Transformed Logit Confidence Intervals for Small Populations in Single Capture–Recapture Estimation

Mauricio Sadinle
Departamento de Estadística, Universidad Nacional de Colombia, Bogotá
Conflict Analysis Resource Center – CERAC, Bogotá

Abstract

The good performance of logit confidence intervals for the odds ratio with small samples is well known. This is true unless the actual odds ratio is very large. In single capture–recapture estimation the odds ratio is equal to 1 because of the assumption of independence of the samples. Consequently, a transformation of the logit confidence intervals for the odds ratio is proposed in order to estimate the size of a closed population under single capture–recapture estimation. It is found that the transformed logit interval, after adding .5 to each observed count before computation, has actual coverage probabilities near to the nominal level even for small populations and even for capture probabilities near to 0 or 1, which is not guaranteed for the other capture–recapture confidence intervals proposed in statistical literature. Thus, given that the .5 transformed logit interval is very simple to compute and has a good performance, it is appropriate to be implemented by most users of the single capture–recapture method.

Key words: Asymptotic normality; Coverage probability; Dual–system estimation; Expected width; Monte Carlo study; Odds ratio; Pivotal statistic; Profile likelihood.

1 Introduction

Capture–recapture techniques are used to estimate population sizes. Particularly, single capture–recapture or dual–system estimation is used to estimate the size of a closed population with two independent samples. This methodology has been widely used in wildlife studies for estimating abundance parameters (e.g. Kekäläinen et al., 2008; Olsson et al., 2006; Pfeiler et al., 2005) and has been proposed for census correction (Wolter, 1986). For example, the United States Census Bureau applied this method by the 1990 Post–Enumeration Survey (see Breiman, 1994; Hogan, 1993; Alho et al., 1993; Mulry and Spencer, 1991) producing estimations by strata attempting
the idea presented by Chandra-Sekar and Deming (1949) in order to deal with differential capture probabilities. Also, this technique has been used for estimating population prevalence in epidemiological studies using data obtained from separate sources or record systems (e.g. Seber et al., 2000; Faustini et al., 2000; Abeni et al., 1994).

For capture–recapture, several confidence intervals for the population size expressed by closed formulae have been proposed in the statistical literature (Jensen, 1989; Chao, 1989; Evans et al., 1996; Borchers et al., 2002). However, those intervals are constructed under the assumption of asymptotic normality of some pivotal statistic and have a bad performance for small populations and for small or high capture probabilities. There are alternatives to find confidence intervals by other more robust methods as Monte Carlo and bootstrap techniques (Efron, 1979) applied to capture–recapture estimation (Buckland, 1984; Buckland and Garthwaite, 1991) and the profile–likelihood based method (see Venzon and Moolgavkar, 1988) which has been proposed to be applied to several capture–recapture models (Cormack, 1992; Evans et al., 1996; Gimenez et al., 2005). Nevertheless, a closed formula is always easiest to apply, specially by researchers without computational skills.

In this paper transformed logit confidence intervals are proposed inspired in the paper of Agresti (1999), where he shows that the logit confidence intervals for the odds ratio are acceptable for small samples unless the true odds ratio is very large. This condition holds true in single capture–recapture estimation given that, under independence of the samples, the odds ratio of the $2 \times 2$ table that classifies the individuals into captured and recaptured equals to 1.

In Section 2 the single capture–recapture method is reviewed. We review in Section 3 confidence intervals for capture–recapture estimation based in the asymptotic normality of some statistic and in Section 4 the Monte Carlo and profile–likelihood based confidence intervals. In Section 5 two new confidence intervals are proposed based in a transformation of the logit confidence intervals for the odds ratio. Later, in Section 6 the performances of those intervals are explored for small populations and a simulation study is presented for large populations.

## 2 Single Capture–Recapture Estimation

The single capture–recapture method proceeds firstly capturing $n_{1+}$ individuals from the whole population, marking and releasing them. Later, in a second capture occasion a sample of size $n_{+1}$ is taken. Let $N = n_{11} + n_{12} + n_{21} + n_{22}$ be the number of individuals in the population, i.e. classified in $n_{11}$ captured in the first sample and recaptured in the second one, $n_{12}$ captured in the first sample and not captured in the second one, $n_{21}$ not captured in the first sample and captured in the second one, and finally $n_{22}$ never captured. The information of individuals sampled in the above scheme can be arranged in a $2 \times 2$ table whose two rows correspond to the first sample and whose two columns correspond to the second sample as in Table 1. Of course, $n_{22}$ is unknown, and consequently $n_{2+}, n_{+2}$ and $N$ are also unknown.

