Statistica Sinica **13**(2003), 243-254

ESTIMATING POPULATION SIZE IN PROPORTIONAL TRAPPING-REMOVAL MODELS

Liping Liu and Paul S. F. Yip

Peking University and University of Hong Kong

Abstract: We examine the inference problem of a proportional trapping model introduced by Good, Lewis, Gaskins and Howell (1979), and consider a similar removal model in continuous time. The existence and uniqueness of the maximum likelihood estimates of the parameters are studied. Corresponding variance estimates are also given. Martingale theory is used to obtain the asymptotic properties, and simulations are conducted to examine the performance of the estimation procedure.

Key words and phrases: Counting process, intensity function, martingale, proportional trapping.

1. Introduction

In wildlife studies, capture-recapture and/or removal methods are often used to estimate the unknown population size. Usually the capture procedure is to set up m traps at the start of each period. It is assumed that the capture probability of an individual in a trapping period remains constant (Moran (1951), Darroch (1958), Good, Lewis, Gaskins and Howell (1979)) and may depend on covariates (Huggins (1989), (1991), Yip, Huggins and Lin (1996), Yip, Zhou, Lin and Fang (1999), Lin and Yip (1999)). For an excellent review on recapture and removal studies, see Otis, Burnham, White and Anderson (1978) and Pollock (1991). However, nearly none of the papers except Good et al. (1979) take into account the number of traps available.

For many species, once an animal is trapped, it occupies the trap until the trap is cleared. The trap can only catch one animal at a time. The capture probability is therefore zero once all traps are occupied (see Good et al. (1979)). Standard estimation procedures make the implicit assumption that there are an infinite number of traps or, equivalently, that individuals who are trapped are released immediately so as to make the traps available for further trapping (Pollock (2000) and Seber (1982)). Such an assumption would be appropriate for a bird banding exercise in which a net is set up and any bird that flies through the net is caught. However, it is not appropriate for most mammal trapping exercises since in practice, traps are only cleared at specified times. For example, a small mammal study was conducted at Dummy Bottom within Browns Park National Wildlife Refuge along the Green River (see Falck (1996)). Falck used a 14×14 trapping grid (196 traps) with $15m$, spacing between traps. Traps were baited in the evening with a rolled oats and peanuts butter mixture and checked the following morning. Once an animal was trapped, the trap could not catch other animals until it was cleared. Here the capture probability is clearly dependent on the availability of traps. The standard models for capture intensity generally do not adjust for the number of traps available at the time. The one exception is the work of Good et al. (1979) who considered a model allowing for a reduction of trapping capacity. They derived a complicated formula to obtain an estimate of the population size. As this formula is ill-conditioned, they developed other estimates. Here we use a counting process to provide a continuous time formulation for the proportional trapping model to estimate the population size, ν.

We introduce our model and the martingale estimating equations, then obtain maximum likelihood estimates in Section 2. Asymptotic properties of estimates are derived in Section 3, and simulations are done to assess the performance of the estimating procedure in Section 4. An example is given in Section 5 to illustrate our method. Proof of the existence and uniqueness of the maximum likelihood estimates is given in the Appendix.

2. Formulation of the Proportional Trapping Model and Estimation Method

In the present study we consider only the situation in which the capture experiment is conducted in a single period: $[0, \tau]$. Suppose that, at the beginning of the time interval, m traps are distributed in a region to investigate the size of a closed population. The following conditions are assumed – similar to those in Good et al. (1979).

(i) Each trap can catch at most one animal, and the capture time is recorded. The caught animal remains in the trap till the end of the experiment, and can be regarded as having been removed from the population.

(ii) The traps are distributed evenly in the region, so that an animal has equal probability of being captured by each trap.

(iii) The animals behave homogeneously, i.e., the risks of capture are the same.

(iv) There is no interference from animals of another species.

