Title	Statistical analysis of removal experiments with the use of auxillary variables
Author(s)	Huggins, RM; Yip, PSF
Citation	Statistica Sinica, 1997, v. 7 n. 3, p. 705-712
Issued Date	1997
URL	http://hdl.handle.net/10722/45345
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Statistica Sinica **7**(1997), 705-712

STATISTICAL ANALYSIS OF REMOVAL EXPERIMENTS WITH THE USE OF AUXILLARY VARIABLES

Richard M. Huggins and Paul S. F. Yip

La Trobe University and University of Hong Kong

Abstract: Conditional likelihood methods are applied to data from removal experiments to both model capture probabilities in terms of auxillary variables, corresponding to observable individual characteristics and environmental conditions, and to estimate the size of the population. Asymptotic properties of the resulting estimators are derived and the methods are applied to a data set which include time dependent covariates. A simulation study is also given.

Key words and phrases: Covariates, log-linear models, removal method.

1. Introduction

The use of auxillary variables in the analysis of removal experiments has been previously considered in Pollock, Hines and Nichols (1984) who consider models which include environmental conditions and catch effort. Their methods also allow the use of individual characteristics which are analysed by constructing categories based on these variables and then estimating the size of each category. For example this allows varying capture probabilities for males and females or large and small individuals. The categories can however become restrictive if there are several characteristics related to capture probabilities. A further difficulty with their method is the somewhat arbitrary rescaling of the variance estimates (see Pollock et al. (1984), p.355). The conditional likelihood approach to capture-recapture experiments of Huggins (1989, 1991) enables the modelling of capture probabilities in terms of observable characteristics of the captured individuals directly and of the environment which removes the need for the construction of categories and provides asymptotically valid estimates of variance. Here this approach is extended to removal experiments by conditioning on the removal of an individual and using information on which occasion the individual was removed to estimate the parameters in the model for the removal probabilities.

In Section 2 the conditional likelihood is constructed and its use in estimating the population size is discussed in Section 3. The asymptotic properties of the resulting estimator is discussed in the Appendix. To illustrate these methods we consider a banding recoveries data set collected at Mai Po Bird Sanctuary in

Hong Kong in 1992 and a recapture data set from East Stuart Gulch Colorado in U.S.A. in Section 4. A simulation study to examine the performance of the estimator is given in Section 5.

2. The Conditional Likelihood

Suppose the removal experiment consists of t distinct removal occasions where individuals are captured and then removed from a closed population. The individuals are supposed to behave independently of one another and the individual covariates are supposed to be independently and identically distributed so that, in particular, the individual covariates are independent observations from a common distribution. Let $\beta = (\beta_1^T, \beta_2^T)^T$ be a vector of parameters. Take x_i to be the vector of characteristics corresponding to the ith individual and let z_j be a vector of environmental characteristics. The probability that individual i is captured and removed from the population on occasion j given he has not been removed before j is denoted by p_{ij} and let $q_{ij} = 1 - p_{ij}$. We assume a log linear model, i.e.

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_1^T x_i + \beta_2^T z_j, \quad \text{for} \quad i = 1, \ldots, \nu, \quad j = 1, \ldots, t,$$

where ν denotes the population size, and assume that individuals act independently of one another.

Let $\{1, \ldots, n\}$ denote the captured individuals and $\{n+1, \ldots, \nu\}$ the uncaptured individuals. Let $c_{ij} = 1$ if individual i is captured on occasion j and be 0 otherwise and $c_i = 1$ if individual i is captured in the course of the experiment and be 0 otherwise.

The full likelihood is given by

$$L = \prod_{i=1}^{n} \prod_{j=1}^{t} (q_{i1} \ldots q_{i,j-1} p_{ij})^{c_{ij}} \prod_{i=n+1}^{\nu} (q_{i1} \ldots q_{it})^{(1-c_i)}$$

which may be rewritten as

$$L = \Big[\prod_{i=1}^{n} \frac{\prod_{j=1}^{t} (q_{i1} \ldots q_{i,j-1} p_{ij})^{c_{ij}}}{(1 - q_{i1} \ldots q_{it})}\Big] \prod_{i=1}^{\nu} (q_{i1} \ldots q_{it})^{1-c_i} (1 - q_{i1} \ldots q_{it})^{c_i}.$$

Since no information about the covariates of the uncaught individual is available, inferences concerning β are based on the conditional likelihood

$$L^* = \prod_{i=1}^{n} \frac{\prod_{j=1}^{t} (q_{i1} \ldots q_{i,j-1} p_{ij})^{c_{ij}}}{(1 - q_{i1} \ldots q_{it})} = \prod_{i=1}^{\nu} \frac{\prod_{j=1}^{t} (q_{i1} \ldots q_{i,j-1} p_{ij})^{c_{ij}}}{(1 - q_{i1} \ldots q_{it})^{c_i}}, \tag{1}$$

where the later form of the conditional likelihood is used in the Appendix below to derive the asymptotic properties of the estimators.