Single capture–recapture estimation must be used for independent samples under the assumption that the population is closed, namely, $N$ is invariant in the time between the first and the second sampling. Also, it is needed to assume all individuals are equally likely to be
captured in each independent sample and identifying individuals captured in both independent samples is possible. Under the assumptions of the single capture-recapture method, the probabilistic model for \((n_{11}, n_{12}, n_{21}, n_{22})\) in the capture-recapture method, unconditional on the values \(n_{1+}\) and \(n_{+1}\), is the multinomial model (e.g. Bishop et al., 1975; Alho, 1990)

\[
P_\phi(n_{11}, n_{12}, n_{21}, n_{22}) = \frac{N!}{n_{11}!n_{12}!n_{21}!n_{22}!}p_{11}^{n_{11}}p_{12}^{n_{12}}p_{21}^{n_{21}}p_{22}^{n_{22}}
\]

where \(n_{22} = N - n_{11} - n_{12} - n_{21}\). This multinomial distribution depends on a vector of parameters \(\phi = (N, p_{11}, p_{12}, p_{21}, p_{22})\), where the probabilities are obtained, from the assumption of independence, as \(p_{11} = p_1p_2, p_{12} = p_1(1 - p_2), p_{21} = (1 - p_1)p_2, p_{22} = (1 - p_1)(1 - p_2) = 1 - p_1 - p_2 - p_1p_2\), where \(p_1\) is the probability to be captured in the first sample and \(p_2\) the probability to be captured in the second sample. Thus, note for the capture-recapture method, the parameters of the multinomial distribution are reduced to \((N, p_1, p_2)\).

The estimated probability to be captured in both samples in the method above is \(n_{11}/N\). If the samples are independent, the estimated probability of being in both samples can be computed as the multiplication of the estimated probabilities of being in each sample as \((n_{1+}/N)(n_{+1}/N)\). Then \(n_{11}/N \approx (n_{1+}/N)(n_{+1}/N)\), and from this, the Lincoln–Petersen estimator is obtained as \(\hat{N} = n_{1+}n_{+1}/n_{11}\) (see Pollock, 2000). Under the assumptions above, conditional on \(n_{1+}\) and \(n_{+1}\), the Lincoln–Petersen estimator is a maximum likelihood estimator (Wittes, 1972) and an estimator of its variance is \(\hat{\sigma}^2(\hat{N}) = n_{1+}n_{+1}n_{12}n_{21}/n_{11}^2\) (Chandra-Sekar and Deming, 1949; Bishop et al., 1975). Nevertheless, Alho (1990) shows that the Lincoln–Petersen estimator is a maximum likelihood estimator, and derives the same estimator of its variance (presented above) unconditional on \(n_{1+}\) and \(n_{+1}\) in a population with homogeneous capture probabilities.

The Lincoln–Petersen estimator lacks finite moments, because \(n_{11}\) may be zero if \(n_{1+} + n_{+1} < N\), and it is biased (Wittes, 1972). Chapman (1951) modified the Lincoln–Petersen estimator to produce a nearly unbiased estimator which possesses all finite moments given as \((n_{1+} + 1)(n_{+1} + 1)/(n_{11} + 1) - 1\), which is obtained adding 1 to the recaptured count \(n_{11}\) in the estimation of the number of individuals never captured. Wittes (1972) demonstrates that the Chapman estimator is unbiased if \(n_{1+} + n_{+1} \geq N\), and when \(n_{1+} + n_{+1} \leq N\), its bias is given by \(-N - n_{1+})!(N - n_{1+}!/(N - n_{1+} - n_{+1} - 1)!N!\). Also Wittes (1972) presents an estimator for the variance of the Chapman estimator given by \(n_{12}n_{21}(n_{1+} + 1)(n_{+1} + 1)/(n_{11} + 1)^2(n_{11} + 2)\).

Through another reasoning Evans and Bonett (1994) modified the Lincoln–Petersen estimator in which \(.5\) is added to each observed frequency resulting \((n_{1+} + 1)(n_{+1} + 1)/(n_{11} + .5) - 1.5\)
with an estimator of its variance given by \( (n_{1+} + 1)(n_{+1} + 1)(n_{12} + .5)(n_{21} + .5)/(n_{11} + .5)^3 \).

This work is just concerned with the single capture–recapture method, however the reader is referred to Pollock (2000) for a general discussion, to Seber (2001) for relatively new advances in capture–recapture methods and to Chao (2001) and Chao and Huggins (2005a,b) for a review of models for closed populations. Also the reader interested in multiple samples or multiple record systems is really encouraged to consult Bishop et al. (1975) where the methodology of estimation for multiple recapture through log-linear models is presented.

3 Closed–Form Confidence Intervals

3.1 Wald–type Confidence Intervals

The simplest \((1 - \alpha)100\%\) confidence interval for \(N\) is the Wald–type

\[
\hat{N} \pm z_{\alpha/2}\hat{\sigma}(\hat{N})
\]

based on the asymptotic normality of the maximum likelihood estimator \(\hat{N} = n_{1+}n_{+1}/n_{11}\), with \(z_{\alpha/2}\) the standard normal quantile \(\alpha/2\). Because of the bias of both the Lincoln–Petersen estimator and the estimator of its variance (Evans and Bonett, 1994), in this paper the interval built with these estimators is not considered. Evans and Bonett (1994) show that adding .5 to each cell in this estimator reduces the mean square error. Here, the method to construct the interval presented by Evans et al. (1996) is followed. It works with the Wald–type interval adding .5 to each observed count before computation of both the estimator and the estimator of its variance, giving a confidence interval based in (2) as

\[
\frac{(n_{1+} + 1)(n_{+1} + 1)}{n_{11} + .5} - 1.5 \pm z_{\alpha/2}\sqrt{\frac{(n_{1+} + 1)(n_{+1} + 1)(n_{12} + .5)(n_{21} + .5)}{(n_{11} + .5)^3}}
\]

