Statistica Sinica 11(2001), 843-853

A NOTE ON NONPARAMETRIC INFERENCE FOR CAPTURE-RECAPTURE EXPERIMENTS WITH HETEROGENEOUS CAPTURE PROBABILITIES

Richard M. Huggins and Paul S. F. Yip

La Trobe University and University of Hong Kong

Abstract: The estimation of the size of a population using capture-recapture data if the capture probabilities are heterogeneous is essentially an unsolved problem in that no nonparametric estimator has been shown to be uniformly applicable. Moreover, often the conditions under which the various nonparametric estimators give reasonable estimates are either unknown or uncheckable. Here a feasible nonparametric extension of a parametric empirical Bayes method used by Huggins (2001) is examined. It is concluded that assumptions concerning the distribution of the capture probabilities are required to successfully estimate the population size. A byproduct of the procedure is a conditional likelihood based test for time dependent capture probabilities in the presence of heterogeneous capture probabilities.

Key words and phrases: Bayes method, capture-recapture, Horvitz-Thompson estimator, logistic model.

1. Introduction

Many methods have been proposed to estimate the size of a population using capture-recapture data when the capture probabilities are heterogeneous. Here we examine an extension of a parametric approach of Huggins (2001) who assumed a Beta distribution for the capture probabilities. He then used empirical Bayes methods to estimate individual capture probabilities by the mean of the posterior distribution of the capture probabilities for each individual, and hence estimated the population size via a Horvitz-Thompson estimator. A parametric bootstrap, resampling from the estimated posterior distribution of the capture probabilities, was used to correct for the bias. The resulting estimators were comparable to the martingale estimators of Lloyd and Yip (1991) and the nonparametric sample coverage estimators of Chao, Yip, Lee and Chu (2001).

The work of Waclawiw and Liang (1993) suggests the possibility of a nonparametric extension of this approach. We explore it here. To do this we apply linear empirical Bayes techniques previously developed for random effects models to capture–recapture experiments. The random effects model for capture-recapture experiments has been examined in the literature (e.g., Sanathanan (1972), Darroch, Fienberg, Glonek and Junker (1993), Agresti (1994)) using different methods that those considered here. Our approach is based on the optimal linear estimators of binomial probabilities developed by Griffin and Krutchoff (1971) and Waclawiw and Liang (1993). This approach allows the estimation of time dependent effects and a random effect related to the individual capture probability for each captured individual. The resulting estimated capture probabilities may then be incorporated in the Horvitz-Thompson (1952) estimator to obtain an estimate of the population size.

From the point of view of estimating the population size, the results are disappointing. However, the transparent nature of the method and the lack of assumptions lead us to conclude that assumptions are required to satisfactorily estimate the size of a population using capture-recapture data. In compensation, the conditional likelihood procedure described in Section 3.1 allows us to test model $\mathcal{M}_{th}$, with heterogeneous capture probabilities and time dependence, versus model $\mathcal{M}_h$, which only assumes heterogeneous capture probabilities, of Otis, Burnham, White and Anderson (1978).

2. The Random Effects Model

One may view capture–recapture data as binary longitudinal or repeated measures data. Random effects models are commonly used in the analysis of longitudinal data to allow for heterogeneity (see Diggle, Liang and Zeger (1995)). Here we adapt these methods to a version of model M_{th} that allows capture probabilities to vary according to time and individual. The model analysed here is a random intercept model.

Consider a population of N individuals and t capture occasions. Let p_{ij} be the probability individual i is captured on occasion j, x_{ij} be a vector of time dependent covariates associated with individual i, and let η_i be an individual effect associated with individual i. Furthermore, let y_{ij} take the value 1 if individual i has been captured on occasion j and zero otherwise, and let c_i denote the number of times individual i is captured.

We consider the logistic model

$$p_{ij} = \frac{\exp(x'_{ij}\beta + \eta_i)}{1 + \exp(x'_{ij}\beta + \eta_i)} \tag{1}$$

and initially regard the η_i as nuisance parameters. This is just a random effects model (e.g., Diggle et al. (1995)) employed in the analysis of longitudinal binary data.

