Environmental and Ecological Statistics **7,** 405–414, 2000

Estimating the population size with a behavioral response in capture-recapture experiment

PAUL S.F. YIP

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

LIQUN XI

Mathematical Section, Zhejiang Institute of Silk and Textile, Hangzhou, China

ANNE CHAO & WEN-HAN HWANG

Institute of Statistics, National Tsing Hua University, Hsin-Chu, Taiwan 30043

Received September 1998; Revised December 1999

A new estimating procedure is suggested to estimate the population size in a capture-recapture experiment. The capture intensities for first-capture and recapture are allowed to be different and time dependent but they are assumed to be proportional. It is shown that the information on the proportionality constant is crucial to the estimation of the population size. Sensitivity analysis with a misspecification of the proportionality constant is conducted. The method has also been extended to the case with an unknown proportionality. A real example is given.

Keywords: behavioral effect, capture-recapture experiment, estimating function, population size estimation

1. Introduction

In a capture-recapture experiment, the capture probabilities of animals would often change after being captured. There are two possibilities: trap-shy or trap-happy. The former suggests the recapture probability is smaller than the probability of being caught first time whereas the latter assumes the recapture probability is larger than the probability being caught first time. It is important to consider the behavioral effect in estimating population size (Otis *et al.*, 1978). Here we allow the capture and recapture intensities to be different and time dependent but assume a proportionality relationship. The usual maximum likelihood estimation method would not be applicable since the intensity functions depend on time. An estimating function approach is used to estimate the population size ν with or without the information about the proportionality, θ. It is assumed that all individuals are catchable and the population is closed throughout the experimental period. Also, a

1352-8505 © 2000 Kluwer Academic Publishers

continuous time framework is assumed such that no two individuals can be caught at the same time. The use of martingale estimating equation in capture-recapture experiment can be found in Becker (1984), Yip (1991), Yip, Fong, and Wilson (1993). However, they assumed the intensities for capture and recapture are the same.

We will first derive estimators for ν with a known θ in Section 2 and the method is extended to an unknown θ in Section 3. The effect of misspecification of θ and a simulation study to assess the performance of the proposed estimators are given in Section 4. A real example of a bird banding data from Mai Po Bird Sanctuary is given in Section 5. The use of capture-recapture method in estimating population size also has application in software reliability, see Yip (1995, 1996), Yang and Chao (1995), Nayak (1988) and van Pul (1992).

2. Martingale estimating equations

Let M_t and K_t be the number of first-captures and recaptures in the population during the time period $[0, t]$. Consider a bivariate counting process $\{M_t, K_t; t\in[0, \tau]\}$ where τ is the termination time. The processes M_t and K_t are assumed to be right-continuous so that both $dM_t = M_t - M_{t-}$ and $dK_t = K_t - K_{t-}$ only takes value 0 or 1. Define $\mathscr{F}_t$ to be the σ-algebra generated by $\{M_s, K_s; 0 \leq s \leq t\}$. Then $\{\mathscr{F}_t; t\in[0, \tau]\}$ is a filtration which contains all the information of the process up to and including time t. There is also a pre-t-σ-algebra $\mathscr{F}_{t-}$ which is the smallest σ-algebra containing all $\mathscr{F}_s$ with $s<t$.

Note that, we have

$$E[dM_u|\mathscr{F}_{u-}] = \lambda_u(\nu - M_{u-})du$$

and

$$E[dK_u|\mathscr{F}_{u-}] = \beta_u M_{u-}du$$

where λ_u and β_u are the capture intensities for newly caught and recapture individuals respectively. Here we assume a proportional relationship between first-capture and recaptures, i.e., $\lambda_t = \theta\beta_t$ for all t where θ is some constant. We consider the martingale difference:

$$d\mathscr{D}_u = M_{u-}dM_u - \theta(\nu - M_{u-})dK_u \tag{1}$$

with $E[d\mathscr{D}_u|\mathscr{F}_{u-}] = 0$ since

$$\begin{aligned} E[d\mathscr{D}_u|\mathscr{F}_{u-}] &= M_{u-}E[dM_u|\mathscr{F}_{u-}] - \theta(\nu - M_{u-})E[dK_u|\mathscr{F}_{u-}] \\ &= M_{u-}\lambda_u(\nu - M_{u-})du - \lambda_u(\nu - M_{u-})M_{u-}du \\ &= 0. \end{aligned}$$

