Science and Technology, 3 (1998) 67-75

Recursive Estimation in Capture-Recapture Methods

Lakhdar Aggoun[*] and Robert J. Elliott[#]

[*]Department of Mathematics and Statistics, College of Science, Sultan Qaboos University, P.O.Box 36, Al-Khod 123, Muscat, Sultanate of Oman, E-mail: aggounl@squ.edu.om; and [#]Department of Mathematical Sciences, University of Alberta, Edmonton, AB Canada, E-mail: relliott@gpu.srv.ualberta.ca.

التقدير الشرطي المتكرر لمجموعة حيوانية

الأخضر عقون ، وروبرت اليوت

خلاصة : هذا المقال يعالج مسألة مهمة في الإحصاء البيئي وهي مسألة حساب التعداد الحيواني . الطريقة المستعملة تفسر كالآتي : تؤخذ عينة عشوائية من المجموعة الحيوانية وتميز بإشارة خاصة ثم يطلق سراحها لتنضم لبقية القطيع ، وبعد فترة تؤخذ عينة عشوائية من القطيع ككل ويحصى منها الأفراد المميزين بالإشارة الخاصة ومن هذه المعلومات تستعمل طرق تغيّر القياس لتقدير التوزيع الشرطي المتكرر لهذه المجموعة الحيوانية .

ABSTRACT: An important problem in statistical ecology is how to determine the size of an animal population. The best known technique is the capture-recapture technique. A random sample of individuals is captured, tagged in some way and released back into the population. After allowing time for the marked and unmarked individuals to mix sufficiently, a second random sample is taken and the marked ones are observed. Using measure change techniques, we estimate recursively conditional distributions of various quantities.

Key Words: Capture Recapture, measure change, unnormalized conditional distributions, parameter estimation, Gaussian noise.

Hidden Markov Models (Elliott *et al*, 1995) have been used extensively in such areas as Engineering, Computer Science, Communications, Medical Imaging etc. In this paper we are using measure change techniques to estimate the hidden number of individuals in an animal population using partial information provided by the so-called capture-recapture technique.

A random sample of individuals are captured, tagged or marked in some way, and then released back into the population. After allowing time for the marked and unmarked individuals to mix sufficiently, a second simple random sample is taken and the marked ones are observed.

At epoch l write N_l for the population size, m_l for the number of marked and released individuals, $\bar{n}_k = \sum_{l=1}^{k} m_l$ for the total number of captured and marked individuals up to time k, M_l for the sample size, n_l for the number of available marked individuals for sampling and y_l for the number of captured (or recaptured) marked individuals.

Recursive estimation and Maximum Posterior estimators are discussed.

The model proposed in Section 2 (see autoregressive equation 1 below) leads to recursions for the unnormalized conditional probability distribution of the hidden number of individual which involves integrations. However the model proposed in Section 3 (a finite state Markov chain) leads to finite dimensional filters involving only finite summations.
Finally the case where tags could be lost between samples is discussed and a Martingale 'noise' present in Markov chains is replaced by a Gaussian noise as suggested by (Krichagina *et al*, 1985).

A First Model

All random variables are defined initially on a probability space (Ω, F, P). All the filtrations defined here will be assumed to be complete.

Write $G_k = \sigma(N_l, n_l, y_l, M_l, l \leq k)$, and $y_k = \sigma(y_l, l \leq k)$. We assume here that

1. The population size N follow the dynamics:

$$N_k = N_{k-1} + s(N_{k-1})v_k \qquad (1)$$

N_0 has distribution π_0 and v_k is a sequence of independent random variables with densities ϕ_k.

2. The n_k are random variables with conditional binomial distributions with parameters

$p_k = p(\tilde{n}_k, y_1, \ldots, y_k, \theta)$ and $\tilde{n}_k$. For example

$$\begin{aligned} p_1 &= \frac{\theta m_1}{m_1} = \theta \\ p_2 &= \frac{\theta m_2 + \theta^2 m_1}{m_1 + m_2} = \frac{\theta m_2 + \theta^2 m_1}{\tilde{n}_2}, \ldots, \\ p_k &= \frac{\sum_{i=1}^{k} m_i \theta^{k-i+1}}{\tilde{n}_k} = \frac{\tilde{n}_{k-1}}{\tilde{n}_k}\theta p_{k-1} + \theta\frac{m_k}{\tilde{n}_k}. \end{aligned} \qquad (2)$$

$0 < \theta \le 1$ is a parameter assumed to be known or it is to be estimated. The powers of θ express our belief that as time goes by early marked individuals are becoming less and less available for recapture due to various causes including deaths, emigration, etc.