3. Inference

3.1. Inference on β

In the random effects model (1), β is estimated after conditioning on the sufficient statistics for the η_i, and the resulting contribution to the conditional likelihood of the ith individual is of the form

$$L_i = \frac{\prod_{j=1}^t p_{ij}^{y_{ij}}(1-p_{ij})^{1-y_{ij}}}{\sum_{r_i \in R_i} \prod_{j=1}^t p_{ij}^{r_{ij}}(1-p_{ij})^{1-r_{ij}}}, \tag{2}$$

where R_i is the set of all possible capture histories r_i with c_i captures, r_{ij} the jth element of r_i (Diggle et al. (1995)). Intuitively, L_i is the probability i is captured on the observed occasions given i is captured c_i times. The r_i are the capture histories with the same number of captures as i. For example, if there are 3 occasions and the observed capture history is 101, then R_i consists of the histories 101, 110, 011. The likelihood is then the probability i is captured on occasions 1 and 3 (101) given i is captured twice (101, 110, or 011).

After some simplification it is seen that

$$L_i = \frac{\exp(\sum_{j=1}^t x_{ij}'\beta y_{ij})}{\sum_{r_i \in R_i} \exp(\sum_{j=1}^t x_{ij}'\beta r_{ij})}, \tag{3}$$

which does not involve the individual effects η_i. Furthermore, note that any other effects which are common to all occasions are not included in the conditional likelihood. Moreover, for individuals that are not captured or are captured t times, $L_i \equiv 1$.

The conditional likelihood is then $L = \prod_{i=1}^n L_i$. We use conditional maximum likelihood estimates and their estimated standard errors arising from the conditional information matrix to test for homogeneity of the capture probabilities over time.

Remark. A special case of time dependent covariates occurring in capture–recapture studies is where the probability of capture varies from occasion to occasion. To model this situation, we take $p_{ij} = \exp(\beta_0 + \beta_j + \eta_i)/(1+\exp(\beta_0 + \beta_j + \eta_i))$ where $\beta_1 = 0$. Then L_i is given by

$$\frac{\exp(\sum_{j=1}^t(\beta_0+\beta_j)y_{ij})}{\sum_{r_i \in R_i} \exp(\sum_{j=1}^t(\beta_0+\beta_j)r_{ij})} = \frac{\exp(c_i\beta_0 + \sum_{j=2}^t \beta_j y_{ij})}{\sum_{r_i \in R_i} \exp(c_i\beta_0 + \sum_{j=2}^t \beta_j r_{ij})}$$
$$= \frac{\exp(\sum_{j=2}^t \beta_j y_{ij})}{\sum_{r_i \in R_i} \exp(\sum_{j=2}^t \beta_j r_{ij})}.$$

Thus, the conditional likelihood allows us to estimate the differences β_j, $j = 2, \ldots, t$ but not β_0.

3.2. Inference on the η_i

Having estimated β in the first step, we treat it as fixed and estimate the individual effects η_i for each captured individual. Let

$$p^*_{ij} = \frac{p_{ij}}{1 - \prod_{k=1}^t (1 - p_{ik})}$$

denote the probability individual i is captured on occasion j given it is captured at least once. Let C_i take the value 1 if individual i is captured at least once in the course of the experiment and 0 otherwise. Let $\mathcal{C} = \sigma\{C_1, \ldots, C_n\}$ be the σ-field generated by the C_i. In what follows we actually estimate $\eta_i + \beta_0$. For simplicity we take $\beta_0 = 0$ but, for example, we may regard the mean of the η_i as an estimate of β_0.

To estimate the η_i we modify an approach of Griffin and Krutchoff (1971) and Waclawiw and Liang (1993). We restrict our attention to estimators of p^*_{ij} of the form $a_j y_{ij} + b$ and choose a_j and b to minimize

$$E\Big(\sum_{j=1}^t \Big(a_j y_{ij} + b - p^*_{ij} \Big)^2 \Big| C_i = 1 \Big). \tag{4}$$

Note that a_j and b do not depend on i, unlike the approach of Waclawiw and Liang (1993) but consistent with that of Griffin and Krutchoff (1971). Taking derivatives of (4) with respect to a_k and setting the result equal to zero yields

$$0 = E\left(y_{ik}\left(a_k y_{ik} + b - p^*_{ik}\right) \Big| C_i = 1 \right) = E\left(a_k p^*_{ik} + b p^*_{ik} - p^{*2}_{ik} \Big| C_i = 1 \right),$$

so that

$$a_k = \frac{E(p^{*2}_{ik} | C_i = 1) - b E(p^*_{ik} | C_i = 1)}{E(p^*_{ik} | C_i = 1)}. \tag{5}$$

