Statistica Sinica **17**(2007), 1215-1236

# POPULATION SIZE ESTIMATION IN PROPORTIONAL TRAPPING REMOVAL AND RECAPTURE MODELS WITH KNOWN CAPTURE TIMES

Paul S. F. Yip[1], Liqun Xi[1] and Liping Liu[2]

[1] *The University of Hong Kong and* [2] *Peking University*

*Abstract:* Estimation methods are suggested to estimate the population size in proportional trapping removal and recapture models when the capture times are observed. With this additional information, the maximum likelihood estimate and the optimal martingale estimation of population size are derived. The ill-conditioning problem in Good's estimator is avoided. The asymptotic properties of the proposed estimators are obtained. The effect of allowing for trap availability and the role of capture times are examined through a simulation study. The asymptotic efficiency for the estimates are compared, and the methods are applied to a data set.

*Key words and phrases:* Capture times, proportional trapping model, recapture experiment, removal experiment.

## 1. Introduction

Capture-recapture and removal methods are often used in ecology to estimate population size (Otis, Burnham, White and Anderson (1978)). Usually, traps are set up to capture animals (Moran (1951), Darroch (1958) and Good, Lewis, Gaskins and Howell (1979)). One of the potential problems of using a limited number of traps is that once an animal is trapped, the trap cannot catch other animals until the trap is cleared. There is always a delay in clearing the traps. Hence, the capture intensity not only depends on the animal abundance but also on the availability of traps over the course of the experiment. Ignoring the availability of traps leads to a biased estimate of population size (Good et al. (1979) and Liu and Yip (2003)). However, in papers on the subject, this issue is seldom considered and nearly all have made the implicit assumption of infinite trap availability.

We consider a removal experiment in discrete time with $K$ trapping periods. For the $i$th trapping period, $m_i$ traps are available in the beginning and $n_i$ animals are trapped and removed ($i = 1, \ldots, K$). The capture density that a specific animal is trapped at a given time is assumed to be proportional to the number of unoccupied traps. This assumption is called proportional trapping of animals and an exact formula for the likelihood is given in Good et al. (1979). However,

the formula is illed-conditioned and requires extremely high precision for its numerical evaluation. No algorithm is available for searching for the maximum of the likelihood function.

In a continuous-time proportional trapping model with known capture times, Liu and Yip (2003) proposed a martingale estimate of population size which avoids ill-conditioning difficulties. The availability of the capture times in a recapture and removal experiment is quite feasible: for example, a timer can be installed in each trap, so that the capture time is recorded automatically when an animal is trapped. By the end of each trapping period, when the trapped animals are cleared from the traps, the capture times are recorded and the timers are reset for the next trapping period. Here we consider a proportional trapping model with the availability of capture times as in Liu and Yip (2003).

In Section 2, the maximum likelihood estimate for the population size is obtained. This is also the optimal martingale estimate. The asymptotic properties of the estimator are given, and simulation studies are conducted to examine its performance. It is shown that the ill-conditioning inherent in the method of Good et al. (1979) is avoided. We also reanalyse the data set on *Rattus* in Good et al. (1979). In Section 3, we extend the experimental setting from removal to recapture, i.e., at the end of each trapping period the trapped animals are marked, released and thus subject to recapture in the next trapping period. Under the two common cases, with or without behavioral response (due to trapping), we obtain maximum likelihood estimates, as well as their asymptotic properties. Comparison is made between the two designs, removal and recapture, by simulation and asymptotic efficiency comparison. We also compare the proposed estimators with the discrete type estimators in the recapture settings. In Section 4, we allow for time-dependent baseline capture hazards that may be different for different trapping periods. Optimal martingale estimators for the population size are obtained. Simulation studies, asymptotic properties and some comparisons of the proposed estimates are given. A conclusion is given in Section 5.

## 2. Removal experiment in proportional trapping model

We deal with the estimation problem in the removal experiment as in Good et al. (1979), i.e. trapped animals are removed from the area, at least over the experimental period. Let $\nu$ denote the unknown population size, the parameter of interest. The capture times in all the trapping periods are assumed known.

### 2.1. Notations and estimation procedure

Let $\tau_i$ denote the duration of the $i$th trapping period and $N_i(t)$ denote the number of trapped animals in $[0, t]$ in the $i$th trapping period ($i = 1, 2 \ldots, K$). So $T_i(t) = m_i - N_i(t-)$ is the number of available traps and $\nu - R_i(t)$ is the

number of untrapped animals just before time $t$, where $R_i(t) = R_i(0) + N_i(t-)$ and $R_i(0) = \sum_{j=1}^{i-1} N_j(\tau_j)$ (or $\sum_{j=1}^{i-1} n_j$) denotes the total number of trapped and removed animals prior to the start of the $i$th trapping period. We define $R_1(0) = 0$. The likelihood function for all trapping occasions is then given by

$$L \propto \prod_{i=1}^{K} \Big\{ \prod_{t\in[0,\tau_i]} [\alpha T_i(t)(\nu - R_i(t))]^{dN_i(t)} \exp\Big[ - \int_0^{\tau_i} \alpha T_i(t)(\nu - R_i(t))dt \Big] \Big\},$$

where $\alpha$ is a constant baseline capture hazard that can be interpreted as hazard rate being captured if there is only one trap and only one animal. For $K = 1$, this reduces to the model of Liu and Yip (2003). Taking derivatives of the log-likelihood function with respect to $\nu$ and $\alpha$, and equating them to zero, we obtain

$$\frac{\partial \log L}{\partial \nu} = \sum_{i=1}^{K} \Big( \int_0^{\tau_i} \frac{dN_i(t)}{\nu - R_i(t)} - \int_0^{\tau_i} \alpha T_i(t)dt \Big) = 0, \tag{1}$$

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^{K} \frac{N_i(\tau_i)}{\alpha} - \sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)(\nu - R_i(t))dt = 0. \tag{2}$$

From (2), we have

$$\hat{\alpha} = \frac{\sum_{i=1}^{K} N_i(\tau_i)}{\sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)(\hat{\nu} - R_i(t))dt}, \tag{3}$$

and by substituting (3) into (1), we obtain an equation for the MLE of $\nu$ as

$$\sum_{j=1}^{Z} \frac{1}{\nu - j + 1} = \frac{Z}{\nu - C}, \tag{4}$$

where $Z = \sum_{i=1}^{K} N_i(\tau_i)$, and $C = (\sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)R_i(t)dt)/(\sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)dt)$. From (4), the MLE of $\nu$ can be readily obtained numerically. The inverse of the observed information matrix can be used to estimate the variance-covariance matrix for the unknown parameters. In addition, the necessary and sufficient condition for the existence and uniqueness of the solution of (4) is:

$$\frac{1}{2}(Z-1) < C \le \Big(1 - \frac{1}{h_Z}\Big) Z, \tag{5}$$

where $h_Z = \sum_{i=1}^{Z} 1/i$. This necessary and sufficient condition is a direct result of Theorem 1 in Liu and Yip (2003).

Using martingale estimation equations with optimal weights (see Godambe (1985)), the estimating equations (1) and (2) can also be obtained, a detailed derivation can be found in Yip, Xi and Liu (2005).

As $\nu \to \infty$, if $m_i$ remains unchanged, then all traps will soon be occupied. Therefore, when considering the asymptotic properties of the proposed estimator, a reasonable assumption is that $m_i/\nu \to r_i$, $i = 1, 2, \ldots, K$, as $\nu \to \infty$. The following theorem describes the asymptotic properties of the proposed estimator.