If the number of captured and marked individuals $\tilde{n}_l$ is kept constant (2) takes the form:

$$p_k(\theta) = \frac{k-1}{k} p_{k-1} + \frac{\theta^k}{k}.$$

3. The observed random variable y_k is assumed to have a conditional binomial distribution:

$$p(y_k = m | G_{k-1}, N_k, M_k) = \binom{M_k}{m}\left(\frac{n_k}{N_k}\right)^m \left(1 - \frac{n_k}{N_k}\right)^{M_k - m}$$

Define $\lambda_0 = 1$. For $l \ge 1$ and for suitable density functions ψ_l write:

$$\lambda_l = \frac{s(N_{l-1})\Psi_l(N_l)}{\phi_l(v_l)} \frac{1}{2^{M_l - \tilde{n}_l}} p_l^{-n_l}(1-p_l)^{-\tilde{n}_l + n_l}\left(\frac{n_l}{N_l}\right)^{-y_l}\left(1 - \frac{n_l}{N_l}\right)^{y_l - M_l}$$

and $\Lambda_k = \Pi_{l=0}^{k} \lambda_l$

LEMMA 2.1: The process Λ_k is a G-martingale

PROOF: $E[\Lambda_k | G_{k-1}] = \Lambda_{k-1} E[\lambda_k | G_{k-1}]$, so we must show that $E[\lambda_k | G_{k-1}] = 1$.

$$\begin{aligned} E[\lambda_k | G_{k-1}] &= E\left[\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)} \frac{1}{2^{M_k + \tilde{n}_k}}\left(\frac{n_k}{N_k}\right) - y_k\left(1 - \frac{n_k}{N_k}\right)^{y_k - M_k} p_k^{-n_k}(1-p_k)^{-\tilde{n}_k + n_k} | G_{k-1}\right] \\ &= \frac{1}{2^{m_k + \tilde{n}_k}} E\left[\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)} p_k^{-n_k}(1-p_k)^{-\tilde{n}_k - n_k} E\left[\left(\frac{n_k}{N_k}\right)^{-y_k}\left(1 - \frac{n_k}{N_k}\right)^{y_k - M_k} | G_{k-1}, N_k, n_k, M_k\right] | G_{k-1}\right] \\ &= \frac{1}{2^{M_k + \tilde{n}_k}} E\left[\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)} p_k^{-n_k}(1-p_k)^{-\tilde{n}_k + n_k} \sum_{m=0}^{M_k} \binom{M_k}{m} | G_{k-1}\right] \\ &= \frac{1}{2^{\tilde{n}_k}} E\left[\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)} p_k^{n_k}(1-p_k)^{-\tilde{n}_k + n_k} E\left[\left(\frac{n_k}{N_k}\right)^{y_k}\left(1 - \frac{n_k}{N_k}\right)^{y_k - M_k} | G_{k-1}\right]\right] \\ &= E\left[\frac{s(N_{k-1})\psi_k(N_{k-1} + s(N_{k-1})v_k)}{\phi_k(v_k)} | G_{k-1}\right] \\ &= \int \frac{s(N_{k-1})\psi_k(N_{k-1} + s(N_{k-1})v)}{\phi_k(v)} \phi_k(v)dv = \int \psi_k(u)du = 1 \end{aligned}$$

A new probability mesure Q can be defined by setting $dQ/dP|_{G_k}=\Lambda_k$. The point here is that:

LEMMA 2.2: Under the new probability mesure Q, N_k, n_k and y_k are three sequences of independent random variables which are independent of each other. Further, N_k has density ψ_k, n_k has distribution bin $\left(\tilde{n}_k,\frac{1}{2}\right)$ and y_k has distribution bin $\left(M_k,\frac{1}{2}\right)$.

PROOF: For any integrable real-valued functions f, g and h and using a version of Bayes' theorem (see Elliott *et al*, 1995) we can write:

$$E_Q[f(N_k)g(n_k)h(y_k)|G_{k-1}]=\frac{E[f(N_k)g(n_k)h(y_k)\Lambda_k|G_{k-1}]}{E[\Lambda_k|G_{k-1}]}$$
$$=E[f(N_k)g(n_k)h(y_k)\lambda_k|G_{k-1}]$$
$$=E[f(N_k)g(n_k)h(y_k)\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)}\frac{1}{2^{\tilde{n}_k}}p_k^{-n_k}(1-p_k)^{-\tilde{n}_k+n_k}$$
$$E\left[h(y_k)\frac{1}{2^{M_k}}\left(\frac{n_k}{N_k}\right)^{-y_k}\left(1-\frac{n_k}{N_k}\right)^{y_k-M_k}|G_{k-1},N_k,n_k,M_k]|G_{k-1}\right]$$
$$=E[f(N_k)g(n_k)h(y_k)\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)}\frac{1}{2^{\tilde{n}_k}}p_k^{-n_k}(1-p_k)^{-\tilde{n}_k+n_k}$$
$$\left[\sum_{m=0}^{M_k}h(m)\left(\frac{n_k}{N_k}\right)^{m}\left(1-\frac{n_k}{N_k}\right)^{-m+M_k}\binom{M_k}{m}\frac{1}{2^{M_k}}\left(\frac{n_k}{N_k}\right)^{-m}\left(1-\frac{n_k}{N_k}\right)^{m-M_k}\right]\Bigg]$$
$$=E\Bigg[f(N_k)g(n_k)h(y_k)\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)}\frac{1}{2^{\tilde{n}_k}}p_k^{-n_k}(1-p_k)^{-\tilde{n}_k+n_k}$$
$$\left[\sum_{m=0}^{M_k}h(m)\binom{M_k}{m}\frac{1}{2^{M_k}}\right]G_{k-1}\Bigg]$$
$$=E_Q[h(y_k)]E\left[f(N_k)g(n_k)\frac{s(N_{k-1})\psi_k(N_k)}{\phi_k(v_k)}\frac{1}{2^{\tilde{n}_k}}(1-p_k)^{-\tilde{n}_k-n_k}|G_{k-1}\right]$$
$$=E_Q[h(y_k)]E_Q[g(n_k)]E[f(N_{k-1}+v_k)\frac{s(N_{k-1})\psi_k(N_{k-1}+s(N_{k-1})v_k)}{\phi_k(v_k)}|G_{k-1}]$$
$$=E_Q[h(y_k)]E_Q[g(n_k)]\int f(N_{k-1}+s(N_{k-1})v)s(N_{k-1})\psi_k(N_{k-1}+s(N_{k-1}v)dv$$
$$=E_Q[h(y_k)]E_Q[g(n_k)]\int f(u)\psi_k(u)du$$
$$=E_Q[h(y_k)]E_Q[g(n_k)]E_Q[f(N_k)]$$

