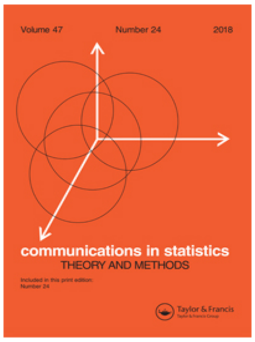

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: https://www.tandfonline.com/loi/lsta20

A new approach to precise interval estimation for the parameters of the hypergeometric distribution

Mark F. Schilling & Alyssa Stanley

To cite this article: Mark F. Schilling & Alyssa Stanley (2020): A new approach to precise interval estimation for the parameters of the hypergeometric distribution, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2020.1737879

To link to this article: https://doi.org/10.1080/03610926.2020.1737879

Published online: 16 Mar 2020.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
https://doi.org/10.1080/03610926.2020.1737879

A new approach to precise interval estimation for the parameters of the hypergeometric distribution

Mark F. Schilling and Alyssa Stanley

Department of Mathematics, California State University, Northridge, Los Angeles, California, USA

ABSTRACT

We study interval estimation for both parameters of the hypergeometric distribution: (i) the number of successes in a finite population and (ii) the size of the population. In contrast to traditional methods that specify intervals via a formula, our approach is to first establish the coverage probability function of an ideal procedure. This in turn determines the set of confidence intervals. In the case when the population size is known and we wish to estimate the number of successes, our approach is superior to existing methods in terms of average interval length and in fact achieves the minimum possible average length. In the case of estimating population size, our procedure also tends to produce shorter intervals than existing methods. Both procedures also possess an attractive coverage property.

ARTICLE HISTORY
Received 22 May 2019
Accepted 18 February 2020

KEYWORDS
Hypergeometric confidence interval; length minimizing; coverage probability

1. Introduction

A fairly common problem in statistics is to estimate the proportion of a finite population having a certain characteristic, which we may refer to as *success*, from a random sample of size n drawn from the population. When sampling is done without replacement, the probability of obtaining any particular number of successes is given by the hypergeometric distribution, which has the probability mass function

$$\mathrm{P}(X=x;\ M,\ N,\ n)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}},\quad \text{where } \max\{0,\ n+M-N\}\le x\le \min\{M,\ n\}$$

From observed data, a confidence interval may be desired (i) for the unknown parameter M if N is known; or (ii) for the unknown parameter N if M is known. Until recently, such interval estimates were based on formulas, see for example Katz (1953) and Thompson (2002). Today we can employ technology to easily generate non-formula based confidence intervals that perform better in certain important ways than do those obtained from formulas.

In the following, we develop confidence procedures for both M and N using an approach similar to that of Schilling and Doi (2014), Choi (2015), and Schilling and Holladay (2017). We provide a comparison to recently introduced high performing procedures due to Wang (2015), and also compare the performance of our procedures with

CONTACT Mark F. Schilling mark.schilling@csun.edu Department of Mathematics, California State University, Northridge, Los Angeles, CA, USA.

those obtainable from Blaker's (2000) recipe. We demonstrate the length superiority of our confidence procedures versus existing methods, show in addition that our procedures achieve high coverage, and also examine to what extent they achieve other desirable criteria.

The procedures we derive are *strict* in the sense that the minimum coverage probability over the parameter space is at least the nominal level 1 – α. However, our procedures can easily be modified to create *approximate* confidence procedures that achieve, for example, a *mean* coverage equal to 1 – α (see Schilling and Holladay 2017). (Note: The term *exact* is often used rather than *strict*; however this can lead to confusion since the term *exact* can also refer to the fact that the confidence procedure is derived from the actual distribution involved—here, the hypergeometric—rather than from, say, a normal approximation.)

2. Estimation of the number of successes

As discussed in Sahai and Khurshid (1995), the hypergeometric distribution plays an important role in modern biological and biomedical applications when it is reasonable to assume a finite population N. For example, suppose a researcher knows that N people have been exposed to a disease and wants to estimate the number of people M that are infected by it. If it is not practical to determine the disease status of every individual in the population, then in order to obtain an estimate for M, the researcher can observe a random sample n of those exposed and determine the number of people x from the sample that are infected. The sample proportion $\hat{p} = x/n$ gives a natural point estimate for the true proportion $p = M/N$, from which M is estimated by Nx/n. Obtaining a valid and precise interval estimate for either M or p is considerably more challenging, due to the discreteness of the hypergeometric distribution.

Another application of the hypergeometric distribution is to *acceptance sampling*. Suppose a company wants to determine whether to accept a lot or reject it for having too many defective items. Perhaps examining the entire lot is too expensive, takes too much time, or involves destructive sampling, so testing only a random sample of n items is feasible. Now suppose that it is acceptable to have x or less defective items in the sample, otherwise the lot cannot be distributed to consumers. The probability of accepting the lot is given by a function of the true number of defective items M:

$$P(X \leq x) = \sum_{i=0}^{x} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

This an example of what we call an *acceptance function*. In the particular case shown in Figure 1 there is a lot of 100 units that needs to be accepted or rejected; 25 items are sampled and at most five of them are allowed to be defective without rejecting the lot. As one would expect, the probability of accepting a lot decreases as the true number of defective units M in the lot increases.

The key to our construction of a high performing confidence procedure is to make suitable choices of acceptance functions for an associated partition of the parameter space.

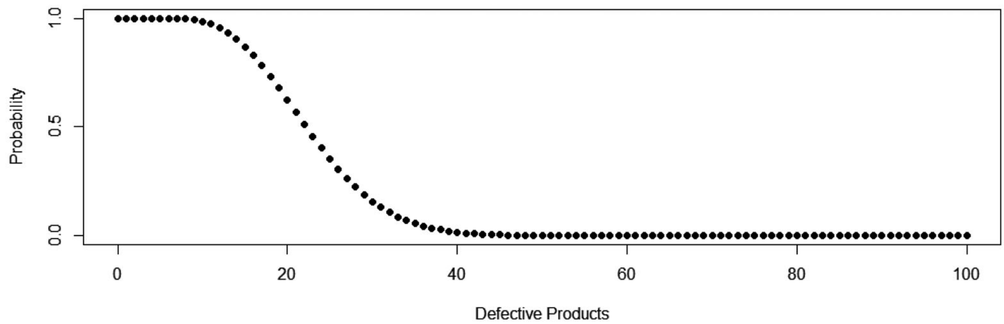

Figure 1. Probability of accepting a lot of size 100, given that five or less defective units in a sample size of 25 units is acceptable.

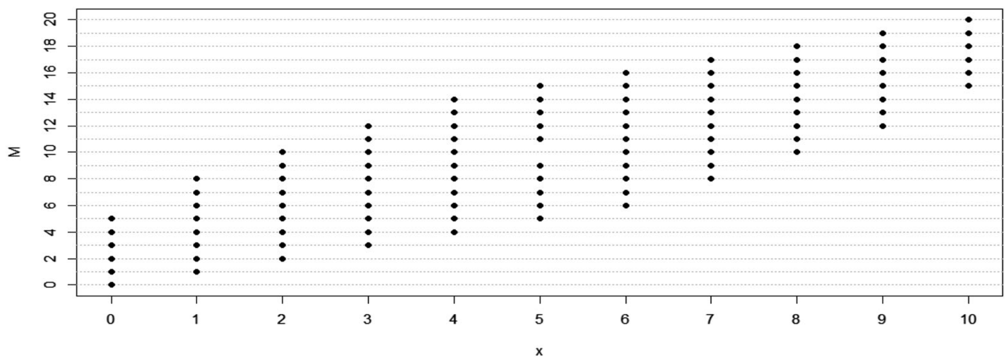

Figure 2. Hypothetical confidence sets.

2.1. Selection of acceptance functions

For a given confidence procedure for M, the coverage probability for a particular M is the probability $\mathrm{CP}(M)$ of observing any number x of successes in the sample for which the associated confidence interval contains M. The set of all $\mathrm{CP}(M)$ for $M \in \{0, 1, 2, \ldots, N\}$ is called the *coverage probability function* (*cpf*), and is comprised of portions of acceptance functions. The confidence procedures for M and N derived below are each *determined* by the choice of an ideal cpf.

Let the *acceptance set* Ω be any subset of $\{0, 1, 2, \ldots, n\}$; the acceptance function associated with Ω is $AF_\Omega(M) = P(X \in \Omega) = \sum_{x \in \Omega} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$. The question is how to choose the constituent acceptance functions to construct an optimal cpf.