The maximum conditional likelihood estimators of β are then obtained from maximising the conditional likelihood in (1). The asymptotic properties of these estimators are easily derived and are outlined in the Appendix. These properties allow the use of a likelihood ratio tests to test between models as in Section 2.4 of Pollock et al. (1984).

3. Estimating the Population Size

To estimate the population size the approach of Huggins (1989) is followed. Essentially it is a Horvitz-Thompson estimator (1952). Let p_i be the probability that individual i is captured in the course of the removal experiment. Then $p_i = 1 - q_{i1}q_{i2}\ldots q_{it}$. If p_i is known, an unbiased estimate of the population size is given by

$$\hat{\nu}(\beta) = \sum_{i=1}^{n} p_i^{-1}$$

which has variance $\sigma^2 = \sum_{i=1}^{\nu}(1-p_i)/p_i$ and an unbiased estimator of σ^2 is $s^2 = \sum_{i=1}^{n}(1-p_i)/p_i^2$. In practice we estimate ν by $\hat{\nu}(\hat{\beta}) = \sum_{i=1}^{n}\hat{p}_i^{-1}$ where $\hat{p}_i$ are evaluated at the estimated β. The variance of $\hat{\nu}(\hat{\beta})$ includes a contribution from the variance of $\hat{\beta}$ and in the Appendix the estimate of the asymptotic variance of $\hat{\nu}(\hat{\beta})$ is shown to be $s^2+D(\hat{\beta})^T I(\hat{\beta})^{-1}D(\hat{\beta})$, where $I(\hat{\beta})$ is the observed information matrix and $D(\hat{\beta}) = d\hat{\nu}(\beta)/d\beta$ evaluated at the estimated value β. Note that it is possible for an experiment to fail and valid estimation of ν not be possible. Seber and Whale (1970) and Otis et al. (1978), p.113 give conditions under which the generalised removal model fails but for our models this is still an open question.

4. Examples

Example 1. We consider bandings of a species *Prinia inornate* collected at the Mai Po Bird Sanctuary in Hong Kong in 1992. During the year 105 birds were banded and there were no recaptures after the first four months suggesting the population size is very close to 105. In fact using the full data set the methods of Huggins (1989, 1991) estimated the popualtion size as 105 with an estimated standard error of 10^{-4}. We regard the first banding as a removal and consider the first four months data. Available covariates were the weight of each bird and an age variable indicating whether the bird was aged less than one year at capture. Usually the age of a bird is determined by its color and length of feathers. In order to introduce occasion effects we allowed different removal probabilities in the last two months compared with the first two. The likelihood ratio test showed no significant relationship between the covariates and the removal probabilities (full model, -log-likelihood=104.9, 4 parameters: constant removal probabilities, -log-likelihood=105.9, 1 parameter) and the resulting estimate of the population size was 107.9 with an estimated standard error of 2.24.

Table 1. Removals of peromyscus maniculatus.

m	y	12	1	m	y	13	1	f	y	5	2	m	sa	16	4
f	y	15	1	f	a	22	1	f	a	20	2	f	y	11	4
m	y	15	1	m	y	14	1	m	y	12	2	m	y	14	5
m	y	15	1	m	y	11	1	f	y	6	3	f	y	11	5
m	y	13	1	f	y	10	1	f	a	22	3	m	a	24	5
m	a	21	1	f	a	23	2	f	y	10	3	m	y	9	6
m	y	11	1	f	y	7	2	f	y	14	3	m	sa	16	6
m	sa	15	1	m	y	8	2	f	a	19	3	f	a	19	6
m	y	14	1	m	a	19	2	f	a	19	3				
m	y	14	1	m	y	13	2	f	a	20	4				

The columns represent, respectively, the sex (m or f), the ages (y = young, sa = semi-adult, a = adult), the weights in grams, and the occasion the individual was removed over 6 trapping occasions.