**Theorem 1.** *Suppose $\nu \to \infty$ and $m_i/\nu \to r_i$. Let $\hat{\nu}$ and $\hat{\alpha}$ denote the maximum likelihood estimates of $\nu$ and $\alpha$, respectively. If $G = \sum_{i=1}^{K} r_i < 1$, then $\hat{\zeta} = (\nu^{-1/2}(\hat{\nu} - \nu),\ \nu^{1/2}(\hat{\alpha} - \alpha))' \xrightarrow{\mathrm{d}} \mathbf{N}_2(\mathbf{0}, \mathbf{\Omega}_1^{-1})$, where*

$$\mathbf{\Omega}_1 = \begin{bmatrix} \dfrac{G}{1-G} & \dfrac{1}{\alpha}\log\dfrac{1}{1-G} \\ \dfrac{1}{\alpha}\log\dfrac{1}{1-G} & \dfrac{G}{\alpha^2} \end{bmatrix}.$$

*If $G \geq 1$, then $\hat{\nu}/\nu \xrightarrow{\mathrm{p}} 1$ with a speed faster than $\nu^{-1/2}$.*

**Proof.** See Appendix 1 in Yip et al. (2005).

### 2.2. Simulation study

Simulation studies were conducted to examine the performance of the proposed estimator. Several combinations of parameter values were used in the simulations. Without loss of generality, all $\tau_i$ were fixed at 1 and all $m_i$ were equal to $m$. In this paper, all simulation results were based on 2,000 repetitions. Let $P_T$ (%) denote the average total removal proportion by the end of experiment, and $pf$ (%) denote the proportion of failures among 2,000 repetitions. A failure means that no solution is obtained from the estimating equations, i.e., (5) is not satisfied. Let $CP$ (%) denote the coverage proportion of the 95% confidence intervals for $\nu$ which are calculated using the log transformation presented in Chao (1987). When a failure occurs, we say that $\nu$ lies outside the 95% confidence interval for $CP$; while the average values are based only on the successful repetitions. Some simulation results are listed in Table 1 where we denote the proposed estimator as $\hat{\nu}^{(\mathrm{rm})}$.

As the number of traps or the number of trapping periods increase, the performance of the proposed estimator is improved. The bias for $\hat{\nu}$ is small, and it diminishes as $P_T$ increases. There were very few failures. Some failures appeared when $P_T$ was less than about 60% and 50% for $\nu = 200$ and 500, respectively. Overall, the coefficient of variation (CV) (i.e $\mathrm{se}(\hat{\nu})/\hat{\nu}$) mainly depends on the value of $P_T$. In order for the CV to be less than 10% the capture proportion $P_T$ need to be increased to about 78% and 70% for $\nu = 200$ and 500, respectively. Although $\hat{\nu}/\nu \xrightarrow{\mathrm{p}} 1$ and $\mathrm{se}(\hat{\nu})/\nu \to 0$ as $\nu \to \infty$ (from Theorem 1), the simulation studies show that, even for a fairly large $\nu$, a high removal proportion is needed to obtain a reliable estimate.

Table 1. Simulation results for the proposed estimator ($\hat{\nu}^{(\mathrm{rm})}$) for removal experiments. with known capture times.

| $\nu = 200,\ \alpha = 0.0025$ | | | | | |
|---|---|---|---|---|---|
| $m$ | $K$ | $P_T$ (%) | $\hat{\nu}^{(\mathrm{rm})}$ | $CP(\%)$ | $pf$ (%) |
| 100 | 5 | 65.8 | 207.3 (53.6, 44.9) | 95.0 | 0.0 |
| 100 | 6 | 72.9 | 202.4 (26.9, 26.5) | 95.9 | 0.0 |
| 100 | 7 | 78.6 | 201.0 (20.1, 18.8) | 94.1 | 0.0 |
| 150 | 5 | 80.9 | 200.1 (16.4, 15.9) | 94.5 | 0.0 |
| 200 | 5 | 89.5 | 199.5 ( 8.3, 8.4) | 95.5 | 0.0 |
| $\nu = 500,\ \alpha = 0.001$ | | | | | |
| $m$ | $K$ | $P_T$ (%) | $\hat{\nu}^{(\mathrm{rm})}$ | $CP(\%)$ | $pf(\%)$ |
| 200 | 5 | 57.1 | 516.0 (114.1, 98.8) | 95.2 | 0.0 |
| 200 | 6 | 64.2 | 504.2 ( 63.7, 62.1) | 95.5 | 0.0 |
| 200 | 7 | 70.3 | 503.2 ( 46.1, 45.4) | 95.6 | 0.0 |
| 300 | 5 | 72.8 | 501.7 ( 39.2, 39.2) | 95.8 | 0.0 |
| 400 | 5 | 83.0 | 500.2 ( 21.8, 21.6) | 95.1 | 0.0 |

The first number in the parentheses is the empirical standard deviation and the second is the average of the estimated standard errors.

In the case of the same duration and same number of traps for all the trapping periods, if the availability of traps is ignored, the capture probabilities for the different trapping periods could be regarded as the same. Here, we apply the MLE estimate without taking the effect of trap into account, i.e., Moran's estimator, $\hat{\nu}_{\mathrm{M}}$ (Moran (1951)). It is to be expected that $\hat{\nu}_{\mathrm{M}}$ would produce a positively biased result since the capture probability for each trapping period would increase since there are the same number of traps but fewer available animals for a later trapping period. Table 2 shows that in the case of a large number of traps and a small $\alpha$, the improvement of using the proportional trapping model is marginal, the bias of using $\hat{\nu}_{\mathrm{M}}$ is not serious (see the first row of Table 2). But as the number of traps decreases and $\alpha$ increases, the effect of traps availability becomes more important. The $\hat{\nu}_{\mathrm{M}}$ which assumes that capture probability remains unchanged produces a positively biased estimate.

Table 2. Simulation results for Moran's estimator ($\hat{\nu}_{\mathrm{M}}$) and the proposed estimate ($\hat{\nu}^{(\mathrm{rm})}$) for the proportional trapping removal model.

| $\nu = 500,\ K = 6,\ \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6 = 1$ | | | | | | |
|---|---|---|---|---|---|---|
| $m$ | $\alpha$ | $P_T$ | $\hat{\nu}^{(\mathrm{rm})}$ | RMSE | $\hat{\nu}_{\mathrm{M}}$ | RMSE |
| 500 | 0.0005 | 75.4 | 500.8 (35.3, 34.0) | 35.3 | 518.5 (39.0, 40.1) | 43.2 |
| 300 | 0.0009 | 76.2 | 501.7 (33.5, 32.8) | 33.5 | 533.7 (39.4, 43.4) | 51.8 |
| 200 | 0.0015 | 77.2 | 501.5 (32.1, 30.9) | 32.1 | 556.4 (42.1, 48.9) | 70.4 |
| 100 | 0.0040 | 76.7 | 502.5 (32.4, 32.0) | 32.5 | 697.7 (66.3, 108.7) | 208.5 |

RMSE: the sample root mean squared error.

Using the capture times only, and without incorporating the information on the availability of traps into the model, gives the counting process $N_i(t)$ with the intensity $\lambda(\nu - R_i(t))$. Simulation results (not reported here) showed that the resulting estimate is strongly negatively biased, due to the decreasing intensity $\lambda(t) = \alpha(m_i - N_i(t-))$.

### 2.3. The *Rattus* example

Good et al. (1979) considered the trap effect but their resulting estimator requires high precision for its numerical evaluation, and no algorithm is available for searching for the maximum of the likelihood function. For this reason, no numerical simulation was given in Good et al. (1979), and only an example was considered. In an area of 22.5 acres in Freetown, Sierra Leone, *Rattus* were captured by traps on 18 successive occasions over a period of 6 weeks, 210 traps being set on each occasion. The trapped animals were not released after being cleared from traps at the end of each trapping occasion. The numbers of captures for each occasion are given in Table 3.

Table 3. Data on the number of *Rattus* trapped on eighteen occasions in Good et al. (1979).

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n_i$ | 49 | 32 | 31 | 34 | 16 | 33 | 22 | 27 | 17 | 19 | 18 | 16 | 18 | 12 | 14 | 12 | 17 | 7 |

In applying Good's estimator to this data set, the arithmetic needs 52 significant digits; see Good et al. (1979). Since the double precision of FORTRAN only provides 16 significant digits, a complex codes was needed. By assuming all the durations of trapping occasions (or trapping periods) were equal to 1, they calculated the values of the log-likelihood for a wide range of possible $(\nu, \alpha)$ and found, for $\nu = 505$ and $\alpha = 0.000422$, the corresponding value of the log-likelihood is the largest among those values considered. The corresponding estimated standard errors for $\hat{\nu}$ and $\hat{\alpha}$ were 29 and 0.000052. Here, we make a comparison with the proposed estimator. Since no capture times were available in the original data set, we simulated capture times by setting the values of $\nu$ and $\alpha$. We set $\nu$ to be 485, 495, 505, 515, 525 respectively, and $\alpha \times 10^4$ to be 3.0, 3.5, 4.0, 4.5, 5.0 respectively. For each combination of the $\nu$ and $\alpha$, with $K = 18$, $m = 210$ and $\tau_i = 1$ (for all $i$), the capture times based on a time-dependent Poisson process depending on the number of capture and traps available over the course of the experiment were simulated. For each given $(\nu, \alpha)$, 100 such data sets were generated with the same number of trapped numbers for each occasion, as in Table 3. The average of $\hat{\nu}$, $\hat{\alpha}$ and the corresponding estimated standard errors were obtained (based on these 100 data sets) and are listed in Table 4. We also list the simulated probability that, given $(\nu, \alpha)$, the trapped numbers in

Table 3 occur once. This simulated probability is the reciprocal of the average repetition numbers to obtain such data once.