That is, under Q the three processes are independent sequences of random variables with the desired distributions.

Using this fact we derive a recursive equation for the unnormalized conditional distribution of N_k given y_k. For any measurable test function f consider:

$$E[f(N_k)|Y_k]=\frac{E_Q[f(N_k)\Lambda_k^{-1}|Y_k]}{E_Q[\Lambda_k^{-1}|y_k]}. \tag{3}$$

The denominator of (3) being a normalizing factor we focus only on the expectation under Q in the numerator. Write

$$E_Q[f(N_k)\Lambda_k^{-1}|Y_k]=\int f(z)q_k(z)dz. \tag{4}$$

In view of Lemma 2.2 the left hand side of (4) is:

$$= E_Q[f(N_k)\Lambda_{k-1}^{-1}\lambda_k^{-1}|Y_k]$$

$$= 2^{\bar{n}_k \cdot M_k} E_Q\Big[\sum_{i=0}^{\bar{n}_k} \int f(z)\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k} \frac{\phi_k\left(\frac{z-N_{k-1}}{s(N_{k-1})}\right)}{s(N_{k-1})\psi_k(z)}\psi_k(z)$$

$$p_k^i(1-p_k)^{\bar{n}_k-i} dz \frac{1}{2^{\bar{n}_k}}\binom{\bar{n}_k}{i}\Lambda_{k-1}^{-1}|Y_k]$$

$$= 2^{M_k}\sum_{i=0}^{\bar{n}_k}\int\int f(z)\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}\frac{\phi_k\left(\frac{z-u}{s(u)}\right)}{s(u)}$$

$$p_k^i(1-p_k)^{\bar{n}_{k}-i}\binom{\bar{n}_k}{i} q_{k-1}(u)dzdu$$

Comparing this last expression with the right hand side of (4) we have:

THEOREM 2.3: The unnormalized conditional probability density function of the hidden Markov model given by (1) follows the recursions:

$$q_k(z) = 2^{M_k}\sum_{i=0}^{\bar{n}_k}\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}\binom{\bar{n}_k}{i}$$

$$\cdot\; p_k^i\left(1-p_k\right)^{\bar{n}_k-i}\int\frac{\phi_k\left(\frac{z-u}{s(u)}\right)}{s(u)}q_{k-1}(u)du \tag{5}$$

(**Note:** *we take* $0^0 = 1$).

REMARK 2.4: The normalized conditional density of N_k is given by $\dfrac{q_k(z)}{\int q_k(u)du}$. The initial (normalized) probability density of N_0, prior to sampling, is $\pi_0(.)$, so:

$$q_0(z) = \pi_0(z)$$

$$q_1(z) = 2^{M_1}\sum_{i=0}^{\bar{n}_1}\left(\frac{i}{z}\right)^{y_1}\left(1-\frac{i}{z}\right)^{M_1-y_1}\binom{\bar{n}_1}{i}p_1^i(1-p_1)^{\bar{n}_1-i}$$

$$\int\frac{\phi_1\left(\frac{z-u}{s(u)}\right)}{s(u)}\pi_0(u)du \tag{6}$$

and further estimates follow from (8).

If the distribution of N_0 is a delta function concentrated at some number A, (9) becomes:

$$q_1(z) = 2^{M_1}\sum_{i=0}^{\bar{n}_1}\left(\frac{i}{z}\right)^{y_1}\left(1-\frac{i}{z}\right)^{M_1-y_1}\binom{\bar{n}_1}{i}p_1^i(1-p_1)^{\bar{n}_1-i}\frac{\phi_1\left(\frac{z-A}{s(A)}\right)}{s(A)} \tag{7}$$

PARAMETRIC ESTIMATION: Our model is function of the parameter p_k, the proportion of the accessible marked individuals at epoch k. Suppose p_k has dynamics given by (2). We also assume that θ will take values in some measurable space (Θ, β, γ). We now derive a recursive joint conditional unnormalized distribution for N_k and θ. We keep working under the probability measure Q.