If there are ν animals altogether, from which k animals are caught at times $0 < t_1 < t_2 < \cdots < t_k < \tau$, respectively, then by our assumptions, the hazard rate for any animal to be caught at time $t \in [0, \tau]$ is proportional to the number of available traps. Let $N_i(t)$ indicate if the ith animal has been caught up to time t and $N(t) = \sum_{i=1}^{\nu} N_i(t)$, which represents the total number of captures up to time t. Then

$$\mathcal{N}_i(t) = N_i(t) - \int_0^t \lambda\{m - N(s-)\}\{1 - N_i(s-)\}ds$$

is a $\mathcal{F}_t^*$ martingale. Here, $1 - N_i(s-)$ indicates if the ith animal is at risk just before time s, λ is a constant baseline hazard, which can be interpreted as hazard rate if there is only one trap and only one animal (see Good et al. (1979)), and $\mathcal{F}_t^* = \sigma\{N_i(s) : 0 \leq s \leq t, i = 1, \ldots, \nu\}$. Summing over i, we obtain an $\mathcal{F}_t$ martingale, $\mathcal{N}(t) = \sum_{i=1}^{\nu} \mathcal{N}_i(t)$, i.e.,

$$\mathcal{N}(t) = N(t) - \int_0^t \lambda\{\nu - N(s-)\}\{m - N(s-)\}ds, \tag{1}$$

where $\mathcal{F}_t = \sigma\{N(s) : 0 \leq s \leq t\}$. The above formulation can be regarded as a generalization of most other removal models. Consider the case in which the number of traps, m, is much bigger than ν. Then $\lambda(m - N(s-))$ is approximately a constant $\tilde{\lambda}$, and so our model becomes the usual one without considering the availability of traps: see Becker (1984), Becker and Heyde (1990), Yip, Fong and Wilson (1993) and Chao (1987).

If we denote by $s_1 = t_1$, $s_i = t_i - t_{i-1}$ $(i = 2, \ldots, k)$ and $s_{k+1} = \tau - t_k$, then equating $\mathcal{N}(t)$ to its mean, namely zero, we obtain

$$k = \lambda \sum_{i=1}^{k} (\nu - i + 1)(m - i + 1)s_i. \tag{2}$$

It is worth noting that $N(t)$ is a birth process so that the s_i are independent and exponentially distributed.

Actually (2) is a score function which can be derived from the likelihood function given below. The intensity function for $N(t)$ is $\lambda\{\nu - N(t-)\}\{m - N(t-)\}$, and the likelihood function for $N(t)$ is given by (see Andersen, Borgan, Gill and Keiding (1993))

$$\begin{aligned} L(\lambda, \nu) = &\prod_{0 \leq t \leq \tau} [\lambda\{m - N(t-)\}\{\nu - N(t-)\}]^{\Delta N(t)} \\ &\times \exp\Big[-\int_0^\tau \lambda\{\nu - N(t-)\}\{m - N(t-)\}dt\Big]. \end{aligned}$$

The derivatives of $\log L(\lambda, \nu)$ with respect to λ and ν are given by

$$\frac{\partial \log L}{\partial \lambda} = \frac{k}{\lambda} - \int_0^\tau \{\nu - N(s-)\}\{m - N(s-)\}ds, \tag{3}$$

$$\frac{\partial \log L}{\partial \nu} = \sum_{i=1}^{k} \frac{1}{\nu - i + 1} - \int_0^\tau \lambda\{m - N(s-)\}ds. \tag{4}$$

Note that (3) is equivalent to (2), but (4) is different from the ones suggested in Darroch (1958) and Yip et al. (1993) where the availability of traps is ignored. Furthermore, (3) and (4) can also be obtained by choosing the corresponding optimal weights for the martingale

$$\mathcal{N}^*(t) = \int_0^t H(s) d\mathcal{N}(s), \tag{5}$$

where $H(s) = \{H_1(s), H_2(s)\}$ is a predictable vector (weight function) with respect to $\mathcal{F}_s$. According to Godambe (1985), the optimal weight functions for λ and ν can be determined as

$$H_1(s) = \frac{\mathrm{E}\{\frac{\partial d\mathcal{N}(s)}{\partial \lambda} | \mathcal{F}_{s-}\}}{\mathrm{E}\{d\mathcal{N}(s)^2 | \mathcal{F}_{s-}\}} \quad \text{and} \quad H_2(s) = \frac{\mathrm{E}\{\frac{\partial d\mathcal{N}(s)}{\partial \nu} | \mathcal{F}_{s-}\}}{\mathrm{E}\{d\mathcal{N}(s)^2 | \mathcal{F}_{s-}\}},$$

which, in this case, are given by

$$H_1(s) = \frac{1}{\lambda} \quad \text{and} \quad H_2(s) = \frac{1}{\nu - N(s-)}, \tag{6}$$

respectively. Putting (6) back to (5), again we get (3) and (4), which are originally derived from the likelihood function.