For different simulated capture times with the different values of $(\nu, \alpha)$, Table 4 shows that $\hat{\nu}$ varies from 500 to 510, the average values of $\hat{\nu}$, $\hat{\alpha}$, se$(\hat{\nu})$ and se$(\hat{\alpha})$ are almost equal to the estimated values obtained by Good et al. (1979), and all the standard deviations are quite small. From the likelihood, the capture times do provide information for estimating $\nu$. Through computing the $C$ in (4) and analysing its form, it was found that with known values of $m_i$ and $\tau_i$, given $n_i$ $(i = 1, \ldots, K)$, $C$ varies not much for different $n_i$ capture times within $\tau_i$. Therefore, given $m_i$, $\tau_i$ and $n_i$ $(i = 1, \ldots, K)$, the capture times provide less information. However, with the additional information of capture times, the ill-conditioning difficulty in Good's estimator is avoided, the estimate is more stable, and easily obtained.

Since the results for various $\nu$ and $\alpha$ values in Table 4 are very close, we turn to the simulated probabilities. At the point (515, 0.0004), the simulated probability is the largest among those listed. This suggests that the true value of $(\nu, \alpha)$ is around (515, 0.0004). Further extensive simulation studies around this point (not reported here) show that it is likely that $\nu$ is between 505 and 515, and $\alpha$ is between 0.0004 and 0.0045. Moran's estimator gives $\hat{\nu}_{\mathrm{M}} = 525$ (se $= 33$) and, as analyzed in the last subsection, it is expected to be positively biased.

## 3. Recapture in proportional trapping model

We extend the proportional trapping model to a recapture setting, i.e., the trapped animals are marked, released and subject to recapture again in the next trapping period. Recapture experiments usually provide more information than that of removal experiments (Lloyd, Yip and Chan (1998)). However, the results are different in the proportional trapping model.

### 3.1. Estimation procedure

In the $i$th trapping period, let $N_i(t)$ be the number of trapped animals during $[0, t]$ $(0 < t \leq \tau_i)$, $N_{i0}(t)$ be the number of trapped animals that were previously untrapped, and $N_{i1}(t)$ be the number of trapped animals that had been previously trapped and marked $(1 \leq i \leq K)$. For $i = 1$, $N_{10}(t) = N_1(t)$ and $N_{11}(t) = 0$. Obviously, $N_i(t) = N_{i0}(t) + N_{i1}(t)$. The number of available traps at time $t$ in the $i$th period is $T_i(t) = m_i - N_i(t-)$. For $i = 2, \ldots, K$, let $M_i(0) = \sum_{j=1}^{i-1} N_{j0}(\tau_j)$, the number of previously trapped and marked animals at the start of the $i$th trapping period. We define $M_1(0) = 0$ so that $M_i(t) = M_i(0) + N_{i0}(t-)$ is the number of trapped and marked animals at time $t$ of the $i$th period. Then the intensity processes of $N_{i0}(t)$ and $N_{i1}(t)$ are $\alpha T_i(t)(\nu - M_i(t))$ and $\beta T_i(t) L_i(t)$ respectively, where $L_i(t) = M_i(0) - N_{i1}(t-)$, and $\beta$ is the baseline capture hazard for the previously captured animals.

Table 4. Estimates for the *Rattus* data using the proposed estimator in which the capture times were simulated by a combination of different values of $\nu$ and $\alpha$.

| | $\nu = 485$ | $\nu = 495$ | $\nu = 505$ | $\nu = 515$ | $\nu = 525$ |
|---|---|---|---|---|---|
| $\alpha \times 10^4$ = 3.0 | * | * | * | * | 506.4 (2.41) 4.20 (0.05)<br>29.4 (0.68) 0.52 (0.00)<br>$1.0\times10^{-27}$ |
| 3.5 | 506.3 (1.96) 4.20 (0.04)<br>29.3 (0.55) 0.52 (0.00)<br>$1.2\times10^{-27}$ | 506.2 (2.40) 4.20 (0.05)<br>29.3 (0.68) 0.52 (0.00)<br>$1.9\times10^{-27}$ | 506.1 (1.94) 4.20 (0.04)<br>29.3 (0.54) 0.52 (0.00)<br>$2.9\times10^{-26}$ | 506.2 (2.06) 4.20 (0.05)<br>29.3 (0.58) 0.52 (0.00)<br>$1.8\times10^{-24}$ | 506.0 (2.31) 4.21 (0.05)<br>29.2 (0.65) 0.52 (0.00)<br>$1.1\times10^{-23}$ |
| 4.0 | 505.9 (2.21) 4.21 (0.05)<br>29.2 (0.62) 0.52 (0.00)<br>$1.4\times10^{-24}$ | 505.8 (2.04) 4.21 (0.05)<br>29.2 (0.57) 0.52 (0.00)<br>$2.6\times10^{-24}$ | 506.2 (2.36) 4.20 (0.05)<br>29.3 (0.66) 0.52 (0.00)<br>★ $2.6\times10^{-22}$ | 506.0 (1.88) 4.21 (0.04)<br>29.3 (0.53) 0.52 (0.00)<br>★ $4.5\times10^{-22}$ | 505.8 (2.05) 4.21 (0.05)<br>29.2 (0.58) 0.52 (0.00)<br>★ $2.2\times10^{-22}$ |
| 4.5 | 505.6 (1.98) 4.22 (0.04)<br>29.1 (0.55) 0.52 (0.00)<br>$1.2\times10^{-23}$ | 505.4 (2.11) 4.22 (0.05)<br>29.1 (0.59) 0.52 (0.00)<br>$4.2\times10^{-23}$ | 505.6 (1.96) 4.22 (0.04)<br>29.1 (0.55) 0.52 (0.00)<br>★ $3.0\times10^{-22}$ | 505.6 (2.28) 4.22 (0.05)<br>29.1 (0.64) 0.52 (0.00)<br>$3.8\times10^{-23}$ | 505.3 (2.15) 4.22 (0.05)<br>29.0 (0.60) 0.52 (0.00)<br>$2.8\times10^{-24}$ |
| 5.0 | 505.4 (1.96) 4.22 (0.04)<br>29.1 (0.55) 0.52 (0.00)<br>$1.5\times10^{-24}$ | 505.4 (2.03) 4.22 (0.05)<br>29.1 (0.57) 0.52 (0.00)<br>$3.8\times10^{-23}$ | 505.0 (2.05) 4.23 (0.05)<br>29.0 (0.57) 0.52 (0.00)<br>$2.2\times10^{-24}$ | 505.2 (2.22) 4.23 (0.05)<br>29.0 (0.62) 0.52 (0.00)<br>$6.3\times10^{-26}$ | 505.1 (2.40) 4.23 (0.05)<br>29.0 (0.67) 0.52 (0.00)<br>$1.1\times10^{-27}$ |

Note: A ( ) C ( ) denotes the respective positions in the cell, where
B ( ) D ( )
E

A: Average of $\hat{\nu}$ based on the 100 data sets.
B: Average of the estimated standard error of $\hat{\nu}$.
C: Average of $\hat{\alpha} \times 10^4$ based on the 100 data sets.
D: Average of the estimated standard error of $\hat{\alpha} \times 10^4$.
E: The simulated probability that, given $(\nu, \alpha)$, the trapped numbers in Table 3 occur once.
★ : when E$> 10^{-22}$
$*$ : the simulated probability is smaller than $1.0\times10^{-28}$, almost impossible to simulate data as in Table 3.

All the values in the parentheses are the corresponding standard deviation for 100 estimates of values.