THEOREM 2.5: Write

$q_k(z,\theta)dzd\theta = E_Q[I(N_k \in dz, \theta \in d\theta)\Lambda_k^{-1}|Y_k]$. Then:

$$q_k(z,\theta) = 2^{M_k}\sum_{i=0}^{\bar{n}_k}\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}p_k(\theta)^i\left(1-p_k(\theta)\right)^{\bar{n}_k-i}\binom{\bar{n}_k}{i}$$

$$\cdot\int\frac{\phi_k\left(\frac{z-u}{s(u)}\right)}{s(u)}q_{k-1}(u,\theta)du \tag{8}$$

PROOF: Let f, g be integrable test functions.

$$E_Q[f(N_k)g(\theta)\Lambda_k^{-1}|y_k] = \int\int f(z)g(v)q_k(z,v)dzd\gamma(v). \tag{9}$$

Using the independence assumption under Q the left hand side of (12) is:

$$= E_Q[f(N_k)g(\theta)\Lambda_{k-1}^{-1}\lambda_k^{-1}|Y_k]$$

$$= 2^{M_k}E_Q\Big[\sum_{i=0}^{\bar{n}_k}\int\int f(z)g(v)\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}\frac{\phi_k\left(\frac{z-N_{k-1}}{s(N_{k-1})}\right)}{s(N_{k-1})\psi_k(z)}\psi_k(z)$$

$$\cdot\; p_k(v)^i(1-p_k(v))^{\bar{n}_k-1}dzd\gamma(v)\binom{\bar{n}_k}{i}\Lambda_{k-1}^{-1}|Y_k\Big]$$

$$= 2^{M_k}\Big[\sum_{i=0}^{\bar{n}_k}\int\int\int f(z)g(v)\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}\frac{\phi_k\left(\frac{z-u}{s(u)}\right)}{s(u)}$$

$$\cdot\; p_k(v)^i(1-p_k(v))^{\bar{n}_k-i}\binom{\bar{n}_k}{i}q_{k-1}(u,v)dzdud\gamma(v)$$

Comparing this last expression with the right hand side of (9) gives (8). If at time 0, θ has density $h(\theta)$ then:

$$q_1(z,\theta)=2^{M_k}\sum_{i=0}^{\tilde{n}_k}\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}$$

$$p_1(\theta)^i\left(1-p_1(\theta)\right)^{\tilde{n}_k-i}\binom{\tilde{n}_k-i}{i}h(\theta)\int\frac{\phi_k\left(\frac{z-u}{s(u)}\right)}{s(u)}\pi_0(u)du$$

and further updates are given by Theorem 2.5

If no dynamics enter the population size and N_k has density $\phi_k(.)$ independently of N_l, $l<k$, the recursion in Theorem 2.5 simplifies to:

$$q_k(z,\theta)=\phi_k(z)q_{k-1}(z,\theta)2^{M_k}\sum_{i=0}^{\tilde{n}_k}\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{1}{z}\right)^{M_k-y_k}$$

$$p_k(\theta)^i\left(1-p_k(\theta)\right)^{\tilde{n}_k-i}\binom{\tilde{n}_k}{i} \qquad (10)$$

MAXIMUM POSTERIOR ESTIMATORS: Quantity (6) (or 7) is a function of the unknown population size and could be maximized with respect to z yielding a critical value $\hat{N}_1$ which is the Maximum Posterior estimate of N at epoch 1 given y_1. Similar maximizations at later times will provide Maximum posterior Estimators (MAP) for the population size at these times.

PATHWISE ESTIMATION: We now derive a recursive equation, which does not involve any integration, for the unnormalized density of the whole path up to epoch k. Write

$$q_k(z_0,...,z_k)dz_0...dz_k := E_Q[I(z_0\in dz_0)...I(z_k\in dz_k)\Lambda_k^{-1}|Y_k].$$

THEOREM 2.6:

$$q_k(z_0,...,z_k)=2^{M_k}\sum_{i=0}^{\tilde{n}_k}\left(\frac{i}{z}\right)^{y_k}\left(1-\frac{i}{z}\right)^{M_k-y_k}\binom{\tilde{n}_k}{i}$$

$$p_k^i(1-p_k)^{\tilde{n}_k-i}\frac{\phi_k\left(\frac{z_k-z_{k-1}}{s(z_{k-1})}\right)}{s(z_{k-1})}q_{k-1}(z_0,...,z_{k-1}) \qquad (11)$$

Again we have:

$$q_0(z)=\pi_0(z)$$

$$q_1(z_0,z_1)=2^{M_1}\sum_{i=0}^{\tilde{n}_1}\left(\frac{i}{z}\right)^{M_1-y_1}\binom{\tilde{n}_1}{i}p_k^i\left(1-p_1\right)^{\tilde{n}_1-i}\frac{\phi_1\left(\frac{z_1-z_0}{s(z_0)}\right)}{s(z_0)}\pi_0(z_0)$$

and further estimates follow from (11).

However no integration is needed in subsequent recursions.

