - \sum_{j=1}^{k-1} p_j)$ are increasing functions of p_{k-1} . $B(p_{k-1})$ is also an increasing function of p_{k-1} . The first and second derivatives with respect to p_{k-1} of $B(p_{k-1})$ and $C(p_{k-1})$ are positive for $p_{k-1} \in (0, 1 - \sum_{j=1}^{k-2} p_j)$. Hence $B(p_{k-1})$ and $C(p_{k-1})$ are two strictly increasing convex functions. The domain of p_{k-1} is $(0, 1 - \sum_{j=1}^{k-2} p_j)$. When p_{k-1} goes to zero, $C(p_{k-1})$ goes to zero, and $B(p_{k-1})$ goes to a positive number. When p_{k-1} goes to $1 - \sum_{j=1}^{k-2} p_j$, both $B(p_{k-1})$ and $C(p_{k-1})$ go to infinity. There are at most two intersections of $B(p_{k-1})$ and $C(p_{k-1})$.

If $B(p_{k-1})$ and $C(p_{k-1})$ have two intersections, and (4) has a local minimum and a local maximum at the two points, respectively, then it contradicts the conditions in (i). Since by the conditions $a_{k-1} > 0$ and $x_1 + \cdots + b_{k-1}(x_1, \dots, x_{k-2}) < N$, (4) goes to 0 as p_{k-1} goes to 0 and p_{k-1} goes to $1 - \sum_{j=1}^{k-2} p_j$, which contradicts that (4) has only one local minimum and one local maximum.

If $B(p_{k-1})$ and $C(p_{k-1})$ have no intersection, $B(p_{k-1})$ is greater than $C(p_{k-1})$ for all $p_{k-1} \in (0, 1 - \sum_{j=1}^{k-2} p_j)$ because $B(p_{k-1})$ is greater than $C(p_{k-1})$ when p_{k-1} goes to zero. Then (16) is positive and (4) is an increasing function.

If $B(p_{k-1})$ and $C(p_{k-1})$ have one intersection at $p_{k-1} = t$, $B(p_{k-1})$ is less than $C(p_{k-1})$ when p_{k-1} goes to $1 - \sum_{j=1}^{k-2} p_j$. Then (16) is increasing for $p \in (0, t)$ and (16) is decreasing for $p \in (t, 1 - \sum_{j=1}^{k-2} p_j)$. Thus, (16) is a unimodal function with a maximum at $p_{k-1} = t$.

Combining the above results, (16) is a unimodal function or an increasing function of p_{k-1} .

(ii) When $a_{k-1} = 0$, (16) is not greater than zero, and (4) is a decreasing function.

(iii) When $x_1 + \dots + x_{k-2} + b_{k-1}(x_1, \dots, x_{k-2}) = N$, (16) is not less than zero, and (4) is an increasing function. \square

Proof of Lemma 3. For a fixed subset, by (ii) in Assumption 1 and the fact there are no other endpoints of the confidence interval in the interior of this subset, there exists common $a_1, b_1, c_i(x_1, \dots, x_{i-1})$ and $b_i(x_1, \dots, x_{i-1}), i = 2, \dots, k - 1$ such that the form of the coverage probabilities at (p_1, \dots, p_{k-1}) in this subset is

$$\sum_{x_1=a_1}^{b_1} \dots \sum_{x_i=c_i(x_1, \dots, x_{i-1})}^{b_i(x_1, \dots, x_{i-1})} \dots \sum_{x_{k-1}=c_{k-1}(x_1, \dots, x_{k-2})}^{b_{k-1}(x_1, \dots, x_{k-2})} \times \frac{N!}{x_1! \dots x_{k-1}!(N - x_1 - \dots - x_{k-1})!} p_1^{x_1} p_2^{x_2} \dots \left(1 - \sum_{j=1}^{k-1} p_j\right)^{(N-x_1-\dots-x_{k-1})}. \quad (17)$$

Let a_i denote $c_i(x_1 = a_1, \dots, x_{i-1} = a_{i-1})$ for $i \geq 2$. That is, $a_2 = c_2(x_1 = a_1)$ and $a_3 = c_3(x_1 = a_1, x_2 = a_2) \dots$ etc. We are going to show $c_2(x_1 = a_1 + 1) = c_2(x_1 = a_1) = a_2$. Note that for a point $(p_1^*, p_2^*, \dots, p_{k-1}^*)$ in the subset, $p_2^* \in (L_2(c_2(x_1 = i)), U_2(c_2(x_1 = i)))$, $i = a_1$ and $i = a_1 + 1$ if $b_1 > a_1$. If $c_2(x_1 = a_1 + 1) < c_2(x_1 = a_1)$, then $a_1 + c_2(x_1 = a_1 + 1) < a_1 + c_2(x_1 = a_1) \leq N$, which contradicts that $c_2(x_1 = a_1)$ is the smallest value of x_2 such that the confidence intervals based on $x_1 = a_1$ and x_2 cover the points in the subset. Thus, we have $c_2(x_1 = a_1 + 1) \geq c_2(x_1 = a_1)$. If $c_2(x_1 = a_1 + 1) > c_2(x_1 = a_1)$, we have $a_1 + 1 + c_2(x_1 = a_1) < a_1 + 1 + c_2(x_1 = a_1 + 1) \leq N$, which contradicts that $c_2(x_1 = a_1 + 1)$ is the smallest value of x_2 such that the confidence intervals based on $x_1 = a_1 + 1$ and x_2 cover the parameters in the subset. Thus, combining the above argument, we have $c_2(x_1 = a_1 + 1) = c_2(x_1 = a_1) = a_2$. By a similar argument, we have $c_2(x_1 = j) = c_2(x_1 = a_1) = a_2$, where $a_1 < j \leq b_1$. Then by a similar argument, we have $c_i(x_1, \dots, x_{i-1}) = a_i$ for $i = 2, \dots, k - 1$. \square