The likelihood function is given by

$$L=\prod_{i=1}^{K}\Bigg\{\prod_{t\in[0,\tau_i]}[\alpha T_i(t)(\nu-M_i(t))]^{dN_{i0}(t)}\exp\Big(-\int_0^{\tau_i}\alpha T_i(t)(\nu-M_i(t))dt\Big)$$
$$\times\prod_{t\in[0,\tau_i]}[\beta T_i(t)L_i(t)]^{dN_{i1}(t)}\exp\Big(-\int_0^{\tau_i}\beta T_i(t)L_i(t)dt\Big)\Bigg\}. \tag{6}$$

In case there is no behavioral response to trapping for the previously trapped animals, $\beta=\alpha$, and we only have two parameters $\nu$ and $\alpha$. The MLE estimates of $\nu$, $\alpha$ are obtained as the solution of the equations,

$$\frac{\partial\log L}{\partial\nu}=\sum_{i=1}^{K}\Big(\int_0^{\tau_i}\frac{dN_{i0}(t)}{\nu-M_i(t)}-\int_0^{\tau_i}\alpha T_i(t)dt\Big)=0, \tag{7}$$

$$\frac{\partial\log L}{\partial\alpha}=\sum_{i=1}^{K}\Big(\frac{N_i(\tau_i)}{\alpha}-\int_0^{\tau_i}T_i(t)(\nu-N_i(t-))dt\Big)=0. \tag{8}$$

From (8), we have

$$\hat\alpha=\frac{\sum_{i=1}^{K}N_i(\tau_i)}{\sum_{i=1}^{K}\int_0^{\tau_i}T_i(t)(\nu-N_i(t-))dt}. \tag{9}$$

Putting (9) into (7), the MLE $\hat\nu$ can be obtained numerically. The inverse of the observed information matrix can be used to estimate the variance-covariance matrix for the unknown parameters.

Using martingale estimation equations with optimal weights, the estimating equations (7) and (8) can also be obtained. The following theorem describes the asymptotic properties of the proposed estimator.

**Theorem 2.** *Suppose* $\nu\to\infty$, $m_i/\nu\to r_i$ *and* $g=\max(r_1,r_2,\ldots,r_K)$. *Let* $\hat\nu$ *and* $\hat\alpha$ *denote the maximum likelihood estimates of* $\nu$ *and* $\alpha$ *in the case of no behavioral response to trapping. If* $g<1$, *then* $\hat\zeta=(\nu^{-1/2}(\hat\nu-\nu),\nu^{1/2}(\hat\alpha-\alpha))'\xrightarrow{\mathrm{d}}\mathrm{N}_2(\mathbf{0},\mathbf{\Omega}_2^{-1})$, *where*

$$\mathbf{\Omega}_2=\begin{bmatrix}\dfrac{1-W}{W} & \frac{1}{\alpha}\log\frac{1}{W}\\ \dfrac{1}{\alpha}\log\dfrac{1}{W} & \frac{G}{\alpha^2}\end{bmatrix}, \tag{10}$$

*with* $W = \prod_{i=1}^{K}(1 - r_i)$ *and* $G = \sum_{i=1}^{K} r_i$. *If* $g \geq 1$, *then* $\hat{\nu}/\nu \xrightarrow{\text{p}} 1$ *with a speed faster than* $\nu^{-1/2}$.

**Proof.** See Appendix 2 in Yip et al. (2005).

In the presence of behavioral response for trapped animals, the unknown parameters are $\nu$, $\alpha$ and $\beta$. The corresponding likelihood equations are

$$\frac{\partial \log L}{\partial \nu} = \sum_{i=1}^{K} \Big( \int_0^{\tau_i} \frac{dN_{i0}(t)}{\nu - M_i(t)} - \int_0^{\tau_i} \alpha T_i(t) dt \Big) = 0, \tag{11}$$

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^{K} \frac{N_{i0}(\tau_i)}{\alpha} - \sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)(\nu - M_i(t)) dt = 0, \tag{12}$$

$$\frac{\partial \log L}{\partial \beta} = \sum_{i=2}^{K} \frac{N_{i1}(\tau_i)}{\beta} - \sum_{i=2}^{K} \int_0^{\tau_i} T_i(t) L_i(t) dt = 0. \tag{13}$$

Note that $\nu$ and $\alpha$ appear in (11) and (12) only; $\beta$ and $N_{i1}(t)$ appear in (13) only. This means that, with an unknown behavioral response, recaptures do not provide any information for estimating $\nu$. We only need to solve (11) and (12) to obtain $\hat{\nu}$. From (12), we have

$$\hat{\alpha} = \frac{\displaystyle\sum_{i=1}^{K} N_{i0}(\tau_i)}{\displaystyle\sum_{i=1}^{K} \int_0^{\tau_i} T_i(t)(\nu - M_i(t)) dt}. \tag{14}$$

Substituting (14) into (11), the equation for obtaining the MLE of $\nu$ is

$$\sum_{j=1}^{Z} \frac{1}{\nu - j + 1} = \frac{Z}{\nu - D}, \tag{15}$$

where $Z = \sum_{i=1}^{K} N_{i0}(\tau_i)$ and

$$D = \frac{\displaystyle\sum_{i=1}^{K} \int_0^{\tau_i} M_i(t)(m_i - N_i(t-)) dt}{\displaystyle\sum_{i=1}^{K} \int_0^{\tau_i} (m_i - N_i(t-)) dt}.$$

From (15), the MLE of $\nu$ can be obtained numerically. The inverse of the observed information matrix can be used to estimate the variance-covariance matrix for the

unknown parameters. Following Liu and Yip (2003), the necessary and sufficient condition for the existence and uniqueness of the solution of (15) is that

$$\frac{1}{2}(Z-1) < D \le \Big(1 - \frac{1}{h_Z}\Big) Z,$$

where $h_Z = \sum_{i=1}^{Z} 1/i$.

Using martingale estimation equations with optimal weights, the estimating equations (11), (12) and (13) can also be obtained. The following theorem describes the asymptotic properties of the proposed estimator.

**Theorem 3.** *Suppose $\nu \to \infty$, $m_i/\nu \to r_i$ and $g = \max(r_1, r_2, \ldots, r_K)$. Let $\hat{\nu}$ and $\hat{\alpha}$ denote the maximum likelihood estimates of $\nu$ and $\alpha$ in the case of an unknown behavioral response to trapping. If $g < 1$, then $\hat{\zeta} = (\nu^{-1/2}(\hat{\nu} - \nu), \nu^{1/2}(\hat{\alpha} - \alpha))' \xrightarrow{\mathrm{d}} \mathrm{N}_2(\mathbf{0}, \mathbf{\Omega}_3^{-1})$, where*

$$\mathbf{\Omega}_3 = \begin{bmatrix} \dfrac{\sum_{i=1}^{K} a_i}{1 - \sum_{i=1}^{K} a_i} & \dfrac{1}{\alpha} \log\left(\dfrac{1}{1 - \sum_{i=1}^{K} a_i}\right) \\ \dfrac{1}{\alpha} \log\left(\dfrac{1}{1 - \sum_{i=1}^{K} a_i}\right) & \dfrac{\sum_{i=1}^{K} a_i}{\alpha^2} \end{bmatrix},$$

*$a_i$ is the solution of $(1 - [a_i/(1 - b_i)])^{\beta} = (1 - [(r_i - a_i)/b_i])^{\alpha}$, and $b_i = \sum_{j=1}^{i-1} a_j$, $a_1 = r_1$. If $g \ge 1$, then $\hat{\nu}/\nu \xrightarrow{\mathrm{p}} 1$ with a speed faster than $\nu^{-1/2}$.*

**Proof.** See Appendix 2 in Yip et al. (2005).

### 3.2. Simulation and comparison

Simulations were conducted to compare the removal setting and recapture settings. All $\tau_i$ are fixed at 1, and the number of traps $m$ is kept the same for each trapping occasion. We denote the proposed estimators of $\nu$ as $\hat{\nu}^{(\mathrm{rc1})}$ and $\hat{\nu}^{(\mathrm{rc2})}$, respectively, for the recapture settings without or with behavioral response. Some results are listed in Table 5.

It is well known that a capture-recapture setting is always better or at least not worse than the corresponding capture-removal setting. For the proportional trapping experiments, because of the limited number of traps, recapture means fewer new captures compared with the corresponding removal setting. The situation becomes tricky and there is a gain-and-loss relationship between the two settings. Our aim is to study the effect on estimating population size of using the information of trap availability and the role of the capture times. Some observations and explanations are given as follows.

- With a significant behavioral response to trapping in the recapture setting, the effect of the recapture setting is worse than that of the removal setting.

Table 5. Simulation results for comparison of the removal and recapture settings with or without behavioral response.

