Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

Asymptotic properties of a capture-recapture population size estimator when s populations are partially uncatchable

Cibele Queiroz da-Silva

Relatório Técnico
RTP-08/2001

Relatório Técnico
Série Pesquisa

Abstract

We present the asymptotic distribution for the estimator of the population size for the case of s partially catchable populations. Our approach is useful for capture-recapture studies with photo-identification data where part of the population does not have any distinctive characteristic which allows unique identification of the individuals. Estimation of bowhead whale (*Balaena mysticetus*) abundance using photo-id data is an example of such a problem since it is not possible/practical to attach an artificial mark to the captured individuals, but the acquired natural marks throughout their lives are useful to allow the analyst to distinguish individuals. This work represents an extension of Theorem 4 in Sanathanan (1972).

Keywords: Capture-recapture; Closed uncatchable populations; Asymptotic properties.

1 Introduction

Sanathanan (1972) derives asymptotic theory for estimating the number of trials of a multinomial distribution from an incomplete observation of the total cells. The author also studies the case of s populations. Estimation of a population size when a capture-recapture experiment is undertaken is an example of such a problem since cell totals are observed only for the cases where individuals are captured at least once over the sampling experiment.

This paper deals with the problem of estimating the size of s partially catchable populations. Our approach is useful for capture-recapture studies where photo-identification data are collected and part of the population does not have any distinctive characteristic which allows unique identification of the individuals. Estimation of bowhead whale (*Balaena mysticetus*) abundance using photo-id data is an example of such a problem. In the case of the bowhead whale it is not possible/practical to attach an artificial mark to the captured individuals, but the acquired natural marks throughout their lives are useful to allow the analyst to distinguish individuals. Contrary to the notion of a marked individual in the capture-recapture studies, a *marked bowhead* means that it has acquired natural marks enough to make reidentification possible.

Since part of the population never acquire any natural marks, classical capture-recapture estimators are not able to estimate the size of the whole population composed of marked and unmarked individuals.

Some previous work has been done on estimating population size when only part of the population is catchable. Seber (1982), p. 72, gave an estimate $\hat{N} = \frac{\hat{N}^m}{\hat{p}^*}$, where $\hat{N}^m$ is the estimated number of individuals in the catchable population and $\hat{p}^*$ is the estimated proportion of the population that is catchable. Using the delta method, he derived a variance expression under the assumption that $\hat{N}^m$ and $\hat{p}^*$ are statistically independent. Williams *et al.* (1993), working with bottlenosed dolphin photo-identification data, used Seber's approach with $\hat{N}^m$ as the estimated number of marked individuals in the population and $\hat{p}^*$ as the proportion of the photographs that were of marked individuals. They used photos from the same studies to obtain $\hat{N}^m$ and $\hat{p}^*$, so the assumption of statistical independence of these estimates on which the delta method variance is based does not hold. To address this issue, da Silva (1999) and da Silva *et al.* (2000) developed alternative interval estimates of population size from photo-identification data when the population includes unmarked animals and compared their approach to the one developed by Williams *et al.* (1993) using simulated bowhead data. The work of da Silva *et al.* (2000) showed that their results had a good agreement with previous works (Raftery and Zeh (1998), Givens (1993) - personal communication).

da Silva *et al.* (2000) used parametric bootstrap methods to draw inferences to the bowhead whale population size. Their likelihood expresions do not belong to a regular family of distributions preventing variances of estimators to be obtained via the large sample theory of maximum likelihood estimators. Besides that, other kinds of approximations were too complicated due to some covariance terms involved in the calculations. However, some simplifications in the likelihood construction and the use of results obtained by Sanathanan (1972) allows the evaluation of the asymptotic distribution of the estimator of the total population size of the bowhead whale allowing for the presence of subpopulations. The subpopulations may represent, for example, classes of maturity which are related to the size of the animal. That may have an impact on the capture probabilities and in the population size estimate.

We are considering a closed population approach since bowhead whale photo-id data that are suitable for capture-recapture estimation is available only for four sampling occasions, spring and summer of 1985 and 1986. It is known from literature that bowhead whales have high survival (they live around 60 years) and low reproduction rates, which make us confident that a closed population model is not an unrealist one for the bowhead population.

In section 2 we introduce some notation. In section 3 we define a conditional likelihood based on good photos which incorporates information about the uncatchable part of the population. In section 4 we present and prove a theorem whith the asymptotic distribution of the estimator of the population size for the case of s partially catchable populations.

2 Notation

Quality of photos and extent of natural marks of a whale are important variables in our model formulation. A capture essentially means that a good quality photo of a whale was taken. In this case, if a natural mark is found then the whale is considered marked. We now introduce some notation. Let s be the number of populations being considered, then for $t = 1, \ldots, s$,

- N_t^u : the total number of unmarked whales in population t.
- N_t^m : the total number of marked whales in population t.
- $N_t = N_t^m + N_t^u$: the total number of whales in population t.
- $N = \sum_{t=1}^{s} N_t$: the total number of whales in population.

- $\mathbf{N^m} = (N_1^m, \ldots, N_s^m)$: the vector of population sizes of marked whales.

- $\Theta = (\theta_1, \ldots, \theta_r)$: the vector of independent parameters.

- $\Psi = (\psi_1, \ldots, \psi_s)$: where $\psi_t = N_t^m / N_t$, the vector of proportions of marked whales.

- X_a^t : the number of good photos of whales in population t at occasion a, $a = 1, \ldots, A$, where good photos are those from which the identification of the whales are possible.