Taking the derivative of (4) with respect to b yields

$$0 = E\Big(\sum_{j=1}^t (a_j y_{ij} + b - p^*_{ij}) \Big| C_i = 1 \Big) = E\Big(\sum_{j=1}^t a_j p^*_{ij} + tb - \sum_{j=1}^t p^*_{ij} \Big| C_i = 1 \Big),$$

so that

$$b = \frac{\sum_{j=1}^t E(p^*_{ij} | C_i = 1)(1 - a_j)}{t}. \tag{6}$$

Substituting (5) into (6) yields

$$b = \frac{\sum_{j=1}^t \Big(E(p^*_{ij} | C_i = 1) - E(p^{*2}_{ij} | C_i = 1) \Big)}{t - \sum_{j=1}^t E(p^*_{ij} | C_i = 1)}. \tag{7}$$

Let n denote the number of distinct individuals captured, and n_j the number of distinct individuals captured on occasion j. Now, given $C_i = 1$, $y_{ij} \sim \text{bin}(1, p^*_{ij})$ so that

$$E(\frac{n_j}{n}|\mathcal{C}) = \frac{1}{n}\sum_{i=1}^{n} E(y_{ij}|C_i) = \frac{1}{n}\sum_{i=1}^{n} E(p^*_{ij}|C_i) = E(p^*_{1j}|C_1),$$

since the η_i are independently and identically distributed as are functions of the η_i, namely $E(p^*_{ij}|C_i)$. Hence n_j/n is an unbiased estimator of $E(p^*_{ij}|C_i)$.

We also need to estimate $E(p^{*2}_{ij}|C_i)$. As $y^2_{ij} = y_{ij}$, direct estimation of p^{*2}_{ij} is not possible. We must therefore make some assumptions. Here we suppose that $p^*_{ik_1} = \cdots = p^*_{ik_{r-1}} = p^*_{ij}$ for some $k_1, \ldots, k_{r-1} \neq j$, $k_s \in \{1, \ldots, m\}$, $s = 1, \ldots, r-1$. Let X_{ijr} denote the number of times i is captured on the r occasions $k_1, \ldots, k_{r-1}, j$. Then

$$E\Big(\frac{X^2_{ijr} - X_{ijr}}{r^2 - r}\Big| C_i, \eta_i\Big) = p^{*2}_{ij},$$

$$E\Big(\frac{1}{n}\sum_{i=1}^{n} \frac{X^2_{ijr} - X_{ijr}}{r^2 - r}\Big|\mathcal{C}\Big) = E(p^{*2}_{ij}|C_i).$$

Suppose that the r occasions $k_1, \ldots, k_{r-1}, j$ have been defined for each j (with possibly a different r for different j). Let $\hat{d}_j = \sum_{i=1}^{n}(X^2_{ijr} - X_{ijr})/(r^2 - r)$. Then we estimate b and a_j by

$$\hat{b} = \frac{\sum_{j=1}^{t}\left(n_j - n\hat{d}_j\right)}{nt - \sum_{j=1}^{t} n_j}, \qquad \hat{a}_j = \frac{n\hat{d}_j - \hat{b}n_j}{n_j},$$

respectively. Note that it is possible for the $\hat{a}_j$ to be negative. In some cases this can result in extremely low estimates of the capture probabilities, and hence large estimates of the population size. This was noted in our simulation study.

Finally, we estimate η_i by choosing $\hat{\eta}_i$ to minimize

$$R(\eta_i) = \sum_{j=1}^{t}\left(\hat{a}_j y_{ij} + \hat{b} - p^*_{ij}(\eta_i)\right)^2, \tag{8}$$

where $\hat{\beta}$ is fixed and p^*_{ij} is regarded as a function of η_i.

Remark. The assumption that that $p^*_{ik_1} = \cdots = p_{ik_{r-1}} = p^*_{ij}$ for some $k_1, \ldots, k_{r-1} \neq j$, $k_s \in \{1, \ldots, m\}$, $s = 1, \ldots, r-1$, necessary to estimate $E(p^{*2}_{ij}|C_i)$, requires some thought in practice. However, some guidance is given by the estimated values of β arising from Section 3.1. We noticed that these estimators were not particularly sensitive to how the grouping was done.

3.3. Estimating the population size

Having estimated β and the random effects η_i in the first two steps, we estimate the probability that individual i is captured on occasion j by

$$\hat{p}_{ij} = \frac{\exp(x'_{ij}\hat{\beta} + \hat{\eta}_i)}{1 + \exp(x'_{ij}\hat{\beta} + \hat{\eta}_i)}. \tag{9}$$

The probability an individual is captured at least once is $P_i = 1 - \prod_{j=1}^{t}(1 - p_{ij})$ and, if the P_i were known, we could estimate the size of the population using the Horvitz–Thompson estimator $\hat{N}_0 = \sum_{i=1}^{n} P_i^{-1}$. We could estimate P_i by $\hat{P}_i = 1 - \prod_{j=1}^{t}(1 - \hat{p}_{ij})$ and subsequently estimate N by $\hat{N}_1 = \sum_{i=1}^{n} \hat{P}_i^{-1}$.

