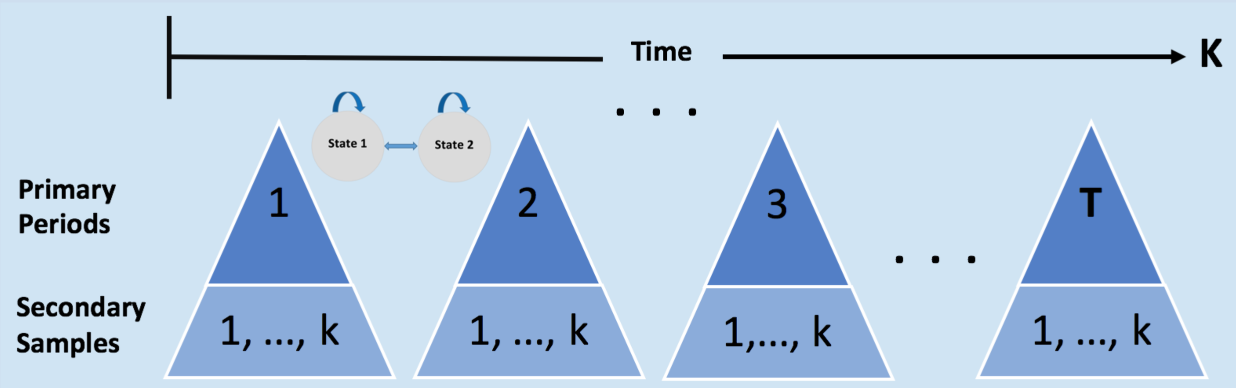
The use of penalised likelihood to improve estimation in removal models

Ming Zhou[1], Eleni Matechou[1], Rachel McCrea[1], Diana Cole[1], Richard Griffiths [2]

[1] *School of Mathematics, Statistics and Actuarial Science, University of Kent*
[2] *The Durrell Institute of Conservation and Ecology, University of Kent*

Introduction

- Data obtained from removal sampling - the permanent movement of protected animals out of the path of development projects as a wildlife management tool - can be used to estimate the abundance of a population within a study area. Removal projects are more **expensive** than any other conservation action for the same species. For example, approximately £100 million are spent on the removal of great crested newts annually (Lewis, 2012; Germano et al., 2015).



- However, the classic removal model (Moran,1951) may give rise to misleading conclusions due to violation of the assumption of no new addition of individuals to the study area.
- The estimation of temporary emigration or population renewal for removal data relies on the use of the robust design (Zhou et al., 2017), where there are at least two secondary occasions between primary occasions.
- The **aim** of the research is to model new arrivals of individuals coming into the study area without the use of the robust design.

Method

- Suppose the total number of sampling occasions is K. The population size is denoted by N. n_k is the number of individuals being removed at the kth sampling occasion, where $k = 1, ..., K$. n_0 is the number of individuals we failed to capture by the end of the study.

- We assume a constant capture probability p over time. β_k represent entry parameters, the proportion of individuals that become available for removal for the first time at the kth sampling occasion. $\sum_{k=1}^{K} \beta_k = 1$.

- The full multinomial likelihood is

$$L(n_0, \beta_k, p|n_k) = \frac{N!}{n_0! \cdot \prod_{k=1}^{K} n_k!} \cdot \alpha_0 \cdot \prod_{k=1}^{K} \alpha_k,$$

where α_k and α_0 are the probability of an individual being removed at the kth sampling occasion and the probability of a individual not being removed by the end of the study respectively.

- This model is parameter redundant without further development, because we have K data points, and we want to estimate $K+1$ parameters.

- In order to overcome the issue of parameter redundnacy, we introduce a penalty for the likelihood, i.e. instead of maximising $\log\{L(n_0, \beta_k, p|n_k)\}$, we maximise the objective function (Tibshirani, 1996)

$$O(n_0, \beta_k, p|n_k) = \log\{L(n_0, \beta_k, p|n_k)\} - \lambda \sum_{b=2}^{K} \beta_b,$$

where $\lambda \geq 0$ is a tuning parameter, to be determined separately.

- We use cross-validation to choose λ that minimises the mean squared error (MSE).