Consider the situation illustrated in Figure 2, in which $N=20$ and $n=10$. Hypothetical acceptance sets are shown horizontally for each value of M. Note that at $M=10$, $\Omega = \{2, 3, 4\} \cup \{6, 7, 8\}$. Sets containing such holes horizontally typically create holes vertically. In this case, the resulting confidence set for M when $x=5$ would be $\{5, 6, 7, 8, 9\} \cup \{11, 12, 13, 14, 15\}$, which is not a proper interval since it is not a set of consecutive integers. Therefore we restrict the choice of acceptance sets to those of the form $\Omega_{lu} = \{l, l+1, \ldots, u-1, u\}$, since any other form of Ω can result in confidence sets for M that are not intervals. Acceptance sets of the form Ω_{lu} are called *acceptance intervals* (Blyth and Still 1983). We use the notation of $AF(l-u)$ to refer to the

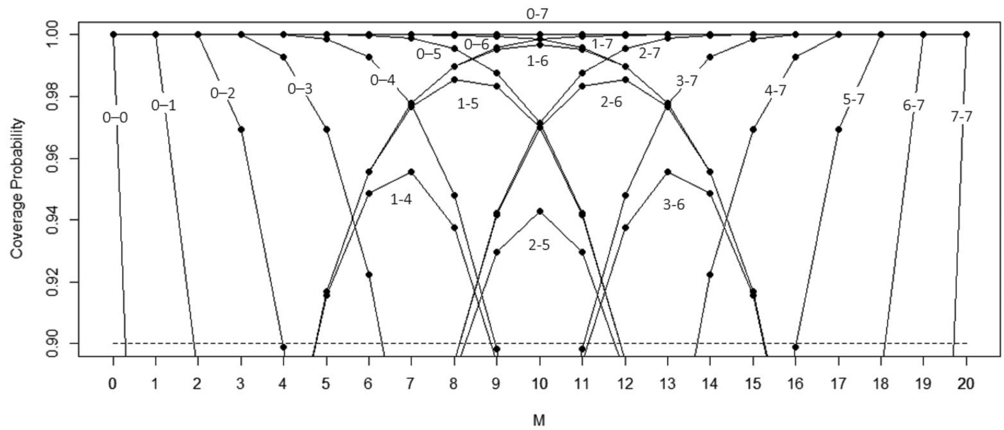

Figure 3. Coverage probabilities of acceptance functions eligible for selection for a 90% confidence procedure for $N = 20$, $n = 7$. Each acceptance function is labeled by its *l*-*u* values.

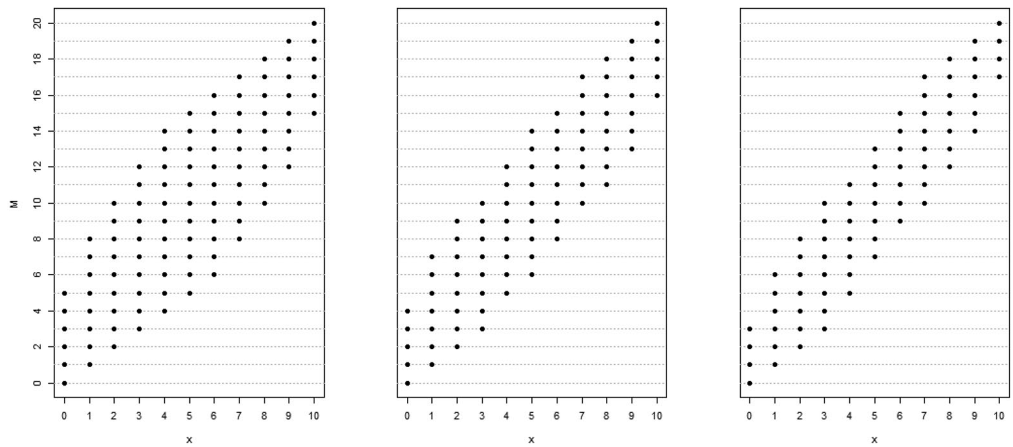

Figure 4. Confidence sets for the case $N = 20$, $n = 10$ for various sets of acceptance intervals.

acceptance function generated by the acceptance interval Ω_{lu} and we call $u - l$ the *span* of the acceptance function.

To illustrate the role of the acceptance sets in the construction of confidence intervals, we provide as an example the case where the population size N is 20 and the sample size n is 7. Figure 3 displays all acceptance functions of the type described above that have coverage probabilities greater than or equal to confidence level $1 - \alpha = .90$ for some values of M. For easier visualization, line segments have been added connecting the consecutive values of each acceptance function.

The goal of creating a confidence procedure with superior length in comparison to existing methods will drive the selection of acceptance functions. By superior length, we mean a procedure for which the mean length of all $n + 1$ confidence intervals is shorter than that of any prevailing strict confidence procedure, where by length we mean the *cardinality* of the confidence interval.

Due to a result from Crow (1956), choosing acceptance functions with minimal span leads to length minimization. To see why, see the example in Figure 4. The horizontal

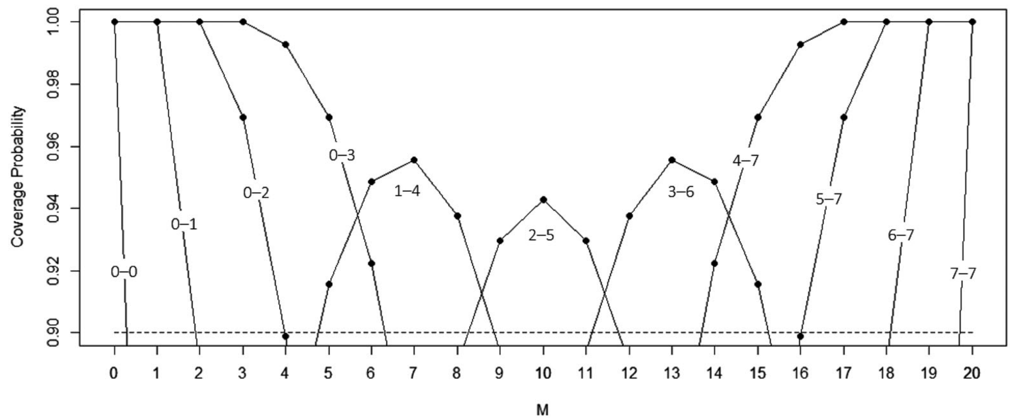

Figure 5. The set of minimal acceptance functions eligible for selection for a strict minimal span confidence procedure for $N=20$, $n=7$ and 90% confidence level.

set at each value of M represents the acceptance interval for that M, while the vertical sets represent the confidence intervals for each x that result. Moving from left to right in Figure 4, as the horizontal sets get shorter, so do the vertical sets. Thus, we expect that acceptance functions of minimal span are ideal for producing short confidence intervals. We state and prove below that this is indeed true.

Theorem: *Let C be a hypergeometric confidence procedure for estimating M constructed by choosing for each M an acceptance function of minimal span. Then C achieves the minimum possible total length of the confidence intervals for $x=0, 1, \ldots, n$.*

Proof: Let $\{A(l(M)-u(M));\ M=0, 1, \ldots, N\}$ be the acceptance intervals of C, and let $\{(LL(x), UL(x));\ x=0, 1, \ldots, n\}$ be the resulting confidence intervals. Since $(LL(x), UL(x)) = \{M{:}\ x \in A(l(M)-u(M))\}$, $\sum_{x=0}^{n}(UL(x)-LL(x)) = \sum_{x=0}^{N}(u(M)-l(M))$. Thus if each term in the latter sum is minimized, the total length of all confidence intervals, represented by the previous sum, achieves its minimum possible value. □

Thus from the collection of all the acceptance functions that lie above $1-\alpha$, we restrict the choice of acceptance functions to those with minimal span. Figure 5 shows the set of all available minimal span acceptance functions for $M \in \{0, 1, \ldots, N\}$ for the example $N=20$, $n=7$ given in Figure 3 above and confidence level 90%. At $M=5$ we have originally eight functions to select from: $AF(0-3)$, $AF(0-4)$, $AF(0-5)$, $AF(0-6)$, $AF(0-7)$, $AF(1-4)$, $AF(1-5)$, and $AF(1-6)$. The shortest span, three, is obtained by the two functions $AF(0-3)$ and $AF(1-4)$.

The question remains as to how to select the acceptance functions when there is a choice between two or more minimal span acceptance functions, as is the case for $M \in \{5, 6, 14, 15\}$ above. For interval estimation for other discrete distributions, various strategies have been used. For the binomial success parameter p, Crow (1956) chose for each value of the parameter the acceptance function with the largest l and u values. If this approach is used for the hypergeometric case illustrated in Figure 5, $AF(1-4)$ is chosen at $M=5$ rather than $AF(0-3)$. On the other hand for estimation of the Poisson parameter, Kabaila and Byrne (2001) use a method that is the opposite of Crow's. With this recipe, $AF(1-4)$ is not selected until $M=7$ for the case shown in Figure 5.

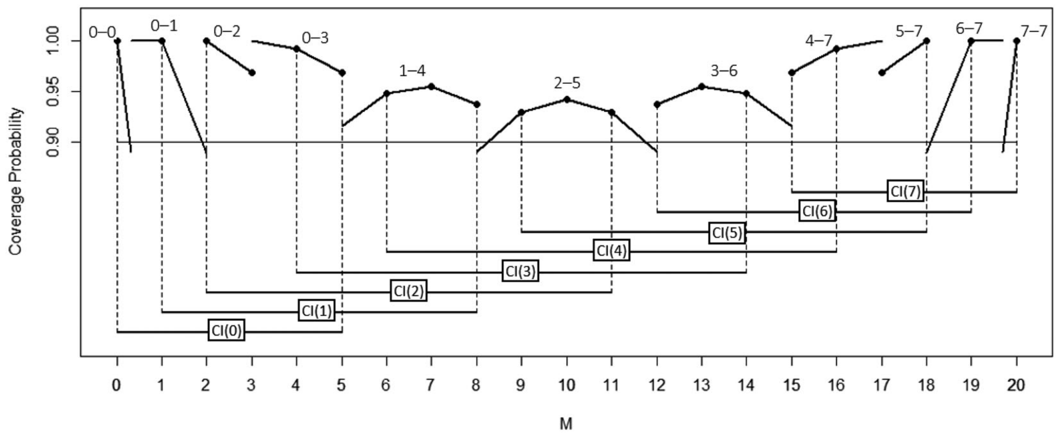

Figure 6. Coverage probability function with associated acceptance functions and confidence intervals for the case $N = 20$, $n = 7$ and 90% confidence level.