Also, the performance of a Wald–type confidence interval using the Chapman estimator is explored. For this approach, the estimator of the variance of the Chapman estimator, presented by Witte (1972), is taken into account and the confidence interval of the type (2) is obtained as (Borchers et al., 2002, pp. 112)

\[
\frac{(n_{1+} + 1)(n_{+1} + 1)}{n_{11} + 1} - 1 \pm z_{\alpha/2}\sqrt{\frac{n_{12}n_{21}(n_{1+} + 1)(n_{+1} + 1)}{(n_{11} + 1)^2(n_{11} + 2)}}
\]

3.2 Chapman Confidence Interval

Jensen (1989) presents a confidence interval based in the work of Chapman (1951), where he presents a nearly unbiased estimator for \(N\), and uses the variance of the Lincoln–Petersen estimator with normal approximation for constructing a confidence interval for \(N\) as

\[
\frac{n_{1+}n_{+1}}{n_{11} + z_{\alpha/2}\sqrt{n_{11}(1 - n_{11}/n_{1+})(1 - n_{11}/n_{1+})}}
\]
suggested to be applied for $n_{11} < .1n_{1+}$ and $n_{1+} > .1N$. Thus, the performance of this interval must be very limited as we shall see. Note this confidence interval can not be computed if $n_{11} = 0$, and here it is taken to be the entire real line in such case. Here this interval is named as the Chapman interval.

### 3.3 Jensen Confidence Interval

Jensen (1989) uses the inversion of the Chapman estimator and the normal approximation to the hypergeometric distribution to give a confidence interval for $1/N$ arguing that the distributions of the reciprocal of this nearly unbiased estimator is more nearly normal. The Jensen confidence interval is presented as

$$\frac{(n_{1+} + 1)(n_{1+} + 1)}{n_{11} + 1 \mp z_{\alpha/2}\sqrt{n_{11}(1 - n_{11}/n_{1+})(1 - n_{11}/n_{1+})}}$$

Jensen (1989) also presents a confidence interval for the capture–recapture experiment with replacement in the second sample, but in this paper we are just concerned with sampling without replacement. The Jensen interval can not be computed if $n_{1+} = 0$ or $n_{1+} = 0$, thus this interval is taken as the entire real line in such cases.

### 3.4 Chao Confidence Interval

Chao (1989) presents a confidence interval for the population size in the multiple capture–recapture method using maximum likelihood estimators. For the single capture–recapture method the Lincoln–Petersen estimator is obtained and an estimator of its variance is presented as $\hat{\sigma}(\hat{N}) = \{((\hat{N} - n_{11} - n_{12} - n_{21})^{-1}) + \hat{N}^{-1} - (\hat{N} - n_{1+})^{-1} - (\hat{N} - n_{1+})^{-1}\}^{-1}$. Chao (1989) builds a confidence interval for the expected unobserved count $\mu_{22}$ as

$$(\hat{N} - n_{11} - n_{12} - n_{21}) \exp \left[ \pm z_{\alpha/2} \sqrt{\log \left\{ \frac{\hat{\sigma}(\hat{N})}{(\hat{N} - n_{11} - n_{12} - n_{21})^2} \right\} \right]$$

This interval is based in the asymptotic normality of log($\hat{N} - n_{11} - n_{12} - n_{21}$). In order to find a confidence interval for $N$, the count of observed individuals $n_{11} + n_{12} + n_{21}$ is added to both the lower and upper bounds of the interval above. Note if $n_{11} = 0$ the Lincoln–Petersen estimator can not be computed. Also, if $n_{12} = 0$, then $\hat{N} = n_{1+}$ or if $n_{21} = 0$, then $\hat{N} = n_{1+}$ and thus the estimator of the variance can not be computed if some of the counts $n_{11}, n_{12}$ or $n_{21}$ is equal to zero. In such cases this interval is taken to be the entire real line.

