Alternative Bayesian analysis of capture recapture data with behavioral effect modelling

Danilo Alunni Fegatelli and Luca Tardella

Abstract In the context of capture-recapture experiments we consider a generalized version of a model framework recently proposed in Chao and Yang (2005) to better understand behavioural pattern in response to trapping experience. A general order Markov structure allows to incorporate in the analysis flexible behavioural effects on the capture probabilities. We point out that the conditional likelihood approach used in the original work to carry out inference yields unbounded estimates with positive probability. We derive conditions under which pathological inference occurs and connect it to a similar problem highlighted within a different restricted removal sampling framework. To overcome the likelihood failure we investigate alternative Bayesian estimators under different non-informative prior distributions and verify with a simulation study their comparative merits in terms of efficiency and interval estimate coverage.

Key words: Capture-recapture experiments, behavioural effect, Markov chain models, likelihood failure, Bayesian approach

1 Capture-Recapture behavioural effect modelling

Consider a discrete-time closed capture-recapture experiment in which population size is assumed to be constant (no birth-death or immigration-emigration) and individuals are observed at $T > 1$ trapping times. Let N be the true population size. The number of distinct units captured during the study is denoted with M. We also suppose that all units act independently and there is no misclassification. Ideally data can be represented as a $N \times T$ binary matrix $\mathbf{X} = [x_{it}]$ where $x_{it} = 1$ if unit i-th is captured in the t-th occasion and $x_{it} = 0$ otherwise. Assume that the units captured during the whole trapping stage are conveniently labeled from 1 to M and those not captured from $M+1$ to N; hence we can observe only the firsts M rows of the matrix $\mathbf{X}$. The space of all possible capture histories is $\mathscr{X}^T = \{0,1\}^T$. In the traditional behavioural model M_b as in [1] capture probabilities vary only once when first capture

Danilo Alunni Fegatelli
Sapienza University of Rome, e-mail: danilo.alunnifegatelli@uniroma1.it

Luca Tardella
Sapienza University of Rome e-mail: luca.tardella@uniroma1.it

occurs. Model M_b is the simplest way to consider a permanent change in behaviour which has been termed in [5] ”enduring effect to capture”; formally

$$M_b : \begin{cases} Pr(x_{i1} = 1) = Pr(x_{ij} = 1 | \sum_{l=1}^{j-1} x_{il} = 0) = p \qquad j \geq 2 \quad \forall i \\ Pr(x_{ij} = 1 | \sum_{l=1}^{j-1} x_{il} > 0) = r \end{cases}$$

Alternative model framework have been recently proposed to consider different aspects of the behaviour. First order Markov chain models have been introduced by [5] to consider ”ephemeral effect to capture”. Other ways to account behavioral effect can be found in [8] and in [2] where other sources of heterogeneity are allowed. In the generic Markov chain model of order k capture probabilities depends only on the capture status in the previous k occasions. For $k = 1, 2$ we have

$$M_{c_1} : \begin{cases} p(x_{i1} = 1) = p(x_{ij} = 1 | x_{ij-1} = 0) = p_0, \quad \forall j > 1 \quad \forall i \\ p(x_{ij} = 1 | x_{ij-1} = 1) = p_1 \end{cases}$$

$$M_{c_2} : \begin{cases} Pr(x_{i1} = 1) = Pr(x_{ij} = 1 | x_{i1} = 0) = Pr(x_{ij} = 1 | x_{ij-2} = 0, x_{ij-1} = 0) = p_{00}, \quad \forall j > 2 \quad \forall i \\ Pr(x_{ij} = 1 | x_{i1} = 1) = Pr(x_{ij} = 1 | x_{ij-2} = 0, x_{ij-1} = 1) = p_{01} \\ Pr(x_{ij} = 1 | x_{ij-2} = 1, x_{ij-1} = 0) = p_{10} \\ Pr(x_{ij} = 1 | x_{ij-2} = 1, x_{ij-1} = 1) = p_{11} \end{cases}$$

