

Modelling and inferential issues for behavioral patterns in capture-recapture data

Modelli e metodi inferenziali per lo studio dell'effetto comportamentale in esperimenti di cattura-ricattura

Danilo Alunni Fegatelli and Luca Tardella

Abstract A comprehensive understanding of the factors that affect capture probability in a capture-recapture study plays a fundamental role in determining the accuracy and precision of population parameter estimates. In this paper we provide an updated account of recent advances in the statistical modelling of capture recapture experiments for closed populations which highlight general frameworks to explore and fit alternative behavioural patterns and point out some key issues which are relevant for developing possibly efficient and flexible inferential tools.

Abstract *Si fornisce una rassegna aggiornata dei recenti avanzamenti introdotti nella modellistica degli effetti comportamentali per esperimenti di tipo cattura-ricattura per popolazioni chiuse. Si evidenziano alcuni nuovi modi di strutturare modelli generali ed alcune problematiche nella derivazione di procedure inferenziali efficienti*

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

1 Introduction

In the context of discrete-time closed capture-recapture modelling for estimating the unknown size of a finite population it is often required a flexible framework for dealing with the behavioural variability due to the change of behaviour of each unit after trapping experience. Many alternative settings have been proposed in the literature to account for the variation of capture probability. Some of them exploit general regression frameworks such as the Poisson loglinear model in [8] or the logistic regression [1] for the classical multinomial model [15]. Other proposals start from more specific perspectives. The most basic way to involve the behavioural effect to capture in the model is to assume that once a unit is captured its probability of being recaptured is modified permanently [15]. In the last decades alternative

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

modelling approaches have been proposed to account for flexible behavioural response to capture [23, 17, 6]. Inference is typically carried out relying on either the so-called conditional likelihood or the full likelihood approach. However, in [2] is highlighted that the maximum likelihood approach may, with positive probability, lead to an inferential pathology called *likelihood failure* (unbounded estimates for the finite size of the population). Such annoying phenomenon has been characterized within a very general class of behavioural effect models.

2 Large subclasses of behavioural models

In the process of modelling the joint probability distribution of all the observed capture outcomes x_{ij} for all the N units in the population including the unobserved ones during the experiment it seems a rather natural approach to model for each unit the sequence of consecutive conditional probabilities of capture which can indeed be specified so that many sources of variability can be easily accommodated

$$p_j(x_{i1}, \dots, x_{ij-1}) = Pr(X_{ij} = 1 | x_{i1}, \dots, x_{ij-1}) \quad \forall j > 1, \forall (x_{i1}, \dots, x_{ij-1}) \in \mathcal{X}^{j-1}(1)$$

In fact it is easy to realize that whenever the conditional probabilities in (1) are constrained to be all equal we get the classical null model M_0 where no source of heterogeneity among units or capture/temporal occasions or behavioral pattern is allowed. On the other hand the specification of suitable subclasses of partial capture histories $H_1, \dots, H_b, \dots, H_B$ partitioning the set of all partial capture histories so that the corresponding conditional probabilities are constrained to be the same whereas the conditioning partial capture history belongs to the same H_r can be a very flexible device to embed heterogeneity which can be interpreted in terms of the common characteristics of the partial capture histories within the same subset. In fact, if in a capture-recapture experiment with t trapping occasions one considers as partition subsets $H_1, \dots, H_j, \dots, H_t$ where the generic H_j comprises all the partial capture histories with length equal to j this partition can be easily interpreted as follows: the change in the conditional capture probability is driven by the number of previous trapping occasions encountered and hence this has to be understood as a temporal source of heterogeneity. Since our main focus is on modelling possible heterogeneity due the behavioural changes we can readily provide few examples of behavioural pattern at different degrees of flexibility which can be studied within the general framework of partitioning into equivalence classes the subset of partial capture histories and the related conditional probabilities. Consider for instance the bipartition $\mathcal{H}(M_b) = \{H_1, H_2\}$ where the first subset contains all the partial capture histories with no capture at all. This partition recovers the classical behavioral model where only after the first trapping experience there is a behavioural effect which may turn its probability to become larger (trap-happiness) or smaller (trap-shyness) according to whether the common conditional probability corresponding to the partial capture history in H_2 is larger or smaller. This kind behavioural pattern is an example of the so called *enduring* effect to capture which implies that once the

unit has modified its capture probability this modification is permanent. A different behavioural pattern called *ephemeral* effect can be set up in terms of Markovian dependence of consecutive capture occurrences as already illustrated in [23] and later on extended to arbitrary k -th order in [6] where the general and flexible framework related to general partition of capture histories was indeed introduced in terms of linear constraints for the saturated parameterization of conditional probabilities. To our knowledge there have been only few other attempts to setup alternative model frameworks for a deeper understanding of behavioural patterns and we refer the reader to [22] for an ephemeral effect specified in terms of autoregressive structure with unknown lag or to the more specific pattern of memory persistence in [17].