Then we can construct a stochastic integral

$$\mathscr{D}_t^* = \int_0^t W_{u-}d\mathscr{D}_u \tag{2}$$

where W_{u-} is any locally bounded and predictable process with respect to $\mathscr{F}_{u-}$, is a zero mean martingale (ZMM). In the case of known θ, a class of estimator for ν can be obtained by equating Equation (2) to zero and evaluated at τ, the termination time and is given by

$$\hat{\nu}_w = \int_0^\tau W_{u-}\{M_{u-}dM_u + \theta M_{u-}dK_u\} \Big/ \theta \int_0^\tau W_{u-}dK_u \tag{3}$$

The weight function W_{u-} can be chosen in an optimal way which minimizes the size of confidence bounds for the estimates as suggested in Godambe (1985). The optimal weight for a given martingale difference is given by

$$W_{u-} = \frac{E[\frac{d\mathscr{D}_u}{d\nu}|\mathscr{F}_{u-}]}{\mathrm{Var}[d\mathscr{D}_u|\mathscr{F}_{u-}]}.$$

Hence we have the optimal weight for Equation (2) in the case of θ is known:

$$\frac{1}{(\nu - M_{u-})\{M_{u-} + \theta(\nu - M_{u-})\}}.$$

The $\hat{\nu}$ is the solution of the following equation:

$$\int_0^\tau \frac{1}{(\nu - M_{u-})\{M_{u-} + \theta(\nu - M_{u-})\}}\{M_{u-}dM_u - \theta(\nu - M_{u-})dK_u\} = 0. \tag{4}$$

The variance process of $\mathscr{D}^*_\tau$ is given by

$$\langle \mathscr{D}^* \rangle_\tau = \int_0^\tau W^2_{u-}\{M^2_{u-}dM_u + \theta^2(\nu - M_{u-})^2 dK_u\}.$$

An estimate of the variance can be obtained by proper substitution of parameters. Under mild regularity condition on existence and behavior of the first and second moments. We have

$$\frac{\mathscr{D}^*_\tau}{\langle \mathscr{D}^* \rangle_\tau^{1/2}} \longrightarrow N(0,1).$$

In the case of $W_{u-} = 1$, an explicit expression for an estimate of ν is available, i.e.,

$$\hat{\nu}^*_1 = \int_0^\tau [M_{u-}dM_u + \theta M_{u-}dK_u] \Big/ \theta K_\tau.$$

In the case of $\theta = 1$ which implies capture and recapture have the same intensity, we obtain the result as in Becker (1984) and Yip (1991). Also, $\hat{\nu}^*_1$ can be rewritten as

$$\hat{\nu}^*_1 = \hat{\nu}_\tau + \left(\frac{1}{\theta} - 1\right) \int_0^\tau M_{u-}dM_u \Big/ K_\tau. \tag{5}$$

where $\hat{\nu}_\tau = \int M_{u-}dN_u/K_\tau$ and $dN_u = dM_u + dK_u$. Note that $\hat{\nu}_\tau$ is the population size estimator for the model suggested in Becker (1984) and Yip *et al.* (1993) under the assumption of $\theta = 1$. Also, in the trap shy situation (i.e., $\lambda_t > \beta_t$, and $\theta > 1$) $\hat{\nu}_\tau$ leads to over-estimation whereas in trap happy situation (i.e., $\lambda_t < \beta_t$, and $\theta < 1$) $\hat{\nu}_\tau$ leads to under-estimation. If the value of θ used in Equation (5) is larger (smaller) than the true value of θ, $\hat{\nu}^*_1$ leads to underestimation (overestimation) of ν. An approximate estimate of the standard error of $\hat{\nu}_w$ is given by

$$se(\hat{\nu}_w) = \frac{\langle \mathscr{D}^* \rangle_\tau}{\int_0^\tau W_{u-}dK_u}.$$