$\nu = 500$, $m = 200$, $\alpha = 0.001$

| Settings | | $K = 6$ | | | | $P_T$ | $K = 10$ | | | | $P_T$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Recapture | | av($\hat{\nu}$) | RMSE | $CP$ | $pf$ | | av($\hat{\nu}$) | RMSE | $CP$ | $pf$ | |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 760.8 | 269.8 | 0.1 | 0.0 | | 670.5 | 173.1 | 0.0 | 0.0 | |
| $\beta = 0.5\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 506.1 | 72.1 | 95.5 | 0.0 | 62.8 | 501.1 | 24.0 | 95.4 | 0.0 | 81.4 |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 373.0 | 128.0 | 0.0 | 0.0 | | 416.9 | 83.7 | 0.0 | 0.0 | |
| $\beta = 2\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 513.3 | 94.8 | 94.4 | 0.0 | 59.6 | 501.0 | 30.1 | 94.7 | 0.0 | 76.9 |
| | $\hat{\nu}^{(\mathcal{M}_0)}$ | 500.5 | 30.0 | 94.8 | 0.0 | | 499.2 | 15.4 | 94.9 | 0.0 | |
| $\beta = \alpha$ | $\hat{\nu}^{(\mathrm{rc1})}$ | 500.5 | 29.8 | 95.2 | 0.0 | 61.6 | 499.2 | 15.1 | 95.1 | 0.0 | 79.9 |
| | $\hat{\nu}^{(\mathrm{rc2})}$ | 509.8 | 82.8 | 95.9 | 0.0 | | 499.6 | 25.8 | 95.0 | 0.0 | |
| | $\hat{\nu}^{(\mathrm{mt1})}$ | 500.7 | 29.9 | 95.2 | 0.0 | | 499.2 | 15.1 | 94.8 | 0.0 | |
| | $\hat{\nu}^{(\mathrm{mt2})}$ | – | – | – | 19.8 | | 502.4 | 45.8 | 94.2 | 0.8 | |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 514.2 | 33.8 | 93.2 | 0.0 | | 509.1 | 18.3 | 89.4 | 0.0 | |
| $\beta = 0.95\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 510.9 | 76.2 | 95.7 | 0.0 | 61.7 | 501.6 | 26.7 | 95.2 | 0.0 | 79.9 |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 487.8 | 30.5 | 92.5 | 0.0 | | 491.6 | 16.6 | 92.4 | 0.0 | |
| $\beta = 1.05\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 508.9 | 72.4 | 96.0 | 0.0 | 61.5 | 498.9 | 26.3 | 95.9 | 0.0 | 79.5 |
| Removal | $\hat{\nu}^{(\mathrm{rm})}$ | 504.2 | 62.8 | 95.5 | 0.0 | 64.2 | 500.7 | 22.3 | 94.5 | 0.0 | 83.1 |

$\nu = 500$, $m = 100$, $\alpha = 0.0025$

| Settings | | $K = 6$ | | | | $P_T$ | $K = 10$ | | | | $P_T$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Recapture | | av($\hat{\nu}$) | RMSE | $CP$ | $pf$ | | av($\hat{\nu}$) | RMSE | $CP$ | $pf$ | |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 769.8 | 279.7 | 0.1 | 0.0 | | 673.9 | 176.7 | 0.0 | 0.0 | |
| $\beta = 0.5\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 508.5 | 79.3 | 95.4 | 0.0 | 62.1 | 501.0 | 25.4 | 95.3 | 0.0 | 80.6 |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 360.6 | 140.5 | 0.0 | 0.0 | | 400.8 | 99.9 | 0.0 | 0.0 | |
| $\beta = 2\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 525.3 | 125.7 | 96.6 | 0.0 | 54.5 | 504.1 | 45.4 | 94.9 | 0.0 | 71.2 |
| | $\hat{\nu}^{(\mathcal{M}_0)}$ | 500.1 | 32.4 | 95.6 | 0.0 | | 499.8 | 16.9 | 94.8 | 0.0 | |
| $\beta = \alpha$ | $\hat{\nu}^{(\mathrm{rc1})}$ | 500.5 | 32.4 | 95.5 | 0.0 | 58.6 | 499.8 | 16.6 | 95.6 | 0.0 | 77.0 |
| | $\hat{\nu}^{(\mathrm{rc2})}$ | 514.4 | 97.7 | 96.4 | 0.1 | | 501.1 | 31.8 | 94.8 | 0.0 | |
| | $\hat{\nu}^{(\mathrm{mt1})}$ | 500.6 | 32.4 | 95.6 | 0.0 | | 499.8 | 16.6 | 95.6 | 0.0 | |
| | $\hat{\nu}^{(\mathrm{mt2})}$ | – | – | – | 24.8 | | 505.3 | 54.2 | 94.8 | 1.5 | |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 515.8 | 37.8 | 92.7 | 0.0 | | 509.7 | 20.0 | 90.5 | 0.0 | |
| $\beta = 0.95\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 518.1 | 92.7 | 95.8 | 0.0 | 58.8 | 502.0 | 31.8 | 94.6 | 0.0 | 77.3 |
| | $\hat{\nu}^{(\mathrm{rc1})}$ | 488.3 | 33.0 | 93.6 | 0.0 | | 491.3 | 18.7 | 92.0 | 0.0 | |
| $\beta = 1.05\alpha$ | $\hat{\nu}^{(\mathrm{rc2})}$ | 516.5 | 96.2 | 95.6 | 0.0 | 58.4 | 501.2 | 31.6 | 95.8 | 0.0 | 76.7 |
| Removal | $\hat{\nu}^{(\mathrm{rm})}$ | 508.0 | 65.5 | 94.7 | 0.0 | 64.4 | 500.3 | 19.4 | 95.2 | 0.0 | 84.9 |

Note: the empirical standard deviations are very close to the averages of the estimated standard errors, we only list RMSE for comparison. The exceptions are for $\hat{\nu}^{(\mathrm{mt2})}$ with $K = 6$, the results are not presented for comparison due to many failures. In the recapture settings without behavioral response, the average trap occupation rates are 36.8% and 68.4%, respectively, for $(\nu, m, \alpha) = (500, 200, 0.001)$ and $(500, 100, 0.0025)$.

In the case of a behavioral response in the recapture setting, those previously trapped do not provide any information for estimating $\nu$ (see (11), (12) and (13)), and worse, they occupy traps so that the capture probability for the untrapped

animals is reduced – the smaller the number of traps, the worse the recapture setting. Besides, applying the estimator ignoring the behavioral response would lead to a serious bias.

- When there is no behavioral response to trapping, in general, the recapture setting is better than that of the removal setting, especially when the number of traps is large or the total capture proportion is low.

When there is no behavioral response, although recapture means fewer new captures, the recaptures provide information for estimating $\alpha$. The total captures (including the recaptures) is larger than that of the removal setting; the larger the number of traps, the less loss from the reduction in new captures due to the recaptures. In the case of a slight behavioral response, the performance of $\hat{\nu}^{(\mathrm{rc1})}$ (assuming no behavioral response) is better than those of $\hat{\nu}^{(\mathrm{rc2})}$ and $\hat{\nu}^{(\mathrm{rm})}$ (based on the corresponding removal setting), in term of a smaller mean square error, especially when the capture proportion is low.

- As the number of trapping occasions $K$ increases and a high capture proportion is achieved for the removal setting, the effect of the removal setting will exceed that of the corresponding recapture setting without behavioral response. The smaller the number of traps, the earlier this exceedance happens; see Table 6.

Table 6. The superiority of the removal setting as $K$ increases.

$\nu = 500$, $m = 200$, $\alpha = 0.001$

| Settings and | Est. | $K = 14$ | $P_T$ | $K = 16$ | $P_T$ | $K = 18$ | $P_T$ |
|---|---|---|---|---|---|---|---|
| Removal | $\hat{\nu}^{(\mathrm{rm})}$ | (10.3, 10.2 ) | 92.2 | (7.38, 7.45) | 94.7 | $(5.68, 5.61)^{\triangle}$ | 96.5 |
| Recapture | $\hat{\nu}^{(\mathrm{rc1})}$ | $( 9.07, 9.20)^{\triangle}$ | 89.3 | (7.29, 7.43) | 92.2 | (6.07, 6.08) | 94.3 |

$\nu = 500$, $m = 100$, $\alpha = 0.0025$

| Settings and | Est. | $K = 11$ | $P_T$ | $K = 12$ | $P_T$ | $K = 13$ | $P_T$ |
|---|---|---|---|---|---|---|---|
| Removal | $\hat{\nu}^{(\mathrm{rm})}$ | (15.1, 15.1) | 88.0 | (12.2, 12.1) | 90.5 | $( 9.98, 9.91)^{\triangle}$ | 92.5 |
| Recapture | $\hat{\nu}^{(\mathrm{rc1})}$ | $(14.6, 14.8)^{\triangle}$ | 80.1 | (12.9, 13.1) | 82.9 | (11.7, 11.7) | 85.2 |

Note: the recapture setting is without behavioral response; we only list the empirical standard deviation and the average of the estimated errors since the averages of $\hat{\nu}$ are almost unbiased. $\triangle$: the smaller of the removal and recapture.