Proof of Theorem 1. The space $\{(p_1, \dots, p_{k-1}) : 0 \leq p_i \leq 1, i = 1, \dots, k - 1\}$ can be separated into $(g + 1)^{k-1}$ subsets by the endpoints, which are between 0 and 1, in each p_i -axis. Thus, the parameter space $\{(p_1, \dots, p_{k-1}) : 0 \leq p_i \leq 1, i = 1, \dots, k - 1, p_1 + \dots + p_{k-1} \leq 1\}$ can be separated into at most $(g + 1)^{k-1}$ subsets by these endpoints. For a fixed subset, by Lemma 3, there exists a_i and $b_i(x_1, \dots, x_{i-1}), i = 1, \dots, k - 1$ such that the coverage probability of $I(X)$ at the parameter $(p_1 \dots p_{k-1})$ in this subset is (3).

We show that (3) attains its minimum value at one of the endpoints of this subset.

(i) Assume that the endpoints of this subset are in the interior of the parameter space. Denote the endpoints of this subset in the p_i -axis as v_{w_i} and v_{w_i+1} . For any fixed $p_1 \dots p_{k-2}$, (4) is a curve on (3). By Lemma 2, the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}}$ or $p_{k-1} = v_{w_{k-1}+1}$ when $a_{k-1} > 0$ and $b_1 + \dots + b_{k-1}(x_1, \dots, x_{k-2}) \leq N$ because (4) is a unimodal function or an increasing function of p_{k-1} , and the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}+1}$ when $a_{k-1} = 0$ because (4) is a decreasing function of p_{k-1} . Thus, to derive the minimum of (3), we only need to consider the cases when $p_{k-1} = v_{w_{k-1}}$ and $p_{k-1} = v_{w_{k-1}+1}$. By a similar argument, when $p_{k-1} = v_{w_{k-1}}$ and $p_1, \dots, p_{h-1}, p_{h+1} \dots p_{k-2}$ are fixed, the minimum value of (10) occurs at $p_h = v_{w_h}$ or $p_h = v_{w_h+1}$. By induction, the minimum value of (3) occurs at $p_1 = v_{w_1}$ or $v_{w_1+1}, p_2 = v_{w_2}$ or v_{w_2+1}, \dots and $p_{k-1} = v_{w_{k-1}}$ or $v_{w_{k-1}+1}$.

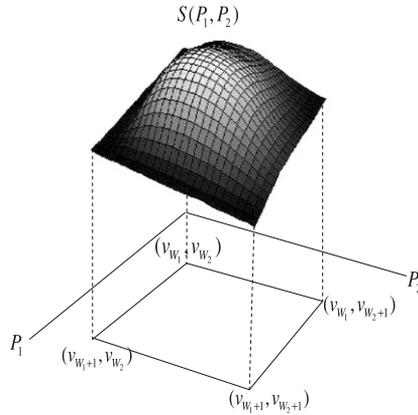


Fig. 2. One of the cases of $S(p_1, p_2)$ in a subset. It is clear that the minimum value of $S(p_1, p_2)$ occurs at $p_1 = v_{w_1}$ or v_{w_1+1} , and $p_2 = v_{w_2}$ or v_{w_2+1} when p is in the subset.

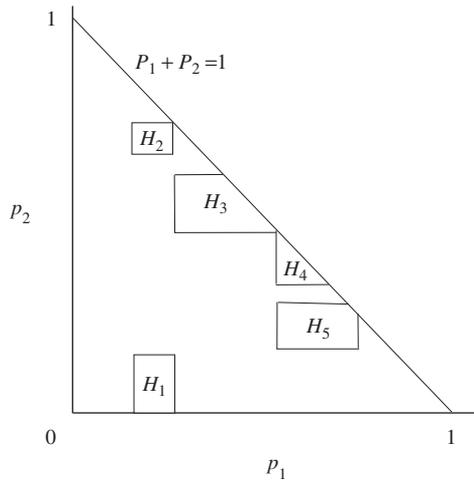


Fig. 3. The subsets separated by the endpoints of the confidence intervals have endpoints outside the interior of the parameter space, such as H_1, H_2, H_3, H_4 and H_5 .