MAXIMUM POSTERIOR ESTIMATORS: Expression (11) is a function of the path $(z_0, \dots, z_k)$ and could be maximized yielding a critical path $(\hat{N}_0,...,\hat{N}_k)$. Since no integration is involved here one could substitute, at a time k say, the sequence of critical values $\hat{N}_0,...,\hat{N}_{k-1}$ and then maximize $q_k(\hat{N}_0,...,\hat{N}_{k-1},z_k)$ with respect to the variable z_k to obtain an estimate for N_k.

A Second Model

Suppose that on a probability space (Ω, F, Q) are given three sequences of independent random variables N_k, n_k and y_k. For $k\in\mathbb{N}$, N_k is uniformly distributed over some finite set $S=\{s_1,\dots,s_L\}\subset\mathbb{N}-\{0\}$, n_k has a binomial distribution with parameters $\left(\tilde{n}_k,\frac{1}{2}\right)$ and y_k has a binomial distribution with parameters $\left(M_k,\frac{1}{2}\right)$ where $M_k\in\mathbb{N}-\{0\}$ is given.

We wish to define a new probability measure P such that y_k has a binomial distribution with parameters $\left(M_k,\frac{n_k}{N_k}\right)$, N_k is a Markov chain with state space S and stochastic matrix $C=\{c_{ij}\}=P[N_{k+1}=s_i|N_k=s_j]$, n_k are random variables with conditional distributions with parameters $(p_k,\tilde{n}_k)$.

Define the G-predictable sequences

$$\alpha_l^i=\sum_{j=1}^{L}I(N_{l-1}=s_j)c_{ij}$$

for $l=1,\dots,L$. In vector notation this is

$$\alpha_l(N_{l-1})=C^TI(N_{l-1})$$

where $I(N_{l-1})=(I(N_{l-1}=s_1),...,I(N_{l-1}=s_L))$

Now write $\lambda_0=1$,

$$\lambda_l=2^{M_l+\tilde{n}_l}p_l^{n_l}(1-p_l)^{\tilde{n}_l-n_l}\left(\frac{n_l}{N_l}\right)^{y_l}\left(1-\frac{n_l}{N_l}\right)^{M_l-y_l}\prod_{i=1}^{L}(LA_l^i)^{I(N_l=s_i)},$$

and $$\Lambda_k=\prod_{l=0}^{k}\lambda_l.$$

The process Λ_k is a G-martingale and a new probability measure P can be defined by setting $\frac{dp}{dQ}|_{G_k} = \lambda_k$.

LEMMA 3.1: Under the probability measure P the above processes obey the desired dynamics, i.e. N_k is a Markov chain with state space S and stochastic matrix C, y_k and n_k are random variables with conditional binomial distributions with parameters $\left(M_k, \frac{n_k}{N_k}\right)$ and $(p_k, \tilde{n}_k)$ respectively.

PROOF: We give proof only for the first statement regarding N_k.

$$
\begin{aligned}
P[N_k = s_j | G_{k-1}] &= E[I(N_k = s_j) | G_{k-1}] \\
&= \frac{E_Q[I(N_k = s_j)\Lambda_k | G_{k-1}]}{E_Q[\Lambda_k | G_{k-1}]} \\
&= E_Q[I(N_k = s_j)\lambda_k | G_{k-1}] \\
&= E_Q[I(N_k = s_j) 2^{M_k + \tilde{n}_k} p_l^{n_k} (1-p_k)^{\tilde{n}_k - n_k} \left(\frac{n_k}{s_j}\right)^{M_k - y_k} L\alpha_k^j | G_{k-1}] \\
&= L\alpha_k^j 2^{M_k + \tilde{n}_k} E_Q[I(N_k = s_j) p_l^{n_k} (1-p_k)^{\tilde{n}_k - n_k} \left(\frac{n_k}{s_j}\right)^{y_k} \left(1 - \frac{n_k}{s_j}\right)^{M_k - y_k} | G_{k-1}] \\
&= \alpha_k^j L 2^{M_k + \tilde{n}_k} \frac{1}{L} \sum_{n=0}^{\tilde{n}_k} \sum_{m=0}^{M_k} \binom{M_k}{m} \left(\frac{n}{s_j}\right)^m \left(1 - \frac{n}{s_j}\right)^{m_k - m} \frac{1}{2^{M_k}} \\
&\quad \binom{\tilde{n}_k}{n} p_k^n (1-p_k)^{\tilde{n}_k - n} \frac{1}{2^{M_k}} \\
&= \alpha_k^j = P[N_k = s_j | N_{k-1}].
\end{aligned}
$$

Working under the probability measure Q, we derive recursive equations for the unnormalized conditional probability distribution of N_k. Write

$$
\begin{aligned}
P[N_k = s_j | Y_k] &= E[I(N_k = s_j) | Y_k] \\
&= \frac{E_Q[I(N_k = s_j)\Lambda_k | Y_k]}{E_Q[\Lambda_k | Y_k]}
\end{aligned}
$$

and $q_k^{s_j} = E_Q[I(N_k = s_j)\Lambda_k | Y_k]$.