- x_a^t : the number of good photos of **marked** whales in population t at occasion a, $a = 1, \ldots, A$.

- Ω: The set with 2^A elements where each element is a sequence of A binary components.

- n_{ti} : the total number of marked whales in population t with capture history i, where i is a label for an element of Ω, with $i = 1, \ldots, l$.

- p_{ti} : the capture probability of whales in population t with capture history i, where i is a label for an element of Ω, with $i = 1, \ldots, l$.

- n_t: the number of different whales in population t that were captured over the experiment. Notice that $n_t = \sum_{i=1}^{l} n_{ti}$.

Let $(n_{t1}, \ldots, n_{tl})$ be distributed according to the multinomial law $M(N_t^m; p_{t1}, \ldots, p_{tl})$, with $p_{tl} = 1 - \sum_{i=1}^{l-1} p_{ti}$. Let

$$p_{ti}(\Theta) = f_{ti}(\Theta), \quad i = 1, \ldots, l$$

where f_{ti} are known functions. For example, $f_{ti}(\Theta)$ may be a logistic function.

Sanathanan (1972) showed that expressing the capture histories in terms of f_{ti} leads to estimability of population size. In the next section we present a conditional likelihood for the model which we are proposing and some definitions.

3 A conditional likelihood based on good photos

The model we are going to discuss involves a combination of s multinomial models factorized in the same fashion described by Sanathanan (1972) and s binomial models. The multinomial models account for the marked (catchable) part of the population while the binomial ones incorporate, through the number of good photos of unmarked individuals, information about the uncatchable part of the population. Next we present a conditional likelihood function based on good photos. The conditional likelihood of $(\mathbf{N^m}, \Psi, \Theta)$, given $\{X_a^t\}$ is

$$\begin{aligned}
\mathcal{L} &= L(\mathbf{N^m}, \Psi, \Theta) = P\left(\{n_{t1}, \ldots, n_{tl-1}\}, \{x_a^t\} \mid \{X_a^t\}, \mathbf{N^m}, \Psi, \Theta\right) \\
&= P(\{n_{t1}, \ldots, n_{tl-1}\} \mid \mathbf{N^m}, \Theta)\, P\left(\{x_a^t\} \mid \{X_a^t\}, \Psi\right) \\
&= \prod_{t=1}^{s} \frac{N_t^m!}{(N_t^m - n_t)! \prod_{i=1}^{l} n_{ti}!} [p_{tl}(\Theta)]^{N_t^m - n_t} \prod_{i=1}^{l-1} [p_{ti}(\Theta)]^{n_{ti}} \prod_{a=1}^{A} \begin{pmatrix} X_a^t \\ x_a^t \end{pmatrix} \psi_t^{x_a^t} (1-\psi_t)^{X_a^t - x_a^t} \qquad (1)
\end{aligned}$$

where $p_{ti}(\Theta) = f_{ti}(\Theta)$, $i = 1, \ldots, l$, $n_{tl} = N_t^m - n_t$, and $n_t = \sum_{j=1}^{l-1} n_{tj}$. Thus, if $A = 3$, there are $l = 2^3 = 8$ possible capture histories: (1,1,1),(0,1,1),. . .,(0,0,0). For $t = 1, \ldots, s$, $N_t^m - n_t$ is the number of individuals in the population with capture history (0,0,0). For example, $f_{ti}(\Theta)$ may be a logistic function.

Let $\mathcal{L} = L(\mathbf{N^m}, \Theta)\, L(\Psi)$. According to Sanathanan (1972), $L(\mathbf{N^m}, \Theta)$ can be written as $L(\mathbf{N^m}, \Theta) = \prod_{t=1}^{s} L_{t1}(N_t^m, p_{tl}(\Theta))\, L_{t2}(\Theta)$. Thus, let us write $\mathcal{L}$ as

$$\mathcal{L} = \left(\prod_{t=1}^{s} L_{t1}(N_t^m, p_{tl}(\Theta))\right) \left(\prod_{t=1}^{s} L_{t2}(\Theta)\right) \left(\prod_{t=1}^{s} L_{t3}(\psi_t)\right) = L_1 \times L_2 \times L_3 \qquad (2)$$

where

$$\begin{aligned}
L_{t1}(N_t^m, p_{tl}(\Theta)) &= (N_t^m!/(n_t!(N_t^m - n_t)!))[1 - p_{tl}(\Theta)]^{n_t} [p_{tl}(\Theta)]^{N_t^m - n_t}, \\
L_{t2}(\Theta) &= (n_{t1}!/(n_t! \ldots n_{tl-1}!))[q_{t1}(\Theta)]^{n_{t1}} \ldots [q_{tl-1}(\Theta)]^{n_{tl-1}}, \quad \text{and} \\
L_{t3}(\psi) &= \prod_{a=1}^{A} \begin{pmatrix} X_a^t \\ x_a^t \end{pmatrix} \psi_t^{x_a^t} (1-\psi_t)^{X_a^t - x_a^t},
\end{aligned}$$

with $q_{ti}(\Theta) = p_{ti}(\Theta)/(1 - p_{tl}(\Theta))$, $i = 1, \ldots, l-1$.