Also, it can be shown that $\hat{\nu}$, the MLE of the population size, is the solution of the following equation

$$\frac{k}{\nu - C} = \sum_{i=1}^{k} \frac{1}{\nu - i + 1}, \tag{7}$$

where $C = \sum_{i=1}^{k+1} (i-1)(m-i+1)s_i / \sum_{i=1}^{k+1} (m-i+1)s_i$. Note that we always have $0 \le C \le k$. The existence and uniqueness conditions for $\hat{\nu}$ are given in the following theorem, the proof is given in the Appendix.

Theorem 1. *If the number of animals caught k is 0 or 1, then (7) has no finite solution; when $k \ge 2$, the necessary and sufficient condition for the existence and uniqueness of the solution of (7) is*

$$\frac{k-1}{2} < C \le \left(\frac{\sum_{i=2}^{k} \frac{1}{i}}{\sum_{i=1}^{k} \frac{1}{i}} \right) k. \tag{8}$$

From Theorem 1 it can be seen that as in most other capture-removal models, there is a positive probability that the MLE does not exist.

Once the estimate $\hat{\nu}$ is available, the estimate of λ is given by $\hat{\lambda} = k / \{\sum_{i=1}^{k+1} (\hat{\nu} - i + 1)(m - i + 1)s_i\}$. The observed information matrix is

$$I = -\begin{bmatrix} \frac{\partial^2 \log L}{\partial \lambda^2} & \frac{\partial^2 \log L}{\partial \lambda \partial \nu} \\ \frac{\partial^2 \log L}{\partial \lambda \partial \nu} & \frac{\partial^2 \log L}{\partial \nu^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\lambda^2} \int_0^\tau dN(s) & \int_0^\tau \{m - N(s-)\} ds \\ \int_0^\tau \{m - N(s-)\} ds & \int_0^\tau \frac{dN(s)}{\{\nu - N(s-)\}^2} \end{bmatrix},$$

and, since the predictable covariation process of $(\partial \log L/\partial\lambda, \partial \log L/\partial\nu)$ at $t = \tau$ takes the value

$$D = \begin{bmatrix} \frac{1}{\lambda}\int_0^\tau \{\nu - N(s-)\}\{m - N(s-)\}ds & \int_0^\tau \{m - N(s-)\}ds \\ \int_0^\tau \{m - N(s-)\}ds & \int_0^\tau \lambda\frac{m - N(s-)}{\nu - N(s-)}ds \end{bmatrix},$$

a reasonable estimate of the covariance matrix of $(\hat{\lambda}, \hat{\nu})$ is $\hat{I}^{-1}\hat{D}\hat{I}^{-1}$, where $\hat{I}$ and $\hat{D}$ are obtained by replacing λ and ν by $\hat{\lambda}$ and $\hat{\nu}$ in I and D, respectively. As a consequence, we get the estimate of the variance of $\hat{\nu}$, $\hat{s}^2_{\hat{\nu}}$. Since D and I are asymptotically equivalent, $\hat{I}^{-1}\hat{D}\hat{I}^{-1}$ can be replaced by $\hat{I}^{-1}$ or $\hat{D}^{-1}$. In the next section we examine the asymptotic properties of the estimator $\hat{\nu}$.

3. Asymptotic Properties

Let $B(t) = \mathrm{E}\{N(t)\}$ and $V(t) = \mathrm{Var}\{N(t)\}$. Taking expectation in (1), we get an integral equation

$$\begin{aligned} B(t) &= \int_0^t \lambda\mathrm{E}\left[\{\nu - N(s-)\}\{m - N(s-)\}\right]ds \\ &= \int_0^t \lambda[\nu m - (\nu + m)B(s) + \mathrm{E}\{N^2(s-)\}]ds \\ &= \int_0^t \lambda\{\nu m - (\nu + m)B(s) + V(s) + B^2(s)\}ds. \end{aligned}$$

The corresponding differential equation is therefore

$$dB(t) = \lambda\{\nu - B(t)\}\{m - B(t)\}dt + \lambda V(t)dt. \tag{9}$$

We study the asymptotic properties when $\nu \to \infty$. If m remains constant, then all traps will soon be occupied when $\nu \to \infty$. Therefore a reasonable assumption is that $m \to \infty$ too. The counting process $N(t)$ does not have the form of i.i.d. summation, but the martingale limit theorems can still be used. Note that if m and ν have the same order when $\nu \to \infty$, the intensity is actually a quadratic function of ν, hence we need a time scale transformation which is expressed in the following lemma.