### 4 Robust Confidence Intervals

#### 4.1 Monte Carlo Confidence Intervals

Following Buckland and Garthwaite (1991), the Monte Carlo method to construct confidence intervals for the population size in single capture–recapture estimation unconditional on $n_{1+}$
and \( n_{+1} \) works as follows. First, estimate the parameters \((N, p_1, p_2)\) of the probabilistic model (1) using some single capture–recapture estimator \( \hat{N} \) and the capture probabilities by \( \hat{p}_1 = n_{1+}/\hat{N} \) and \( \hat{p}_2 = n_{+1}/\hat{N} \). Replace \((N, p_1, p_2)\) by \((\hat{N}, \hat{p}_1, \hat{p}_2)\) in the multinomial model (1), rounding \( \hat{N} \) to the nearest integer. Later, generate a random vector from the fitted multinomial distribution and compute the value of the single capture–recapture estimator for the random vector obtained. Repeat the last step \( b \) times, construct the empirical distribution of the single capture–recapture estimator from those replications and find the quantiles \( \alpha/2 \) and \( 1 - \alpha/2 \) of this empirical distribution. These quantiles are taken as the lower and upper bounds of the Monte Carlo confidence interval. Of course the above method depends on the single capture–recapture estimator from those replications and find the quantiles \( \alpha/2 \) and \( 1 - \alpha/2 \) of this empirical distribution. These quantiles are taken as the lower and upper bounds of the Monte Carlo confidence interval of the single capture–recapture estimator and on the number of replications \( b \). In this paper the Chapman estimator is used because of its properties referenced in Section 2 and the number of replications is fixed at 10000. On the other hand, as presented in Buckland (1984), if confidence intervals would be computed conditional on \( n_{1+} \) and \( n_{+1} \), the model associated to the single capture–recapture method is the hypergeometric model and the random variable is the count \( n_{11} \). Thus, random values \( n_{11}^* \) are generated from the hypergeometric distribution with parameters \((N, n_{1+}, n_{+1})\) and the Monte Carlo interval can be computed similarly as stated above, but such is not the case here.

Bootstrap methods (Efron, 1979) for capture–recapture estimation for closed populations require the estimation of the unobserved count \( n_{22} \) to complete the random multinomial vector \((n_{11}, n_{12}, n_{21}, n_{22})\). The bootstrap samples are taken from the \( \hat{N} \) individual capture histories and thus Buckland and Garthwaite (1991) explain how this method coincides with the multinomial Monte Carlo method explained above and implemented in this work.

### 4.2 Profile–Likelihood Based Confidence Intervals

Let \( \hat{\phi} \) be the maximum likelihood estimator of the vector of parameters \( \phi \) of the multinomial model, which belongs to some parameter space \( \Phi \). In single capture–recapture estimation with the model (1) \( \hat{\phi} \) is given by \((\hat{N}, \hat{p}_1, \hat{p}_2) = (n_{1+}/n_{11}, n_{11}/n_{+1}, n_{11}/n_{+1})\) (Bishop et al., 1975, pp. 232). For this vector we obtain the maximum log–likelihood as \( l(\hat{\phi}) = \max_{\phi \in \Phi} l(\phi) = \log P_{\hat{\phi}}(n_{11}, n_{12}, n_{21}, n_{22}) \). Now replace the parameter \( N \) by some value \( \beta \geq n_{11} + n_{12} + n_{21} \). With this restriction the parameter space is reduced to \( \Phi_N(\beta) = \{ \phi \in \Phi | N = \beta \} \). Now compute the profile likelihood for \( \beta \) as \( \bar{I}_N(\beta) = \max_{\phi \in \Phi_N(\beta)} l(\phi) = \log P_{\hat{\phi}_N(\beta)}(n_{11}, n_{12}, n_{21}, n_{22}) \), i.e. the log–likelihood obtained with the maximum likelihood estimator of \( \beta \) constrained to \( N = \beta \). In this case it is \( \hat{\phi}_N(\beta) = (\beta, n_{1+}/\beta, n_{+1}/\beta) \) (Bishop et al., 1975, pp. 446). Now, note that the value \( \hat{\phi} \) is obtained from the hypotheses of independence. For each value \( \beta \) we obtain a possible multinomial vector for which we can test the hypotheses of independence with the log–likelihood ratio statistic \( 2[l(\hat{\phi}) - \bar{I}_N(\beta)] \) and we reject the hypothesis if this is greater than the quantile \( 1 - \alpha \) of a chi–squared distribution with 1 degree of freedom \( \chi^2_{1,(1-\alpha)} \), for some nominal confidence level \( (1 - \alpha)100\% \). Thus it is natural to think about a confidence interval for \( N \) as the set of values \( \beta \) for which the hypothesis is not rejected, i.e. \( \{ \beta | 2[l(\hat{\phi}) - \bar{I}_N(\beta)] \leq \chi^2_{1,(1-\alpha)} \} \). The bounds of this interval can be obtained finding the values of \( \beta \) such that \( 2[l(\hat{\phi}) - \bar{I}_N(\beta)] - \chi^2_{1,(1-\alpha)} = 0 \). Note that if \( n_{11} = 0 \) the maximum likelihood estimators can not
be computed and consequently this interval is uninformatative. Thus it is taken to be the entire real line in such case. However, due to its flexibility Cormack (1992) proposes this method for closed capture–recapture models and Evans et al. (1996) present this method for capture–recapture models using the method to find the profile–likelihood intervals proposed by Venzon and Moolgavkar (1988) which requires less computation. In order to make this method more efficient, recently a modification was proposed by Gimenez et al. (2005). Recently, Baillargeon and Rivest (2007) present the implementation of this method in the package \texttt{Rcapture} of the software \texttt{R} (R Development Core Team, 2008), based in the results of Cormack (1992). In this paper we study the performance of the profile–likelihood confidence interval using the most recent version (1.2–0) of this package (Baillargeon and Rivest, 2009).