For $t = 1, \ldots, s$, $\quad \hat{\psi}_t = \left(\sum_{a=1}^{A} x_a^t / \sum_{a=1}^{A} X_a^t\right)$, with $\hat{\psi}_t \rightarrow_{a.s.} \psi_t$.
Following the same lines of lemma 1 in Sanathanan (1972), for any p_t,

$$\hat{N}_t^m = n_t/(1 - p_{tl}), \quad \hat{N}_t = n_t/(\hat{\psi}_t(1 - \hat{p}_{tl})) = \hat{N}_t^m/\hat{\psi}_t, \quad \text{and} \quad \hat{N} = \sum_{t=1}^{s} \hat{N}_t.$$

An unconditional MLE of N is obtained when there exists $\hat{N}_t^m$ of N_t^m and $\hat{\psi}_t$ of ψ_t for $t = 1, \ldots, s$, which simultaneously maximize $\mathcal{L}$ over all admissible values of $(\mathbf{N^m}, \Psi, \Theta)$. A conditional MLE of N, $\hat{N}_c$, is obtained when we find $\hat{N}_t^m$ maximizing $L_{t1}(N_t^m, \hat{p}_{tc})$ where $\hat{p}_{tc} = p_{tl}(\hat{\Theta}_c)$ and $\hat{\Theta}_c$ is the value of Θ maximizing $L_{t2}(\Theta)$.

In the following section we enunciate and prove a theorem which is an extension of theorem 4 in Sanathanan (1972). Such theorem incorporates the uncatchable part of the population, making possible that asymptoptic properties of the MLE's be evaluated for the whole population size estimator of N. Now we need to introduce some more notation.

Let $\mathbf{N}_o^m$, Ψ_o, Θ_o, and $\mathbf{N}_o$ respectively be the true values of $\mathbf{N}^m$, Ψ, Θ, and $\mathbf{N}$. For $t = 1, \ldots, s$, and $i = 1, \ldots, l$, let $p_{ti}(\Theta)$ be denoted by p_{ti}^o when $\Theta = \Theta_o$, and denoted by $\hat{p}_{ti}$ when $\Theta = \hat{\Theta}$. Similarly let the partial derivatives of $p_{ti}(\Theta)$ with respect to θ_j be denoted by $p_{ti,j}^o$ when $\Theta = \Theta_o$, and denoted by $\hat{p}_{ti,j}$ when $\Theta = \hat{\Theta}$. Let $L_{1j} = \frac{\partial \log L_1}{\partial \theta_j}$, $L_{2j} = \frac{\partial \log L_2}{\partial \theta_j}$, and $\hat{L}_j = \hat{L}_{1j} + \hat{L}_{2j}$.

4 Case of s partially catchable populations

Let the $p_{ti}(\Theta)$'s admit first order partial derivatives which are continuous at every admissible value Θ. Let $\hat{\mathbf{N}}^m = (\hat{N}_1^m, \ldots, \hat{N}_t^m)$, $\hat{\Psi} = (\hat{\Psi}_1, \ldots, \hat{\Psi}_t)$, $\hat{\Theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_r)$, and $\hat{\mathbf{N}} = (\hat{N}_1, \ldots, \hat{N}_t)$ be the estimates of $\mathbf{N}_o^m$, Ψ_o, Θ_o, and $\mathbf{N}_o$ respectively such that

(i) $\hat{\Theta} \rightarrow_{a.s.} \Theta_o$

(ii) $\hat{\Psi} \rightarrow_{a.s.} \Psi_o$

(iii) $\left\langle (N^m_{ot})^{-1/2}\Big(\hat{N}^m_t - n_t/(1-\hat{p}_{tl})\Big)\right\rangle \rightarrow_{a.s.} 0$

(iv) $\left\langle (N^m_{ot})^{-1/2}\Big(\hat{N}_t - n_t/(\hat{\psi}_t(1-\hat{p}_{tl}))\Big)\right\rangle \rightarrow_{a.s.} 0$

(v) $(N^m_T)^{-1/2}\hat{L}_j \rightarrow_{a.s.} 0, \quad j=1,\ldots,r$

where $\langle e_t \rangle$ denotes the vector $(e_1,\ldots,e_s)$, and $N^m_T = \sum_{t=1}^{s} N^m_{ot}$. Let $\bar{\Sigma}^{-1} = (\sigma^{i,j})$ be the $(r+2s)\times(r+2s)$ matrix given by

$$
\begin{aligned}
\sigma^{j,h} &= \sum_{t=1}^{s} c_t \sum_{i=1}^{l} [p^o_{ti}]^{-1} p^o_{ti,j} p^o_{ti,h}, \quad j=1,\ldots,r \ \ ; h=1,\ldots,r \\
\sigma^{j,r+t} &= -(c_t)^{1/2}\psi_{ot}[p^o_{tl}]^{-1} p^o_{tl,j}, \quad j=1,\ldots,r \ \ ; t=1,\ldots,s \\
\sigma^{j,r+s+t} &= -(c_t)^{1/2} N_{ot}[p^o_{tl}]^{-1} p^o_{tl,j}, \quad j=1,\ldots,r \ \ ; t=1,\ldots,s \\
\sigma^{r+t,r+u} &= \delta_{tu}\psi^2_{ot}[p^o_{tl}]^{-1}(1-p^o_{tl}), \quad t=1,\ldots,s \ \ ; u=1,\ldots,s \\
\sigma^{r+t,r+s+w} &= \delta_{tw}\psi_{ot}N_{ot}[p^o_{tl}]^{-1}(1-p^o_{tl}), \quad t=1,\ldots,s \ \ ; w=1,\ldots,s \\
\sigma^{r+s+t,r+s+v} &= \delta_{tv}N^2_{ot}[p^o_{tl}]^{-1}(1-p^o_{tl}), \quad t=1,\ldots,s \ \ ; v=1,\ldots,s
\end{aligned}
$$