Lemma 1. *Let ν tend to ∞ and m/ν tend to a constant $r > 0$. Then $N(\frac{t}{\nu})/\nu$ converges in probability to a continuous, strictly increasing deterministic function*

$$A(t) = \begin{cases} r \cdot \dfrac{\exp\{\lambda(1-r)t\} - 1}{\exp\{\lambda(1-r)t\} - r} & \textit{if } r < 1; \\ \dfrac{\lambda t}{1 + \lambda t} & \textit{if } r = 1; \\ r \cdot \dfrac{\exp\{\lambda(r-1)t\} - 1}{r \cdot \exp\{\lambda(r-1)t\} - 1} & \textit{if } r > 1. \end{cases}$$

And the martingale $\mathcal{M}(t) = \nu^{-\frac{1}{2}}[N(\frac{t}{\nu}) - \int_0^t \frac{\lambda}{\nu}\{\nu - N(\frac{u}{\nu}-)\}\{m - N(\frac{u}{\nu}-)\}du]$ *converges weakly to a Gaussian process with mean* 0 *and predictable variation process* $A(t)$.

Proof. It can be proved that $\lim_{\nu\to\infty} B(\frac{t}{\nu})/\nu$ exists. Denote it by $A(t)$. It is well known that the predictable variation process of $\mathcal{N}(t)$ is just the compensator of $N(t)$, i.e., the process $\mathcal{N}^2(t) - \int_0^t \lambda\{\nu - N(s-)\}\{m - N(s-)\}ds$ is a $\mathcal{F}_t$martingale. Therefore by using the fact that $\mathrm{E}\,[\{N(t) - x\}^2]$ is minimized when $x = E\{N(t)\} = B(t)$, we have

$$\begin{aligned} V(t) &= \mathrm{E}\,[\{N(t) - B(t)\}^2] \le \mathrm{E}\left[\{N(t) - \int_0^t \lambda(\nu - N(s-))(m - N(s-))ds\}^2\right] \\ &= \mathrm{E}\,[\mathcal{N}^2(t)] = \mathrm{E}\left[\int_0^t \lambda\{\nu - N(s-)\}\{m - N(s-)\}ds\right] = B(t) \le \nu. \end{aligned}$$

For any fixed $t \in [0, \tau]$, by applying Chebyshev's Inequality to $N(\frac{t}{\nu})/\nu$, it converges in probability to $\lim_{\nu\to\infty} B(\frac{t}{\nu})/\nu = A(t)$. From (9) we have

$$\frac{dB(\frac{t}{\nu})}{\nu} = \lambda\left\{1 - \frac{B(\frac{t}{\nu})}{\nu}\right\}\left\{\frac{m}{\nu} - \frac{B(\frac{t}{\nu})}{\nu}\right\}dt + \lambda\frac{V(\frac{t}{\nu})}{\nu^2}dt.$$

Let $\nu \to \infty$ to obtain a differential equation $dA(t) = \lambda\{1 - A(t)\}\{r - A(t)\}dt$ with initial condition $A(0) = 0$. The solution is expressed in Lemma 1.

The predictable variation process of $\mathcal{M}(t)$ is

$$\begin{aligned} \langle\mathcal{M}\rangle(t) &= \nu^{-1}\int_0^t \frac{\lambda}{\nu}\{\nu - N(\frac{u}{\nu}-)\}\{m - N(\frac{u}{\nu}-)\}du \\ &= \int_0^t \lambda\{1 - \frac{N(\frac{u}{\nu}-)}{\nu}\}\{\frac{m}{\nu} - \frac{N(\frac{u}{\nu}-)}{\nu}\}du, \end{aligned}$$

which converges in probability to $\int_0^t \lambda\{1 - A(u)\}\{r - A(u)\}du = A(t)$. The weak convergence of $\mathcal{M}(t)$ is then a direct consequence of Theorem II.5.1 of Andersen et al. (1993).

Theorem 2. *Suppose* ν *tends to* ∞ *and* m/ν *tends to* $r < 1$. *Let* $\hat{\lambda}$ *and* $\hat{\nu}$ *denote the maximum likelihood estimates of* λ *and* ν, *respectively. Then* $\hat{\zeta} = \{\nu^{\frac{1}{2}}(\hat{\lambda} - \lambda), \nu^{-\frac{1}{2}}(\hat{\nu} - \nu)\}$ *converges in distribution to a bivariate normal distribution with mean zero and covariance matrix* Ω^{-1}, *where*

$$\Omega = \begin{bmatrix} \dfrac{r}{\lambda^2} & \dfrac{1}{\lambda}\log\dfrac{1}{1-r} \\ \dfrac{1}{\lambda}\log\dfrac{1}{1-r} & \dfrac{r}{1-r} \end{bmatrix}.$$