3. θ unknown

According to Godambe and Heyde (1987) we have two optimal weight functions for the martingale difference in Equation (1) with respect to ν and θ respectively:

$$\frac{1}{(\nu - M_{u-})\{M_{u-} + \theta(\nu - M_{u-})\}} \quad \text{and} \quad \frac{1}{\theta\{M_{u-} + \theta(\nu - M_{u-})\}}$$

for estimating ν and θ respectively. Substituting these optimal weights into (2), we have two estimating functions:

$$\mathscr{D}_1^*(\tau) = \int_0^\tau \frac{M_{u-}dM_u - \theta(\nu - M_{u-})dK_u}{(\nu - M_{u-})\{M_{u-} + \theta(\nu - M_{u-})\}} \tag{6}$$

$$\text{and} \quad \mathscr{D}_2^*(\tau) = \int_0^\tau \frac{M_{u-}dM_u - \theta(\nu - M_{u-})dK_u}{\theta\{M_{u-} + \theta(\nu - M_{u-})\}}. \tag{7}$$

Estimates for ν and θ denoted by $\hat{\nu}$ and $\hat{\theta}$ can be obtained by finding the roots of the Equation (6) and Equation (7).

To measure the precision of the estimates, we proceed as follows: Let $\boldsymbol{v} = (\nu, \theta)'$ and $P(\mathrm{v}) = (\mathscr{D}_1^*(\tau), \mathscr{D}_2^*(\tau))'$. Then, with obvious notation, by Taylor's series expansion of $P(\boldsymbol{v})$ at $\boldsymbol{v} = \hat{\boldsymbol{v}}$, we have

$$P(\boldsymbol{v}) \approx P(\hat{\boldsymbol{v}}) + P'(\hat{\boldsymbol{v}})(\boldsymbol{v} - \hat{\boldsymbol{v}})$$

where $P'(\hat{\mathrm{v}})$ denotes the first derivatives of $P(\boldsymbol{v})$ evaluated at $\hat{\boldsymbol{v}}$. Or equivalently,

$$(\boldsymbol{v} - \hat{\boldsymbol{v}}) \approx [P'(\hat{\boldsymbol{v}})]^{-1} P(\boldsymbol{v}).$$

Thus, a measure of the dispersion of $(\boldsymbol{v} - \hat{\boldsymbol{v}})$ can be obtained as

$$[P'(\hat{\boldsymbol{v}})]^{-1} V(\hat{\boldsymbol{v}}) [P'(\hat{\boldsymbol{v}})]^{-T} \tag{8}$$

where A^{-T} denotes the transpose of the inverse of a matrix A and $V(\hat{\boldsymbol{v}})$ is the dispersion matrix of $P(\hat{\boldsymbol{v}})$. Note that, we have

$$V(\boldsymbol{v}) = \begin{bmatrix} \langle \mathscr{D}_1^*, \mathscr{D}_1^* \rangle(\tau) & \langle \mathscr{D}_1^*, \mathscr{D}_2^* \rangle(\tau) \\ \langle \mathscr{D}_2^*, \mathscr{D}_1^* \rangle(\tau) & \langle \mathscr{D}_2^*, \mathscr{D}_2^* \rangle(\tau) \end{bmatrix}$$

where

$$\langle \mathscr{D}_1^*, \mathscr{D}_1^* \rangle(\tau) = \int_0^\tau \frac{M_{u-}^2 dM_u}{(\nu - M_{u-})^2 \{M_{u-} + \theta(\nu - M_{u-})\}^2} + \frac{\theta^2 dK_u}{\{M_{u-} + \theta(\nu - M_{u-})\}^2},$$

$$\langle \mathscr{D}_1^*, \mathscr{D}_2^* \rangle(\tau) = \int_0^\tau \frac{M_{u-}^2 dM_u}{\theta(\nu - M_{u-}) \{M_{u-} + \theta(\nu - M_{u-})\}^2} + \frac{\theta(\nu - M_{u-}) dK_u}{\{M_{u-} + \theta(\nu - M_{u-})\}^2}$$

and

$$\langle \mathscr{D}_2^*, \mathscr{D}_2^* \rangle(\tau) = \int_0^\tau \frac{M_{u-}^2 dM_u}{\theta^2 \{M_{u-} + \theta(\nu - M_{u-})\}^2} + \frac{(\nu - M_{u-})^2 dK_u}{\{M_{u-} + \theta(\nu - M_{u-})\}^2}.$$