Blythe and Still (1983) use an intermediate solution for the binomial case that is between the previous two extremes: when two minimal span acceptance functions are available they transition between them at the midpoint of the interval for which both are available. However this method does not have a direct parallel for estimating the hypergeometric parameter M as the parameter space is discrete here.

We choose to transition in the same fashion as did Schilling and Doi (2014) for estimating the binomial success parameter. Contrary to the three previously mentioned methods, this strategy looks at the acceptance functions from top to bottom rather than from left to right. When faced with a choice between two or more minimal span acceptance functions, the one with the highest coverage probability is selected.

A rationale for this approach is as follows: When a researcher desires an interval estimate for an unknown parameter, the two primary objectives are precision and correctness. Thus a confidence procedure that tends to minimize interval length is paramount, and for discrete distributions, where coverage necessarily fluctuates rather than achieving the given confidence level exactly for all values of the parameter, maximizing the probability that the interval will capture the parameter is also desirable—particularly since end users may instinctively regard the confidence interval as containing the parameter despite this not being certain.

Using the described method, we choose $AF(0-3)$ at $M = 5$ in Figure 5 since it has a higher coverage probability than $AF(1-4)$. However, $AF(1-4)$ has a higher coverage probability than $AF(0-3)$ at $M = 6$, and thus is selected there. Figure 6 shows the resulting coverage probability function, along with the associated acceptance functions and the confidence intervals produced for the case $N = 20$, $n = 7$, and $1 - \alpha = 90\%$ after choosing minimal span acceptance functions with the highest coverage probability at each M.

Table 1 displays the 90% confidence intervals produced for this case as well as the intervals for confidence levels 95% and 99%. This example demonstrates several characteristics we would prefer that a confidence procedure possesses. Note that at each confidence level, both the lower limits and the upper limits of the confidence intervals increase as x increases. This is known as *monotonicity in x* and is a natural outcome—

Table 1. Confidence intervals for $N = 20$, $n = 7$ and $1 - \alpha = 90\%$, 95%, and 99% using the choice of minimal span acceptance functions with highest coverage probability.

	Confidence Level		
x	90%	95%	99%
0	[0, 5]	[0, 6]	[0, 8]
1	[1, 8]	[1, 10]	[1, 11]
2	[2, 11]	[2, 12]	[2, 13]
3	[4, 14]	[4, 14]	[3, 15]
4	[6, 16]	[6, 16]	[5, 17]
5	[9, 18]	[8, 18]	[7, 18]
6	[12, 19]	[11, 19]	[9, 19]
7	[15, 20]	[14, 20]	[12, 20]

as the number of successes in the sample increases, so does the estimated number of successes in the population.

Another feature that can be seen in Table 1 is *nesting*. A confidence procedure possesses the nesting property if for each x, every confidence interval contains any confidence interval with a lower confidence level that is computed from the same data. This property is satisfied at the confidence levels shown in Table 1. In Section 2.5 we explore the extent to which monotonicity and nesting hold in general for our confidence procedure for M.

Still another desirable property is *equivariance*: If x generates the confidence interval $[l_x,\ u_x]$ then $n - x$ yields the confidence interval $[N - l_x,\ N - u_x]$ (Blyth and Still 1983). Equivariance is appropriate when estimating the success parameter of the hypergeometric distribution, since the choice of "success" is arbitrary and can just as easily be regarded as "failure."

The intervals shown in Table 1 exhibit equivariance at the 90% and 99% confidence levels, as well as at the 95% level with one exception—the lower confidence limit for $x = 6$ is not the equivariant complement of the upper confidence limit for $x = 1$. Here is why this violation occurs: For $M \neq N/2$, equivariance follows from the fact that the acceptance function for M is the symmetric complement of the acceptance function for $N - M$ (see for example Figure 6). However for $M = N/2 = 10$ there are two minimal span acceptance functions available for 95% confidence, $AF(1-5)$ and $AF(2-6)$. Whichever one is selected, an asymmetry results. In general, equivariance will not hold for one pair of intervals when N is even and if two minimal span acceptance functions are available for $M = N/2$, as they will not be symmetric around $n/2$. A simple alternative approach that obeys equivariance for the present example is to use the acceptance function $AF(1-6)$ at $M = N/2 = 10$, but the resulting confidence procedure would then not have minimal length. It would represent a minimal length *equivariant* confidence procedure, however.

2.2. Adjustment for gaps

Although the selection of acceptance functions in the example $N = 20$, $n = 7$, 90% confidence considered above produces an acceptable confidence procedure, occasionally for other $(N,\ n)$ pairs and confidence levels the proposed method results in *gaps* in some confidence sets. Such gaps occur when the sequences *of* acceptance functions $\{AF(l-u)\}$ selected fail to be nonincreasing in both l and u. Figure 7 shows such a

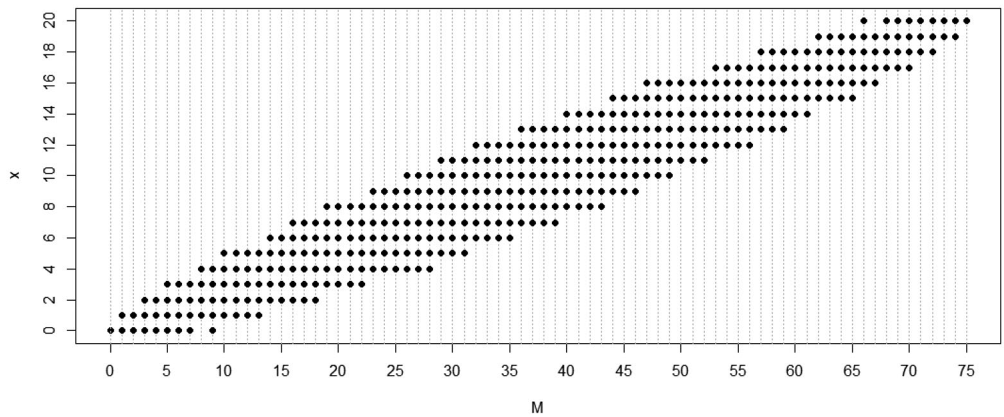

Figure 7. Confidence sets for the case when $N = 75$, $n = 20$ and 90% confidence level.

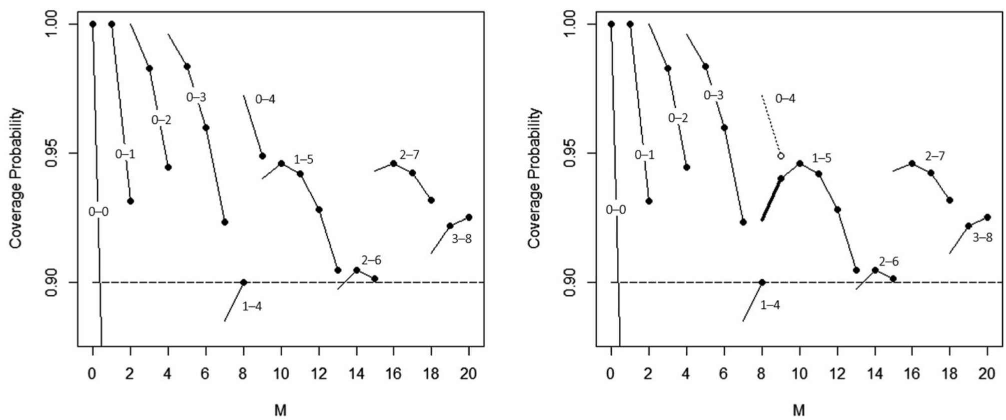

Figure 8. Acceptance functions selected for the case $N = 75$, $n = 20$ and 90% confidence level. (a) The initial procedure leads to a gap at $M = 9$. (b) Resolution to the gap issue.

case ($N = 75$, $n = 20$), and demonstrates the consequence of selecting a sequence of acceptance functions having a decreasing l and/or u value. This time the vertical sets of x represent the acceptance intervals for M and the confidence intervals are displayed horizontally. Notice that the minimum value in the acceptance interval for $M = 9$ is smaller than the minimum value for $M = 8$. This results in a gap in the confidence interval for $x = 0$; a similar pathology for maximum acceptance interval values when M moves from 67 to 68 results in a gap in the confidence interval for $x = 20$. This illustrates why the l and u sequences need to be nondecreasing.

To learn how we resolve such gaps, consider for instance the selection of acceptance functions for the case $N = 75$, $n = 20$ at the 90% confidence level based on the method established so far. Figure 8(a) shows that the cpf jumps from $AF(1-4)$ at $M = 7$ to $AF(0-4)$ at $M = 8$, consequently the value of l changes from 1 to 0 and therefore the l sequence is not monotonic increasing, producing a gap. The resulting confidence set for $x = 0$ in this case is the integers $M \in [0, 7] \cup [9]$ rather than a proper interval.