Proof. Let $U = (\nu^{-\frac{1}{2}}(\partial \log L/\partial\lambda), \nu^{\frac{1}{2}}(\partial \log L/\partial\nu))$. Then, by Taylor's series expansion,

$$U = J_\nu \hat{\zeta} + o_p(1), \text{ where } J_\nu = \begin{bmatrix} \nu^{-1}\displaystyle\int_0^\tau \frac{dN(t)}{\lambda^2} & \displaystyle\int_0^\tau \{m - N(t-)\}dt \\ \displaystyle\int_0^\tau \{m - N(t-)\}dt & \nu\displaystyle\int_0^\tau \frac{dN(t)}{\{\nu - N(t-)\}^2} \end{bmatrix}.$$

Using Lemma 1, direct calculation (see Andersen et al. (1993)) shows that $J_\nu \xrightarrow{\mathrm{P}} \Omega$ as $\nu \to \infty$, and since $U \xrightarrow{\mathrm{d}} \mathrm{N}_2(\mathbf{0}, \Omega)$, it follows that $\hat{\zeta} \xrightarrow{\mathrm{d}} \mathrm{N}_2(\mathbf{0}, \Omega^{-1})$.

From the theorem we see that if ν and m are large (with $m < \nu$), then replacing r with $m/\hat{\nu}$ in Ω, the asymptotic variance of $\hat{\nu}$ can also be estimated by $\tilde{s}^2_{\hat{\nu}} = \{m(\hat{\nu} - m)\hat{\nu}\}/\{m^2 - \hat{\nu}(\hat{\nu} - m)\{\log(1 - \frac{m}{\hat{\nu}})\}^2\}$.

When $r \geq 1$, it is found that the variance of $\nu^{-\frac{1}{2}}(\hat{\nu} - \nu)$ tends to 0. The intuitive explanation for this is that if the baseline hazard λ remains constant and there are not less traps than animals, then the proportion of animals caught will tend to 1.

It seems that the condition $m/\nu \to r$ is too strong, but it is not. When m and ν are large, there could be three situations: ν is much larger than m; m is much larger than ν; or m and ν are comparable. In the first case, since the number of caught animals cannot exceed m, the capture proportion will be too low to provide enough information for estimating ν. In the second case, almost all animals will be caught very soon, the MLE is certainly consistent (see the previous paragragh). In the third case we need the asymptotic properties given in Theorem 2.

4. Simulations

Different combinations of population size ν, number of traps m, and baseline hazard rate λ are chosen for the simulation study. The choice of τ is always fixed to be 1, since its influence can be reflected in the choice of λ. The combinations are: for $\nu = 100$, we take $m = 80, 100, 120$ and $\lambda = 0.05, 0.10$; for $\nu = 500$, we take $m = 300, 400, 500, 600$ and $\lambda = 0.003, 0.01, 0.02$. The overall probabilities of being captured are from 40% to 97.5%. For each combination of ν, m and λ, 10,000 replications were simulated to estimate the mean and standard error of the population size estimator $\hat{\nu}$, the mean of the standard error estimator $\hat{s}_{\hat{\nu}}$, and the coverage percentage of the 95% confidence interval $(\hat{\nu} - 1.96\hat{s}_{\hat{\nu}}, \hat{\nu} + 1.96\hat{s}_{\hat{\nu}})$. The simulation results are summarized in Table 1.

These results show that the proposed inference procedures perform well when m and ν are large and the overall capture proportion is not too low (above 60%). The bias for $\hat{\nu}$ (and for $\hat{s}_{\hat{\nu}}$) is usually not large, and it diminishes quickly as the capture proportion increases. The coefficient of variation for $\hat{\nu}$ is fairly small, and the confidence interval has proper coverage probability. For small values of m and ν the performance is still satisfactory for a high capture proportion, whereas the estimate would be very unstable for low capture proportion. There are very few unsuccessful simulations, which means that the probability of the corresponding model failure (see Section 2) is fairly small. Bias is positive when m is smaller than ν, or if the overall capture proportion is low. The standard error estimate $\tilde{s}_{\hat{\nu}}$, obtained from the asymptotic result (not presented here), is close to $\hat{s}_{\hat{\nu}}$ when m and ν are large, but is apparently an underestimate otherwise.

Table 1. Simulation results for various combination of ν, m and λ.