Example 2. The data of Table 1 is extracted from Appendix 1 of Huggins (1991) who reports a data set on the captures of *peromyscus maniculatus* collected by V. Reid at East Stuart Gulch Colorado and distributed with the program CAPTURE of Otis et al. (1978). Here the first captures are treated as removals and the capture probabilities are modelled as functions of the sex, age and weight of the individuals. Using this full model the population size was estimated at 43.69 with a standard error of 3.04 which is quite close to the estimates given by Otis et al. (1978) using their model $M_b(\hat{\nu} = 41, \text{se}(\hat{\nu}) = 3.01)$. However the estimate and its standard error are somewhat less than that of 47.1 with a standard error of 7.2 obtained in Huggins (1991) using the full capture histories. A likelihood ratio test revealed that in the removal model, unlike the capture-recapture model based on the full capture histories, none of the covariates were significantly related to the first capture probability. Using this reduced model the estimate of the population size and the associated standard error model remained essentially the same.

The discrepancy between the estimate using the full capture histories and the removal model which only uses the first capture occasion is of some interest as one would expect the estimate which uses the most information to have the lesser variance. The estimates of the population size and its associated standard error of Huggins (1991) were inflated by contributions from two young females with low weights which give them a leverage effect. Thus, it appears the use of removal methods applied to first captures may aid in the detection of influential individuals.

5. Simulations

In order to examine the performance of the estimators we conducted a small simulation study. We considered a population of 300 individuals whose removal probabilities depended upon the sex (SEX) and weight (WT) of the individual and allowed an environmental covariate (ENV) for each occasion. The simulated individuals were assigned their sex with probability $\frac{1}{2}$ and the weights were normally distributed with mean 3 and variance 1. The environmental covariates were normal with mean 2 and variance 1. All covariates were recomputed for each simulation. We used the logistic model for the removal probabilities,

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = -7 - 1 \times SEX + 1.5 \times WT + 1 \times ENV.$$

We considered 4, 6 and 8 removal occasions. On several occasions the maximum likelihood algorithm did not converge but stopped due to roundoff error, which we detected by the information matrix not being positive definite and these simulations were excluded. Only 1% of the number of simulation for each trial is excluded. The percentage was even less when t is 6 or 8. Further, the simulations did not include the modelling procedure which a statistician would employ and we noted that the use of an over parameterized model can result in grossly inflated estimates of the population size. This largely contributes to the large tails observed in the simulations. The results of the 1000 simulations are reported in Table 2 and Figure 1. It is clear from Table 2 and Figure 1 that whilst the bulk of the estimates are close to the true population size, the distribution of the estimator is not normal even for the moderately large population sizes considered in the simulations. Note that despite the non-normality of the estimators the coverage probabilities of the resulting confidence intervals are reasonable. However, in practice a bootstrap procedure to construct confidence intervals may be preferable.

Table 2. Simulation results for $\nu = 300$.

Number of occasions:	4	6	8
Number of captures:	206.3 (27.4)	232.1 (20.2)	246.6 (16.2)
$\hat{\nu}^{+}$:	836.8 (6663.4)	340.5 (207.1)	319.2 (84.0)
$[\hat{\nu} < 1000^{+}]$:	336.8 (126.6)	324.3 (79.4)	316.6 (59.4)
Median ($\hat{\nu}$):	304.8	279.5	286.2
$se(\hat{\nu})^{\sharp}$:	3188	112.1	49.5
$(< 1000)^{\sharp}$:	137.7	72.0	45.6
Coverage* (%):	88	90	92

Note: $^{+}$ We give the mean over the simulations, with the standard deviation in parentheses. $^{\sharp}$ The mean of the estimated standard errors. * Coverage is the percentage of the 95% confidence intervals constructed using the asymptotic distribution of the estimators which contains the true value of $\hat{\nu}$.