To eliminate such gaps, we follow the method of Schilling and Doi (2014): If the choice of a particular minimal span acceptance function produces a gap, use the minimal span acceptance function with the *next* highest coverage probability in its place. Following this method, since $AF(0-4)$ is the function that disrupts the monotonicity in the l sequence, we substitute $AF(1-5)$ as the acceptance function at $M=9$. Figure 8(b) shows this replacement. From this modification we obtain a confidence *interval* for $x=0$, namely $M \in [0, 7]$, at the negligible expense of a slightly lower coverage probability at $M=9$. The effect on the confidence procedure is to transfer $M=8$ to the confidence set (now an interval) for $x=0$ and remove it from the confidence interval for $x=5$.

Since the confidence intervals for M of the hypergeometric distribution are nearly always equivariant, we perform the above modification only for (i) $M \in [0, N/2-1/2]$ if N is odd or (ii) $M \in [0, N/2]$ if N is even and apply the equivariance property to solve for the remaining intervals. (It can be shown that no gaps occur for $M=N/2$, the case which sometimes yields a pair of nonequivariant intervals.)

Upon investigating all 15,150 $\{N, n\}$ cases with population size $N \in \{1, 2, \dots, 100\}$ and samples size $n \in \{1, 2, \dots, N\}$ at the commonly used confidence levels 90%, 95%, and 99%, we found that gaps occur for some x in only 54 instances (0.36%). No gaps occurred when the sample size n was more than 37% of the population size N. Furthermore, in 50 of the 54 cases with gaps, there is only a single equivariant pair of gaps. Overall only 114 of the 530,250 individual confidence sets found contain gaps. Thus the gap remedy described above is very rarely necessary.

The extreme uncommonness of gaps for hypergeometric interval estimation is in contrast to the binomial case, in which Schilling and Doi (2014) found that gaps occur using the same method for some x in approximately 40% of the 300 confidence level-sample size combinations investigated. The rarity of gaps for the hypergeometric problem is due to the fact that the parameter M is restricted to positive integers, whereas the binomial success parameter can take on all real numbers $p \in [0, 1]$.

Each gap found consisted only of a singleton set. Gaps for $M \in [0, [N/2]]$ were all analogous to the situation in Figure 8(a), where the confidence set had the form $M \in [a, b] \cup [b+2]$ for some integers a and b. Equivariance provides the obvious counterpart for $M \in [[N/2]+1, N]$.

In summary, our method of producing confidence intervals for the hypergeometric parameter M having superior length and coverage performance chooses the acceptance function of minimal span with the highest coverage probability, except in the extremely rare case where there is a need to resolve a gap, in which case the minimal span acceptance function that causes the gap is replaced with the one having the next highest coverage. We refer to our confidence procedure and the resulting intervals as *Length/Coverage Optimal*, or *LCO* for short, since it is a length minimizing procedure that maximizes coverage.

2.3. Formal description of the procedure for estimating M

The following is a formal description of the algorithm used to find the LCO procedure for estimating the hypergeometric parameter M:

Step 1: Beginning with $M=0$, for each $M \in [0, N]$ let $AF_M\ (l-u)$ denote the acceptance function achieving the highest coverage probability above $1-\alpha$ among all acceptance functions of minimal span at M. If more than one acceptance function achieves the highest coverage probability, select the function with the largest values of l and u.

Step 2: Whenever a transition from $AF_M\ (l-u)$ to $AF_{M+1}(l'-u')$, each obtained from Step 1, are such that $l'<l$ and/or $u'<u$, let k be the largest integer that produces $AF_{M+k}\ (l'-u')$. Reassign $M+1,\ \ldots,\ M+k$ to $x \in [l'+1, u'+1]$ *if* $M \le N/2$ and to $[l'-1,\ u'-1]$ $M \le N/2$ if $M > N/2$.

Step 3: For each resulting $AF_M\ (l-u)$, assign M to the confidence intervals for those $x \in [l,\ u]$.

2.4. Comparison to existing methods

Traditionally, most confidence intervals have been provided by formulas. There exist a few variations of Wald-type intervals for the hypergeometric parameter M; for example the formula

$$C_{Wald}(x) = \frac{Nx}{n} \pm t_{1-\frac{\alpha}{2}, n-1} \left(\frac{(N-n)\ Nx\ (n-x)}{n^2(n-1)} \right)^{1/2}$$

is provided in Thompson (2002). Intervals constructed by formulas may achieve short intervals, but typically violate strictness—the requirement that the coverage probability remains at least equal to the nominal level $1-\alpha$ for all values of the unknown parameter. For instance the confidence procedure that uses the above formula can be shown to have coverage probabilities as low as n/N (Wang 2015).

Strict confidence procedures for the above hypergeometric problem have been developed by Cochran (1977), Konijn (1973), Buonaccorsi (1987), and Wang (2015). Buonaccorsi shows that his confidence intervals are the same as Konijn's, which were shown to be superior to Cochran's in terms of length.

Wang's Method: The method recently developed by Wang (2015) is a modification to the intervals produced by Konijn (1973) that is superior to the other three procedures above in terms of length. Since Wang's procedure outperforms the previously studied methods, we provide a brief explanation of his method, along with a detailed comparison of performance to that of our LCO procedure.

Wang begins by determining the shortest level $1-\alpha/2$ one-sided intervals $[0,\ U(X)]$ and $[L(X),\ N]$. These are the same one-sided intervals that our method would produce. He next forms all level $1-\alpha$ confidence intervals $[U(X),\ L(X)]$, which are in fact the intervals obtained from Konijn's procedure. Wang then uses an iterative procedure to progressively shorten the intervals when possible, beginning with the interval(s) for the central value(s) of x and working outward to $x=0$ and N. Whenever shortening preserves strict coverage, that shortening is accepted. Wang's confidence procedure has the property that if any interval is replaced with a proper subinterval, the resulting confidence procedure will have level less than $1-\alpha$. Since our procedure uses only minimal span acceptance intervals, it shares this property.

In the event that either N or α is very small, the resulting two-sided intervals obtained from our procedure uniquely achieve minimal length and are the same as Wang's. This is not generally the case otherwise. Among all cases at confidence levels 90%, 95%, and 99% where $N \in \{1, 2,\ \ldots,\ 100\}$ and $n \in \{1, 2,\ \ldots,\ N\}$, it is common for the LCO method to produce different confidence intervals than those obtained from Wang's algorithm. When they are different, Wang's procedure tends to produce shorter intervals than ours toward the middle of the range of possible values of x, while the

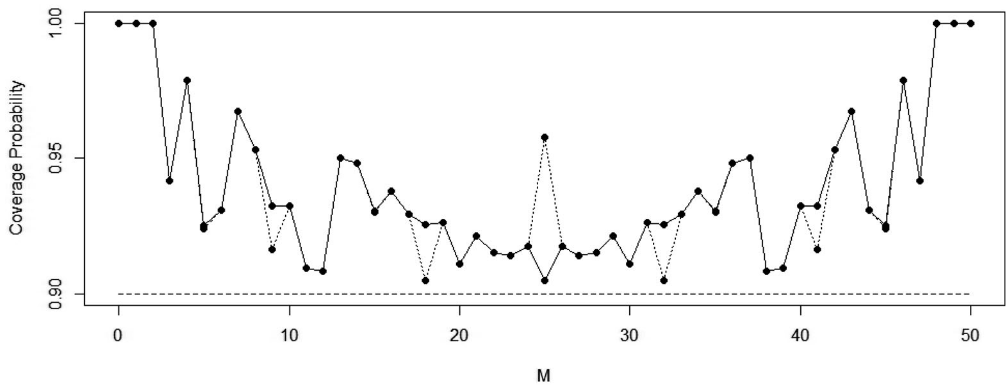

Figure 9. Coverage probability of the minimal span procedure, along with the coverage probability of Wang's method shown in dotted line segments when it differs, for the case $N = 50$, $n = 20$ and 90% confidence level.

LCO method tends to produces shorter intervals than Wang's at or near the ends. That Wang's intervals may be shorter near $x = n/2$ is intuitively reasonable since that is where his method first attempts to shorten them.

Blaker's Method: Blaker (2000) proposed an entirely different approach for generating confidence intervals for parameters of discrete distributions, which can be described as follows: Given the observed value x of the data, for each possible value of the parameter find $\alpha^* = \min\{P(X \leq x), P(X \geq x)\}$. Whichever tail this is determined from, find the largest tail probability α^{**} in the opposite tail that does not exceed α^*. The set of parameter values for which $\alpha^* + \alpha^{**} \geq \alpha$ constitutes the confidence interval for that x. Blaker's approach is guaranteed to achieve nesting.

While Wang's and Blaker's algorithms usually result in the use of minimal span acceptance functions, they do not always do so. Take for example the case where $N = 50$, $n = 20$ and $1 - \alpha = 0.90$. Figure 9 shows the cpf of our method, along with the cpf of Wang's method, shown in dotted line segments where it differs from ours. Consider the peak in Wang's cpf at $M = 25$. From Table 2 one can deduce that the acceptance function corresponding to Wang's algorithm at that value is $AF(7 - 13)$, giving a span of six, whereas we use $AF(8 - 13)$ there—a span of five.