THEOREM 3.2:

$$
q_k^{s_i} = 2^{M_k} \sum_{n=0}^{\tilde{n}_k} \left(\frac{n}{s_i}\right)^{y_k} \left(1 - \frac{n}{s_i}\right)^{M_k - y_k} \binom{\tilde{n}_k}{n} p_k^n (1-p_k)^{\tilde{n}_k - n} \cdot \sum_{j=1}^{L} c_{ij} q_{k-1}^{s_j} \tag{12}
$$

If at time 0, $q_0 = \pi = (\pi_1, \ldots, \pi_L)$

$$
q_1^{s_i} = 2^{M_1} \sum_{n=0}^{\tilde{n}_1} \left(\frac{n}{s_i}\right)^{y_1} \left(1 - \frac{n}{s_i}\right)^{M_1 - y_1} \binom{\tilde{n}_1}{n} p_1^n (1-p_1)^{\tilde{n}_1 - n} \sum_{j=1}^{L} c_{ij}\pi_j \tag{13}
$$

If $N_0 = s_\alpha$ with probability one:

$$
q_1^{s_i} = 2^{M_1} \sum_{n=0}^{\tilde{n}_1} \left(\frac{n}{s_i}\right)^{y_1} \left(1 - \frac{n}{s_i}\right)^{M_1 - y_1} \binom{\tilde{n}_1}{n} p_1^n (1-p_1)^{\tilde{n}_1 - n} c_{i\alpha} \tag{14}
$$

and further updates are given by (12).

MAP estimators of $N_1, \ldots, N_k$ are provided by:

$$
\begin{aligned}
\hat{N}_1 &= \arg\max\{q_1^{s_1}, q_1^{s_2}, \ldots, q_1^{s_L}\}, \ldots, \\
\hat{N}_k &= \arg\max\{q_k^{s_1}, q_k^{s_2}, \ldots, q_k^{s_L}\}.
\end{aligned}
$$

RECURSIVE PARAMETER ESTIMATION: The previous model is a function of the parameters p_k and $C = c_{ij}$. Let $p_k = p_k(\theta_1)$ and $C = C(\theta_2) = c_{ij}(\theta_2)$ and $\theta = (\theta_1, \theta_2)$. Suppose θ belongs to some measurable space (Θ, β, γ).

Working again under the probability measure Q write

$$
q_k^{s_i}(\theta)\, d\theta = E_Q[I(N_k = s_i) I(\theta \in d\theta)\Lambda_k | Y_k]
$$

THEOREM 3.3:

$$
\begin{aligned}
q_k^{s_i}(\theta) &= 2^{M_k} \sum_{n=0}^{\tilde{n}_k} \left(\frac{n}{s_i}\right)^{y_k} \left(1 - \frac{n}{s_i}\right)^{M_k - y_k} \binom{\tilde{n}_k}{n} p_k^n(\theta) \\
&\quad \cdot (1 - p_k(\theta))^{\tilde{n}_k - n} \sum_{j=1}^{L} c_{ij}(\theta) q_{k-1}^{s_j}(\theta)
\end{aligned}
$$

If θ_1 has density $h(.)$ and θ_2 has density $g(.)$:

$$
\begin{aligned}
q_1^{s_i}(\theta) &= 2^{M_1} \sum_{n=0}^{\tilde{n}_1} \left(\frac{n}{s_i}\right)^{y_1} \left(1 - \frac{n}{s_i}\right)^{M_1 - y_1} \binom{\tilde{n}_1}{n} p_1^n(\theta_1) \\
&\quad \cdot (1 - p_1(\theta_1))^{\tilde{n}_1 - n} h(\theta_1) g(\theta_2) \sum_{j=1}^{L} c_{ij}(\theta_2)\pi_j
\end{aligned}
$$

A Tags Loss Model

In this section we propose a model where the marks or tags are not permanent. In this situation the marking is done using double tagging where each individual is marked with two tags. For simplicity we assume that the two tags on each individual are nondistinguishable and that individuals can retain or lose their tags independently.

We start again with a probability space (Ω, F, Q) on which are given two sequences of independent random variables N_k and y_k. For $k \in \mathbb{N}$, N_k is uniformly distributed over some finite set $S = \{s_1, \ldots, s_L\} \subset \mathbb{N} - \{0\}$, and y_k has a trinomial distribution with parameters $\left(M_k, \frac{1}{3}\right)$ where $M_k \in \mathbb{N} - \{0\}$ is given.

At any epoch l each individual in the population is in any of three states, namely unmarked, marked with only one tag, or marked with two tags which states we shall call 0, 1, 2 respectively. We suppose that each individual behaves like an independent time homogenous Markov chain with transition matrix $\{p_{ij}\}$.