5 Transformed Logit Confidence Intervals

Agresti (1999, 2002, pp. 71) presents the general idea of a logit confidence interval for the odds ratio taking exponentials of

\[
\log \hat{\theta}_c \pm z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)
\]

(3)

where \(\hat{\theta}_c = (n_{11} + c)(n_{22} + c)/(n_{12} + c)(n_{21} + c)\) denotes an estimator for the odds ratio \(\theta = \mu_{11}\mu_{22}/\mu_{12}\mu_{21}\) for a 2 × 2 contingency table of expected frequencies \(\mu_{ij}, i = 1, 2, j = 1, 2\).

In (3) \(\hat{\sigma}(\log \hat{\theta}_c)\) is given by

\[
\sqrt{\frac{1}{n_{11} + c} + \frac{1}{n_{12} + c} + \frac{1}{n_{21} + c} + \frac{1}{n_{22} + c}}
\]

(4)

The value \(c\) in (3) is a constant, and depending on this value Agresti (1999) refers to this confidence interval as the Woolf interval for \(c = 0\) and as the Gart interval for \(c = .5\).

From (3) it is possible to state that

\[
P[\log \hat{\theta}_c - z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c) \leq \log \theta \leq \log \hat{\theta}_c + z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)] \approx 1 - \alpha
\]

(5)

For \(c = .5\) with small samples this is true unless the true log odds ratio is large, say \(|\log \theta| > 4\) (Agresti, 1999). Under the assumption of independence of the samples in the capture–recapture method it is known that \(\theta \approx 1\), and then here \(\log \hat{\theta}_c = 0\) is fixed. However, we estimate \(\hat{\sigma}(\log \hat{\theta}_c)\) as we shall see to take into account the variability of the single capture–recapture method. Thus, taking exponentials of the members of the inequalities presented in (5) it is obtained

\[
P[\exp\{-z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)\} \leq \theta \leq \exp\{z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)\}] \approx 1 - \alpha
\]

The odds ratio is given by \(\theta = \mu_{11}\mu_{22}/\mu_{12}\mu_{21}\), and from this it is found

\[
P\left[\frac{\mu_{12}\mu_{21}}{\mu_{11}} \exp\{-z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)\} \leq \mu_{22} \leq \frac{\mu_{12}\mu_{21}}{\mu_{11}} \exp\{z_{\alpha/2} \hat{\sigma}(\log \hat{\theta}_c)\}\right] \approx 1 - \alpha
\]
However $\mu_{11}, \mu_{12}$ and $\mu_{21}$ must be estimated, and here we propose to estimate these expected frequencies following the method based on adding $c$ to each cell of the observed counts $n_{ij}$ giving an approximate $(1 - \alpha)100\%$ confidence interval for $\mu_{22}$ as

$$\frac{(n_{12} + c)(n_{21} + c)}{n_{11} + c} \exp\{\pm z_{\alpha/2}\hat{\sigma}(\log \hat{\theta})\} - c$$

(6)

Of course, the resulting transformed logit confidence interval for $N$ is found by adding $n_{11} + n_{12} + n_{21}$ to each limit of the confidence interval for $\mu_{22}$. In (6) the standard error is still unspecified because in (4) the value of $n_{22}$ is required. Here it is replaced by an estimation given by $\hat{\mu}_{22} = \frac{(n_{12} + c)(n_{21} + c)}{(n_{11} + c)} - c$ to give $\hat{\sigma}(\log \hat{\theta})$ as

$$\sqrt{\frac{1}{n_{11} + c} + \frac{1}{n_{12} + c} + \frac{1}{n_{21} + c} + \frac{n_{11} + c}{(n_{12} + c)(n_{21} + c)}}$$

In the case that we take $c = 0$, if $n_{11} = 0$, $n_{12} = 0$ or $n_{21} = 0$, the confidence interval is taken to be the entire the real line.

This paper refers to the standard transformed logit interval for $c = 0$ and to the .5 transformed logit interval for $c = .5$, despite the fact that following Agresti (1999), the last one must be named as the Haldane–Anscombe–Gart transformed interval, but for capture–recapture also some credit must be given to Evans and Bonett (1994) because of their estimators correction using $c = .5$. However, such long name is not practical.

6 Comparison of Performances

6.1 Coverage Probabilities with Small Populations

The actual coverage probabilities of the confidence intervals are computed from the formula

$$\sum_{n_{11}=0}^{N} \sum_{n_{12}=0}^{N-n_{11}} \sum_{n_{21}=0}^{N-n_{11}-n_{12}} [I(n_{11}, n_{12}, n_{21}, \alpha)P(\omega)(n_{11}, n_{12}, n_{21}, n_{22})]$$

(7)

where $I(n_{11}, n_{12}, n_{21}, \alpha)$ is computed as a function of the observed counts $(n_{11}, n_{12}, n_{21})$ and the nominal coverage value $(1 - \alpha)100\%$, here stated as 95%. This function equals 1 if the confidence interval contains the true value $N$ and equals 0 if it does not contain $N$.