It is also possible to consider both ephemeral and enduring effects together considering the marked status (marked-no marked) in the previous occasions; obviously if a unit is captured in the previous occasions it is also marked. In [5] a conditional likelihood approach is used to estimate the parameters involved in the model. The procedure is based on the factorization of the likelihood function as in [10]; denote with **p** the vector of transition probabilities involved in the model, the likelihood can be factorized

$$L(N, \mathbf{p}) \propto L_r(N, \mathbf{p}) \times L_c(\mathbf{p})$$

where $L_c(\mathbf{p})$ is the conditional likelihood and it represents the conditional distribution of the M capture histories conditionally on the fact that each unit is observed and $L_r(N, \mathbf{p})$ is the residual likelihood. The procedure consists of 2 steps: first compute $\hat{\mathbf{p}}$ maximizing $L_c(\mathbf{p})$ and then maximize $L_r(N, \hat{\mathbf{p}})$ with respect to N. In [6] it is pointed out for the first time the conditional likelihood failure for a removal model. One can easily show that the likelihood structure of the removal model is indeed equivalent to the likelihood structure of model M_b which is relevant for the estimation of N; in the same article the authors characterized the condition under which the estimate of N is unbounded. Indeed we draw the attention on the fact that the same problem occurs for Markovian models and similar conditions can be obtained (see [3]). The likelihood failure problem is often neglected in the literature, although it is not infrequent especially when the population size N is not big and the capture probability is small. In [7] it is shown that the failure problem affect also UMLE and the condition for occurrence is derived. To overcome the likelihood failure problem the authors proposed a method based on weighted (integrated) likelihood. It is natural to regard more generally the weighted likelihood approach within the Bayesian approach where the weight function acts as a prior distribution on the capture probability. A theoretical justification is given in [4] where the advantages of using the integrated likelihood are illustrated. In this short paper we discuss alternative implementations of a fully Bayesian approach where one has to specify prior distributions on all unknown parameters. In particular we will consider 4 different non-informative prior distributions on N: Uniform, one over N, one over N^2 and the Rissanen's prior; when conjugate distributions are used for the capture probabilities the marginal posterior for N can be obtained in closed form integrating the joint distribution over the parameters $\mathbf{p}$ as follow

$$\pi(N|\mathbf{X}) \propto \int L(N,\mathbf{p})\pi(N)d\mathbf{p}$$

2 Simulation study

In order to evaluate the performance of our Bayesian analysis and compare it with the classical CMLE we set up a small simulation study where the true population size N is fixed to be 100 and the number of trapping occasions is $T = 5$. Alternative capture probability configurations are considered corresponding to moderately high and low expected sample coverage as in the following Table 1.

Table 1 Description of simulation experiments.

Trial	**Model**	**Capture probabilities**	$\bar{M}$
1	M_b	$p = 0.2; r = 0.4$	67.49
2	M_b	$p = 0.1; r = 0.3$	41.06
3	M_{c_1}	$p_0 = 0.2; p_1 = 0.4$	67.49
4	M_{c_1}	$p_0 = 0.1; p_1 = 0.3$	41.06
5	M_{c_2}	$p_{00} = 0.2; p_{10} = 0.3; p_{01} = 0.35; p_{11} = 0.4$	67.49
6	M_{c_2}	$p_{00} = 0.1; p_{10} = 0.2; p_{01} = 0.3; p_{11} = 0.4$	41.06

[a] $\bar{X}$ is the empirical mean of the number of sampled distinct individuals which approximates the underlying expected sample coverage.