where $\delta_{tz} = 1_{\{t=z\}}$. Then,

$$
\begin{aligned}
U^t = \Big(& (N^m_T)^{1/2}(\hat{\Theta}-\Theta_o), (N^m_{o1})^{-1/2}(\hat{\psi}_1-\psi_{o1}),\ldots,(N^m_{os})^{-1/2}(\hat{\psi}_s-\psi_{os}), \\
& (N^m_{o1})^{-1/2}(\hat{N}_1-N_{o1}),\ldots,(N^m_{os})^{-1/2}(\hat{N}_s-N_{os})\Big)
\end{aligned}
$$

is asymptotically $N(0,\bar{\Sigma})$.

Proof: Since the proof for s populations was not presented in details by Sanathanan (1972), we are going to fill some of the gaps in order to make our proof more comprehensible. The general organization of this proof follows largelly the same steps developed by Sanathanan (1972): (I) prove that a built random vector Z^* follows a multivariate $N(0,\Sigma^{-1})$ distribution, and a random vector U, representing the difference between the vector of parameters' estimators and their true population values multiplied by the square root of an adequate constant are such that $U-\Sigma Z^*$ converges in probability to 0; (II) build random vectors $V_k^{(t)}$, with $t=1,\ldots,s$ and $k=1,\ldots,N_t$ such that Z^* may be expressed as a function of the sum of $V_k^{(t)}$. Then, by (i), U follows a multivariate normal distribution $N(0,\Sigma)$.

Step I:

Consider $N_T^m = \sum_{t=1}^{s} N_{ot}^m$, and $\lim_{N_T^m\to\infty} \frac{N_{ot}^m}{N_T^m} = c_t$ with $\sum_{t=1}^{s} c_t = 1$. Then, since $n_{tl} = N_t^m - n_t$, and by adding and subtracting adequate quantities involving $\hat{N}_t^m$ we have

$$\begin{aligned}
\frac{1}{\sqrt{N_T^m}}\frac{\partial \log \mathcal{L}}{\partial \theta_j}\bigg|_{\Theta=\hat{\Theta}} &= \frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\sum_{i=1}^{l}\frac{n_{ti}}{p_{ti}(\Theta)}\frac{\partial p_{ti}(\Theta)}{\partial \theta_j} \\
&= \frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\left[\frac{n_{tl}}{p_{tl}(\Theta)}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} + \sum_{i=1}^{l-1}\frac{n_{ti}}{p_{ti}(\Theta)}\frac{\partial p_{ti}(\Theta)}{\partial \theta_j}\right] \\
&= -\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}(\hat{N}_t^m - N_t^m)\left[p_{tl}(\Theta)\right]^{-1}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} \\
&\quad -\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\left[\frac{n_t}{1-p_{tl}(\Theta)} - \frac{\hat{N}_t^m - n_t}{p_{tl}(\Theta)}\right]\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} \\
&\quad +\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\left[\frac{n_t}{1-p_{tl}(\Theta)}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} + \sum_{i=1}^{l-1} n_{ti}\left[p_{ti}(\Theta)\right]^{-1}\frac{\partial p_{ti}(\Theta)}{\partial \theta_j}\right]\Bigg|_{\Theta=\hat{\Theta}}.
\end{aligned} \tag{3}$$

Since

$$\begin{aligned}
L_{2j} = \frac{\partial \log L_2}{\partial \theta_j} &= \sum_{t=1}^{s}\frac{n_t}{1-p_{tl}(\Theta)}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} + \sum_{t=1}^{s}\sum_{i=1}^{l-1} n_{ti}\left[p_{ti}(\Theta)\right]^{-1}\frac{\partial p_{ti}(\Theta)}{\partial \theta_j} \quad \text{and} \\
L_{1j} = \frac{\partial \log L_1}{\partial \theta_j} &= -\sum_{t=1}^{s}\frac{n_t}{1-p_{tl}(\Theta)}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j} + \sum_{t=1}^{s}(N_t^m - n_t)\left[p_{tl}(\Theta)\right]^{-1}\frac{\partial p_{tl}(\Theta)}{\partial \theta_j},
\end{aligned}$$

for

$$N_t^m = N_{ot}^m, \quad \hat{L}_{2j} = L_{2j}\Big|_{\Theta=\hat{\Theta}}, \quad \text{and} \quad \hat{L}_{1j} = L_{1j}\Big|_{\Theta=\hat{\Theta}, N_t^m=\hat{N}_t^m},$$

we have

$$\frac{1}{\sqrt{N_T^m}}\frac{\partial \log \mathcal{L}}{\partial \theta_j}\Big|_{\Theta=\hat{\Theta}} = -\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}(\hat{N}_t^m - N_{ot}^m)\,[p_{tl}(\Theta)]^{-1}\,\frac{\partial p_{tl}(\Theta)}{\partial \theta_j}\Big|_{\Theta=\hat{\Theta}} + \frac{1}{\sqrt{N_T^m}}\left(\hat{L}_{1j}+\hat{L}_{2j}\right). \quad (4)$$