The actual coverage probabilities (7) where evaluated for the nine intervals presented in Section 3, 4 and 5, for two small population sizes $N = 20, 50$, for capture probabilities of the second sample fixed at $p_2 = .1, .5, .9$, and for 99 uniformly distributed values of the probability of being captured in the first sample, $p_1$, between .01 and .99. Also the coverage probabilities were computed for $N = 10, 30$ but the results are not reported here because of the similarity with those obtained for $N = 20$. All figures here presented show the actual coverage probabilities as a function of $p_1$. For pedagogic purposes, the coverage probabilities for the nine intervals are presented separately in three groups in Figures 1, 2 and 3. In Figure 1 we present the coverage of the .5 Wald–type interval by the dot–dashed line, the Chapman Wald–type interval by the gray dot–dashed line and the Chapman interval by the gray dashed line. In
Figure 2 the coverage of the Jensen interval is presented by the dashed line, the standard logit interval by the gray solid line and the Chao interval by the dotted line. Finally in Figure 3 the coverage probability level of the .5 logit transformed interval is represented by the solid line, the Monte Carlo interval by the dot–dot–dashed line and the profile–likelihood interval by the gray dot–dot–dashed line. In addition, in these three figures a horizontal gray dotted line represents the nominal 95% coverage level.

In Figure 1 it is shown that the .5 Wald–type interval presents an under coverage for \( p_2 = .1 \) for most values of \( p_1 \) and it approaches to the nominal level as \( p_1 \) increases. Also it presents a small under coverage for \( p_2 = .5 \). The .5 Wald–type interval seems as it improves its coverage probability for large sums of capture probabilities under the conditions presented. This is due partly to the large bias of the estimators used to build this interval with these small populations (Evans and Bonett, 1994) even though those estimators are corrected. In fact, for \( N = 50 \) with very small capture probabilities, \( p_1 = p_2 = .1 \), Evans and Bonett (1994) show that their modification presents a mean squared error greater than the Lincoln–Petersen estimator. However, Evans and Bonett (1994) also show that the estimators presented by them becomes less biased for moderately large capture probabilities, which coincides with the improvement of the coverage probabilities of the .5 Wald–type interval. The performance of the .5 Wald–type interval was studied also by Evans et al. (1996) through a Monte Carlo study for \( N = 25, 50, 100, 800 \) and they also find that the performance of this interval is improved as the population size or the capture probabilities increase.

Figure 2 shows that the standard transformed logit and Chao intervals perform very similar for the configurations presented. Both perform poorly for \( p_2 = .9 \) for most values of \( p_1 \), and for large values of \( p_1 \) with \( p_2 = .5 \) even though that in general they present an over coverage when \( p_2 = .1 \). For these intervals the actual coverage probability decreases as \( p_1 \) increases for the configurations presented in Figure 2 until \( p_1 \) is near to 1. It seems as these intervals loose their coverage probability for medium sums of capture probabilities with these small populations and they increase their coverage probabilities only for very large capture probabilities. Let us remind that these intervals are taken to be the entire real line if \( n_{11} = 0, n_{12} = 0 \) or \( n_{21} = 0 \). The probability to obtain those values for those small \( N \) with small probabilities is not negligible, in fact is very large. However, the probability to obtain \( n_{11} = 0, n_{12} = 0 \) or \( n_{21} = 0 \) approaches to zero as \( p_1 \) and \( p_2 \) increases. Thus, the good performance of these intervals for small capture probabilities is due to take the entire real line when some of the observed counts is zero. This is clue to say that these intervals should not be used for small populations because for non large capture probabilities they become uninformative with a large probability. On the other hand, for large capture probabilities with small populations these intervals do not have good coverage probabilities.

In Figure 3 it can be seen that the performance of the profile–likelihood based confidence interval is quite irregular, as it is not improved by the increment of \( N \) from 20 to 50, as it is shown for \( p_2 = .1 \) and \( p_2 = .9 \). Nevertheless, there is an improvement of this interval from \( N = 20 \) to \( N = 50 \) with \( p_2 = .5 \). Also, in Figure 3 it is shown that the .5 transformed logit interval has an over coverage for most values of \( p_1 \), for all \( N \) and \( p_2 \) reported. This can be explained partly by the performance of the Gart version of the logit interval for the odds
The coverage probability level of the nominal 95% (gray dotted), .5 Wald–type (dot–dashed), Chapman Wald–type (gray dot–dashed) and Chapman (gray dashed) confidence intervals.