Table 2 provides a comparison of the LCO, Konijn, Wang and Blaker confidence procedures for this example ($N = 50$, $n = 20$, 90% confidence level). As expected, Wang's intervals are subsets of Konijn's, being proper subsets for 15 of the 21 intervals. Table 2 is marked to highlight the differences between the LCO intervals and those of the other procedures. A competing method's intervals are shown in **bold** when they are longer than what is produced by the LCO procedure, while intervals shown in *italics* are shorter than the corresponding LCO interval. All but two of Konijn's intervals are longer than ours. Seven of Wang's intervals are longer than the corresponding LCO intervals, while five are shorter. Blaker's intervals are longer in three cases and shorter in two.

Considering all 21 confidence intervals, the LCO confidence procedure yields shorter average confidence interval length than does any of the competing procedures for this particular case, at 11.00 versus 11.05 for both Wang's and Blaker's methods and 12.10 for Konijn's procedure. In order to assess whether the LCO procedure tended to give

Table 2. 90% confidence intervals for LCO, Konijn's, Wang's and Blaker's procedures for $N = 50$, $n = 20$.

x	LCO Procedure	Konijn's Procedure	Wang's Procedure	Blaker's Procedure
0	[0, 4]	**[0, 5]**	**[0, 5]**	[0,4]
1	[1, 8]	**[1, 9]**	**[1, 9]**	[1,8]
2	[2, 11]	**[2, 12]**	[2, 11]	[2,11]
3	[4, 14]	**[3, 15]**	**[4, 15]**	[4,14]
4	[5, 17]	**[5, 18]**	[6, 18]	[5,17]
5	[7, 20]	[7, 20]	[7, 20]	*[7,19]*
6	[9, 22]	**[9, 23]**	*[10, 22]*	[9,22]
7	[12, 24]	**[11, 25]**	**[12, 25]**	**[12,25]**
8	[13, 27]	**[13, 28]**	[13, 27]	[13,27]
9	[15, 29]	**[15, 30]**	*[16, 29]*	**[15,30]**
10	[18, 32]	**[17, 33]**	*[19, 31]*	[18,32]
11	[21, 35]	**[20, 35]**	*[21, 34]*	**[20,35]**
12	[23, 37]	**[22, 37]**	[23, 37]	[23,37]
13	[25, 38]	**[25, 39]**	[25, 38]	[25,38]
14	[28, 41]	**[27, 41]**	*[28, 40]*	[28,41]
15	[30, 43]	[30, 43]	[30, 43]	*[31,43]*
16	[33, 45]	**[32, 45]**	[32, 44]	[33,45]
17	[36, 46]	**[35, 47]**	**[35, 46]**	[36,46]
18	[39, 48]	**[38, 48]**	[39, 48]	[39,48]
19	[42, 49]	**[41, 49]**	**[41, 49]**	[42,49]
20	[46, 50]	**[45, 50]**	**[45, 50]**	[46,50]

Intervals shown in **bold** are longer than LCO intervals, while those shown in *italics* are shorter.

Table 3. Ratio of average length of LCO procedure relative to competing procedures for cases when average length differs, at confidence levels 90%, 95% and 99%.

	Confidence Level		
	90%	95%	99%
LCO/Konijn's	91.29%	93.25%	95.51%
LCO/Wang	99.42%	99.58%	99.73%
LCO/Blaker	99.13%	99.40%	99.68%

the shortest intervals across a comprehensive range of population and sample size values, we compared average confidence interval lengths for the LCO, Konijn's, Wang's and Blaker's procedures at confidence levels 90%, 95% and 99% for all 15,150 $\{N, n\}$ pairs for which $N \in \{1, 2, \ldots, 100\}$ and $n \in \{1, 2, \ldots, N\}$. In the majority of cases the average interval lengths for Wang's, Blaker's and the LCO procedures are equal. Table 3 provides a summary of how the average lengths of the alternative procedures compare to LCO's average length when they differ.

Due to the exclusive use of minimal span acceptance functions, the LCO confidence procedure obtains a smaller average length overall than competing procedures for each confidence level. Thus our method provides shorter average interval length than any existing strict method. Moreover, when the cpfs of the LCO method and an alternative procedure differ, the cpf of the former is usually higher than that of the latter. Thus our procedure tends to provide greater certainty that the parameter will be contained within the confidence interval.

2.5. Additional properties

Monotonicity in x: As the number of successes in a sample increases for fixed n, we expect the lower and upper confidence limits to increase. Upon investigating all sample

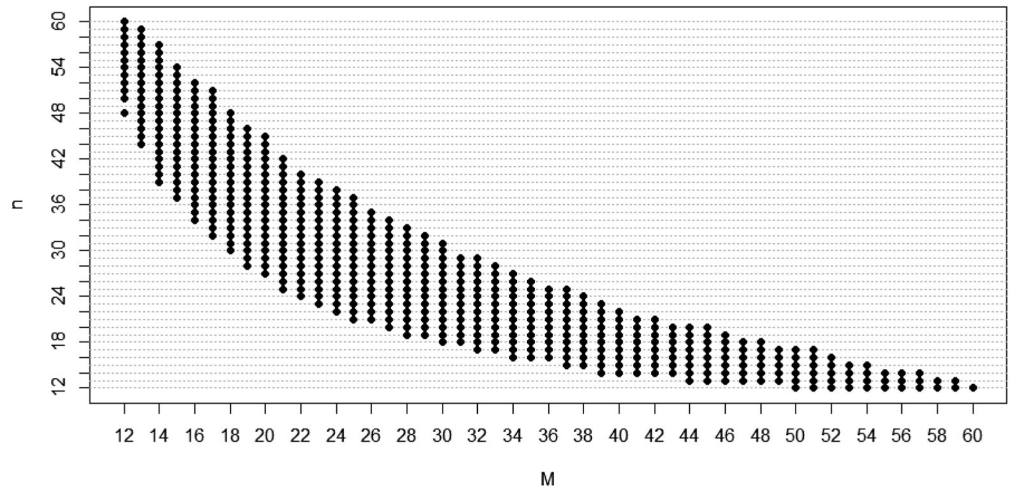

Figure 10. 90% confidence intervals for $N=60$, $x=12$ at each $n \leq N$. A violation to monotonicity occurs between the lower confidence limits for $n=48$ and $n=49$.

sizes $n \leq N$ for each population size $N \in \{1, 2, \ldots, 100\}$ at confidence levels 90%, 95%, and 99%, we found that for the LCO method, both lower and upper confidence limits increase as expected when x increases in all but 54 of the 15,150 cases tested (0.36%), in which they share either the same lower or same upper confidence limit.

Monotonicity in n: An increase in n that does not include a change in x should result in nonincreasing confidence limits. To investigate the extent to which the LCO method achieves such monotonicity in n, we determined for each $N \in \{1, 2, \ldots, 100\}$ at the 90%, 95%, and 99% confidence levels the behavior of the endpoints of the confidence intervals for each fixed $x \in \{0, 1, 2, \ldots, n\}$ at all sample sizes $n \leq N$. We only considered the case when an additional trial results in a failure, since the opposite case is essentially equivalent due to the fact that equivariance holds almost without exception.

Looking at the same 15,150 cases as before, we found only 58 violations to monotonicity in n (0.38%). Figure 10 displays a representative one of the rare cases where a monotonicity violation occurs, as the lower confidence limit for $x=12$ increases from 12 to 13 when n increases from 48 to 49. Every other value of n behaves the way we would expect—as n increases, the lower and upper limits of the confidence intervals for $x=12$ are monotonically nonincreasing.

Nesting: Checking all 353,500 confidence interval comparisons at levels 90%, 95% and 99% for all $n \leq N$ for each $N \in \{1, 2, \ldots, 100\}$, we found only two violations of the nesting property, both of which occurred when $N=97$, $n=29$ at the pair of confidence levels 0.90 and 0.95, as a result of fixing a gap at the 0.90 confidence level.

2.6. Estimation of functions of M

Aside from estimation of the hypergeometric parameter M itself, there may sometimes be interest in estimating some function of M such as the proportion of successes in the population M/N or the odds of success $M/(N-M)$. Confidence intervals for such functions, shown in Table 4 for the previously examined example $N=20$, $n=7$ are easily

Table 4. 90% confidence intervals when $N = 20$, $n = 7$ for the number of successes, proportion of successes, and odds of success.

x	M	M / N	Odds ($M/(N-M)$)
0	[0, 5]	[0.00, 0.25]	[0.00, 0.33]
1	[1, 8]	[0.05, 0.40]	[0.05, 0.67]
2	[2, 11]	[0.10, 0.55]	[0.11, 0.56]
3	[4, 14]	[0.20, 0.70]	[0.25, 2.33]
4	[6, 16]	[0.30, 0.80]	[0.43, 4.00]
5	[9, 18]	[0.45, 0.90]	[0.83, 9.00]
6	[12, 19]	[0.60, 0.95]	[1.50, 19.0]
7	[15, 20]	[0.75, 1.00]	[3.00, ∞]

obtained simply by substituting the confidence limits for M into the function of interest. For example, the confidence intervals for the success proportion are obtained by simply dividing the lower and upper limits of the confidence intervals for M by the population size N.