Models combining behavioural patterns and heterogeneity have been considered in [14, 9] and [3] although with more emphasis on the the flexibility in accounting for unobserved heterogeneity rather than on understanding behavioural effects.

Indeed a possible general strategy for designing and fitting a flexible class on nested behavioural models would require the definition of a suitable sequence of nested partitions $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_K$ such that each one represents, with a suitably increasing number of parameters, the complexity of some behavioural pattern. Of course this task has a two-fold difficulty of searching within the subset of all possible partitions which is actually made of a formidable number of alternatives and providing an appropriate behavioural interpretation. In fact we will show how both goals can be hit starting from the definition of a suitably interpretable quantification of partial capture histories and embedding the model building within a regression framework.

3 Flexible regression settings

When we acknowledge the key role of the reparametrization of the joint probability of all the binary capture occurrences of each unit in terms of the product of conditional probabilities we then have to understand how the saturated model can be simplified in terms of possible hierarchical/nested reduced parameterization which can balance the usual trade-off between goodness of fit and parsimony. A very natural and conventional way of addressing this issues could rely for instance on the use of arbitrary order Markovian models or on the arbitrarily lagged autoregressive model. The first approach has the disadvantage of requiring a number of parameters which grows exponentially with the order whereas in the second approach the progression of the number of parameters is linear. In both cases the conditional dependence on the past is limited to most recent occasions and the behavioural pattern would be of ephemeral nature.

Alternatively, we could follow and extend the ideas in [7] where the use of a suitable meaningful behavioural covariate is first introduced as a way of summarizing the partial capture history which can be relevant in affecting the conditional probability of being trapped in the next occasion. We could for instance consider for each unit i the quantification at each occasion j given by $z_{ij} = g(x_{i1}, \dots, x_{ij-1}) = \sum_{s=1}^{j-1} x_{is}$. This covariate may convey the information of the possible impact that previous cap-

tures in the past can modify the probability of being trapped in the next capture occasion. We may refer to z_{ij} as a *memory-effect* covariate.

The availability of a univariate covariate could be exploited within the framework of linear logistic regression in order to embed behavioural patterns in terms of few regression parameters. In our previous example modelling $\text{logit} p_j(x_{i1}, \dots, x_{ij-1}) = \alpha + \beta z_{ij}$ could help inferring an enduring (and possibly reinforced) behavioural effect. The flexibility of this device can be arbitrarily extended in terms of non-linear regression whose flexible shape could account for possible non linear decay of memory effects or more complex conflicting patterns of trap-happiness followed by trap shyness. One of the most remarkable results achieved in [7] is the fact that using another suitable definition of the memory-effect covariate $z_{ij} = g(x_{i1}, \dots, x_{ij-1})$ and an appropriate non-linear step functions with jumps at prescribed values of an alternative new behavioural covariate are able to recover Markovian models of arbitrary order. This opens up the possibility of exploring within a regression context larger and possibly more parsimonious classes of behavioural patterns by exploiting the meaningful covariate and searching the model space of regressions with arbitrary step functions. On one hand rephrasing the search of a suitable behavioural model within such a large classes as a suitable selection of a regression function of one behavioural covariate allows to exploit all the standard devices to fit non linear regressions with possible integration of auxiliary information such as environmental or individual covariates.

Unfortunately the task of model selection within large class of non linear functions can be computationally demanding and alternative efficient approaches for exploring this kind of model space should be carefully selected. One can use simple classical AIC or BIC selection criteria and easily carry out a grid search for step functions with few jumps possibly extending the search to more jumps using greedy forward selection. However we briefly mention the possibility of implementing alternative selection strategies using regression trees or more sophisticated statistical learning techniques.

4 Likelihood issues and alternative inference

After the modelling phase we need to select an appropriate methodology to infer on the unknown parameters of the model with particular emphasis on the size N of the population. Several alternative strategies have been proposed in the literature and successfully implemented in the currently available software [4]. One of the most popular is still the conditional likelihood advocated for its simplicity in the seminal works of [21], later advocated in [10] and, more recently in [23, 11, 6]. However we need to stress here the fundamental relevance of the behavioural component of the model in inferring the unknown N and the role of the inferential approach. In fact in [2] it has been shown that for a large class of behavioural models represented in terms parameters associated to subsets of partitions of conditional probabilities the use of conditional likelihood as well as of the complete likelihood can lead to an annoying unboundedness of the estimate $\hat{N}$ (*likelihood failure*) and it is also shown

how the use of an integrated likelihood or a fully Bayesian analysis circumvent this problem and provide a more accurate inference. The class of models suffering this likelihood pathology contains the most basic behavioural model denoted as M_b and all the other models where the partial capture histories corresponding to no capture belongs to the same partition set so that the model prescribes a common conditional probability of first capture p_{H_1} . For this general class of models the likelihood factorization $L = L^c \times L^r$ is such that the residual likelihood L^r is a function of N and p_{H_1} only. The conditional maximum likelihood estimator is

$$\hat{N}_{CMLE} = \frac{M}{1 - \hat{P}_0} = \frac{M}{1 - (1 - \hat{p}_{H_1})^t}$$

where P_0 is the probability of never being captured during all the experiment. When failure conditions are met we have that $\hat{p}_{H_1} = 1$ and hence $\hat{N}_{CMLE} = \infty$ [2].