	True Values					Simulation Results			
Trials	ν	m	λ	P_{ot}	P_τ	$\tilde{n}$	$\hat{\nu}$	$\hat{s}_{\hat{\nu}}$	P_c
1	100	80	0.05	89.58	71.66	9997	108.16 (160.81)	161.92 (12676.5)	88.89
2	100	80	0.10	96.96	77.57	9998	104.84 (27.36)	19.92 (94.89)	90.22
3	100	100	0.05	83.33	83.33	10000	108.40 (13.20)	10.39 (5.66)	90.26
4	100	100	0.10	90.91	90.91	10000	100.07 (5.65)	5.60 (2.20)	90.85
5	100	120	0.05	75.97	91.16	10000	99.66 (5.18)	5.32 (1.89)	91.21
6	100	120	0.10	81.22	97.46	9988	99.54 (2.01)	2.62 (0.62)	93.64
7	500	300	0.003	67.27	40.36	9892	711.76 (2484.09)	5694.10 (185287.25)	86.31
8	500	300	0.01	94.11	56.46	10000	520.91 (112.14)	103.05 (90.70)	91.08
9	500	300	0.02	99.26	59.56	10000	521.24 (99.28)	88.89 (71.81)	92.76
10	500	400	0.003	63.63	50.90	10000	535.97 (399.31)	274.13 (11940.38)	89.61
11	500	400	0.01	89.58	71.66	10000	504.76 (43.22)	42.61 (14.14)	93.54
12	500	400	0.02	96.96	77.57	10000	503.40 (31.64)	30.85 (8.84)	94.37
13	500	500	0.003	60.00	60.00	10000	512.10 (85.77)	80.70 (49.22)	91.52
14	500	500	0.01	83.33	83.33	10000	500.87 (21.39)	21.32 (4.50)	94.22
15	500	500	0.02	90.91	90.91	10000	500.24 (11.95)	11.85 (1.97)	94.24
16	500	600	0.003	54.44	67.73	10000	503.77 (53.34)	51.64 (20.08)	92.56
17	500	600	0.01	75.97	91.16	10000	499.84 (11.47)	11.50 (1.79)	94.12
18	500	600	0.02	81.22	97.46	10000	499.59 (4.48)	4.59 (0.58)	94.29

Notes: 1. P_τ: capture probability;
2. P_{ot}: averaged proportion of occupied traps;
3. P_c: 95% coverage probability;
4. Simulated standard errors in parentheses;
5. $\tilde{n}$: no. of successes from 10,000 simulations.

If the number of traps is neglected, an inappropriate model gives the counting process $N(t)$ the intensity $\tilde{\lambda}\{\nu - N(t-)\}$, with $\tilde{\lambda}$ being the corresponding baseline hazard, and the maximum likelihood estimate for ν again as the solution of (7), but with constant C replaced by $\tilde{C} = \frac{1}{\tau}\sum_{i=1}^{k+1}(i-1)s_i$. Simulation shows that the estimator from the inappropriate model is strongly biased negatively. For example, for Trials 2–4 in Table 1, if the inappropriate model is used, the corresponding means of the estimate $\hat{\nu}$ are 78.76 (1.11), 84.58 (2.63) and 92.81 (1.60) respectively. The simulated standard errors are misleading small. Estimates from the proportional trapping model give the values 104.84 (27.36), 108.40 (13.2) and 100.07 (5.65). Table 2 shows that when the number of traps increases and ν remains at 100, the underestimation via the inappropriate model is less severe. The overall capture probabilities for each of the trials in Table 2 is around 85%.

Table 2. The effect of different number of traps (m) in using the inappropriate model for $\nu = 100$.

	True Values		Simulation Results	
Trials	m	λ	$\tilde{n}$	$\tilde{\nu}$
1	100	0.057	8077	86.16 (2.36)
2	200	0.0135	10000	91.80 (5.06)
3	400	0.0055	10000	95.46 (6.87)
4	600	0.0035	10000	96.94 (7.44)
5	1000	0.002	10000	98.04 (8.31)

Notes: 1. Standard errors in parentheses;
2. $\tilde{n}$: no. of successes from 10,000 simulations.