At each time l the population size N_l is distributed or partitioned into three groups $N_l(2)$, $N_l(1)$, and $N_l(0) = N_l - N_l(2) - N_l(1)$ among the three states and we would like to define the set of all such partitions as the states of a three dimensional Markov chain $(N_l(0), N_l(1), N_l(2))$. Recall that at each epoch l, $0 \le N_l(2), N_l(1) \le \tilde{n}_l$.

$$p_{(i_0,i_1,i_2),(j_0,j_1,j_2)} = P[(N_k(0),N_k(1),N_k(2)) = (i_0,i_1,i_2) \mid (N_{k-1}(0),N_{k-1}(1),N_{k-1}(2)) = (j_0,j_1,j_2)]$$

Write and for any real numbers x_0, x_1, x_2 define the function

$$F(x_0,x_1,x_2,j_0,j_1,j_2) = \left(\sum_{l=0}^{2} p_{l0}x_l\right)^{j_0}\left(\sum_{l=0}^{2} p_{l1}x_l\right)^{j_1}\left(\sum_{l=0}^{2} p_{l2}x_l\right)^{j_2}$$

Then $p(j_0 j_1 j_2, j_0, j_1 j_2)$ is the coefficient of $x_0^{i_0} x_1^{i_1} x_2^{i_2}$ in $F(x_0, x_1, x_2, j_0, j_1 j_2)$.

We wish to define a new probability measure P such that y_k has a conditional trinomial distribution with parameters $\left(M_k, \frac{N_k(0)}{N_k}, \frac{N_k(1)}{N_k}, \frac{N_k(2)}{N_k}\right)$, N_k is a Markov chain with state space S and stochastic matrix $C = \{c_{ij}\}$. The Markov chain $(N_l(0), N_l(1), N_l(2))$ is the same under both probability measures.

$$\alpha_l^i = \sum_{j=1}^{L} I(N_{l-1} = s_j)c_{ij}, \; i = 1, \ldots, L.$$

Define again the G-predictable sequences. Now write $\lambda_0 = 1$,

$$\lambda_l = 3^{M_l}\left(\frac{N_l(0)}{N_l}\right)^{y_l(0)}\left(\frac{N_l(1)}{N_l}\right)^{y_l(1)}\left(\frac{N_l(2)}{N_l}\right)^{y_l(2)} \prod_{i=1}^{L} (L\alpha_l^i)^{I(N_l = s_i)}$$

and

$$\Lambda_k = \prod_{l=0}^{L} \lambda_l.$$

The process Λ_k is a G-martingale and a new probability measure P can be defined by setting $\frac{dP}{dQ}\Big|_{Gk} = \Lambda_k$. It can be checked that under P the above processes have the desired distributions. Working under the probability measure Q, we derive recursive equations for the unnormalized conditional joint probability distribution of N_k and $(N_k(0), N_k(1), N_k(2))$. Write

$$\begin{aligned} P[N_k = s_i, (N_k(0),N_k(1),N_k(2)) = (i_0,i_0,i_2) \mid Y_k] &= E[I(N_k = s_i)I[(N_k(0),N_k(1),N_k(2)) = (i_0,i_1,i_2)] \mid Y_k] \\ &= \frac{E_Q[I(N_k = s_i)I[(N_k(0),N_k(1),N_k(2)) = (i_0,i_1,i_2)]\Lambda_k \mid Y_K]}{E_Q[\Lambda_k \mid Y_k]} \end{aligned}$$

and

$$q_k(s_i, i_1, i_2) = E_Q[I(N_k = s_i, N_k(2) = i_2, N_k(1) = i_1, N_k(0) = s_i - i_1 - i_2)\Lambda_k \mid Y_k].$$

It can be shown that $q_k(S_1, i_1, i_2)$ is given by the following recursions:

THEOREM 4.1:

$$q_k(s_i, i_1, i_2) = 3^{M_k}\left(\frac{s_i - i_1 - i_2}{s_i}\right)^{y_k(0)}\left(\frac{i_1}{s_i}\right)^{y_k(1)}\left(\frac{i_2}{s_i}\right)^{y_k(2)} \sum_{j=1}^{L} c_{ij} \cdot \sum_{j_1+j_2=0}^{\tilde{n}_k - 1} p_{(s_i - i_1 - i_2, i_1, i_2)(s_j - j_1 - j_2, j_1, j_2)} q_{k-1}(s_j, j_1, j_2)$$

The expected value of N_k given the observations Y_k is given by:

$$E[N_k|Y_k] = \frac{\sum_{i=1}^{L} s_i \sum_{i_1+i_2=0}^{\bar{n}_k} q_k(s_i, i_1, i_2)}{\sum_{i=1}^{L} \sum_{i_1+i_2=0}^{\bar{n}_k} q_k(s_i, i_1, i_2)}$$

Another way of looking at the problem is by considering only the subpopulation of tagged individuals in the definition of the Markov chain $(N_k(0), N_k(1), (N_k(2))$. In this case the state space is the set of all the partitions of the totality of tagged individuals into three groups: the one with two tags, the ones with one tag and the ones who lost both tags. Hence we write the total number of tagged individuals as $\bar{n}_k = \bar{n} = N_k(2) + N_k(1) + N_k(0)$. Note that, when sampling, we cannot observe directly members belonging to the group of individuals who lost their two tags as they become undistiguishable from the unmarked ones in the sample.

Now we assume that under Q the observation process is multinomial with parameters $\left(M_k, \frac{1}{4}\right)$ and under P it is (conditional) multinomial with parameters $\left(M_k, \frac{N_k(0)}{N_k}, \frac{N_k(1)}{N_k}, \frac{N_k(2)}{N_k}, \frac{N_k(u)}{N_k}\right)$.

