Supplements to Borchers and Efford Biometrics 2007

Updated 19/11/2007

These notes extend Borchers and Efford (2007) (‘B&E’), and provide further background for the implementation of spatially explicit capture–recapture (SECR) methods in the software DENSITY 4.1 (Efford 2007).

Contents

Parameters for simple (within-session) models

From B&E p.2,

> “The likelihood, or equivalently here, the joint distribution of the number of animals captured n, and their [spatial] capture histories $\omega_1, \ldots, \omega_n$ can be written in terms of the marginal distribution of n and the conditional distribution of $\omega_1, \ldots, \omega_n$ given n, as
>
> $$L(\phi, \theta \mid n, \omega_1, \ldots, \omega_n) = \Pr(n \mid \phi, \theta) \Pr(\omega_1, \ldots, \omega_n \mid n, \theta, \phi) \qquad (1)$$
>
> where θ is the vector of capture function parameters and ϕ is a vector of parameters of the spatial point process governing animal density and distribution.”

In all analyses and coding to date, ϕ has been a single parameter, the homogeneous Poisson population density, and we take this no further here. For simplicity, ϕ and θ are concatenated in software to form a single vector over which the likelihood is maximised numerically.

The detection parameters θ conceal considerable complexity because the detection function may take several forms each with multiple parameters having differing scales and probably different link functions, and there are several possible types of covariate. Mixture models also add complexity. Here we suggest one way to organize this complexity, the one implemented in DENSITY 4.1. We consider only a single closed-population sample, deferring discussion of the complications of product multinomial models as used for between-year trend in the red-eyed vireo example of B&E.

To evaluate the probabilities on the right-hand side of B&E equation (1), we must specify the probability of each possible detection event (P_{iks} for animal i at trap k on occasion s), conditional on other synchronous and previous events. For a finite mixture model, the latent class must also be specified (P_{iksu} for the fraction of the population in class u).

Primary detection parameters

The set of core parameters depends upon the chosen detection function (examples in Efford, Borchers and Byrom in press), but it always includes a magnitude component, usually as the intercept (g_0), and a spatial scale (σ). The hazard detection function has an additional shape parameter (b). Variation in the probability of detection events may be modelled as a function of these parameters (e.g., $P_{iksu} = f(g_{0iksu}, \sigma_{iksu}, b)$). Here we assume that b is constant.

Each primary parameter is manipulated on an appropriate transformed scale. The scale is chosen mostly for numerical convenience i.e. so that all possible values ($-\infty < x < \infty$) on the transformed scale map to valid values of the parameter. We use the term ‘link’ function for the transformation by analogy with generalised linear modelling, and following established practice in capture–recapture (Lebreton et al. 1992, Cooch and White 2006). The most generally useful link functions are logit(x) = log(x/(1–x)) ($0 < x <$

1) and log(x) ($x > 0$). Their inverses are $\mathrm{logit}^{-1}(x) = e^x/(1+e^x)$ and $\log^{-1}(x) = e^x$ (both $-\infty < x < \infty$). In DENSITY the default link functions are logit for g_0 and log for σ and b.

Parameters for finite mixture models

Mixture models with U latent classes may be specified independently for each of the primary parameters g_0 and σ (with some restrictions). Only 2-part and 3-part mixtures are coded in Density 4.1 (i.e. $U \in \{2,3\}$). Parameters for the mixture proportions $\psi_u(g_0)$ and $\psi_u(\sigma)$ ($u \in \{1,\ldots,U\}$) are grouped functionally with the primary parameters and have their own link function (default logit, constraining values between 0 and 1). Class membership is treated as a discrete covariate (below) coded either 0,1 ($U = 2$) or (0,0),(1,0),(0,1) ($U = 3$). The mixture likelihood includes a weighted sum over the U classes.

Covariates of detection

Covariates of detection may relate to the sampling occasion (s), previous experience of capture, a permanent attribute of the individual (z_i), or the trap site (k).

The same mapping property that makes the link function attractive for numerical maximization also makes it a suitable additive scale for combining the effects of several covariates (i.e. all combinations map to meaningful values of the parameter). There have been many previous applications in capture–recapture (e.g. Huggins 1989, Lebreton et al. 1992, Pledger 2000, Cooch and White 2006).