4. An Approximate Standard Error for $\hat{N}$

The three step nature of the estimator makes an analytic derivation of the standard error of the estimator difficult. A bootstrap approach is feasible in theory, but the iterative procedures employed to estimate β and the η_i largely precludes its use. Here we give an approximate standard error. Let $\theta = (\beta', \eta_1, \ldots, \eta_n)$. Then applications of the Mean Value Theorem yields, for large N,

$$N(\hat{\theta}) \approx N(\theta) - S(\beta)' \left(\frac{dS(\beta)}{d\beta}\right)^{-1} \frac{dN(\beta)}{d\beta} - \sum_{i=1}^{n}(\hat{\eta}_i - \eta_i)P_i^{-2}\frac{dP_i}{d\eta_i}.$$

The Central Limit Theorem shows that asymptotically each of the terms in this approximation has a normal distribution (we have supposed the $\hat{\eta}_i$ are independent to approximate the distribution of the third term). A heuristic argument that the three terms are uncorrelated is that the second term has zero expectation given the sufficient statistics for the individual effects, and hence is uncorrelated with the other two terms. The individual effects are estimated conditional on capture at least once and hence are uncorrelated with the first term. The variances of the three terms may be estimated by $v_1 = \sum_{i=1}^{n}(1 - \hat{P}_i)/\hat{P}_i^2$, $v_2 = (dN/d\beta)'(dS(\beta)/d\beta)^{-1} = (dN/d\beta)$, $v_3 = \sigma_\eta^2 \sum_{i=1}^{n} P_i^{-4}(dP_i/d\eta_i)^2$. We then take an approximate standard error of $\hat{N}$ as $\sqrt{v_1 + v_2 + v_3}$.

5. Examples

5.1. Example 1

Our first example is a data set previously discussed in Otis et al. (1978, p.62) and Huggins (1989) concerning 173 animals captured on 10 occasions, capture occasions alternating between morning and evening. The methods of Otis et al. lead to a model where the capture probabilities are heterogeneous and depend on time. Huggins examined this data set and modelled the individual heterogeneity as a function of the sex and age category of the individuals. He found a difference in the capture probabilities between morning and evening with an increased probability of capture in the mornings, no evidence of a behavioural response to capture, and estimated the population size as 176.9 with a standard error of 2.01. Applying the methods of Chao et al. (1992) for model M_{th} to this data set using the program CAPTURE yielded an estimated population size of 170 with a standard error of 0.

Table 1 gives the estimated occasion effects arising from the first part of our procedure. As noted above, we need to make assumptions about the capture probabilities. We supposed $p_{i1} = p_{i3}$, $p_{i2} = p_{i4}$, $p_{i5} = p_{i7}$, $p_{i6} = p_{i8}$, $p_{i7} = p_{i9}$, and $p_{i8} = p_{i10}$. In Table 2 we give the mean and standard deviation of the estimated capture probabilities for each occasion. The estimated population size was 176.4, the approximate standard error was 2.47.

Table 1. Estimate occasion effects for Example 1, where E denotes evening and M denotes morning capture occasions.

Occasion	2	3	4	5	6	7	8	9	10
Time of day	E	M	E	M	E	M	E	M	E
effect	-0.17	-0.15	-0.41	0.14	-0.75	0.19	-0.92	0.19	-0.81
s.e.	0.23	0.24	0.23	0.19	0.24	0.21	0.25	0.22	0.24

Table 2. Mean and standard deviation of the estimated capture probabilities on each occasion for Example 1.

Occasion	1	2	3	4	5	6	7	8	9	10
average	0.38	0.34	0.34	0.28	0.41	0.22	0.42	0.19	0.42	0.21
St.Dev.	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

5.2. Example 2

We also consider an example concerning the capture of deer mice by S. Hoffmann, originally examined in Otis et al. (1978). A total of 110 animals were captured on 5 occasions. The data is included with the program CAPTURE. Otis et al. determined that the most appropriate model for this data is M_{tbh} which allows capture probabilities to vary by time, capture history and individual. Using model M_b the program CAPTURE estimated the populations size as 142 with standard error 16.4. Model M_t gave 113 with standard error 1.76. Fitting model M_{th}, according to Chao (1992), gave an estimate of 139 with a standard error of 9.6.