In the case of no behavioral response, the recaptures provide information on estimating $\alpha$. Only new captures provide direct information on estimating $\nu$. As $K$ increases, the number of recaptures increases, that of the new captures decreases (due to the availability of traps). Eventually, the gain from the recaptures on estimating $\alpha$ is marginal, and the gain from recaptures cannot compensate to the loss of information from the reduction in the new captures.

An asymptotic efficiency of the recapture setting (without behavioral response) relative to the removal setting is considered here. Take the asymptotic

efficiency to be

$$\text{a.eff.} = \sqrt{\frac{\text{a. var}(\hat{\nu}^{(\text{rm})})}{\text{a.var}(\hat{\nu}^{(\text{rc1})})}}.$$

For convenience of expression, let $m_1 = \cdots = m_K = m$ and $m/\nu \to r$. From Theorems1 and 2, for $Kr < 1$,

$$\text{a.var}(\hat{\nu}^{(\text{rm})}) = \frac{\nu Kr}{K^2 r^2 (1-Kr)^{-1} - \log^2(1-Kr)},$$
$$\text{a.var}(\hat{\nu}^{(\text{rc1})}) = \frac{\nu Kr}{Kr[(1-r)^{-K} - 1] - K^2 \log^2(1-r)};$$

for $Kr \geq 1$, the asymptotic efficiency is zero. Figure 1 plots the asymptotic efficiency against $K$ when $r = 0.1$.

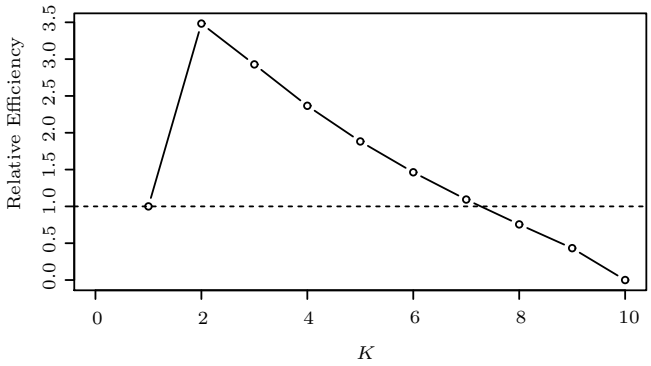



Figure 1. Asymptotic efficiency of $\hat{\nu}^{(rc1)}$ relative to $\hat{\nu}^{(rm)}$ against $K, r = 0.1$.

When $K = 1$, there is no difference between the two settings: as $K$ increases, the efficiency is much larger than 1 at first, but it decreases and eventually goes to zero.

Now we check the role of the availability of capture times on the estimation. Without capture times, the traditional discrete-type estimation methods are also available for the recapture setting considered in this section. Intuitively, the proposed estimators should be better than the discrete-type estimators due to availability of the capture times. First, we consider the case that the trapping duration and the number of traps are fixed for all the trapping occasions, and there is no behavioral response to trapping. Then the capture probabilities for all the trapping ocassions are the same for all the individuals, whether previously trapped or not. This is the model $\mathcal{M}_0$; see Otis et al. (1978). $\hat{\nu}^{(\mathcal{M}_0)}$, the MLE of $\nu$ based on the model $\mathcal{M}_0$, and the proposed estimator $\hat{\nu}^{(\text{rc1})}$ are applied to such a

setting for the sake of comparison. Simulation results show that the performances of the two estimators are almost the same, whether $\nu$ and $P_\tau$ are large or small; see Table 5.

We check the asymptotic efficiency of $\hat{\nu}^{(\mathcal{M}_0)}$ relative to $\hat{\nu}^{(\mathrm{rc1})}$. The asymptotic variance of $\hat{\nu}^{(\mathcal{M}_0)}$ is given by

$$\mathrm{a.var}(\hat{\nu}^{(\mathcal{M}_0)}) = \frac{\nu}{[(1-p)^{-K}-1]-pK/(1-p)},$$

where $p$ is the capture probability for all the individuals on all trapping occasions. In the asymptotic situation $p$ is just $r$, and the asymptotic efficiency of $\hat{\nu}^{(\mathcal{M}_0)}$ relative to $\hat{\nu}^{(\mathrm{rc1})}$ is

$$\mathrm{a.eff.} = \sqrt{\frac{(1-r)^{-K}-1-Kr/(1-r)}{(1-r)^{-K}-1-K[\log(1-r)]^2 r^{-1}}}.$$

Figure 2 plots the asymptotic efficiency against $K$ for $r = 0.1$, 0.3 and 0.5 respectively.

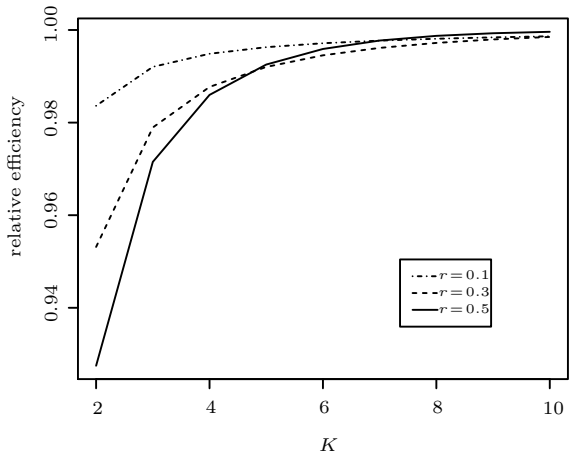



Figure 2. Asymptotic efficiency of $\hat{\nu}^{(\mathcal{M}_0)}$ relative to $\hat{\nu}^{(rc1)}$ against $K$, for $r = 0.1$, 0.3 and 0.5.

It can be seen that the asymptotic efficiency is smaller than 1 but very close to 1, especially for large $K$. In fact, the MLE for the model $\mathcal{M}_0$ makes use of the information that the capture probability $p$ is unchanged for all the individuals over all occasions, but the proposed estimator $\hat{\nu}^{(\mathrm{rc1})}$ does not. There is no difference for $\hat{\nu}^{(\mathrm{rc1})}$ when the number of traps and the trapping duration vary from occasion to occasion, but as a discrete model it becomes the model $\mathcal{M}_\mathrm{t}$; see Otis et al. (1978). The performance of the MLE based on the discrete model

$\mathcal{M}_{\rm t}$ is worse than that based on $\mathcal{M}_0$. When there is a behavioral response to trapping, even the trapping duration and the number of traps stay unchanged for all trapping occasions, due to the limited number of traps, the capture probability for the previously untrapped individuals changes from occasion to occasion, as well as the capture probability for the previously trapped individuals. As a discrete model, this is the general form of the model $\mathcal{M}_{\rm tb}$ described in Otis et al. (1978). Without further assumption, there is no available estimator. But with the capture times, the proposed estimator $\hat{\nu}^{(\rm rc2)}$ performs satisfactorily as long as the capture proportion is not too small.

In summary, when there is a significant behavioral response, the removal setting is strongly recommended, and the information on capture times is crucial; when there is no or slight behaviorial response to trapping, and the number of traps is large, the recapture setting is strongly recommended. Here, keeping the trapping duration and the number of traps unchanged for all trapping occasions, the capture times are unnecessary. Otherwise, the capture times are important. In the next section where arbitrary time-dependent baselines are considered, the availability of capture times on estimation is shown to be crucial.

## 4. Time-dependent baseline capture hazard

In practice, the baseline capture hazards for each trapping occasion may be different and even changing with time, for example different weather conditions. For the removal setting with arbitrary time-dependent baselines, it is impossible to give an estimate for $\nu$ due to an identifiability problem. In the recapture settings, we assume that the baseline capture hazard for the $i$th trapping occasion is $\alpha\lambda_i(t)$ for those previously uncaptured, and $\beta\lambda_i(t)$ for those previously captured. We have, for $i = 1, \ldots, K$, $dN_{i0}(t) = \alpha\lambda_i(t)T_i(t)(\nu - M_i(t))dt$, and $dN_{i1}(t) = \beta\lambda_i(t)T_i(t)L_i(t)dt$.