Then, from (3) and (4), and denoting $\hat{L}_j = \hat{L}_{1j} + \hat{L}_{2j}$, and since $\sum_i \frac{\partial p_{ti}(\Theta)}{\partial \theta_j} = 0$ for $t = 1, \ldots, s$, we have

$$\begin{aligned} &\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\sum_{i=1}^{l}[p_{ti}(\Theta)]^{-1}\,(n_{ti} - N_{ot}^m p_{ti}^o)\frac{\partial p_{ti}(\Theta)}{\partial \theta_j}\Big|_{\Theta=\hat{\Theta}} \\ =\ &\sum_{t=1}^{s}\frac{N_{ot}^m}{\sqrt{N_T^m}}\sum_{i=1}^{l}[\hat{p}_{ti}]^{-1}\,(\hat{p}_{ti}-p_{ti}^o)\,\hat{p}_{ti,j} - \frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}(\hat{N}_t^m - N_{ot}^m)\,[\hat{p}_{tl}]^{-1}\,\hat{p}_{tl,j} + \frac{1}{\sqrt{N_T^m}}(\hat{L}_j). \end{aligned}$$

From the mean value theorem,

$$p_{ti}(\hat{\Theta}) - p_{ti}(\Theta_o) = \sum_{h=1}^{r}(\hat{\theta}_h - \theta_{oh})p_{ti,h}(\Theta^i), \quad \theta_h^i \in (\hat{\theta}_h, \theta_{oh}).$$

Thus,

$$\begin{aligned} &\sum_{t=1}^{s}\frac{N_{ot}^m}{\sqrt{N_T^m}}\sum_{i=1}^{l}[\hat{p}_{ti}]^{-1}\,\hat{p}_{ti,j}\sum_{h=1}^{r}(\hat{\theta}_h - \theta_{oh})p_{ti,h}(\Theta^i) + \frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}[\hat{p}_{tl}(\Theta)]^{-1}\,\hat{p}_{tl,j}(\hat{N}_t^m - N_{ot}^m) \\ =\ &\frac{1}{\sqrt{N_T^m}}\sum_{t=1}^{s}\sum_{i=1}^{l}[\hat{p}_{ti}]^{-1}\,\hat{p}_{ti,j}(n_{ti} - N_{ot}^m p_{ti}^o) - \frac{1}{\sqrt{N_T^m}}\hat{L}_j. \end{aligned}$$

Therefore,

$$\sum_{t=1}^{s}\frac{N_{ot}^m}{N_T^m}\sum_{h=1}^{r}b_{t,j,h}\sqrt{N_T^m}(\hat{\theta}_h - \theta_{oh}) + \sum_{t=1}^{s}\left[\frac{N_{ot}^m}{N_T^m}\right]^{1/2} b_{j,r+t}(N_{ot}^m)^{-1/2}(\hat{N}_t^m - N_{ot}^m) - Z_j \to_p 0\ , \quad (5)$$

where

$$b_{t,j,h} = \sum_{i=1}^{l}p_{ti,h}\,[\hat{p}_{ti}]^{-1}\,\hat{p}_{ti,j},$$

$$b_{j,r+t} = -\,[\hat{p}_{tl}]^{-1}\,\hat{p}_{tl,j}, \quad \text{and}$$

$$Z_j = \sum_{t=1}^{s}c_t^{1/2}\,(N_{ot}^m)^{-1/2}\sum_{i=1}^{l}[p_{ti}^o]^{-1}\,p_{ti,j}^o n_{ti}, \quad j = 1, \ldots, r.$$

This result is similar to the one given by Sanathanan (1972) in page 150. The term $(N_{ot}^m)^{-1/2}(\hat{N}_t^m - N_{ot}^m)$ in expression (5) can be written as

$$(N_{ot}^m)^{-1/2}(\hat{\psi}_t \hat{N}_t - \psi_{ot} N_{ot}) = (N_{ot}^m)^{-1/2}(\hat{N}_t(\hat{\psi}_t - \psi_{ot}) + \psi_{ot}(\hat{N}_t - N_{ot})).$$

Then (5) implies

$$\begin{aligned}
\mathbf{B}^*{}_t &= \sum_{t=1}^{s} \frac{N_{ot}^m}{N_T^m} \sum_{h=1}^{r} b_{t,j,h} (N_T^m)^{1/2}(\hat{\theta}_h - \theta_{oh}) \\
&+ \sum_{t=1}^{s} \left[\frac{N_{ot}^m}{N_T^m}\right]^{1/2} \hat{N}_t b_{j,r+t} (N_{ot}^m)^{-1/2}(\hat{\psi}_t - \psi_{ot}) \\
&+ \sum_{t=1}^{s} \left[\frac{N_{ot}^m}{N_T^m}\right]^{1/2} \psi_{ot} b_{j,r+t} (N_{ot}^m)^{-1/2}(\hat{N}_t - N_{ot}) - Z_j^* \rightarrow_p 0
\end{aligned}$$

Let $Z_j^* = Z_j$ and $b_{t,j,h}^* = b_{t,j,h}$, and consider $\lim_{N_T^m \to \infty} \frac{N_{ot}^m}{N_T^m} = c_t$, $t = 1, \ldots, s$. Thus, for $j = 1, \ldots, r$, $h = 1, \ldots, r$, $u = 1, \ldots, s$, and $t = 1, \ldots, s$,

$$\begin{aligned}
\sum_{t=1}^{s} c_t b_{t,j,h}^* &\rightarrow_{a.s.} \sigma^{j,h}, \\
c_t^{1/2} b_{j,r+t}^* &= c_t^{1/2} \psi_{ot} b_{j,r+t} \rightarrow_{a.s.} \psi_{ot} \sigma^{j,r+t}, \text{ and} \\
c_t^{1/2} b_{j,r+s+t}^* &= c_t^{1/2} \hat{N}_t b_{j,r+t} \rightarrow_{a.s.} N_t \sigma^{j,r+t}.
\end{aligned}$$