3. Estimation of population size

The reverse case is also sometimes of interest: estimating the population size N when M is known. This situation occurs for example when the *capture-recapture* method is used. Suppose an ecologist wants to estimate the population of a particular species. Using the capture-recapture method, he/she would randomly *capture*, mark, and release M units from an unknown population of N units. After a sufficient amount of time has passed, a *recapture* of a random sample of size n is performed from the same population and the number of units, x, that were previously marked is observed. The proportion of the marked sample is then matched to the proportion of the marked population, $x/n = M/N$ to obtain a point estimate of N, namely $\hat{N} = Mn/x$.

We provide below a strict confidence procedure that yields interval estimates for N. Again our primary goal is to obtain short intervals. We use a similar approach as in the estimation of M, making use of minimal span acceptance functions.

3.1. Selection of acceptance functions

The acceptance functions for estimating N are given by the same formula as those used for estimating M, however now M is fixed and they become functions of N:

$$AF_{\Omega}(N) = P(X \epsilon \Omega) = \sum_{x \epsilon \Omega} \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

where Ω is any subset of $\{0, 1, \ldots, n\}$ of the form $\Omega = \{l, l+1, \ldots, u\}$.

These acceptance functions behave much differently as functions of N. Now, small values of x yield large values of N. Therefore, the order of the acceptance functions is switched in comparison to the case of estimating M. In particular, acceptance functions of the form $\mathrm{AF}_{\Omega_{0l}}$ move from right to left as l increases. For large N, only acceptance functions of this type are available for construction of the cpf, as all other acceptance functions fall below the confidence level.

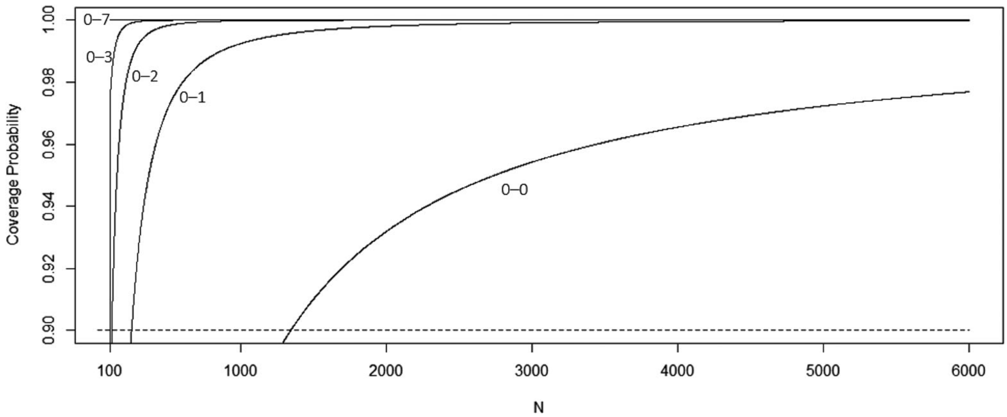

Figure 11. Acceptance functions eligible for selection for $N \geq 100$ for the case when $M = 20$, $n = 7$ and 90% confidence level.

Figure 11 shows the available acceptance functions for $100 \leq N \leq 6{,}000$ for a 90% confidence procedure when $M = 20$ and $n = 7$. We display these functions as continuous curves for simplicity although in fact they are only defined on the integers. We see that $AF(0-0)$ is the acceptance function that crosses the confidence level at the largest value of N, which occurs here for $N = 1{,}342$. Thus, $AF(0-0)$ is the unique minimal span acceptance function for all N exceeding this value.

Our selection of acceptance functions for the construction of an ideal cpf for the estimation of N begins with large values of N and moves to the left until we reach $\max\{n, M\}$. This determines confidence intervals in increasing order of x. As before, our procedure chooses acceptance functions with minimal span and highest coverage probability, except where necessary to resolve a gap in the confidence set for some x. In what follows, we restrict attention to the situation where $n \leq M$ since the results for when $n > M$ are identical to those when n and M are interchanged. The proof of this fact is in the Appendix.

Figure 12a shows for the same case as in Figure 11 the section where $N \in \{20, 21, \ldots, 100\}$, with points where the acceptance functions are defined. Figure 12b shows only the functions of minimal span.

Only for $N = 27$, 28, 35, 36, 43, 44, 45, and 46 are more than one minimal span function available. For each of these values we choose the higher of the two available acceptance functions in order to maximize coverage. The resulting confidence intervals are shown in Table 5.

3.2. Resolving gaps

Gaps occur for the parameter N for the same reason as they do in the estimation of M and are resolved in the same manner. We find that gaps occur much more often in the case of estimating N than for estimating M. Testing for gaps was performed for all $n \leq M$ for each $M \in \{1, 2, \ldots, 50\}$ at the $1 - \alpha = 0.90$, 0.95, and 0.99 levels; gaps occurred for some x in 262 out of the total 3,825 cases (6.85%). Gaps occur more often

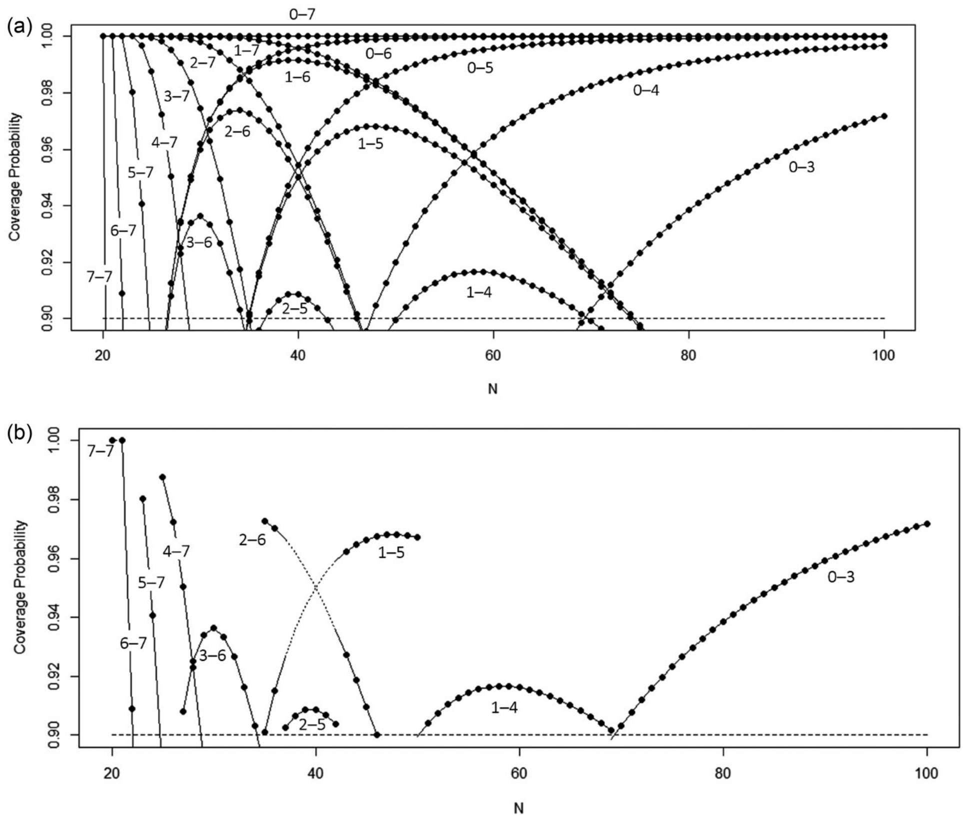

Figure 12. (a) Acceptance functions above the 90% confidence level for all $N \epsilon \{20, 21, \ldots, 80\}$ for the case $M = 20$ and $n = 7$. (b). Minimal span acceptance functions above the 90% confidence level for all $N \epsilon \{20, 21, \ldots, 100\}$ for the case $M = 20$ and $n = 7$.

Table 5. 90% confidence intervals for N when $M = 20$ and $n = 7$.

x	Lower Limit	Upper Limit
0	70	$+\infty$
1	43	1,341
2	35	250
3	28	114
4	25	69
5	23	50
6	21	36
7	20	27

in the estimation of N due to the greater asymmetry of the acceptance functions, and because the acceptance functions generally span a large set of x values.

Recall that in our initial confidence procedure for M, decreasing values of either l or u in the sequences of acceptance functions used resulted in confidence sets that are not intervals; this led to a slight modification in the procedure in order to avoid such decreasing values. In the case of estimating N, we observe each l and u value starting at $N_0 = \min\{N\text{: } AF(0\text{-}0) \geq 1-\alpha\}$ and move in a decreasing order of N until $\max\{n, M\}$ is reached. We again require that the resulting l and u sequences of the acceptance

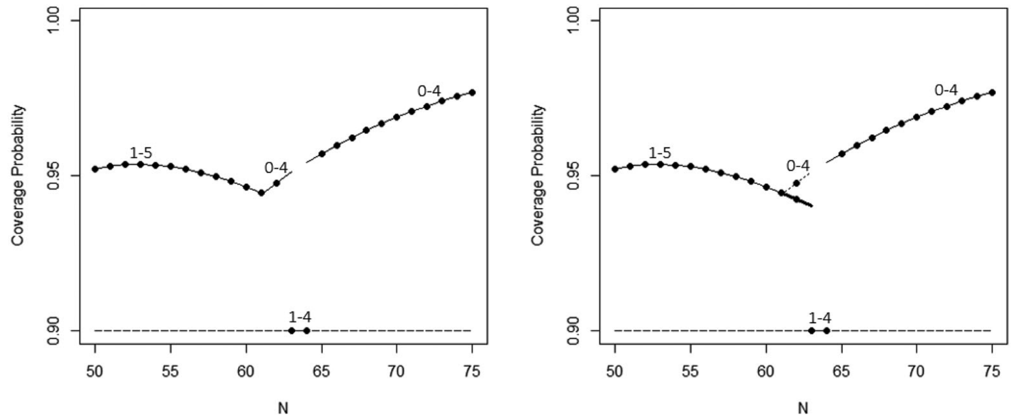

Figure 13. Acceptance functions selected for the case $M = 15$, $n = 10$ and 90% confidence level. (a) The initial procedure leads to a gap at $N = 62$. (b) Resolution of the gap issue.

functions selected are nondecreasing as N decreases, otherwise the same phenomenon occurs: confidence sets containing a gap.