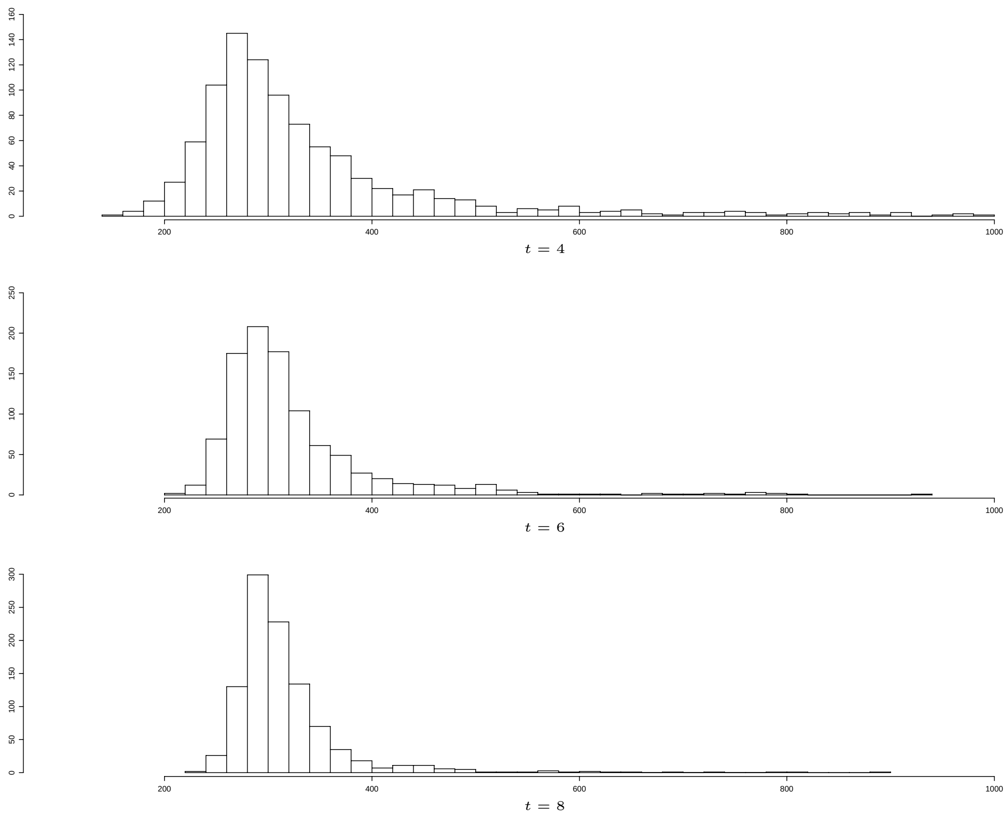

Figure 1. Histograms of estimated population sizes for $\hat{\nu} < 1000$.

Acknowledgements

The authors would like to thank the associate editor and two referees for very useful comments. This work is supported by Research Grant Council of Hong Kong Government. Part of the work was done when R.M. Huggins visited the Department of Statistics of the University of Hong Kong. The authors are also grateful to the World Wild Fund for Nature Hong Kong for providing the Mai Po data and David Melville for helpful discussion.

Appendix. Asymptotic Properties of the Estimates

As the methods are similar to those of Huggins (1989) we only briefly outline the derivation of our results. Let

$$S(\beta) = \frac{d \log L^*}{d\beta} = \sum_{i=1}^{\nu} \Big[\sum_{j=1}^{t} \frac{c_{ij}}{(q_{i1} \ldots q_{i,j-1} p_{ij})} \frac{d(q_{i1} \ldots q_{i,j-1} p_{ij})}{d\beta} + \frac{c_i}{(1 - q_{i1} \ldots q_{it})} \frac{d(q_{i1} \ldots q_{it})}{d\beta} \Big]$$

be the derivative of the log-likelihood with respect to β. Then, under assumptions that the individuals behave independently of one another and that the individual covariates are independently and identically distributed, $S(\beta)$ is the sum of independently and identically distributed random variables.

Now $E(c_{ij}) = q_{i1}\dots q_{i,j-1}p_{ij}$, $E(c_i) = 1 - q_{i1}\dots q_{it}$ and $p_{ij} = 1 - q_{ij}$, and we may write

$$\sum_{j=1}^{t}\Big[\frac{d(q_{i1}\dots q_{i,j-1})}{d\beta} - \frac{d(q_{i1}\dots q_{ij})}{d\beta}\Big] = -\frac{d(q_{i1}\dots q_{it})}{d\beta}.$$

Thus,

$$E(S(\beta)) = \sum_{i=1}^{\nu}\sum_{j=1}^{t}\Big[\frac{d(q_{i1}\dots q_{i,j-1}p_{ij})}{d\beta} + \frac{d(q_{i1}\dots q_{it})}{d\beta}\Big] = 0,$$