Now, considering the s population extension of condition (ii) of theorem 1 by Sanathanan (1972) (see p. 150), and applying similar arguments to the ones used to describe $\mathbf{B}^*{}_t$, let

$$\begin{aligned}
\mathbf{C}_t &= \sum_{h=1}^{r} \left[\frac{N_{ot}^m}{N_T^m}\right]^{1/2} b_{r+t,h} (N_T^m)^{1/2}(\hat{\theta}_h - \theta_{oh}) \\
&+ b_{r+t,r+t} \hat{N}_t (N_{ot}^m)^{-1/2}(\hat{\psi}_t - \psi_{ot}) \\
&+ b_{r+t,r+t} \psi_{ot} (N_{ot}^m)^{-1/2}(\hat{N}_t - N_{ot}) - Z_{r+t} \rightarrow_p 0 \ ,
\end{aligned}$$

where

$$b_{r+t,h} = -\frac{p_{tl,h}}{p_{tl}}, \quad b_{r+t,r+t} = \frac{1 - p_{tl}}{p_{tl}}$$

and

$$Z_{r+t} = [p_{tl}^o]^{-1} (N_{ot}^m)^{1/2}((n_t/N_{ot}^m) - (1 - p_{tl}^o)), \quad t = 1, \ldots, s.$$

Let $\mathbf{C}^*{}_t = \psi_{ot}\mathbf{C}_t$, then $Z^*_{r+t} = \psi_{to} Z_{r+t}, \quad t = 1, \ldots, s.$, and

$$\begin{aligned} c_t^{1/2} b^*_{r+t,h} &= c_t^{1/2} \psi_{ot} b_{r+t,h} \rightarrow_{a.s.} \psi_{ot}\sigma^{r+t,h}, \\ b^*_{r+t,r+t} &= \psi_{ot}^2 b_{r+t,r+t} \rightarrow_{a.s.} \psi_{ot}^2 \sigma^{r+t,r+t}, \text{ and} \\ b^*_{r+t,r+s+t} &= \psi_{ot}\hat{N}_t b_{r+t,r+t} \rightarrow_{a.s.} \psi_{ot} N_t \sigma^{r+t,r+t}. \end{aligned}$$

Now, from condition (iv), let

$$\begin{aligned} \mathbf{D}_t &= (N^m_{ot})^{-1/2}(\hat{\psi}_t \hat{N}_t (1-\hat{p}_{tl}) - n_t) \\ &= \hat{N}_t(1-\hat{p}_{tl})(N^m_{ot})^{-1/2}(\hat{\psi}_t - \psi_{ot}) + (1-\hat{p}_{tl})\psi_{ot}(N^m_{ot})^{-1/2}(\hat{N}_t - N_{ot}) \\ &\quad -\psi_{ot}N_{ot}(N^m_{ot})^{-1/2}(\hat{p}_{tl} - p^o_{tl}) - (N^m_{ot})^{-1/2}(n_t + p^o_{tl}\psi_{ot}N_{ot} - \psi_{ot}N_{ot}) \end{aligned}$$

Thus, using the mean value theorem,

$$\begin{aligned} \mathbf{D}_t &= \hat{N}_t(1-\hat{p}_{tl})(N^m_{ot})^{-1/2}(\hat{\psi}_t - \psi_{ot}) + (1-\hat{p}_{tl})\psi_{ot}(N^m_{ot})^{-1/2}(\hat{N}_t - N_{ot}) \\ &\quad -\psi_{ot}N_{ot}(N^m_{ot})^{-1/2}\sum_{h=1}^{r}(\hat{\theta}_h - \theta_{oh})p_{tl,h} - (N^m_{ot})^{-1/2}(n_t - \psi_{ot}N_{ot}(1-p^o_{tl})). \end{aligned}$$

Let $\mathbf{D}^*{}_t = (N_{ot}/p_{tl})\mathbf{D}_t$ and since $N^m_{ot} = \psi_{ot}N_{ot}$,

$$\begin{aligned} \mathbf{D}^*_t &= -\sum_{h=1}^{r}\left[\frac{N^m_{ot}}{N^m_T}\right]^{1/2} N_{ot}\left[\frac{p_{tl,h}}{p_{tl}}\right](N^m_T)^{1/2}(\hat{\theta}_h - \theta_{oh}) + N_{ot}\psi_{ot}\left[\frac{1-\hat{p}_{tl}}{p_{tl}}\right](N^m_{ot})^{-1/2}(\hat{N}_t - N_{ot}) \\ &\quad + N_{ot}\hat{N}_t\left[\frac{1-\hat{p}_{tl}}{p_{tl}}\right](N^m_{ot})^{-1/2}(\hat{\psi}_t - \psi_{ot}) - \frac{N_{ot}(N^m_{ot})^{-1/2}}{p_{tl}}(n_t - \psi_{ot}N_{ot}(1-p^o_{tl})) \rightarrow_p 0. \end{aligned}$$