To illustrate the gap issue in the case of estimating N, consider Figure 13, which shows the acceptance functions that lie above $1 - \alpha = 0.90$ for the case $n = 10$ and $M = 15$. Figure 13(a) displays the acceptance functions selected by our initial procedure for the values of N such that $75 \geq N \geq 50$. Looking at the graph from right to left, we see that $AF(1-4)$ is selected for $N = 63$ and 64. The next acceptance function that is selected is $AF(0-4)$, which violates the required monotonicity in l, leading to a gap at $N = 62$.

When the choice of an acceptance function results in a gap, we again use the minimal span acceptance function with the next highest coverage probability in substitution. Figure 13(b) shows the resolution to the gap issue, where the dotted portion of $AF(0-4)$ represents the acceptance function in violation of monotonicity; an extension of $AF(1-5)$ has been added in bold, which now becomes the assigned acceptance function.

3.3. Formal description of the procedure for estimating N

The following is a formal description of the algorithm used to find the $1 - \alpha$ level minimal span confidence procedure for the hypergeometric parameter N:

Step 1: Locate the starting point $N_0 = \min\{N \mid AF(0-0) \geq 1 - \alpha\}$for evaluating acceptance functions.

Step 2: For each N beginning with N_0 and continuing in a decreasing fashion until $\max\{n, M\}$ is reached, let $AF_N(l-u)$ denote the acceptance function achieving the highest coverage probability above $1 - \alpha$ among all acceptance functions of minimal span at N. If more than one acceptance function achieves the highest coverage probability, select the function $AF_N(l-u)$ with the largest value of l. Assign N to the confidence intervals for each $x \in [l, u]$, except in the case described by Step 3.

Step 3: Whenever $AF_N(l-u)$ and $AF_{N-1}(l'-u'$ from Step 3 are such that $l' < l$ and/or $u' < u$, let k be the largest integer that produces $AF_{N-k}(l'-u'$. Reassign $N - 1, \ldots, N - k$ to $x \in [l' + 1, u' + 1]$.

Table 6. 90% confidence intervals for the minimal span procedure and competing procedures when $M = 50$ and $n = 25$.

x	Minimal Span Procedure	Wang's Procedure	Blaker's Procedure
0	[495, ∞)	*[521, ∞)*	*[521, ∞)*
1	[312, 11901]	*[324, 11901]*	*[322, 11901]*
2	[228, 2316]	*[240, 2316]*	*[232, 2316]*
3	[193, 1102]	*[198, 1102]*	*[195,1102]*
4	[162, 688]	[162, 688]	[162, 688]
5	[140, 494]	**[141, 520]**	**[140, 526]**
6	[126, 422]	*[127, 422]*	[126, 422]
7	[114, 311]	**[115, 323]**	**[114, 321]**
8	[105, 269]	[105, 269]	[105, 269]
9	[97, 227]	**[97, 239]**	**[97, 231]**
10	[91, 192]	**[91, 197]**	**[91, 194]**
11	[87, 168]	[87, 168]	[87, 168]
12	[81, 155]	*[82, 155]*	[81, 155]
13	[76, 139]	[77, 140]	[76, 139]
14	[72, 125]	**[72, 126]**	[72, 125]
15	[70, 113]	**[70, 114]**	[70, 113]
16	[66, 104]	*[67, 104]*	[66, 104]
17	[63, 96]	[63, 96]	[63, 96]
18	[61, 86]	[61, 86]	[61, 86]
19	[59, 80]	**[59, 81]**	[59, 80]
20	[57, 75]	**[57, 76]**	[57, 75]
21	[55, 71]	[55, 71]	[55, 71]
22	[54, 65]	**[54, 66]**	[54, 65]
23	[52, 62]	[52, 62]	[52, 62]
24	[51, 58]	[51, 58]	[51, 58]
25	[50, 54]	[50, 54]	[50, 54]

Bold represents when the minimal span procedure's intervals are shorter than those produced by competing procedures. *Italics* represent when competing intervals are shorter than those produced by the minimal span procedure.

3.4. Comparison to existing methods

There are several formula-based confidence procedures for estimating N. For example Thompson (2002) provides the formula

$$C(x) = \frac{nM}{x} \pm z_{\alpha/2}\left(\frac{Mn(M-x)(n-x)}{x^3}\right)^{1/2}$$

Note that the confidence interval for $x = 0$ is not defined. Again as is typical of Wald-type intervals for parameters of discrete distributions, coverage is often poor. Many attempts have been made to improve confidence intervals for the parameter N, but none of them come close to achieving the desired minimum coverage probability (see Wang 2015, Figure 1).

Wang uses a similar approach for estimating N as he does for estimating M. Starting with the confidence interval $[U(X),\ L(X)]$ where $[0,\ U(X)]$ and $[L(X),\ N]$ are the shortest level $1-\alpha/2$ one-sided intervals, his algorithm then shortens the length of the intervals where possible while maintaining coverage probability above $1 - \alpha$, beginning with the confidence interval for $x = 0$ and ending with the one for $x = \min\{n, M\}$.

As an example of how confidence intervals for the minimal span method typically compare to those obtained from other high performing procedures, see Table 6, which displays the 90% intervals for our method along with those determined from Wang's and Blaker's procedures for the case $n = 25$, $M = 50$. Intervals are shown in **bold**

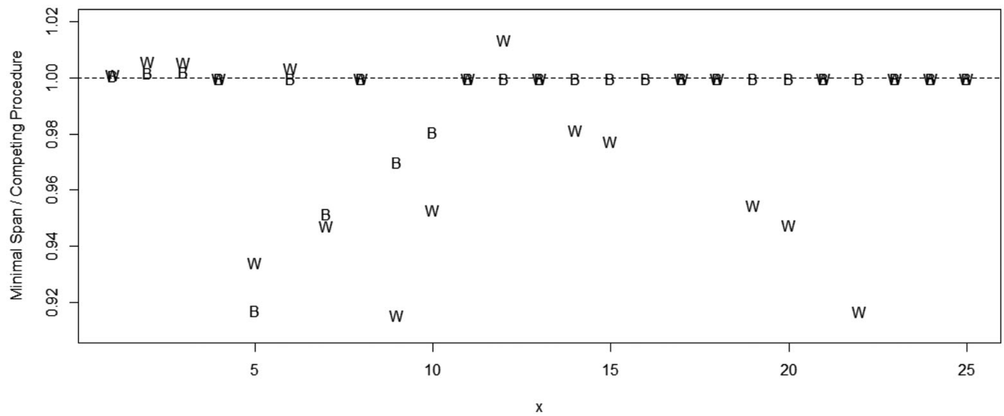

Figure 14. Ratio of lengths of the minimal span intervals relative to Wang's (W) and Blaker's (B) intervals for $x > 0$ for the case when $M = 50$, $n = 25$ and 90% confidence level.

whenever the one produced by the minimal span method is shorter than the competing interval, and in *italics* when the competing interval is shorter than the minimal span interval. Most intervals are the same for all three methods, but there are several differences.

The above example is representative of the situation for (n, M) pairs in general: Wang's and Blaker's methods produce shorter length intervals for several very small values of x, while the minimal span procedure results in shorter intervals more often otherwise. We do not regard the fact that the minimal span intervals are longer for a few small values of x as a significant drawback, as one would rarely attempt to estimate N from such a small x in practice, especially since the resulting intervals are so wide.

Since the lengths of the intervals for x varies significantly, there becomes an issue of how to properly compare the two methods. In particular, the intervals for $x = 0$ are infinitely wide, so comparing the average length of all intervals is not appropriate here. Instead, we compare the ratio of confidence interval lengths for the minimal span procedure and Wang's and Blaker's procedures for $1 \le x \le \min\{n, M\}$. The results are shown in Figure 14 for the case when $M = 50$, $n = 25$, and $1 - \alpha = 90\%$, with a reference line for a ratio equal to one. When points lie above the line, our method produces wider intervals than does the competing procedure, and when points lie below the line ours intervals are shorter. When our method produces wider intervals than Wang's or Blaker's intervals they are only slightly wider, whereas when they are shorter they are sometimes considerably shorter—as much as 8.5% in this case. The average ratio of interval lengths for the minimal span method compared to (i) Wang's procedure and (ii) Blaker's procedure for this example are 98.3% and 99.3%, respectively.