In the first part of our procedure, the estimated occasion effects, with standard errors in parentheses, were 0.66 (0.28), 0.81 (0.28) 1.07 (0.29) and 0.73

(0.28), suggesting that the capture probabilities increased after the first occasion. We supposed $p_{i1} = p_{i2}$ and $p_{i3} = p_{i4} = p_{i5}$. Our procedure resulted in an estimate of the population size of 114.6 with a standard error of 3.1, similar to the estimate arising from model M_t.

6. Simulations

Our simulations were based on a population of 100 individuals and 5 capture occasions. We let $\beta_2, \ldots, \beta_5$ take the values 0.0, -0.3, -0.3, -0.3 for all simulations. The capture probabilities were given by (1). The random effects η_i were independent observations from $N(\mu, \sigma^2)$ distributions for various values of μ and σ^2. At each step we simulated a population of 100 individuals and then generated a capture experiment for each population. To estimate the $\hat{\eta}_i$ we pooled occasions 1 and 2 and occasions 3, 4 and 5. We conducted 100 simulations and report the mean of $\hat{N}_1$, the standard deviation of the estimates and the average of the estimated standard errors. We also give the average and standard deviations of the probabilities of being captured at least once, which are computed separately by simulating 1000 random effects and taking the average of the corresponding probabilities.

Table 3. Models for the simulations using normal random effects, average and standard deviation of the probability an individual is captured at least once, and the means of the estimators from 100 simulations. The model for the capture probabilities is a Beta distribution with no occasion effect to simulate the captures. N_H is the parametric empirical Bayes estimator with a bootstrap bias correction of Huggins (2001). (* 4 extremely large estimates omitted **. In 18 cases the variance could not be computed.)

α,β	10,10	5,5	3,5	3,10
av(P_i), (s.e.(P_i))	0.95 (0.05)	0.94 (0.08)	0.84 (0.16)	0.67 (0.20)
av($\hat{N}_1$)	98.45	96.10	90.60	88.50*
s.d.($\hat{N}_1$)	2.32	2.53	4.02	11.8
av(s.e.($\hat{N}_1$))	2.13	2.04	3.76	8.97**
av($\hat{N}_H$)	100.1	99.9	97.7	103.6
s.d.($\hat{N}_H$)	3.4	3.9	13.6	22.0
av(s.e.($\hat{N}_H$))	5.18	6.4	11.8	17.2

The models from which the data was simulated are summarized in Table 3(a), as are the results of the simulations. We also simulated capture probabilities according to a beta distribution. In this case we set $\beta_j = 0$, $j = 1, \ldots, 5$ for the simulations but estimated it in the inference procedure. The results are given in Table 3(b). For comparison we also give some results concerning the estimator of

Huggins (2001). We note that the performance of the estimator is disappointing when the mean capture probability decreases and the variability increases. The proposed estimator of the standard error tends to slightly underestimate the standard deviation of the simulated population size estimates.

7. Discussion

The poor performance of the estimator when the capture probabilities are low is common amongst nonparametric estimators of population size (for example, see the simulations of Chao et al. (2001)). However, the performance of the current estimator is disappointing when compared with the estimator of Huggins (2001). The performance of the latter estimator was consistent with that of a range of estimators including the optimal estimating function estimators of Chao et al. (2001) and the martingale estimators of Lloyd and Yip (1991). In the Chao et al. simulations, the percent bias in the four cases (Beta(10,10), Beta(5,5), Beta(3,5) and Beta(3,10)) examined were: for the jackknife, 7.0%, 7.2%, 5.5%, 7.5%; for the estimator of Chao et al., −0.7%, −0.5%, −2.7% and −4.0%; for the estimator of Huggins, 0.01%, −0.01%, −2.7% and 3.6%. The percent biases for the method examined here were −1.6%, −3.9%, −9.4% and −11.5%. Thus our nonparametric approach with minimal assumptions compares poorly with these other methods.

The difficulties with our approach appear to arise from assigning all individuals with the same capture histories the same capture probabilities: in our case, the mean of the posterior distribution. In reality, the conditional distributions of the capture probabilities given the capture history may be quite variable. The result of this is that, unlike the parametric bootstrap approach of Huggins (2001) where it was possible to resample from the posterior distribution rather than just the posterior means, it is not possible to use a bootstrap procedure to estimate and correct for the bias. This was also evident in further simulations that are not reported here. Thus it appears that to improve the estimator, assumptions that allow more detailed estimation of the posterior distribution of the capture probabilities are required.