Let

$$\begin{aligned} c_t^{1/2} b^*_{r+s+t,j} &= c_t^{1/2} N_{ot}\frac{p_{tl,j}}{p_{tl}} \rightarrow_{a.s.} N_{ot}\sigma^{r+t,j}, \\ b^*_{r+s+t,r+t} &= N_{ot}\psi_{ot}\frac{1-\hat{p}_{tl}}{p_{tl}} \rightarrow_{a.s.} N_{ot}\psi_{ot}\sigma^{r+t,r+t}, \text{ and} \\ b^*_{r+s+t,r+s+t} &= N_{ot}\hat{N}_t\frac{1-\hat{p}_{tl}}{p_{tl}} \rightarrow_{a.s.} N^2_{ot}\sigma^{r+t,r+t}, \quad \text{and} \\ Z^*_{r+s+t} &= \frac{N_{ot}(N^m_{ot})^{-1/2}}{p_{tl}}(n_t - \psi_{ot}N_{ot}(1-p^o_{tl})). \end{aligned}$$

Summarizing, let

$$
\begin{aligned}
Z_j^* &= \sum_{t=1}^{s} c_t^{1/2} (N_{ot}^m)^{-1/2} \sum_{i=1}^{l} [p_{ti}^o]^{-1}\, p_{ti,j}^o n_{ti}, \quad j = 1, \ldots, r, \\
Z_{r+t}^* &= \psi_{ot}\, [p_{tl}^o]^{-1}\, (N_{ot}^m)^{1/2} ((n_t / N_{ot}^m) - (1 - p_{tl}^o)), \quad t = 1, \ldots, s, \\
Z_{r+s+t}^* &= \frac{N_{ot} (N_{ot}^m)^{-1/2}}{p_{tl}} (n_t - \psi_{ot} N_{ot} (1 - p_{tl}^o)), \quad t = 1, \ldots, s.
\end{aligned}
$$

where $N_{ot} = N_{ot}^u + N_{ot}^m$, $\psi_{ot} = (N_{ot}^m / N_{ot})$, $n_t = \sum_{i=1}^{l-1} n_{ti}$, $p_{tl} = 1 - \sum_{i=1}^{l-1} p_{ti}$, and $N_T^m = \sum_{t=1}^{s} N_{ot}^m$. Let

$$
\begin{aligned}
U^t &= \Big((N_T^m)^{1/2} (\hat{\Theta} - \Theta_o), (N_{o1}^m)^{-1/2} (\hat{\psi}_1 - \psi_{o1}), \ldots, (N_{os}^m)^{-1/2} (\hat{\psi}_s - \psi_{os}), \\
&\quad (N_{o1}^m)^{-1/2} (\hat{N}_1 - N_{o1}), \ldots, (N_{os}^m)^{-1/2} (\hat{N}_s - N_{os}) \Big), \quad \text{and} \\
Z^{*t} &= \Big(z_1^*, \ldots, z_r^*, z_{r+1}^*, \ldots, z_{r+s}^*, z_{r+s+1}^*, \ldots, z_{r+2s}^* \Big)
\end{aligned}
$$

From expressions $\mathbf{B}^*{}_t$, $\mathbf{C}^*{}_t$, and $\mathbf{D}^*{}_t$ we have

$$
U - \bar{\Sigma} Z^* \to_p 0. \tag{6}
$$

Step II:

We will show next that $Z \to_d N(0, \bar{\Sigma}^{-1})$.

Following Sanathanan (1972) p. 147, for $t = 1, \ldots, s$, let $V_k^{(t)}$ be a random vector

$$
V_k^{(t)} = \Big(V_{k,1}^{(t)}, \ldots, V_{k,r}^{(t)}, V_{k,r+1}^{(t)}, \ldots, V_{k,r+s}^{(t)}, V_{k,r+s+1}^{(t)}, \ldots, V_{k,r+2s}^{(t)} \Big)
$$

such that:

(i) when the $k^{(t)}$th trial from population t ($t = 1, \ldots, s$) and $k^{(t)} = 1, \ldots, N_t^m$, results in the ith category, $i = 1, \ldots, l-1$, then, with probability p_{ti}

$$
\begin{aligned}
V_{k,j}^{(t)} &= (c_t)^{1/2}\, [p_{ti}^o]^{-1}\, p_{ti,j}^o, \quad j = 1, \ldots, r \\
V_{k,r+u}^{(t)} &= \begin{cases} \psi_{ot} & \text{if } u = t,\ u = 1, \ldots, s \\ 0 & \text{otherwise} \end{cases} \\
V_{k,r+s+v}^{(t)} &= \begin{cases} N_{ot} & \text{if } v = t,\ v = 1, \ldots, s \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

(ii) when the $k^{(t)}$th trial from population t $(t = 1, \dots, s)$ and $k^{(t)} = 1, \dots, N_t^m$, results in the lth category, then, with probability p_{tl}

$$\begin{aligned}
V_{k,j}^{(t)} &= (c_t)^{1/2} \left[p_{tl}^o\right]^{-1} p_{tl,j}^o, \quad j = 1, \dots, r \\
V_{k,r+u}^{(t)} &= \begin{cases} -\psi_{ot} \left[p_{tl}^o\right]^{-1} (1 - p_{tl}^o) & \text{if } u = t,\ u = 1, \dots, s \\ 0 & \text{otherwise} \end{cases} \\
V_{k,r+s+v}^{(t)} &= \begin{cases} -N_{ot} - \psi_{ot} \left[p_{tl}^o\right]^{-1} (1 - p_{tl}^o) & \text{if } v = t,\ v = 1, \dots, s \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Notice that