### 4.1. Estimation procedure

Let $\theta = \alpha/\beta$, and consider the martingale difference, for $i = 2, \ldots, K$,

$$dD_i(t) = L_i(t)dN_{i0}(t) - \theta(\nu - M_i(t))dN_{i1}(t),$$

which is such that $\mathrm{E}(dD_i(t)|\mathcal{F}_{t-}) = 0$, where $\mathcal{F}_t = \sigma\{N_{i0}(u),\ N_{i1}(u) : 0 \le u \le t\}$. The stochastic integral $\int_0^{\tau_i} W_i(t-)dD_i(t)$ is a zero mean martingale, where $W_i(t-)$ is any locally bounded and predictable process with respect to $\mathcal{F}_{t-}$. Equating this integral to zero and evaluating it at time $\tau_i$, gives an estimator of $\nu$. Obviously, only if the capture proportion for the $i$th trapping occasion is high will the estimator perform satisfactorily (see Yip, Xi, Fong and Hayakawa (1999)). But in practice, it is almost impossible to obtain a high capture proportion on

one trapping occasion. So we combine the $K-1$ trapping occasions (from the second occasion to the $K$th occasion). The resulting estimating equation is

$$\sum_{i=2}^{K}\int_0^{\tau_i} W_i(t-)dD_i(t)=0. \tag{16}$$

In the case of no behavioral response to trapping, i.e., $\theta=1$, there is only one unknown parameter $\nu$. The weight function $W_i(t-)$ can be chosen in an optimal way to minimize the width of the confidence interval based on the estimate $\hat{\nu}$, as suggested in Godambe (1985). The optimal weight is

$$\begin{aligned} W_i(t-)=\frac{\mathrm{E}(\frac{dD_i}{d\nu}|\mathcal{F}_{t-})}{\mathrm{var}(dD_i|\mathcal{F}_{t-})} &= \frac{1}{(\nu-M_i(t))[L_i(t)+\theta(\nu-M_i(t))]} \\ &= \frac{1}{(\nu-M_i(t))(\nu-N_i(t-))}, \end{aligned}$$

so the optimal martingale estimation equation is

$$\sum_{i=2}^{K}\left\{\int_0^{\tau_i}\frac{L_i(t)dN_{i0}(t)}{(\nu-M_i(t))(\nu-N_i(t-))}-\int_0^{\tau_i}\frac{dN_{i1}(t)}{\nu-N_i(t-)}\right\}=0, \tag{17}$$

and $\hat{\nu}$ can be obtained numerically. Using a Taylor's series expansion, the standard error of $\hat{\nu}$ can be estimated by $\sqrt{\Psi_1(\hat{\nu})}/|\Psi_2(\hat{\nu})|$ where $\Psi_1(\nu)$ is the predictable variance process of the optimal martingale, $\Psi_2(\nu)$ is the derivative of the optimal martingale with respect to $\nu$.

In the case of a behavioral response we have two unknown parameters, $\nu$ and $\theta$. The corresponding optimal weight functions for the martingale difference in (16), for estimators $\hat{\nu}$ and $\hat{\theta}$, are given by

$$\begin{aligned} W_i^{(\nu)}(t-)&=\frac{\mathrm{E}(\frac{dD_i}{d\nu}|\mathcal{F}_{t-})}{\mathrm{var}(dD_i|\mathcal{F}_{t-})}=\frac{1}{(\nu-M_i(t))[L_i(t)+\theta(\nu-M_i(t))]},\\ W_i^{(\theta)}(t-)&=\frac{\mathrm{E}(\frac{dD_i}{d\theta}|\mathcal{F}_{t-})}{\mathrm{var}(dD_i|\mathcal{F}_{t-})}=\frac{1}{\theta[L_i(t)+\theta(\nu-M_i(t))]}, \end{aligned}$$

respectively. Substituting these into (16), we obtain

$$\sum_{i=2}^{K}\left\{\int_0^{\tau_i}\frac{L_i(t)dN_{i0}(t)}{(\nu-M_i(t))[L_i(t)+\theta(\nu-M_i(t))]}-\theta\int_0^{\tau_i}\frac{dN_{i1}(t)}{[L_i(t)+\theta(\nu-M_i(t))]}\right\}=0, \tag{18}$$

$$\sum_{i=2}^{K}\left\{\int_0^{\tau_i}\frac{L_i(t)dN_{i0}(t)}{L_i(t)+\theta(\nu-M_i(t))}-\theta\int_0^{\tau_i}\frac{(\nu-M_i(t))dN_{i1}(t)}{[L_i(t)+\theta(\nu-M_i(t))]}\right\}=0. \tag{19}$$

We can then obtain $\hat{\nu}$ and $\hat{\theta}$ by solving (18) and (19) simultaneously. To measure the precision of the estimates, let $\vartheta=(\nu,\theta)$, and let $\boldsymbol{R}(\vartheta)=(\sum_{i=2}^{K} R_{i1}, \sum_{i=2}^{K} R_{i2})$, where $\sum_{i=2}^{K} R_{i1}$ and $\sum_{i=2}^{K} R_{i2}$ denote the left sides of (18) and (19), respectively. Using a Taylor's series expansion of $\boldsymbol{R}(\vartheta)$ at $\hat{\vartheta}$, a measure of the dispersion of $\hat{\vartheta}$ is given by $\dot{\boldsymbol{R}}^{-1}(\hat{\vartheta})\,\boldsymbol{V}(\hat{\vartheta})\dot{\boldsymbol{R}}(\hat{\vartheta})^{-\mathsf{T}}$, where $\dot{\boldsymbol{R}}(\hat{\vartheta})$ denotes the first derivative of $\boldsymbol{R}(\vartheta)$ evaluated at $\hat{\vartheta}$, $\boldsymbol{A}^{-\mathsf{T}}$ denotes the transpose of the inverse of a matrix $\boldsymbol{A}$ and $\boldsymbol{V}(\vartheta)$ is the dispersion matrix of $\boldsymbol{R}(\vartheta)$, given by

$$\boldsymbol{V}(\vartheta) = \begin{bmatrix} \sum_{i=2}^{k}\langle R_{i1}, R_{i1}\rangle & \sum_{i=2}^{k}\langle R_{i1}, R_{i2}\rangle \\ \sum_{i=2}^{k}\langle R_{i2}, R_{i1}\rangle & \sum_{i=2}^{k}\langle R_{i2}, R_{i2}\rangle \end{bmatrix}.$$

The following theorems describe the asymptotic properties of the proposed estimators in this section.

**Theorem 4.** *Suppose $\nu \to \infty$, $m_i/\nu \to r_i$ and $g = \max(r_1, \ldots, r_K)$. Let $\hat{\nu}$ denote the martingale estimate of $\nu$ in the case of no behavioral response, and assume that $\phi_i = \lambda_i(0+) > 0$. If $g < 1$, then $\hat{\zeta} = \nu^{-1/2}(\hat{\nu} - \nu) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \omega^{-1})$, where*

$$\omega = \sum_{i=2}^{K} \frac{b_i r_i}{(1-b_i)(1-r_i)} \quad \text{and} \quad b_i = 1 - \prod_{j=1}^{i-1}(1-r_j).$$

*If $g \geq 1$, then $\hat{\nu}/\nu \xrightarrow{\mathrm{p}} 1$ with a speed faster than $\nu^{-1/2}$.*

**Proof.** See Appendix 3 in Yip et al. (2005).

**Theorem 5.** *Suppose $\nu \to \infty$, $m_i/\nu \to r_i$ and $g = \max(r_1, \ldots, r_K)$. Let $\hat{\nu}$ and $\hat{\theta}$ denote the martingale estimates of $\nu$ and $\theta$ in the case of a behavioral response, and assume that $\phi_i = \lambda_i(0+) > 0$. If $g < 1$, then $\hat{\zeta} = (\nu^{-1/2}(\hat{\nu}-\nu), \nu^{1/2}(\hat{\theta}-\theta))' \xrightarrow{\mathrm{d}}$ $\mathrm{N}_2(\mathbf{0}, \boldsymbol{\Omega}_5^{-1})$, where $\boldsymbol{\Omega}_5 = (\omega_{ij})_{2\times 2}$, with*

$$\omega_{11} = \sum_{i=2}^{K} \int_0^{\infty} \frac{\alpha\phi_i(r_i - A_i(t))(b_i - A_{i1}(t))}{(1-b_i-A_{i0}(t))[(b_i - A_{i1}(t)) + \theta(1-b_i-A_{i0}(t))]} dt,$$

$$\omega_{12} = \omega_{21} = \sum_{i=2}^{K} \int_0^{\infty} \frac{\alpha\phi_i(r_i - A_i(t))(b_i - A_{i1}(t))}{\theta[(b_i - A_{i1}(t)) + \theta(1-b_i-A_{i0}(t))]} dt, \text{ and}$$

$$\omega_{22} = \sum_{i=2}^{K} \int_0^{\infty} \frac{\alpha\phi_i(r_i - A_i(t))(1-b_i-A_{i0}(t))(b_i - A_{i1}(t))}{\theta^2[(b_i - A_{i1}(t)) + \theta(1-b_i-A_{i0}(t))]} dt,$$

*$a_i$ and $b_i$ being the same as in Theorem 3. Here $A_{ij}(t) = \lim_{\nu\to\infty} N_{ij}(t/\nu)/\nu$ satisfies the differential equations*

$$\begin{cases} dA_{i0}(t) = \alpha\phi_i(r_i - A_i(t))(1-b_i-A_{i0}(t))dt, \\ dA_{i1}(t) = \beta\phi_i(r_i - A_i(t))(b_i - A_{i1}(t))dt, \end{cases}$$

*where* $A_i(t) = A_{i0}(t) + A_{i1}(t)$.