Furthermore, we need $P'(\boldsymbol{v})$ which is

$$P'(\boldsymbol{v}) = \begin{bmatrix} \frac{\partial \mathscr{D}_1^*(\tau)}{\partial \nu} & \frac{\partial \mathscr{D}_1^*(\tau)}{\partial \theta} \\ \frac{\partial \mathscr{D}_2^*(\tau)}{\partial \nu} & \frac{\partial \mathscr{D}_2^*(\tau)}{\partial \theta} \end{bmatrix}$$

where

$$\frac{\partial \mathscr{D}_1^*(\tau)}{\partial \nu} = -\int_o^\tau \frac{M_{u-}\{M_{u-} + 2\theta(\nu - M_{u-})\}dM_u}{(\nu - M_{u-})^2\{M_{u-} + \theta(\nu - M_{u-})\}^2} + \int_0^\tau \frac{\theta^2 dK_u}{\{M_{u-} + \theta(\nu - M_{u-})\}^2},$$

$$\frac{\partial \mathscr{D}_1^*(\tau)}{\partial \theta} = -\int_0^\tau \frac{M_{u-}}{\{M_{u-} + \theta(\nu - M_{u-})\}^2}[dK_u + dM_u] = \frac{\partial \mathscr{D}_2^*(\tau)}{\partial \nu}$$

$$\frac{\partial \mathscr{D}_2^*(\tau)}{\partial \theta} = -\int_0^\tau \frac{M_{u-}\{M_{u-} + 2\theta(\nu - M_{u-})\}dM_u}{\theta^2\{M_{u-} + \theta(\nu - M_{u-})\}^2} + \int_0^\tau \frac{(\nu - M_{u-})^2 dK_u}{\{M_{u-} + \theta(\nu - M_{u-})\}^2}.$$

Thus the estimated variance-covariance matrix for $(\boldsymbol{v} - \hat{\boldsymbol{v}})'$ can be computed from Equation (8) by substituting all the unknown parameters with the corresponding estimates.

4. Sensitivity analysis and simulation

Here we give a Monte Carlo study for the case of a known θ and to examine the effect of a misspecification of the value of θ. Without loss of generality, the capture and recapture intensity functions are assumed to follow a homogeneous Possion process. Several combinations of the values are used in the simulation. The capture proportion ($P = 1 - \exp^{-\lambda\tau}$) are assumed to be 0.3, 0.5 and 0.9 and θ values vary from 0.5 to 1.5. The population size is 400 and the size of simulation is 2000 for each combination of the θ and P values. Table 1 give the results of using the optimal weight estimate of the averages and standard deviations of the 1000 estimates ave($\hat{\nu}$), and sd($\hat{\nu}$) respectively. Also, average of the standard errors of the estimates is also given, Ave(se($\hat{\nu}$)). The Ave(se($\hat{\nu}$)) should be close to sd($\hat{\nu}$). The coverage is determined by the proportion of the estimate lies between (± 2se($\hat{\nu}$)) constructed in each simulation. The performance of the estimator improved as P increased. With a correct value of θ, the estimator derived from Equation (4) performs satisfactorily with a proper 95% coverage. However, a misspecified θ leads to an overestimation or underestimation of ν in agreement with the observations in Equation (5). For a larger than expected value of θ gives a misleadingly smaller standard error and the coverage is very unsatisfactory.

Since θ is usually unknown, Table 2 gives results of the optimal weight $\hat{\nu}$ with an unknown θ. The simulation results for a known value of θ is also given for comparison. The capture proportion is assumed to be 0.3, 0.5, 0.7 and 0.9. There was a high proportion of breakdown with small capture proportion. The iterative procedure failed to find the roots of ν and θ from Equations (6) and (7). The columns in Table 2 are similar to Table 1. The last column ''failure'' represents the proportion of failure in the 2000 simulations. The simulated results were based on successful simulations when $P = 0.3$, about 42% of the simulations failed to provide estimates for θ and ν in the case of θ=0.5. Nevertheless the standard error of the estimate were much larger and sd(ν) and Ave(se($\hat{\nu}$)) were quite

Table 1. Simulation results with a known value of θ and its effect of misspecification ($\nu = 400$).