For example, when $k = 3$, Fig. 2 shows one of the cases of $S(p_1, p_2)$ in a subset. It is clear that the minimum value of $S(p_1, p_2)$ occurs at $p_1 = v_{w_i}$ or v_{w_i+1} and $p_2 = v_{w_2}$ or v_{w_2+1} when p belongs to the subset $\{(p_1, p_2) : v_{w_1} < p_1 < v_{w_1+1}, v_{w_2} < p_2 < v_{w_2+1}\}$.

(ii) Assume that some endpoints of this subset are not in the interior of the parameter space, see Fig. 3 for the $k = 3$ case.

If some endpoints of this subset satisfy $p_i = 0$ for some i , like H_1 in Fig. 3 for the $k = 3$ case, by (iii) in Assumption 1, a_i is equal to zero. By (ii) of Lemma 2, $S(p_i | p_1 \cdots p_{i-1}, p_{i+1}, p_k)$ will not attain its minimum at the endpoints satisfying $p_i = 0$. Thus, the points satisfying $p_i = 0$ for some i do not need to be considered.

If some endpoints or interior points of this subset satisfying $p_1 + \dots + p_{k-1} = 1$, like H_2, H_3, H_4 or H_5 in Fig. 3 for the $k = 3$ case, it implies $p_k = 0$ for these points. In this case, there exists an observation $(x_1, \dots, b_{k-1}(x_1, \dots, x_{k-2}))$, satisfying $x_1 + \dots + b_{k-1}(x_1, \dots, x_{k-2}) = N$, such that the simultaneous confidence intervals based on this observation cover these point, otherwise, it contradicts (iv) of Assumption 1. Note that the coverage probability of $I(X)$ at these points is equal to

$$\sum_{x_1 + \dots + x_{k-1} \in F} \frac{N!}{x_1! \dots (N - x_1 \dots - x_{k-1})!} p_1^{x_1} \dots p_{k-1}^{x_{k-1}} \tag{18}$$

for the points with $p_k = 0$, where F is a subset of some points satisfying $x_1 + \dots + x_{k-1} = N$. Note that (18) is an increasing function in p_i for fixed $p_j, j \neq i$. Therefore, the minimum coverage probability of the set of points $\{(p_1, \dots, p_{k-1}) : p_1 + \dots + p_{k-1} = 1, a \leq p_i \leq b\}$ occurs at $p_i = a$. By Lemma 2, for p in this subset, when (4) is an increasing function, the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}}$. When (4) is a decreasing function, the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}+1}$ when $v_{w_{k-1}+1} = 1 - p_1 - \dots - p_{k-2}$, and the minimum value of (4) occurs at $p_{k-1} = 1 - p_1 - \dots - p_{k-2}$ if $v_{w_{k-1}+1} > 1 - p_1 - \dots - p_{k-2}$. When (4) is a unimodal function, the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}}$, or $v_{w_{k-1}+1}$ if $v_{w_{k-1}+1} < 1 - p_1 - \dots - p_{k-2}$, and the minimum value of (4) occurs at $p_{k-1} = v_{w_{k-1}}$ or $1 - p_1 - \dots - p_{k-2}$ if $v_{w_{k-1}+1} > 1 - p_1 - \dots - p_{k-2}$. Then by a similar argument as in (i) and the fact that (18) is an increasing function of p_i , the minimum coverage probability of the subset occurs at the intersection of this subset and $(E_1 \cup E_2)$.

Combining the above results, the confidence coefficient is the minimum value of the coverage probabilities at the points in $E_1 \cup E_2$. \square

References

[1] C.R. Blyth, H.A. Still, Binomial confidence intervals, J. Amer. Statist. Assoc. 78 (1983) 108–116.
 [2] S. Fitzpatrick, A. Scott, Quick simultaneous confidence intervals for multinomial proportions, J. Amer. Statist. Assoc. 82 (1987) 875–878.
 [3] R.Z. Gold, Tests auxiliary to chi-square tests in a Markov chain, Ann. Math. Statist. 34 (1963) 56–74.
 [4] L.A. Goodman, On simultaneous confidence intervals for multinomial proportions, Technometrics 7 (1965) 247–254.
 [5] E.L. Lehmann, W.Y. Loh, Pointwise versus uniform robustness of some large-sample tests and confidence intervals, Scand. J. Statist. 17 (1990) 177–187.
 [6] B. Levin, A representation for multinomial cumulative distribution functions, Ann. Statist. 9 (1981) 1123–1126.
 [7] R.G. Miller, Simultaneous Statistical Inference, second ed., Springer, New York, 1981.
 [8] C.P. Quesenberry, D.C. Hurst, Large-sample simultaneous confidence intervals for multinomial proportions, Technometrics 6 (1964) 191–195.
 [9] C.P. Sison, J. Glaz, Simultaneous confidence intervals and sample size determination for multinomial proportions, J. Amer. Statist. Assoc. 90 (1995) 366–369.
 [10] H. Wang, Exact confidence coefficients of confidence intervals for a binomial proportion, Statist. Sinica 17 (2007) 361–368.