Results

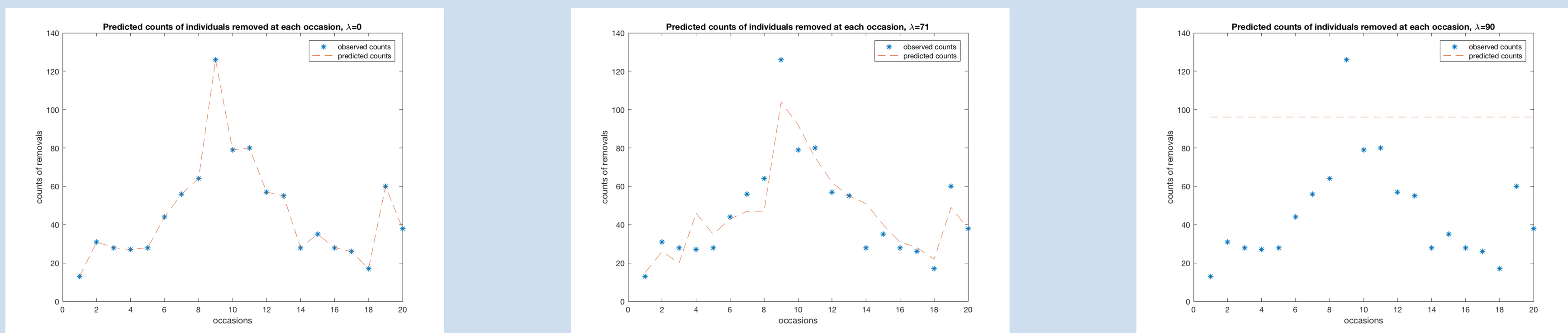


Figure 1: Estimated counts for individuals removed at each occasion when $\lambda = 0$ (left), $\lambda = 71$ (middle) and $\lambda = 90$ (right), where the data is simulated for $K = 20$, $p = 0.3$ and $N = 1000$.

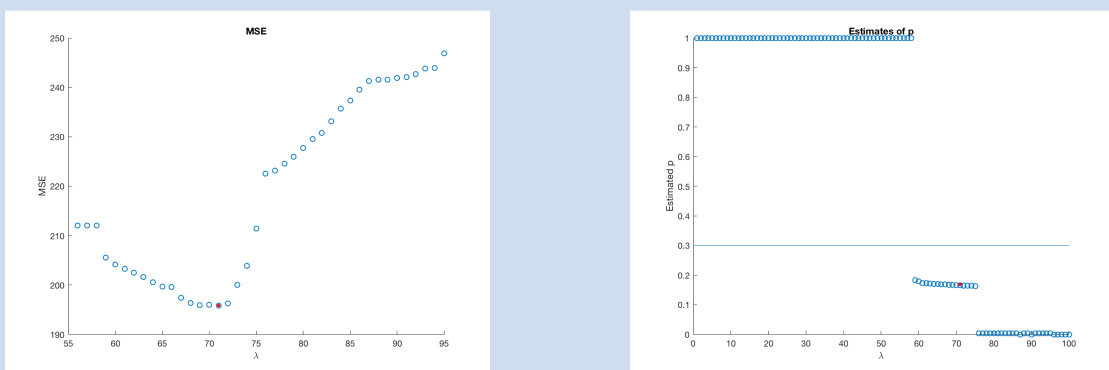


Figure 2: Cross-validation results for the simulated dataset with $K = 20$, $p = 0.3$, and $N = 1000$. MSE (left) and the estimates of capture probabilities (right) are shown.

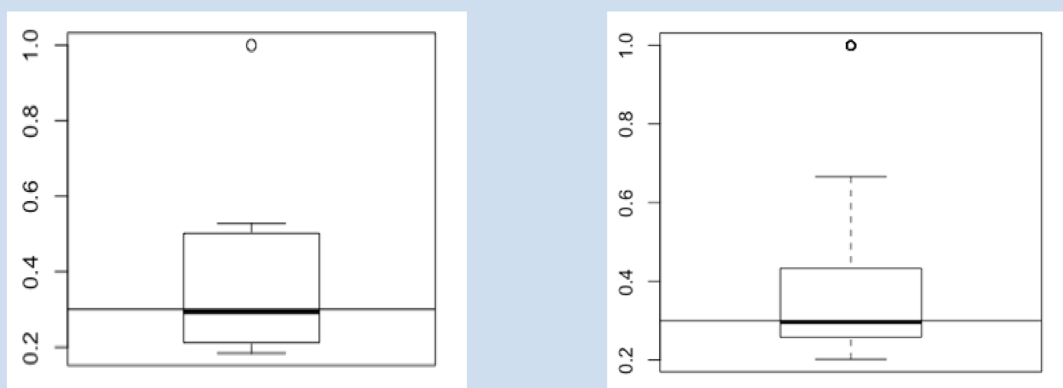


Figure 3: Estimated p for 100 simulated data with $K = 20$, $p = 0.3$, $N = 1000$ under scenarios where there are no zeros and six zeros in β_k on the left and right column respectively.

Conclusion

- The use of maximum penalised likelihood estimation can overcome the issues of parameter redundancy for removal data collected under the standard sampling protocol.
- LASSO can improve the estimation of parameters when the model is not identifiable.
- Cross-validation can be used to determine the tuning parameter λ for the penalised maximum likelihood estimation.
- However, the estimation of penalised likelihood can be time-consuming due to cross-validation.
- **Other things are not discussed today ...**
  - Other penalty functions.
  - Small population sizes.

References

1. Germano, J. M., Field, K. J., Griffiths, R. A., Clulow, S., Foster, J., Harding, G., Swaisgood. R. R. 2015. *Mitigation-driven translocations: Are we moving wildlife in the right direction? Frontiers in Ecology and the Environment* **13**:101âĂŞ105.
2. Lewis B. (2012) An evaluation of mitigation actions for great crested newts at development sites (PhD dissertation). Canterbury, UK: University of Kent.
3. Moran, P.A.P. (1951) A mathematical theory of animal trapping, *Biometrika*, **38**, 307-311.
4. Tibshirani, R. (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*, **58**, 267-288.
5. Zhou, M., McCrea, R. S., Matechou, E., Cole, D. and Griffith, R. A. (2017) Removal Models Accounting for Temporary Emigration, *revision invited.*