True value	θ	$Ave(\hat{\nu})$	$sd(\hat{\nu})$	$Ave(se(\hat{\nu}))$	Coverage
$P = 0.3$					
	0.5	1081.3	283.0	290.9	0.09
	1.0	581.0	142.5	145.2	0.99
$\theta = 1.5$	1.5	415.2	95.6	96.7	0.93
	1.7	376.4	84.5	85.3	0.85
	2.0	332.9	72.1	72.5	0.67
	0.5	745.8	168.6	161.8	0.37
	0.8	496.6	105.9	101.0	0.97
$\theta = 1.0$	1.0	413.9	85.0	80.7	0.94
	1.5	304.6	57.1	53.8	0.48
	2.0	250.8	43.2	40.4	0.13
	0.3	619.4	100.0	99.0	0.35
	0.4	484.6	75.6	74.2	0.89
$\theta = 0.5$	0.5	404.0	61.0	59.4	0.95
	1.0	245.1	31.9	29.8	0.03
	1.5	194.2	22.4	20.0	0.00
$P = 0.5$					
	0.5	924.7	121.3	128.3	0.00
	1.0	530.6	61.4	63.6	0.47
$\theta = 1.5$	1.5	402.1	41.4	42.2	0.96
	1.7	372.5	36.7	37.2	0.82
	2.0	339.7	31.5	31.6	0.46
	0.5	660.1	71.5	73.3	0.00
	0.8	464.8	45.6	45.6	0.79
$\theta = 1.0$	1.0	401.0	36.9	36.5	0.95
	1.5	318.3	25.6	24.4	0.15
	2.0	279.1	20.1	18.4	0.00
	0.3	571.5	46.7	49.6	0.02
	0.4	462.9	36.0	37.2	0.62
$\theta = 0.5$	0.5	398.7	29.6	29.7	0.94
	1.0	276.1	17.3	15.1	0.00
	1.5	240.1	13.7	10.4	0.00
$P = 0.9$					
	0.5	614.5	21.4	26.5	0.00
	1.0	445.3	11.3	12.4	0.02
$\theta = 1.5$	1.5	399.3	8.4	8.0	0.92
	1.7	390.2	7.8	7.1	0.68
	2.0	381.0	7.2	6.0	0.18
	0.5	502.0	13.3	16.1	0.00
	0.8	422.0	9.1	9.7	0.35
$\theta = 1.0$	1.0	399.5	7.9	7.7	0.93
	1.5	375.4	6.7	5.2	0.01
	2.0	367.0	6.4	4.1	0.00
	0.3	468.0	10.0	12.4	0.00
	0.4	422.9	8.3	9.1	0.26
$\theta = 0.5$	0.5	399.2	7.4	7.2	0.95
	1.0	365.7	6.3	3.9	0.00
	1.5	360.8	6.2	2.9	0.00

Table 2. Simulation results with an unknown value of $\theta(\nu = 400)$.