For the .5 transformed logit interval, the under coverage presented in some of the reported configurations is small, i.e. not as extreme as the presented by the other intervals. This under coverage can be seen for \( N = 50 \), with \( p_2 = .1 \) and \( p_1 \approx 1 \) and for \( N = 20 \), with \( p_2 = .9 \) with some small values of \( p_1 \). Nevertheless, in general, the performance of this interval seems to be acceptable. Note for \( N = 50 \), for some configurations of capture probabilities the actual coverage probability of this interval is nearly the nominal level, as it is shown for \( p_2 = .5 \) with \(.3 < p_1 < .9\) and for \( p_2 = .9 \) with \(.1 < p_1 < .7\). In the configurations of the parameters reported in Figure 3, 25% of the actual coverage probabilities of the .5 transformed logit interval are over 99.4%, 50% over 97%, 75% over 94.6%. Also, the coverage level of this interval never falls down 90% for the configurations presented. Another advantage of this interval, for example over the standard transformed logit and Chao intervals, is that it is always informative because using \( c = .5 \) ensures that it can be always computed.

Figures 1, 2 and 3 show that the Chapman Wald–type, Jensen, Chapman and Monte Carlo intervals perform very poorly for small populations. None seems to have a general acceptable performance. For the Chapman interval, remember that it was suggested to be applied, under hypergeometric sampling, for \( n_{11} < .1n_{+1} \) and \( n_{1+} > .1N \) (Jensen, 1989). However, the actual coverage probabilities of those four intervals approach the nominal level for \( N = 50, p_2 = .5 \)
Figure 2: The lines denote the coverage probability level of the nominal 95\% (gray dotted), Jensen (dashed), standard logit (gray solid) and Chao (dotted) confidence intervals.

with \(0.2 < p_1 < 0.8\). In general it seems as the performance of those four intervals is better for capture probabilities near to \(0.5\). Also, for these four intervals, the performance is improved as the population size increases, as it can be seen from the decreasing under coverage from \(N = 20\) to \(N = 50\). Jensen (1989) presents a simulation study to compare the efficiency of the Jensen and Chapman intervals, but it is carried out conditional on \(n_{1+}\) and \(n_{+1}\), i.e. following the hypergeometric model. Unfortunately, the coverage probabilities of these intervals are not computed in that work.

### 6.2 Coverage Probabilities with Large Populations

For each configuration of \((N, p_1, p_2)\) the number of members of the summation (7) required to find the actual coverage probabilities is equal to \((N + 1)(N + 2)(N + 3)/6\). Thus, for large populations, computing (7) is not practical because it would be unreasonably long. Instead, the performance of the intervals presented in Section 3, 4 and 5 can be explored through a Monte Carlo study. Following the scheme of generating random multinomial vectors for each configuration of \((N, p_1, p_2)\), 10000 random vectors were generated from the multinomial distribution for the combinations of \(N = 500, 1000, p_2 = .1, .5, .9\), and for 99 uniformly distributed values of \(p_1\) between \(0.01\) and \(0.99\). For the profile-likelihood interval only 1000 random vectors
Figure 3: The lines denote the coverage probability level of the nominal 95% (gray dotted), .5 logit (solid), Monte Carlo (dot–dot–dashed) and profile–likelihood (gray dot–dot–dashed) confidence intervals.

were generated because the computation of this interval is very time consuming. For each simulated random vector, each 95% nominal confidence interval was computed and determined if it contained the true value $N$. Later, the proportion of the intervals containing the true value was calculated. This value is taken as an estimation of the true actual coverage probability. The simulation study was carried out in the software R (R Development Core Team, 2008).

The results obtained for $N = 1000$ were quite similar to those with $N = 500$. As usual, the performance of those intervals is improved as the population size increases. Thus, only the coverage probabilities for $N = 500$ are presented. In this section we only present graphically the results of those intervals which does not have a very poor performance for small populations, i.e. the .5 Wald–type, the profile–likelihood and .5 transformed logit confidence intervals. In Figure 4 the estimated actual coverage probabilities of these three intervals are also presented as a function of $p_1$.

Figure 4 shows that for small values of $p_1$ with $p_2 = .1$ the .5 Wald–type interval continues presenting a great under coverage. It can be seen that the profile–likelihood interval also presents a great under coverage for $p_1 \approx 1$ for $p_2 = .1$ and $p_2 = .5$. Both the .5 Wald–type and profile–likelihood intervals present a small under coverage for $p_1 \approx 0$ with $p_2 = .5$. The profile–likelihood interval also presents a very great under coverage for $p_2 = .9$ with $p_1$ near to 0 or 1. The performance of the profile–likelihood interval is quite irregular as it was shown
in Section 6. For example for $N = 20$ with $p_2 = .9$ this interval do not have very large under coverage. In general for an increment of the population size we can say that the coverage probabilities of the profile–likelihood interval for single capture–recapture is improved for non extreme capture probabilities but it is deteriorated for large capture probabilities. Also Figure 4 for $p_2 = .1$ shows that the .5 transformed logit interval performs well, even for the values of $p_1$ very near to 0 or 1, where it just presents a small under coverage when $p_1 \approx 1$. For the large populations explored here, it also presents a good performance for $p_2 = .5$ and $p_2 = .9$ even for large and small values of $p_1$, with the exception of an over estimation for $p_1 = .99$ when $p_2 = .9$ and for $p_1 = .01$ when $p_2 = .5$.

