Ecology and Evolution
Open Access

Estimation of capture probabilities using generalized estimating equations and mixed effects approaches

Md. Abdus Salam Akanda[1,2] & Russell Alpizar-Jara[1]

[1]Department of Mathematics, Research Center in Mathematics and Applications, University of Évora, 7000-671 Évora, Portugal
[2]Department of Statistics, Biostatistics & Informatics, University of Dhaka, Dhaka 1000, Bangladesh

Keywords
Closed population, generalized linear models, generalized linear mixed models, heterogeneity, population size estimation.

Correspondence
Md. Abdus Salam Akanda, Department of Mathematics, Research Center in Mathematics and Applications, University of Évora, 7000-671 Évora, Portugal.
Tel: +351965155536; Fax: +351266745393;
E-mail: akanda_du@yahoo.com

Funding Information
This research was funded by EMMA in the framework of the EU Erasmus Mundus Action 2 and Fundacao Nacional para a Ciencia e Tecnologia (FCT), Portugal under the project PEst-OE/MAT/UI0117/2011.

Received: 14 January 2014; Accepted: 23 January 2014

Ecology and Evolution **2014; 4(7): 1158–1165**

doi: 10.1002/ece3.1000

Abstract

Modeling individual heterogeneity in capture probabilities has been one of the most challenging tasks in capture–recapture studies. Heterogeneity in capture probabilities can be modeled as a function of individual covariates, but correlation structure among capture occasions should be taking into account. A proposed generalized estimating equations (GEE) and generalized linear mixed modeling (GLMM) approaches can be used to estimate capture probabilities and population size for capture–recapture closed population models. An example is used for an illustrative application and for comparison with currently used methodology. A simulation study is also conducted to show the performance of the estimation procedures. Our simulation results show that the proposed quasi-likelihood based on GEE approach provides lower SE than partial likelihood based on either generalized linear models (GLM) or GLMM approaches for estimating population size in a closed capture–recapture experiment. Estimator performance is good if a large proportion of individuals are captured. For cases where only a small proportion of individuals are captured, the estimates become unstable, but the GEE approach outperforms the other methods.

Introduction

Many estimation methods have been developed for the analysis of closed population capture–recapture data. For comprehensive material on the subject see, for instance, Otis et al. (1978), Seber (2002), Williams et al. (2002) and Amstrup et al. (2005). The most general capture–recapture closed population model, considered by Otis et al. (1978) was denoted by M_{tbh} where (h) is used to denote inherent individual heterogeneity, (t) time effect, and (b) behavioral response to capture. In this work, we are interested in estimating the population size and SE of a submodel of the type M_h, where individual heterogeneity can be modeled as a function of covariates. Development of capture–recapture models dealing with individual heterogeneity in capture probabilities has been one of the most challenging tasks. Failure to account for such heterogeneity has long been known to cause substantial bias in population estimates (Otis et al. 1978; Lee and Chao 1994; Hwang and Huggins 2005). Moreover, Link (2003) showed that without strong assumptions on the underlying distribution, estimates of population size under model M_h are fundamentally nonidentifiable.

The use of covariates (or auxiliary variables), if available, has been proposed as an alternative way to partially cope with the problem of heterogeneous capture probabilities (Pollock et al. 1984; Huggins 1989; Alho 1990). The idea is to model capture probabilities as a function of individual (i.e., age, sex, and weight) and environmental (i.e., temperature, rainfall, and location) covariates, using a generalized linear modeling (GLM) approach, such as logistic regression. The method of Huggins (1989, 1991), based on a conditional likelihood to estimate population size, has become very popular, but it assumes