$P = 0.3$									
θ	$Ave(\hat{\nu})$	$sd(\hat{\nu})$	$Ave(se(\hat{\nu}))$	Cov	$Ave(\hat{\theta})$	$sd(\hat{\theta})$	$Ave(se(\hat{\theta}))$	Cov	$Failure$
1.5	270.6 (388.9)	113.4 (85.8)	516.7 (84.8)	0.77 (0.87)	3.2	1.41	5.39	1.00	0.67
1.0	245.1 (404.0)	105.9 (81.6)	326.4 (76.4)	0.60 (0.97)	2.8	1.43	3.33	1.00	0.51
0.5	239.5 (402.3)	96.7 (60.1)	222.5 (58.0)	0.53 (0.94)	1.5	0.92	1.25	0.97	0.42
$P = 0.5$									
θ	$Ave(\hat{\nu})$	$sd(\hat{\nu})$	$Ave(se(\hat{\nu}))$	Cov	$Ave(\hat{\theta})$	$sd(\hat{\theta})$	$Ave(se(\hat{\theta}))$	Cov	$Failure$
1.5	370.7 (399.8)	142.8 (40.9)	264.0 (41.4)	0.72 (0.96)	2.4	1.26	1.68	0.98	0.25
1.0	368.9 (399.3)	125.5 (35.8)	219.2 (35.8)	0.75 (0.95)	1.5	0.79	0.93	0.97	0.24
0.5	366.7 (398.2)	102.5 (30.2)	157.6 (29.4)	0.77 (0.94)	0.7	0.29	0.36	0.97	0.24
$P = 0.7$									
θ	$Ave(\hat{\nu})$	$sd(\hat{\nu})$	$Ave(se(\hat{\nu}))$	Cov	$Ave(\hat{\theta})$	$sd(\hat{\theta})$	$Ave(se(\hat{\theta}))$	Cov	$Failure$
1.5	420.3 (400.0)	94.4 (21.3)	106.0 (20.8)	0.88 (0.93)	1.6	0.60	0.63	0.97	0.03
1.0	412.1 (399.5)	81.0 (18.2)	84.5 (18.8)	0.88 (0.94)	1.1	0.34	0.36	0.96	0.02
0.5	405.4 (399.7)	66.0 (16.4)	64.6 (16.5)	0.90 (0.95)	0.5	0.14	0.15	0.96	0.03
$P = 0.9$									
θ	$Ave(\hat{\nu})$	$sd(\hat{\nu})$	$Ave(se(\hat{\nu}))$	Cov	$Ave(\hat{\theta})$	$sd(\hat{\theta})$	$Ave(se(\hat{\theta}))$	Cov	$Failure$
1.5	400.5 (399.3)	17.2 (8.4)	16.5 (8.0)	0.92 (0.93)	1.5	0.26	0.26	0.95	0.00
1.0	400.2 (399.5)	15.9 (8.0)	15.2 (7.7)	0.91 (0.93)	1.0	0.16	0.15	0.94	0.00
0.5	399.9 (399.2)	14.1 (7.4)	13.8 (7.2)	0.94 (0.95)	0.5	0.06	0.07	0.95	0.00

Note: Values inside the brackets represent the simulation results for ν when θ is known.

different from one another. When P increased the performance of the estimation improved as well. As expected, the estimation procedure for ν with unknown θ performed not as good as in the situation of known θ. The coverages are less than 95%. In the case of θ

known, the standard errors are much smaller and has proper 95% coverage. The information of θ is very crucial in determining the availability and the precision of $\hat{\nu}$. The proportion of failure in the simulations decreased when P increased.

5. Mai Po bird example

To illustrate the proposed methodology, we consider some banding data collected at the Mai Po Bird Sanctuary in Hong Kong for the year 1991. The Mai Po reserve is one of the very few wetlands in South China, which provides a haven to a wide variety of water birds and land birds. The particular bird we are interested in is a resident bird in that area called *Prinia subflava*. The banding exercise took place in the early mornings of the dry days. The same number of nets were set every time. The birds were ringed at the times of their first captures and then released immediately. A total of 237 birds were observed over the year. Of those, 176 were captured once, 44 twice, 11 three times, 5 four times and 1 five times. The behavioral response of the bird is expected to be minimal since banding was done by qualified ringers and every care is taken not to disturb their normal behavior before and after capture. The ring (with a unique number) was extremely light and small and was attached to its leg. For the recaptures, their identifications were noted by their rings. After initial examination the birds were released immediately.

Table 3 displays the optimal estimate weight for the population size associated with standard error for various values of θ which are assumed to be known. The estimates are quite sensitive for different values of θ.

In the case of an unknown θ, using Equations (6) and (7) we obtained the estimates for ν and θ, 493 (79.4) and 0.989 (0.29) respectively. The values insides the brackets represent the standard errors associated with the estimates. The behavioral effect was not statistically significant different from 1. The Mai Po biologists confirmed that the estimate around 500 is a reasonable estimate and the behavioral effect should be minimal too.

6. Discussions

The real example demonstrated that the effect of a misspecification of θ towards $\hat{\nu}$. Ideally, we would like to be able to illustrate the behavioral effect in the example, however, the effect of capture for *Prinia subflava* is minimal. Since θ is not significantly different from one, the estimator in Yip *et al*. (1993) can be used which is equivalent to Equation (2) after substituting $\theta = 1$. The optimal weight estimate for the population size for the Mai Po example would be 499.1 with a standard error 45.9 as indicated in Table 3.

The Monte Carlo study in Section 4 suggested the information of θ is crucial to

Table 3. Estimates and standard errors of ν when θ is assumed known for Mai Po Data.