To summarize the posterior distribution and provide a point estimate we consider 4 different statistics: posterior mean, posterior median, posterior mode and a minimizer, denoted as *mR*, of the posterior expected loss corresponding to the quadratic relative error as in [9]. We display in Table 2 the root relative mean square error estimated with a Monte Carlo estimate using 1000 replicated data for each parameter configuration. Not only do we experience the advantages of the Bayesian approach over more traditional conditional likelihood in the more general Markov model framework but we also verify the substantial dependence of the Bayesian performance on the specific choice of prior input and posterior summary and suggest that a more formal investigation on the possibility of defining a robust efficient default Bayesian analysis. In our little experience the choice of a Rissanen prior or $1/N^2$ for N and the relative error posterior loss minimizer *mR* as a posterior estimate seem to be preferred in terms of efficiency and robustness.

Table 2 Estimated root relative mean square error and empirical coverage and average length in simulated data of alternative interval estimates with nominal confidence level 0.95.

Prior		**Tr. 1**	**Tr. 2**	**Tr. 3**	**Tr. 4**	**Tr. 5**	**Tr. 6**
$1/N^2$	Mean	0.374	0.356	0.163	0.463	0.271	0.454
	Median	0.216	**0.323**	0.146	0.313	0.195	0.295
	Mode	**0.167**	0.421	**0.132**	0.286	**0.150**	0.345
	mR	**0.167**	0.350	0.134	**0.262**	0.159	0.291
	HPD I.	95.3%	88.0%	94.7%	90.6%	94.8%	89.8%
		(120.9)	(163.9)	(57.1)	(144.4)	(81.9)	(172.7)
Rissanen	Mean	0.688	0.807	0.177	0.895	0.410	1.109
	Median	0.293	0.342	0.155	0.445	0.241	0.407
	Mode	0.170	0.407	0.135	0.285	0.156	0.330
	mR	0.194	0.327	0.140	0.273	0.178	**0.280**
	HPD I.	96.0%	92.8%	95.1%	97.3%	95.7%	93.0%
		(194.5)	(342.0)	(59.9)	(204.1)	(97.3)	(319.4)
	CMLE	**0.704***	**1.284***	**0.176**	**0.642***	**0.337***	**0.542***
	C.I.	95.1%	94.5%	94.5%	94.5%	94.9%	97.2%
		(2388.9)	(7201.6)	(69.4)	(488.8)	(175.1)	(855.2)

[b] the * sign denotes the presence of likelihood failure.

References

1. White, G. C. and Anderson, D. R. and Burnham, K. P. and Otis, D. L. : Capture-recapture and Removal Methods for Sampling Closed Populations, Los Alamos National Laboratory, (1982)
2. Bartolucci, F. and Pennoni, F. : A class of latent Markov models for capture-recapture data allowing for time, heterogeneity and behavior effects, Biometrics, 63, 568–578. (2007)
3. Tardella, L. and Alunni Fegatelli, D. : Alternative Bayesian Analysis of Capture Recapture Data with Behavioral Effect Modelling, Working Paper (2012)
4. Berger, J.O., Liseo, B. and Wolper, R.L. : Integrated Likelihood Methods for Eliminating Nuisance Parameters. Statistical Science, **14 (1)**, 1–28 (1999)
5. Chao, A and Yang, H.C. : Modeling animals' behavioral response by Markov chain models for capture-recapture experiments. Biometrics, **61**,1010–1017 (2005)
6. Seber, G.A.F. and Whale, J.F. : The removal methods for two and three sample. Biometrics, **26(3)**, 393–400 (1970)
7. Carle, F.L. and Strub, M.R. : A new method for estimating population size from removal data . Biometrics, **34**,621–630 (1978)
8. Farcomeni, A. : Recapture models under equality constraints for the conditional capture probabilities. Biometrika **98(1)**,237–242 (2011)
9. Tardella, L. : A new Bayesian method for nonparametric capture-recapture models in presence of hterogeneity. Biometrika 89, 807–817 (2002)
10. Sanathanan L. : Estimating the size of a multinomial population. The Annals of Mathematical Statistics, **43**,142-152. (1972)