Indeed there are many other general approaches and model extensions to fit capture-recapture data in the discrete setting where the number of capture occasions are fixed. The log-linear framework originally proposed in [8] deserves a special mention as well as later developments in order to account for individual unobserved heterogeneity within loglinear as well as logistic reparameterizations [5, 16, 3]. Although the Bayesian approach has witnessed a successful development especially in the last two decades with only few contributions specifically devoted to a better understanding, modelling and fitting behavioural effects [13, 12, 9]. Only more recently it has become a routinely viable option due to the possibility of embedding flexible models within general all purpose software such as WinBUGS which yield an MCMC approximation of the posterior distribution. One relevant contribution to this aim is [19, 20] and the monograph [18].

On the other hand we must acknowledge that little effort has been made in order to provide an easy-to-use interface which can guide both the process of elicitation of a suitable prior distribution and select (or average among) competing models especially for behavioural features.

References

1. J. M. Alho (1990). Logistic Regression in Capture-recapture Models. *Biometrics* 46, pp. 623–635.
2. D. Alunni Fegatelli and L. Tardella (2013). Improved inference on capture recapture models with behavioural effects. *Statistical Methods & Applications* 22, 1, pp. 45–66.
3. F. Bartolucci and F. Pennoni (2007). A Class of Latent Markov Models for Capture-Recapture Data Allowing for Time, Heterogeneity, and Behavior Effects. *Biometrics* 63, 2, pp. 568–578.
4. J. A. Bunge (2013). A survey of software for fitting capture-recapture models. *Wiley Interdisciplinary Reviews: Computational Statistics* 5, 2, pp. 114–120.

5. B. A. Coull and A. Agresti (1999). The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics* 55, pp. 294–301.
6. A. Farcomeni (2011). Recapture models under equality constraints for the conditional capture probabilities. *Biometrika*.
7. D. A. Fegatelli and L. Tardella (2014). Flexible behavioral capture-recapture modelling. *submitted*. eprint: [arXiv/1401.6433](https://arxiv.org/abs/1401.6433).
8. S. E. Fienberg (1972). The Multiple Recapture Census for Closed Populations and Incomplete 2^k Contingency Tables. *Biometrika* 59, pp. 591–603.
9. S. K. Ghosh and J. L. Norris (2005). Bayesian Capture-recapture Analysis and Model Selection Allowing for Heterogeneity and Behavioral Effects. *Journal of Agricultural, Biological, and Environmental Statistics* 10, 1, pp. 35–49.
10. R. M. Huggins (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* 47, pp. 725–732.
11. R. Huggins and W. Hwang (2011). A Review of the Use of Conditional Likelihood in Capture-Recapture Experiments. *International Statistical Review* 79, 3, pp. 385–400.
12. S.-M. Lee, W.-H. Hwang, and L.-H. Huang (2003). Bayes Estimation of Population Size from Capture-recapture Models with Time Variation and Behavior Response. *Statistica Sinica* 13, 2, pp. 477–494.
13. S.-M. Lee and C. W. S. Chen (1998). Bayesian Inference of Population Size for Behavioral Response Models. *Statistica Sinica* 8, pp. 1233–1248.
14. I. Norris James L. and K. H. Pollock (1995). A Capture-recapture Model with Heterogeneity and Behavioural Response. *Environmental and Ecological Statistics* 2, pp. 305–313.
15. D. L. Otis et al. (1978). *Statistical Inference From Capture Data on Closed Animal Populations*. Wildlife Monographs.
16. S. Pledger (2000). Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics* 56, 2, pp. 434–442.
17. F. Ramsey and P. Severns (2010). Persistence models for mark-recapture. *Environmental and Ecological Statistics* 17 (1), pp. 97–109.
18. J. A. Royle and R. M. Dorazio (2008). *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Academic Press.
19. J. A. Royle (2009). Analysis of Capture-Recapture Models with Individual Covariates Using Data Augmentation. *Biometrics* 65, 1, pp. 267–274.
20. J. Royle and R. Dorazio (2012). Parameter-expanded data augmentation for Bayesian analysis of capture-recapture models. *J. Ornith.* 152, pp. 521–537.
21. L. Sanathanan (1972). Estimating the Size of a Multinomial Population. *The Annals of Mathematical Statistics* 43, pp. 142–152.
22. C. L. Tyrrell et al. (2009). Evaluation of trap capture in a geographically closed population of brown treesnakes on Guam. *J. of Appl. Ecol.* 46, 1, pp. 128–135.
23. H.-C. Yang and A. Chao (2005). Modeling Animals' Behavioral Response by Markov Chain Models for Capture-recapture Experiments. *Biometrics* 61, 4, pp. 1010–1017.