Similar arguments show that

$$E\Big(\frac{dS(\beta)}{d\beta}\Big) = -\sum_{i=1}^{\nu}\Big[\sum_{j=1}^{t}(q_{i1}\dots q_{i,j-1}p_{ij})^{-1}\Big[\frac{d(q_{i1}\dots q_{i,j-1}p_{ij})}{d\beta}\Big]^{\otimes 2} - (1 - q_{i1}\dots q_{it})^{-1}\Big[\frac{d(q_{i1}\dots q_{it})}{d\beta}\Big]^{\otimes 2}\Big],$$

where for a vector a, $a^{\otimes 2} = aa^T$ and it is straightforward to show that $E(S(\beta)^{\otimes 2}) = -E(dS(\beta)/d\beta) = I(\beta)$. To obtain asymptotic properties of our estimators we proceed as in Huggins (1989) and suppose we have an increasing sequence of populations. Then as in Huggins (1989) after noting that for β^+ between $\hat\beta$ and β, $(\hat\beta - \beta) = -[dS(\beta)/d\beta]^{-1}_{\beta^+}S(\beta)$, we may conclude that the maximum conditional likelihood estimators, β are consistent and asymptotically normal with variance and covariance matrix $I(\beta)^{-1}$. Next, note that

$$\begin{aligned}\mathrm{Cov}\,(S(\beta), \hat\nu(\beta)) &= E\Big[\sum_{i=1}^{\nu}\sum_{j=1}^{t}\Big(\frac{c_i c_{ij}}{(q_{i1}\dots q_{i,j-1}p_{ij})p_i}\frac{d(q_{i1}\dots q_{i,j-1}p_{ij})}{d\beta} + \frac{c_i^2}{(1 - q_{i1}\dots q_{it})p_i}\frac{d(q_{i1}\dots q_{it})}{d\beta}\Big)\Big]\\ &= \sum_{i=1}^{\nu}\frac{1}{p_i}\Big[\sum_{j=1}^{t}\Big(\frac{d(q_{i1}\dots q_{i,j-1}p_{ij})}{d\beta} + \frac{d(q_{i1}\dots q_{it})}{d\beta}\Big)\Big] = 0,\end{aligned}$$

arguing as above, where $E(c_i c_{ij}) = E(c_{ij})$ and $E(c_i^2) = E(c_i)$. Hence $S(\beta)$ and $\hat\nu(\beta)$ are uncorrelated. Now $\hat\nu(\beta)$ is the sum of independently and identically distributed random variables so that its asymptotic distribution is easily seen to be normal with mean ν and variance σ^2. Next, for some β^* and β^+ both between β and $\hat\beta$ we may write

$$\hat\nu(\hat\beta) = \hat\nu(\beta) + \left[\frac{d\hat\nu(\beta)}{d\beta}\right]^T_{\beta^*}(\hat\beta - \beta) = \hat\nu(\beta) - \left[\frac{d\hat\nu(\beta)}{d\beta}\right]^T_{\beta^*}\left[\frac{dS(\beta)}{d\beta}\right]^{-1}_{\beta^+}S(\beta).$$

We have shown that $\hat\nu(\beta)$ and $S(\beta)$ are uncorrelated and hence, we have asymptotically, $\mathrm{var}(\hat\nu(\hat\beta)) = \sigma^2 + D(\beta)^T I^{-1}(\beta)D(\beta)$, where $D(\beta) = E(d\hat\nu(\beta)/d\beta)$. In practice we estimate σ^2 by s^2 , $I(\beta)$ by the observed information matrix $-[dS(\beta)/d\beta]_{\hat\beta}$ and $D(\beta)$ by $[d\hat\nu(\beta)/d\beta]_{\hat\beta}$.

References

Harding, C. M., Heathwood, A. W., Hunt, R. G., and Read, K. L. Q. (1984). The estimation of animal population size by the removal method. *Appl. Statist.* **33**, 196-202.

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. (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 experiments. *Biometrics* **47**, 725–732.

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

Pollock, K. H., Hines, J. E. and Nichols, J. D. (1984). The use of auxillary variables in capture-recapture and removal experiments. *Biometrics* **40**, 329-340.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance.* Charles Griffin & Co. London.

Seber, G. A. F. and Whale, J. F. (1970). The removal method for two and three samples. *Biometrics* **26**, 393-400.

School of Statistics, La Trobe University, Bundoora, Australia.

E-mail: rhuggins@uranus.stats.latrobe.edu.au

Department of Statistics, The University of Hong Kong, Hong Kong.

E-mail: sepyip@hkucc.hku.hk

(Received February 1995; accepted August 1996)