We note that assumptions are implicit in other approaches. The sample coverage estimators of Chao (e.g., Chao et al. (1992), Lee and Chao (1994), Chao et al. (2001)) depend on the validity of an approximation. Huggins and Chao (2000) show that for some Beta distributions the bias in the sample coverage estimators of Chao et al. (2001) can be quite large. The jackknife estimator of Burnham and Overton (1978) requires conditions but Cormack (1989) pointed out that there is no theoretical advantage in using their estimator. The approaches of Lloyd and Yip (1991), Yip (1991) and Huggins (2001) make the explicit assumption that the capture probabilities have a Beta distribution. Norris and Pollock (1996) place lower bounds on the capture probabilities and an upper bound on the estimated population size.

Combined with the results of Huggins (2001), this work emphasizes that in order to estimate the size of a population using capture–recapture data some assumptions on the capture probabilities are required. As noted above our approach does yield, through inference on β in Section 3.1, a test for Model $\mathcal{M}_{th}$ against model $\mathcal{M}_h$ of practical importance. Moreover, the variability of the posterior means provides information on the degree of heterogeneity in the capture probabilities.

Acknowledgements

This work was supported by the Research Grant Council of Hong Kong (Ref. 7261/97M). Part of the work was done when R. M. Huggins was visiting the Department of Statistics and Actuarial Science of the University of Hong Kong. The authors are grateful to the editor, associate editor and the referee for their constructive criticisms and suggestions.

References

Agresti, A. (1994). Simple capture-recapture models permitting unequal catchability and variable sampling effort. *Biometrics* **50**, 494-500.

Burnham, K. P. and Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary amongst animals. *Biometrika* **65**, 625-633.

Chao, A., Lee, S.-M. and Jeng, S.-L. (1992). Estimating population size for capture–recapture data when capture probabilities vary by time and individual animal. *Biometrics* **48**, 201-216.

Chao, A., Yip, P. S. F., Lee, S.-M. and Chu, W. (2001). Population size estimation based on estimating functions for closed capture-recapture models. *J. Statist. Plann. Infer.* **92**, 213-232.

Cormack, R. M. (1989). Log-linear models for capture–recapture. *Biometrics* **45**, 396-413.

Darroch, J. N., Fienberg, S. E., Glonek, G. F. V. and Junker, B. W. (1993). A three-sample multiple-recapture approcah to census popualtion estimation with heterogeneous catchability. *J. Amer. Statist. Assoc.* **88**, 1137-1148.

Diggle, P. J., Liang, K. Y. and Zeger, S. L. (1995). *Analysis of Longitudinal Data.* Oxford University Press.

Griffin, B. S. and Krutchoff, R. G. (1971). Optimal linear estimators: an empirical Bayes version with application to the binomial distribution. *Biometrika* **58**, 195-201.

Horvitz, D. G. and Thompson, D. J. (1952) A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47**, 663-685.

Huggins, R. M. (2001). A parametric empirical Bayes approach to the analysis of capture–recapture experiments. To appear in *Aust. New Zealand J. Statist.*

Huggins, R. M. and Chao, A. (2000). Asymptotic properties of an optimal estimating function approach to the analysis of mark recapture data. *Submitted.*

Lee, S.-M. and Chao, A. (1994). Estimating population size via sample coverage for closed capture–recapture models. *Biometrics* **50**, 88-97.

Lloyd, C. J. and Yip, P. (1991). A unification of inference from capture-recapture studies through martingale estimating functions. In *Estimating Functions* (Edited by V. P. Godambe), 65-88. Oxford University Press.

Norris, J. L. and Pollock, K. H. (1996). Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* **52**, 639-649.

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

Sanathanan, L. (1972). Models and estimation methods in visual scanning experiment. *Technometrics* **14**, 813-829.

Waclawiw, M. A. and Liang, K. Y. (1993). Prediction of random effects in the generalized linear model. *J. Amer. Statist. Assoc.* **88**, 171-178.

Yip, P. (1991). An inference procedure for a capture–recapture experiment in discrete time with variable capture probabilities. *Stochastic Models* **7**, 343-362.

School of Statistical Science, La Trobe University, Bundoora 3083, Australia.

E-mail: huggins@latrobe.edu.au

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

E-mail: sfpyip@hku.hk

(Received March 2000; accepted February 2001)