$$Z^{*t} = \sum_{t=1}^{s} (N_{ot}^m)^{-1/2} \sum_{k=1}^{N_{ot}^m} V_k^{(t)}.$$

Since $\sum_{i=1}^{l-1} p_{ti} = (1 - p_{tl})$ and $\sum_{i=1}^{l} p_{ti,j} = 0, \quad j = 1, \dots, r$, then $E\left(V_k^{(t)}\right) = 0$, and we also have Σ_t^{-1} given by

$$\begin{aligned}
Cov\left(V_{k,u}^{(t)}, V_{k,v}^{(t)}\right) &= (c_t)^{1/2} \sum_{i=1}^{l} [p_{ti}^o]^{-1} p_{ti,u}^o p_{ti,v}^o, \quad 1 \le u, v \le r \\
Cov\left(V_{k,j}^{(t)}, V_{k,r+u}^{(t)}\right) &= \begin{cases} -(c_t)^{1/2} \psi_{ot} \left[p_{tl}^o\right]^{-1} p_{tl,j}^o & \text{if } u = t,\ u = 1, \dots, s \\ 0 & \text{if u} \neq \text{t} \end{cases} \\
Cov\left(V_{k,j}^{(t)}, V_{k,r+s+u}^{(t)}\right) &= \begin{cases} -(c_t)^{1/2} N_{ot} \left[p_{tl}^o\right]^{-1} p_{tl,j}^o & \text{for } u = t,\ u = 1, \dots, s \\ 0 & \text{if u} \neq \text{t} \end{cases} \\
Cov\left(V_{k,r+v}^{(t)}, V_{k,r+u}^{(t)}\right) &= \begin{cases} \psi_{ot}^2 (1 - p_{tl}^o) [p_{tl}^o]^{-1} & \text{if } u = t = v,\ 1 \le v, u \le s \\ 0 & \text{otherwise} \end{cases} \\
Cov\left(V_{k,r+v}^{(t)}, V_{k,r+s+u}^{(t)}\right) &= \begin{cases} \psi_{ot} N_{ot} (1 - p_{tl}^o) [p_{tl}^o]^{-1} & \text{if } u = t = v,\ 1 \le v, u \le s \\ 0 & \text{otherwise} \end{cases} \\
Cov\left(V_{k,r+s+v}^{(t)}, V_{k,r+s+u}^{(t)}\right) &= \begin{cases} N_{ot}^2 (1 - p_{tl}^o) [p_{tl}^o]^{-1} & \text{if } u = t = v,\ 1 \le u, v \le s \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

For a given t, $t = 1, \dots, s$, by the Central Limit Theorem,

$$(N_{ot}^m)^{-1/2} \sum_{k=1}^{N_{ot}^m} V_k^{(t)} \to_d N(0, \Sigma_t^{-1})$$

Thus, since $V_k^{(t)}$ is independent from $V_k^{(u)}$ for $u \neq t$,

$$Z^{*t} = \sum_{t=1}^{s} (N_{ot}^m)^{-1/2} \sum_{k=1}^{N_{ot}^m} V_k^{(t)} \rightarrow_d N(0, \bar{\Sigma}^{-1})$$

where $\bar{\Sigma}^{-1} = \sum_{t=1}^{s} \Sigma_t^{-1}$. Thus, by (5) and (6),

$$U \rightarrow_d N(0, \bar{\Sigma}).$$

This completes the proof of the Theorem.

Now, since $\hat{N} = \sum_{t=1}^{s} \hat{N}_t$ and $\lim_{N_T^m \rightarrow \infty} \frac{N_{ot}^m}{N_T^m} = c_t$,

$$\begin{aligned} (N_T^m)^{-1/2}(\hat{N} - N_o) &= \sum_{t=1}^{s} \left(\frac{N_{ot}^m}{N_T^m} \right)^{1/2} (N_{ot}^m)^{1/2} (\hat{N}_t - N_{ot}) \\ &\rightarrow_d \sum_{t=1}^{s} (c_t)^{1/2} N(0, \sigma_t^2) \ =_d \ N(0, \sum_{t=1}^{s} c_t \sigma_t^2) \end{aligned}$$

where $\Sigma_t = (\sigma_t^{i,j})$, and $\sigma_t^2 = \sigma^{r+s+t, r+s+t}$, $\ t = 1, \ldots, s$.

Bibliography

da Silva, C.Q. (1999). Capture-recapture estimation of bowhead whale population size using photo-identification data. Ph.d. Dissertation, Universtiy of Washinghton, Dept. of Statistics.

da Silva, C.Q., Zeh, J., Madigan, D., Laake, J., Rugh, D., Baraff, L., Koski, W. and Miller, G. (2000). Capture-recapture estimation of bowhead whale population size using photo-identification data. *J. Cetacean Research Management* **2** 1, pp. 45-61.

Raftery, A.E., and Zeh, J.E. (1998). Estimating bowhead whale population size and rate of increase from the 1993 census. *Journal of the American Statistical Association* **93**, pp. 451-463.

Sanathanan, L. (1972). Estimating the size of a multinomial population. *The Annals of Mathematical Statistics* **43** 1, pp. 142-152.

Seber, G.A.F. (1982). Estimation of Animal Abundance, *2nd edition.* London: Griffin.

Williams, J.A., Dawson, S.M. and Slooten, E. (1993). Distribution of bottlenosed dolphins (*Tursiops truncatus*) in Doubtful Sound, New Zealand. *Can. J. Zool.* **71**, pp. 2080-2088.