Each primary parameter is modelled as an additive function of the covariates on the link scale. Thus

$$g_0 = \mathrm{logit}^{-1}(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6)$$

$$\sigma = \log^{-1}(\gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3 + \gamma_4 x_4 + \gamma_5 x_5 + \gamma_6 x_6)$$

where the values $x_1,\ldots,x_6$ code levels of the covariates (see below), and $\boldsymbol{\beta} = \beta_0,\ldots,\beta_6$ and $\boldsymbol{\gamma} = \gamma_0,\ldots,\gamma_6$ are fitted coefficients. This might also be expressed in matrix form as

$$g_0 = \mathrm{logit}^{-1}(\mathbf{X}\boldsymbol{\beta})$$

$$\sigma = \log^{-1}(\mathbf{X}\boldsymbol{\gamma})$$

where each vector of covariates $\mathbf{x} = x_0,...,x_6$ ($x_0 = 1$) is a row of the design matrix **X**. β_0 and γ_0 are intercept terms, so for the null model $g_0[.]\sigma[.]$ we have

$$g_0 = \text{logit}^{-1}(\beta_0)$$
$$\sigma = \log^{-1}(\gamma_0).$$

Similar expressions apply for each ψ_u, and for b, because neither is allowed to be a function of within-session covariates. Implicitly, all mixture models in DENSITY 4.1 are $\psi_u[.]$ models, and all hazard-function models are $b[.]$ models, and there is no need to specify these components when describing the within-session detection model. Of course, one should state the order of the mixture model (e.g. $U = 2$), and the type of detection function (e.g., hazard or halfnormal), and report relevant parameter estimates.

Coding of x-vector

Covariates are coded either as continuous variables or indicator (0/1) variables. The particular coding (and the choice of columns in the design matrix) is fixed in DENSITY 4.1 as in the following table.

x	Value	Description	Effect
x_0	1	Intercept	all
x_1	Continuous	Occasion	t
x_2	Indicator 0/1	Previous* capture of individual i in any trap	b, b1
x_3	Indicator 0/1	Latent class 2	h2, h3
x_4	Indicator 0/1	Latent class 3	h3
x_5	Continuous	Permanent attribute of individual i (z_i)	h
x_6	Continuous	Permanent attribute of trap k	k

* used both for a permanent learned response (b) or for a Markov one-step response (b1); in the latter case ‘previous capture’ is defined as capture on the immediately preceding occasion.

Users of the Density software specify the indicator covariates (x_2, x_3, x_4) implicitly when they select a model (Options | ML SECR), and no further action is needed. Input of the continuous covariates (x_1, x_5, x_6) is described in the online help; although nominally continuous, these might also take discrete values (e.g. 0 = cloudy days, 1 = sunny days for x_1). Measured individual attributes (x_5) are relevant only with the conditional likelihood option, when there is a close analogy to the closed-population method of Huggins (1989, see also Chao and Huggins 2005).

The listed effects do not exhaust the possibilities. A 'time effect' might be fitted with a distinct level for each occasion, as in the conventional closed-population model Mt (Otis et al. 1978). An interesting addition would be a trap-specific behavioural response b(*k*) for the change in detection probability of individual *i* in trap *k*, after being caught in that particular trap. For example, birds may avoid sites where they have been caught in mist nets, rather than developing a general ability to avoid nets.

Asymptotic variance of *D* estimated by maximising the conditional likelihood

The conditional likelihood estimate of density is $\hat{D}(\hat{\theta}) = na(\hat{\theta})^{-1}$ (equivalently, $\hat{D}(\hat{\theta}) = \sum_{i=1}^{n} a(\hat{\theta}_i)^{-1}$ when the a_i depend on individual covariates). Following Huggins (1989: 136), we assume the asymptotic sampling variance of $\hat{D}$ has the form

$$\mathrm{var}(\hat{D}(\hat{\theta})) = s^2 + \hat{G}_\theta^{\mathrm{T}} \hat{I}_\theta^{-1} \hat{G}_\theta,$$

where s^2 is the variance of $\hat{D}(\theta)$ when θ is known, *I* is the information matrix (inverse Hessian), and *G* is a vector containing the gradients of $\hat{D}(\theta)$ with respect to the elements of θ, evaluated at the maximum likelihood estimates.