For the other intervals not represented in Figure 4 it is found that the Chao and standard transformed logit intervals present a great under coverage when $p_2 = .9$ and $p_1$ is large ($p_1 > .8$ for $N = 500$). The Chapman Wald–type interval continues presenting a poor performance specially for values of $p_1$ near to 0 or 1, and for $N = 500$ it approaches but never takes the 95% nominal level. For $N = 500$ the Jensen, Chapman and Monte Carlo intervals perform poorly for $p_2 = .1, .5, .9$ when $p_1 > .8$ or $p_1 < .2$ and their under coverage increase as $p_1$ approaches to 0 or 1. Particularly, the Chapman interval presents a very low coverage probability for $p_1$ near 0 when $p_2 = .1$.

Finally, it was found that for central values of $p_2$, i.e. $p_2 = .5$, the actual coverage is nearly the nominal 95% for all the confidence intervals considered here, for most values of $p_1$, with the exception of the Chapman and Chapman Wald–type intervals for $p_1$ near to 0.

### 6.3 Expected Width with Large Populations

The expected width of a confidence interval is an useful measure of its performance. However, a short confidence interval with a low actual coverage probability is not useful. In fact, comparing the expected widths of two intervals with very different actual coverage probabilities is not appropriate. Thus, the expected relative width of the nine intervals studied above is explored.
Figure 5: The lines denote the expected relative width of the .5 Wald–type (dot–dashed), Chapman Wald–type (gray dot–dashed), Chapman (gray dashed), Jensen (dashed), standard logit (gray solid), Chao (dotted), .5 logit (solid), Monte Carlo (dot–dot–dashed) and profile–likelihood (gray dot–dot–dashed) confidence intervals.

for some configurations for which the actual coverage probabilities are nearly the nominal 95%: \(0.2 < p_1 < 0.8, p_2 = 0.1, 0.5, 0.9\) and \(N = 1000\). For each configuration of \((N, p_1, p_2)\) 1000 random multinomial vectors were generated using R and the mean of the widths of the intervals obtained is reported. This mean width is taken to be an estimation of the expected width. For each confidence interval the expected width is computed conditional on those confidence intervals that are not the entire real line. Figure 5 illustrates the relative widths as a function of \(p_1\). As the reader shall see, for these configurations, the expected widths are nearly the same for the nine intervals, and thus they are presented in the same plot. For small capture probabilities, the largest width is presented by the Chapman interval and the shortest width by the Chapman Wald–type interval. The .5 transformed logit appears near to the medium of the expected widths as well as the others. As a general overview, for the nine intervals the expected widths decrease and become the same as the coverage probabilities increase.

7 Discussion

The .5 Wald–type confidence interval has a good performance for large population sizes, but it has a great under coverage for very small record probabilities. Also, it performs poorly for small population sizes for relatively small record probabilities. The standard transformed logit and Chao intervals present an under coverage for large values of the capture probabilities and for small capture probabilities the probability to obtain an uninformative standard transformed logit or Chao confidence interval is great, and thus these intervals should not be taken as a good alternative for small populations in single capture–recapture estimation.

The profile–likelihood confidence interval present a performance quite irregular as it improves its performance as the population size increases only for medium capture probabilities but its coverage is deteriorated for large capture probabilities. Although this interval can be
used for a wide range of models (not only capture-recapture models), it is limited for single capture-recapture estimation because it cannot be computed for the count of recaptures equal to zero and it has a great under coverage for some configurations.

As a result of this study it is found that the .5 transformed logit confidence interval has an exceptionally good performance even for extreme capture probabilities, i.e. near to 0 or 1. For small populations this interval has a conservative performance, showing coverage probabilities greater than the nominal 95% for most of the configurations reported, and small under coverage for some configurations.

For the Jensen, Chapman Wald–type, Chapman and Monte Carlo intervals, their performances do not seem to be acceptable even for large populations with extreme capture probabilities, and are aberrant for small populations, because they systematically exhibit coverage probabilities lower than the nominal. Though the Monte Carlo interval is quite general and robust its poor performance for small populations restricts its implementation for single capture-recapture estimation.

As a general conclusion, it was shown, for large population sizes and for not extreme combinations of probabilities, all the confidence intervals reported here seem to have a coverage level near to the nominal 95% reported and approximately the same expected widths. This general overview lets the .5 transformed logit confidence interval as the best of the intervals reported, because it is better even for small populations and extreme capture probabilities. Also, the .5 transformed logit interval has the advantage to be easily computed and can be always calculated, even for some count equal to zero. Finally, this interval presents an advantage over intervals such as the Wald–type, which can have lower limits below $n_{11} + n_{12} + n_{21}$. This is undesirable because it is known that $N > n_{11} + n_{12} + n_{21}$. On the contrary, the lower limit of the .5 transformed logit interval has the property of never being less than $n_{11} + n_{12} + n_{21}$.

Acknowledgements

The author thanks to B. P. Urdinola, L. M. González, S. T. Buckland, A. Agresti and an anonymous reviewer for reading the preliminary versions of this work and their helpful comments and suggestions. Also to C. E. Pardo, A. Irlande, S. Bailargeon and L.–P. Rivest for their help about programming the simulations and computations in this work.

References