5. A Mammal Capture Study

Since the model of Good et al. (1979) and the proposed one deal with different types of data, it is impossible to use their examples for comparison. We use a data set from a small mammal capture-recapture study (Falck (1996)) to illustrate our method. Captured animals were given unique marks and released at a capture location. The study was designed by using Pollock's robust design (Pollock, Nichols, Brownie and Hines (1990)) with primary and secondary sessions. We use the capture data at the 5th secondary session in the 3rd primary session, in which there were altogether 62 animals caught. The captured animals are regarded as having been removed, since they had to stay in the traps until the end of this secondary session (i.e., one day). In order to apply the proposed method we make the discrete data continuous by generating the unknown capture times based on the proposed model with parameter values $\nu = 75$ and $\lambda = 0.0009$. The capture times of the individuals are given in Table 3. Since the experiment was carried out from the evening to early morning of the next day, τ is taken to be 12 hours.

The estimated population size is 77.26, with a standard error 11.12. If we use the model which ignores the information on the number of traps, then the maximum likehood estimate is 70.53, with a standard error 6.45. Remembering that estimates from this model are usually negatively biased, one might guess that the true value should be greater than 70.53, and close to 77.26. The 95%

confidence intervals for population size ν, constructed by using asymptotic normality, are $(55.46, 99.06)$ and $(57.89, 83.17)$, respectively, according to the two models.

Table 3. Simulated capture times for the removal data from Falck (1996) using $\nu = 75$ and $\lambda = 0.0009$.

No.	Time	No.	Time	No.	Time	No.	Time	No.	Time
1	0.04	14	1.15	27	3.15	40	4.89	53	7.84
2	0.07	15	1.76	28	3.25	41	5.14	54	7.95
3	0.13	16	1.80	29	3.26	42	5.48	55	8.44
4	0.24	17	1.83	30	3.63	43	5.50	56	8.59
5	0.35	18	1.87	31	3.84	44	5.73	57	8.64
6	0.37	19	1.88	32	3.91	45	5.75	58	9.11
7	0.48	20	1.88	33	4.03	46	5.84	59	9.72
8	0.50	21	2.14	34	4.05	47	6.28	60	10.91
9	0.68	22	2.32	35	4.39	48	6.56	61	11.12
10	0.79	23	2.60	36	4.40	49	6.73	62	11.37
11	0.83	24	2.85	37	4.54	50	6.87	–	–
12	0.85	25	2.86	38	4.73	51	6.99	–	–
13	0.89	26	2.86	39	4.83	52	7.02	–	–

In this example, about 1/3 (62/196) of the traps were occupied. This means that at the end of the secondary session, for any animal still at risk, the hazard rate of being caught is about 2/3 of that at the start of the session. The reduction of traps available is significant, which reinforces the necessity of using the proposed proportional trapping model.

Acknowledgements

This work was supported by a RGC grant and the National Natural Science Foundation of China (10071004), and was mainly done while the first author was visiting the University of Hong Kong. We are grateful to Prof. Ray Watson, the referees, an associate editor and the Co-Editor for valuable comments.

Appendix: Proof of Theorem 1

Lemma 2. Let $f(x) = \frac{k}{x-C} - \sum_{i=1}^{k} \frac{1}{x-i+1}$, *for* $x \in [k, \infty)$. *If* $f(x_1) = 0$ *for some* $x_1 \in [k, \infty)$, *then* $f^{'}(x_1) > 0$.

Proof. If $f(x_1) = 0$, then $\frac{1}{x_1-C} = \frac{1}{k}\sum_{i=1}^{k} \frac{1}{x_1-i+1}$. Since $g_1(x) = x^2$ is a convex function, by Jensen's Inequality we have

$$\frac{1}{k}\sum_{i=1}^{k} \frac{1}{(x_1-i+1)^2} > \Big(\frac{1}{k}\sum_{i=1}^{k} \frac{1}{x_1-i+1}\Big)^2 = \frac{1}{(x_1-C)^2},$$

i.e., $f^{'}(x_1) > 0$.