**Proof.** See Appendix 3 in Yip et al. (2005).

### 4.2. Simulation and comparison

We denote the proposed martingale estimators of $\nu$ as $\hat{\nu}^{(\mathrm{mt1})}$ and $\hat{\nu}^{(\mathrm{mt2})}$, respectively, for the settings without or with behavioral response. They do not depend on the form of $\lambda_i(t)$, and only use the sequence of the captures, not the capture times. (But, given the capture times, the capture sequence is available). For convenience of comparison, we choose the scenarios with a constant baseline and apply $\hat{\nu}^{(\mathrm{mt1})}$, $\hat{\nu}^{(\mathrm{mt2})}$, $\hat{\nu}^{(\mathrm{rc1})}$ and $\hat{\nu}^{\mathrm{rct2})}$. Some results are listed in Table 5.

With a slight behavioral response, the performance of $\hat{\nu}^{(\mathrm{mt1})}$ is better than that of $\hat{\nu}^{(\mathrm{mt2})}$. Only with a high capture proportion is the martingale estimator $\hat{\nu}^{(\mathrm{mt2})}$ reliable (e.g., for $\nu = 500$, $P_T$ should be greater than 75%). In the case of no behavioral response, the performance of the martingale estimator $\hat{\nu}^{(\mathrm{mt1})}$ is almost the same as that of $\hat{\nu}^{(\mathrm{rc1})}$, the MLE. This means the proposed martingale estimator is highly efficient when there is no behavioral response. The explanation is given as follows.

Each trapping occasion alone is a continuous-time removal experiment. After the first occasion, there are two types of individuals: previously captured or not. Therefore, starting at the second occasion, each trapping occasion is a seeded continuous-time removal experiment where the individuals previously captured are regarded as seeds. When the seeds have the same capture intensity (i.e., no behavioral responses), it was shown in Lloyd et al. (1998) that such a martingale estimator (using only the sequence of captures) is highly efficient relative to the MLE with a known form baseline. For our case, that is starting at the second occasion, the constant baseline provides little information for estimating $\nu$. The main difference is at the first occasion, a continuous-time removal experiment without seeds. The MLE uses the information of the constant baseline through the capture times, but the martingale estimator cannot. It is for arbitrary time-dependent baselines, and without seeds, there is a identifiability problem in the first occasion.

In the case of no behavioral response, with Theorems 2 and 4, the asymptotic efficiency of the martingale estimator relative to the MLE is

$$\sqrt{\frac{\omega GW}{(1-W)G - W(\log W)^2}}.$$

Letting all the $r_i$ be 0.1 except for $r_1$ which can be chosen as 0.3, 0.2 and 0.1, respectively, Figure 3 plots the efficiencies.

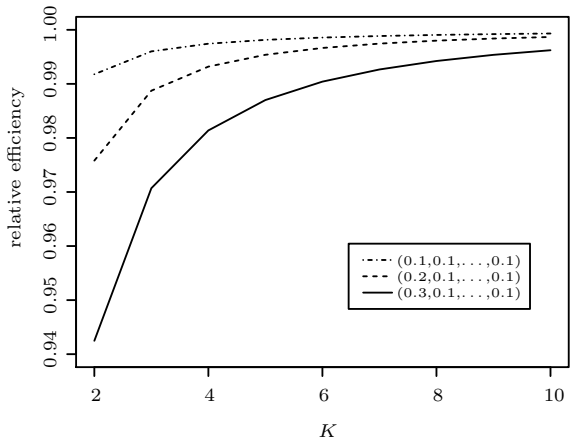



Figure 3. Asymptotic efficiency of $\hat{\nu}^{(mt1)}$ relative to $\hat{\nu}^{(rc1)}$ against $K$, all the $r_i$ are equal to 0.1 except for $r_1$ which is chosen as 0.3 0.2 and 0.1, respectively.

For $K = 2$, the efficiency with $(r_1, r_2) = (0.3, 0.1)$ is the smallest among the three because the proportion of the information from the first occasion is the largest. As $K$ increases, the proportion from the first occasion is reduced, and all efficiencies increase strictly and tend to 1. Letting $K = 2$, $r_1 = 0.6$ and $r_2 = 0.2$, the relative efficiency is 0.73, there is a significant difference between the performances of the two estimators. Simulation results (not reported here) also confirm this point.

## 5. Conclusion

The non-linear relationship of the capture intensity function in the proportional trapping model has complicated inference procedure. Without the capture times, there is no algorithm to search for the solution for the MLE. There would not be any ill-conditioning problem if capture times were available. It is important to record the capture times, especially in genuine continuous setting in which the capture intensity does change with traps availability over the course of the experiment.

Here we have extended the original model in Good et al. (1979) to recapture model with possible behavioral response. In the case of a behavioral response, the removal setting is strongly recommended and the information on capture times is essential for estimating the population size.

In the case of no behavioral response and a plentiful number of traps, the recapture setting is strongly recommended. Keeping trap number and trapping duration unchanged for each occasion, the information on capture times is not

necessary for estimating the population size. Otherwise, capture times are very important.

When baselines are time-dependent and arbitrary, the removal setting is not applicable due to an identifiability problem. Only the recapture setting can provide information on population size, and the capture times are essential. In the presence of a behavioral response, a high capture proportion is needed for obtaining a reliable estimate.

## Acknowledgement

The authors are grateful for the many useful comments of an Associate editor and two referees. The research was supported by a Research Grant Council Grant.

## References

Chao, A. (1987). Estimating the population size for capture-recapture data with unequal catchability. *Biometrics* **43**, 783-791

Darroch, J. N. (1958). The multiple-recapture census: I. Estimation of a closed population. *Biometrika* **45**, 343-359.

Godambe, V. P. (1985). The foundations of finite-sample estimation in stochastic processes. *Biometrika* **72**, 419-428.

Good, I. J., Lewis, B. C., Gaskins, R. A. and Howell, L. W. (1979). Population estimation by the removal method assuming proportional trapping. *Biometrika* **66**, 485-494.

Liu, L. and Yip, P. S. F. (2003). Estimating population size in proportional trapping-removal models. *Statist. Sinica* **13**, 243-254.

Lloyd, C. J., Yip, P. S. F. and Chan, K. S. (1998). A comparison of recapturing, removal and resighting design for population size estimation. *J. Statist. Plann. Inference* **71**, 363-375.

Moran, P. A. P. (1951). A mathematical theory of animal trapping. *Biometrika* **38**, 307-311.

Otis, D. L., Burnham, K. P., White, G. C. and Anderson, D. R. (1978). Statistical Inference from capture data on closed animal populations. *Wildlife Monograph* **62**. The Wildlife Society, Washington, D. C.

Yip, P. S. F., Xi, L., Fong, Y. T. D. and Hayakawa, Y. (1999). Sensitivity-analysis and estimating number-of-faults in removal debugging. *IEEE Trans. Reliability* **48**, 300-305.

Yip, P. S. F., Xi, L. and Liu, L. (2005). Population size estimation in proportional trapping removal and recapture models with known capture times. *Research Report* **419**. Department of Statistics and Actuarial Science, The University of Hong Kong.

Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam, Hong Kong.

E-mail: sfpyip@hku.hk

Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam, Hong Kong.

E-mail: lxiaa@graduate.hku.hk

School of Mathematical Science, Peking University, Beijing 100871, China.
E-mail: liping@math.pku.edu.cn

(Received August 2004; accepted December 2005)