θ	*0.5*	*1.0*	*1.5*
$\hat{\nu}$	819.4	499.1	396.0
$se(\hat{\nu})$	88.8	45.9	31.6

determine precision and availability of estimates. The existing estimating procedure for the behavioural model in discrete time is also very unstable for a small capture proprotion (Chao, 1987; Seber, 1982). Also, most of the capture-recapture data are recorded in discrete time owing to the lack of appropriate methods for analysing continuous time data. It may reduce efficiency and induce bias (Lin and Yip, 1999). The proposed procedure is another attempt to provide robust estimator for estimating population size in continuous time. It is true that in the case of an unknown θ, estimation of θ and ν would not be stable for small capture proportion. If we can provide the information about θ from some other sources, the procedure works well even for small capture proportion. Suggestions about performing auxillary experiment to obtain estimate for θ can be found in Lloyd and Chiaypiong (1997); Lloyd *et al.* (1998) and Yip (1995, 1996).

Acknowledgments

We would like to thank the referee for useful comments; Michael and Elizabeth Levens of the Mai Po Bird Sanctuary for the use of their data and for helpful discussions. This work was supported by the Research Grant Council of Hong Kong.

References

Becker, N.G. (1984) Estimating population size from capture-recapture experiments in continuous time. *The Australian Journal of Statistics*, **26**, 1–7.

Chao, A. (1987) Estimating the population size for capture-recapture data with unequal capturability. *Biometrics*, **43**, 783–91.

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

Godambe, V.P. and Heyde, C.C. (1987) Quasi-likelihood with optimal estimation. *International Statistical Review*, **55**, 231–44.

Lin, D.Y. and Yip, P.S.F. (1999) Parametric rgression models for continuous time removal and recapture studies. *Journal of Royal Statistical Society, Series B*, **61**, 401–11.

Lloyd, C.J. and Chaiyapong, C. (1997) Improved inference for recapture experiments with behavioural response through modelling and auxiliary experimentation. *The Australian Journal of Statistics*, **38**, 317–32.

Lloyd, C.J., Yip, P.S.F., and Chan, K.S. (1998) A comparison of recapturing, removal and weighting design for population size estimation. *Journal of Statistical Planning and Inference*, **71**, 363–75.

Nayak, T.K. (1988) Estimating population size by recapture sampling. *Biometrika*, **75**, 113–20.

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**, 1–135.

Pul, M.C. van (1992) Asymptotic properties of a class of statistical models in software reliability. *Scandinavian Journal of Statistics*, **19**, 235–53.

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

Yang M.C.K. and Chao, A. (1995) Reliability estimation and stopping rules for software testing based on repeated appearance. *IEEE Transaction on Reliability*, **44**, 315–21.

Yip, P. (1991) A martingale estimating equation for capture-recapture experiment in discrete time. *Biometrics*, **47**, 1081–88.

Yip, P., Fong, D.Y.T., and Wilson, K. (1993) Estimating population size by recapturing sampling via estimating function. *Stochastic Models*, **9**, 179–93.

Yip, P. (1995) Estimating the number of errors in a system – Using a Martingale Approach. *IEEE Transaction on Reliability*, **44**, 322–6.

Yip, P.S.F. (1996) Effect of plant-capture in a capture-recapture experiment. *Communications in Statistics*, **25**, 2025–38.

Biographical sketch

Paul Yip is a Senior Lecturer in Department of Statistics and Actuarial Science of the University of Hong Kong. This research is supported by a RGC research grant awarded to him in developing robust estimators for capture-recapture experiments with applicaton in ecology and epidemiology. His interests are in developing and applying statistical techniques in modeling ecological and biological systems.

Liqun Xi has a master degree in Mathematics from Zhejiang University, China. Currently he is a lecturer in Zhejiang Institute of Silk and Textile in China. His interest is in applying martingale theory to model the ecological and reliability systems.

Anne Chao is Professor, Institute of Statistics, National Tsing Hua University, Taiwan. Her interest in ecological statistics began with analyzing birds' data in estuaries of Taiwan. Specific methodological interests include species estimation, population size estimation and related biological samplings.

W. H. Hwang received PhD degree in statistics from National Tsing Hua University in 1997 under the supervision of Anne Chao. He is now in compulsory military service.