The average ratio of interval lengths was computed for all $n \le M$ for each associated $M \in \{1, 2, \ldots, 50\}$ at the $1 - \alpha = 0.90$ confidence level. When judged by this measure, the minimal span procedure yielded shorter intervals for 1,110 of the 1,275 cases tested (87.1%), while Wang's procedure yielded shorter intervals for only four cases. In the remaining cases, interval lengths of the two procedures were the same when judged by average ratio. Across all cases investigated, when assessed by the average ratio of interval lengths, the minimal span procedure was shorter than Wang's by a maximum of

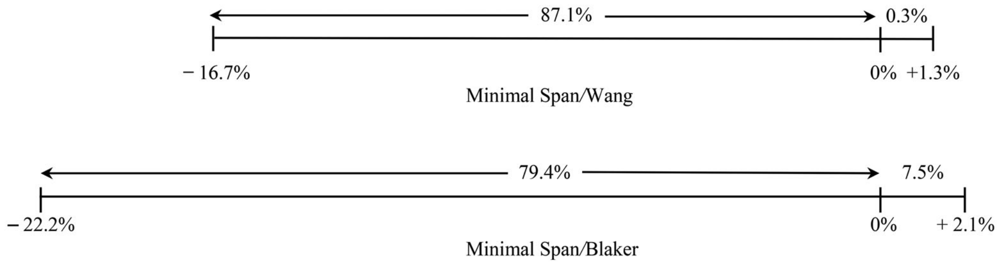

Figure 15. Relative Average Ratio of Interval Lengths of Minimal Span 90% Confidence Intervals vs. Wang's and Blaker's Procedures. Minimal span results were shorter than Wang's in 87.1% of the cases examined and shorter than Blaker's in 79.4% of those cases, longer than Wang's in 0.3% of the cases examined and longer than Blaker's in 7.5% of those cases, and equal in the remaining cases.

16.67%, while Wang's procedure was shorter than our procedure by a maximum of 1.36%.

We also explored the performance of Blaker's method in the case of estimating N. The calculation of average ratio of interval lengths was performed for the same 1,275 cases described above. The minimal span procedure yielded shorter intervals for 1,013 (79.4%) cases, while Blaker's procedure yielded shorter intervals for 96 cases. In the remaining cases, average interval lengths of the two procedures were the same when measured by ratio. Blaker's procedure was shorter than the minimal span procedure by a maximum of 2.13%, while the minimal span procedure was shorter than Blaker's by a maximum of 22.2%. See Figure 15.

3.5. Additional properties

Monotonicity in x: Since small values of x yield large interval estimates of N and vice versa, we would expect the lower and upper confidence limits for each x of a given confidence procedure to decrease as x increases. This property was tested for all $n \leq M$ for each $M \in \{1, 2, \ldots, 50\}$ at the 90%, 95%, and 99% confidence levels; no violations to monotonicity were observed.

Monotonicity in n: We would also like the limits of the confidence intervals for N to change monotonically when an additional observation is added to the sample. There are two situations: (i) the additional trial results in a success and (ii) the additional trial results in failure.

If an additional trial adds a success, we would expect both limits of the confidence interval to decrease or at least not increase. We tested whether the confidence limits for $x+1$ successes in $n+1$ trials are no larger than the confidence limits for x successes in n trials for all $n \leq M$ for each $M \in \{1, 2, \ldots, 50\}$ at the $1 - \alpha = 0.90$, 0.95, and 0.99 levels; violations to monotonicity occurred for just 19 (0.029%) of the 66,150 pairs of confidence intervals compared.

On the other hand, if the additional trial results in failure, we would expect both limits of the confidence interval to increase or at least not decrease. Testing this form of monotonicity for all $n \leq M$ for each $M \in \{1, \ldots, 50\}$ at the 90%, 95%, and 99% confidence levels, we found violations for 237 out of the 66,150 pairs of confidence intervals compared (0.36%).

Nesting: To determine how often the minimal span method achieves nesting, confidence intervals were computed for each $x \in \{0, 1, 2, \ldots, n\}$ for all $n \le M$ for $M \in \{1, 2, \ldots, 50\}$, at each confidence level $1 - \alpha \in \{0.900, 0.901, 0.902, \ldots, 0.997, 0.998, 0.999\}$. Just as the case of estimating M, violations to nesting only occurred after gaps had been resolved. However, resolving a gap does not usually lead to a violation to nesting. Of the 46,750 interval comparisons at the 90%, 95% and 99% confidence levels, only 31 nesting violations were found.

4. Conclusion

The idea of first constructing an ideal coverage probability function and deducing the confidence intervals from this function is an approach to interval estimation that has found substantial success. See, for example, Schilling and Doi (2014), Choi (2015) and Schilling and Holladay (2017).

The minimal span procedures presented above provide (gapless) confidence intervals for the parameters of the hypergeometric distribution that are shorter than those of existing strict procedures. In addition they maintain maximal coverage probability across the parameter space among confidence procedures that tend to produce short intervals through the use of minimal span acceptance functions. We observe violations to monotonicity and nesting only very infrequently; consequently we advocate the use of these procedures for precise interval estimation in hypergeometric situations.

Technical Note: All computations were performed using the R software. R code for determining confidence intervals (i) for M when N is known and (ii) for N when M is known is available at https://github.com/mfschilling/HGCIs. Also at this site are links to Shiny web apps that generate these confidence intervals without the need for the user to run code.

References

Blaker, H. 2000. Confidence curves and improved exact confidence intervals for discrete distributions. *Canadian Journal of Statistics* 28 (4):783–98. doi:10.2307/3315916.

Blyth, C. R., and H. A. Still. 1983. Binomial confidence intervals. *Journal of the American Statistical Association* 78 (381):108–16. doi:10.1080/01621459.1983.10477938.

Buonaccorsi, J. 1987. A note on confidence intervals for proportions in finite populations. *The American Statistician* 41 (3):215–8. doi:10.2307/2685108.

Choi, S. A. 2015. Confidence intervals for the success parameter of the negative binomial random variable. Master's Thesis, California State University, Northridge.

Cochran, W. G. 1977. *Sampling techniques*. 3rd ed. New York: Wiley.

Crow, E. L. 1956. Confidence intervals for a proportion. *Biometrika* 43 (3–4):423–5. doi:10.2307/2332920.

Kabaila, P., and J. Byrne. 2001. Exact short Poisson confidence intervals. *Canadian Journal of Statistics* 29 (1):99–106. doi:10.2307/3316053.

Katz, L. 1953. Confidence intervals for the number showing a certain characteristic in a population when sampling is without replacement. *Journal of the American Statistical Association* 48 (262):256–61. doi:10.1080/01621459.1953.10483471.

Konijn, H. S. 1973. *Statistical theory of sample survey design and analysis*. Amsterdam: North-Holland.

Sahai, H., and A. Khurshid. 1995. A note on confidence intervals for the hypergeometric parameter in analyzing biomedical data. *Computers in Biology and Medicine* 25 (1):35–8. doi:10.1016/0010-4825(95)98883-F.
Schilling, M. F., and B. Holladay. 2017. Length minimization for Poisson confidence procedures. *Communications in Statistics - Theory and Methods* 46 (2):861–73. doi:10.1080/03610926.2015.1006782.
Schilling, M. F., and J. A. Doi. 2014. A coverage probability approach to finding an optimal binomial confidence procedure. *The American Statistician* 68 (3):133–45. doi:10.1080/00031305.2014.899274.
Thompson, S. K. 2002. *Sampling*. 2nd ed. New York: Wiley.
Wang, W. 2015. Exact optimal confidence intervals for hypergeometric parameters. *Journal of the American Statistical Association* 110 (512):1491–9. doi:10.1080/01621459.2014.966191.

Appendix

Proof of equivalence of estimation of N for given n and M with the case when n and M are interchanged:

Consider two scenarios in which we wish to estimate N. The first case has a capture size $M_1 = M$ and a recapture size $n_1 = n$. The second case has a capture size $M_2 = n$ and a recapture size $n_2 = M$. We start by showing that the probability mass function for both cases are equal:

For all $N \epsilon$ [max{n, M}, ∞) and $x \epsilon$ [0, min{n, M}], the pmf for the first case is

$$\frac{\binom{M_1}{x}\binom{N-M_1}{n_1-x}}{\binom{N}{n_1}} = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Expanding this formula, one obtains

$$\frac{\frac{M!}{x!(M-x)!}\frac{(N-M)!}{(n-x)!(N-M-n+x)!}}{\frac{N!}{n!(N-n)!}} = \frac{M!(N-M)!n!(N-n)!}{x!(M-x)!(n-x)!(N-M-n+x)!N!}$$

For all $N \epsilon$ [max{n, M}, ∞) and $x \epsilon$ [0, min{n, M}], the pmf for the second case is

$$\frac{\binom{M_2}{x}\binom{N-M_2}{n_2-x}}{\binom{N}{n_2}} = \frac{\binom{n}{x}\binom{N-n}{M-x}}{\binom{N}{M}}$$

Expanding as before, the following formula is obtained:

$$\frac{\frac{n!}{x!(n-x)!}\frac{(N-n)!}{(M-x)!(N-n-M+x)!}}{\frac{N!}{M!(N-M)!}}$$

which clearly simplifies to the same expression as obtained for the first case. Now since a confidence procedure is determined by its acceptance functions, each of which is simply the sum of particular pmf values, the first case and the second case yield identical sets of confidence intervals. □