Under the binomial (fixed-N) model, we can simply substitute a_i/A for p_i in Huggins formula for var($\hat{N}$) (1989: 136) and scale var($\hat{N}$) by A^{-2} to obtain

$$\begin{aligned} s^2 &= \sum_{i=1}^{n} (1 - p_i)\, p_i^{-2} A^{-2} \\ &= \sum_{i=1}^{n} (1 - a_i / A)\, a_i^{-2}. \end{aligned} \tag{1}$$

Under the Poisson model it is not yet certain what expression to use for s^2. As $A \rightarrow \infty$, the binomial distribution approaches a Poisson, but the consequences for (1) are uncertain.

Putting aside individual variation (i.e., all $a_i = a$, where a is known), so $\hat{D} = n / \hat{a}$:

$$\begin{aligned} s^2 &= \mathrm{var}(n / \hat{a}) \\ &= \hat{\mathrm{var}}(n)\, \hat{a}^{-2} \\ &= n\, \hat{a}^{-2}. \end{aligned}$$

We conjecture that $\sum_{i=1}^{n} \hat{a}_i^{-2} + \hat{G}_\theta^{\mathrm{T}} \hat{I}_\theta^{-1} \hat{G}_\theta$ is an asymptotically unbiased estimator of $\mathrm{var}(\hat{D}(\hat{\theta}))$. The second term is estimated numerically and poses no problems. This is the basis for the standard errors for $\hat{D}$ reported by DENSITY 4.1 when a Poisson model is fitted by maximising the conditional likelihood.

Bootstrap variance estimation

B&E suggested “Bootstrapping of capture histories is potentially useful, but for the moment prohibitively slow”. The SECR algorithm is now faster owing to improvements in coding, so we re-visit the issue.

For each bootstrap replicate, a sample of size n is taken from the n observed capture histories, with replacement. Any of the original capture histories may appear more than once in the bootstrap sample, or not at all. The 0.025 and 0.975 quantiles of the bootstrap estimates provide a 95% confidence interval, but coverage may be poor. Coverage is improved by using quantiles of the studentized values ($\frac{\hat{\theta}^* - \hat{\theta}}{\sqrt{v^*}}$) to estimate the limits on the studentized scale, and then applying these limits to the particular estimates of $\hat{\theta}$ and SE($\hat{\theta}$).

Bootstrapping provides intervals for detection parameters (g_0, σ). When used as described here, it does *not* provide an interval for density D because the bootstrap samples all use the same n, whereas variation in n is an important source of uncertainty in $\hat{D}$.

Log likelihood for saturated full model when *n* is binomial

The log likelihood of the saturated model is needed to calculate model deviance, used in a Monte Carlo goodness-fit-fit test. B&E give the saturated likelihood when the number of animals caught n is Poisson. To complete the picture we need the binomial ('fixed-N') saturated likelihood. The saturated likelihood for the binomial model is:

$$L_{sat} = n\log(\frac{n}{N}) + (N-n)\log(\frac{N-n}{N}) + \log(\frac{N!}{(N-n)!}) - \sum\nolimits_{\omega} \log(n_{\omega}!) + \sum\nolimits_{\omega} n_{\omega} \log(\frac{n_{\omega}}{n})$$

where N is the population in the area A and for evaluation we use an estimate ($\hat{N} = \hat{D}A$). (Note that terms -log(n!) and +log(n!) have cancelled).

References

Borchers DL, Efford MG 2007. Spatially explicit maximum likelihood methods for capture–recapture studies. Biometrics OnlineEarly doi:10.1111/j.1541-0420.2007.00927.x.

Chao A, Huggins RM 2005b. Modern closed population models. In: Amstrup SC, McDonald TL, Manly BFJ (eds) Handbook of capture–recapture methods. Princeton University Press, pp 58–87.

Cooch E, White GC 2006. Program MARK: A gentle introduction. 4th edition. www.phidot.org/software/mark/docs/book/

Efford MG 2007. DENSITY 4.1: software for spatially explicit capture–recapture. Department of Zoology, University of Otago, Dunedin, New Zealand. http://www.otago.ac.nz/density.

Efford MG, Borchers DL, Byrom AE In press. Density estimation by spatially explicit capture–recapture: likelihood-based methods. Environmental and Ecological Statistics.

Huggins RM 1989. On the statistical analysis of capture experiments. Biometrika 76: 133–140.

Lebreton J-D, Burnham KP, Clobert J, Anderson DR 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62: 67–118.

Pledger SA 2000. Unified maximum likelihood estimates for closed capture-recapture models using mixtures. Biometrics 56: 434–442.