Proof of Theorem 1. Since $f(x)$ defined above is a continuous function, if $f(x_0) > 0$ for some $x_0 \geq k$, then $f(x) > 0$ for all $x \geq x_0$, otherwise the smallest x_1 with $f(x_1) = 0$ would satisfy $f^{'}(x_1) \leq 0$, which contradicts Lemma 2. From this we know that if x_1 is a solution of (7), then $f^{'}(x_1) > 0$ and $f(x)$ is positive for all $x > x_1$, so (7) can have at most one solution. If $f(k) > 0$, which is equivalent to $C > (\sum_{i=2}^{k} \frac{1}{i} / \sum_{i=1}^{k} \frac{1}{i})k$, then for all $x \in [k, \infty)$, $f(x) > 0$, and equation (7) has no solution. Notice that $f(x) = h(x)/\{(x-C)\prod_{i=1}^{k}(x-i+1)\}$, where $h(x) = k\prod_{i=1}^{k}(x-i+1) - (x-C)\sum_{i=1}^{k}[\prod_{j=1,j\neq i}^{k}(x-j+1)]$ is a polynomial function of degree $(k-1)$, with the coefficient of x^{k-1} being $-k\sum_{i=1}^{k}(i-1) - \sum_{j=1}^{k}[-\sum_{i=1}^{k}(i-1)+(j-1)-C] = kC - \{k(k-1)\}/2$. Therefore, if $C < (k-1)/2$ then for large enough x, $h(x)$, and hence $f(x)$, is less than 0. Since $f(k) < 0$, from Lemma 2 we know that $f(x)$ is always negative, so the solution does not exist. When $C = (k-1)/2$, the coefficient of x^{k-1} is 0, and it can be shown that the coefficient of x^{k-2} is

$$k\sum_{1\leq i<j\leq k}(i-1)(j-1) - \sum_{i=1}^{k}\sum_{1\leq j<l\leq k, j\neq i, l\neq i}(j-1)(l-1) - C\sum_{i=1}^{k}\sum_{j=1,j\neq i}^{k}(j-1),$$

which is equal to $-k(k+1)(k-1)/12$, and is negative. From Lemma 2 again we have $f(x) < 0$ for all $x \geq k$, which means the solution of (7) does not exist. Finally, when condition (8) is met, we have $f(k) \leq 0$, and for large enough x, $f(x) > 0$, so there is a unique solution. Apply these conclusions to the cases $k = 0, 1$ and $k \geq 2$ to obtain Theorem 1.

References

Andersen, P. K., Borgan, Ø, Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes.* Springer-Verlag, New York.

Becker, N. G. (1984). Estimating population size from capture-recapture experiments in continuous time. *Austral. J. Statist.* **26**, 1-7.

Becker, N. G. and Heyde, C. C. (1990). Estimating population-size from multiple recapture experiments. *Stochastic Process. Appl.* **36**, 77-83.

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

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

Falck, M. (1996). Small mammal population dynamics in riparian zones of regulated versus unregulated rivers in northwestern Colorado. M.S. Thesis. Colorado State University.

Godambe, V. P. (1985). The foundations of finite-sample estimation in stochastic processes. *Biometrika* **72**, 419-28.

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**, 485-94.

Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika* **76**, 133-140.

Huggins, R. M. (1991). Some practical aspects of a conditional likelihood approach to capture-recapture models. *Biometrics* **47**, 725-732.

Lin, D. Y. and Yip, P. S. F. (1999). Parametric regression models for continuous time removal and recapture studies. *J. Roy. Statist. Soc. Ser. B* **61**, 401-411.

Moran, P. A. P. (1951). A mathematical theory of animal trapping. *Biometrika* **38**, 307-11.

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.

Pollock, K. H., Nichols, J. D., Brownie, C. and Hines, J. E. (1990). Statistical inference for capture-recapture experiments. *Wildlife Monographs* **107**, 1-97.

Pollock, K. H. (1991). Modeling capture, recapture and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present and future. *J. Amer. Statist. Assoc.* **86**, 225-238.

Pollock, K. H. (2000). Capture-recapture models. *J. Amer. Statist. Assoc.* **95**, 293-296.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters.* 2nd edition. Macmillan, New York.

Yip, P. S. F. (1989). An inference procedure for a capture and recapture experiment with time dependent capture probabilities. *Biometrics* **45**, 471-479.

Yip, P. S. F. (1991). A martingale estimating equation for capture-recapture experiment in discrete time. *Biometrics* **47**, 1081-1088.

Yip, P. S. F., Fong, D. Y. T. and Wilson, K. (1993). Estimating population size by recapturing sampling via estimating function. *Stochastic Models* **9**, 179-193.

Yip, P. S. F., Huggins, R. and Lin, D. Y. (1996). An inference procedure for capture-recapture experiments in continuous time with variable capture rates. *Biometrika* **83**, 477-483.

Yip, P. S. F., Zhou, Y., Lin, D. Y. and Fang, X. Z. (1999). Estimation of population size based on additive hazards models for continuous-time recapture experiments. *Biometrics* **55**, 904-908.

LMAM, School of Mathematical Sciences, Peking University, Beijing, 100871, P.R. China.

E-mail: liping@math.ualberta.ca

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

E-mail: sfpyip@hku.hk

(Received May 2001; accepted July 2002)