Here $N_k(u)$ is the number of unmarked individuals in the population. Again note that $N_k(u)$ is not $N_k(0)$. Given $y_k(1)$, $y_k(2)$, the unobserved component $y(0)$, under the probability measure Q, is binomial with parameters $(M_k - y_k(1) - y_k(2), \frac{1}{2})$.

Write

$$\begin{aligned} P[N_k = s_i, (N_k(0), N_k(1), N_k(2)) = (i_0, i_0, i_2) \mid Y_k] &= E[I(N_k = s_i) I[(N_k(0), N_k(1) N_k(2)) = (i_0, i_1, i_2)] \mid Y_k] \\ &= \frac{E_Q[I(N_k = s_i) I[(N_k(0), N_k(1), N_k(2)) = (i_0, i_1, i_2)] \Lambda_K \mid Y_K]}{E_Q[\Lambda_k \mid Y_k]} \end{aligned}$$

and

$$\begin{aligned} q_k(s_i, i_0, i_1, i_2) &= E_Q[I(N_k = s_i) I((N_k(0), N_k(1), N_k(2)) = (i_0, i_1, i_2)) \Lambda_k \mid Y_k] \\ &= E_Q[I(N_k = s_i) I((N_k(0), N_k(1), N_k(2)) = (i_0, i_1, i_2)) 4^{M_k} \left(\frac{i_0}{s_i}\right)^{y_i(0)} \\ &\bullet \left(\frac{i_1}{s_i}\right)^{y_i(1)} \left(\frac{i_2}{s_i}\right)^{y_i(2)} \left(\frac{i_u}{s_i}\right)^{y_i(u)} L \sum_{j=1}^{L} I(N_{k-1} = s_j) c_{ij} \Lambda_{k-1} \mid Y_k] \\ &= 4^{M_k} \left(\frac{i_1}{s_i}\right)^{y_i(1)} \left(\frac{i_2}{s_i}\right)^{y_i(2)} \\ &\bullet E_Q\left[\left(\frac{i_0}{s_i}\right)^{y_i(0)} \left(\frac{s_i - i_2 - i_1 - i_0}{s_i}\right)^{M_k - y_i(2) - y_i(1) - y_i(0)}\right. \\ &\bullet \sum_{j=1}^{L} I(N_{k-1} = s_j) c_{ij} I((N_k(0), N_k(1), N_k(2)) = i_0, i_1, i_2)) \\ &\bullet \left. \sum_{j_0 + j_1 + j_2 = \bar{n}} I((N_{k-1}, N_{k-1}(1), N_{k-1}(2)) = (j_0, j_1, j_2)) \Lambda_{k-1} \mid Y_k\right] \\ &= 4^{M_k} \left(\frac{i_1}{s_i}\right)^{y_i(1)} \left(\frac{i_2}{s_i}\right)^{y_i(2)} \\ &\bullet E_Q\left[\sum_{m=0}^{M_k - y_i(2) - y_i(1)} \left(\frac{i_0}{s_i}\right)^m \left(\frac{s_i - i_2 - i_1 - i_0}{s_i}\right)^{M_k - y_i(2) - y_i(1) - m}\right. \\ &\bullet \binom{M_k - y_i(2) - y_i(1)}{m} \left(\frac{1}{2}\right)^{M_k - y_k(2) - y_k(1)} \sum_{j=1}^{L} I(N_{k-1} = s_j) c_{ij} \\ &\bullet \left. p_{(i_0, i_1, i_2)(j_0 j_1 j_2)} \Lambda_{k-1} \mid Y_k\right] \end{aligned}$$

Using the definition of q we have:

THEOREM 4.2:

$$q_k(s_i,i_0,i_1,i_2)=2^{M_k+y_k(1)+y_k(2)}\left(\frac{i_1}{s_i}\right)^{y_k(1)}\left(\frac{i_2}{s_i}\right)^{y_k(2)}\sum_{m=0}^{M_k-y_k(2)-y_k(1)}\left(\frac{i_0}{s_i}\right)^{m}$$

$$\bullet\left(\frac{s_i-i_2-i_1-i_0}{s_i}\right)^{M_k-y_k(2)-y_k(1)-m}\binom{M_k-y_k(2)-y_k(1)}{m}$$

$$\bullet\sum_{j=1}^{L}c_{ij}\sum_{j_0+j_1+j_2=\tilde{n}}p_{(i_0,i_1,i_2)(j_0,j_1,j_2)}q_{k-1}(s_j,j_0,j_1,j_2)$$

Gaussian Noise Approximation

An approximate but simpler form of the recursion in Theorem 4.2 is to use a suggestion proposed by [2] where the martingale increment "noise" present in the representation of a Markov chain is replaced by Gaussian noise. To this effect, let's identify, as it is explained in [1], the three states 0, 1, 2 with the standard unit (column) vectors e_1, e_2, e_3 of $\Re^3$. Write $X_k^n\in\{e_1,e_2,e_3\}$ for the state of the n - th individual at time, $k, 1\le n\le\tilde{n}$. Then each individual behaves like a Markov chain on (Ω, F, P) with transition matrix P.

Define $X_k=\frac{1}{\tilde{n}}\sum_{n=1}^{\tilde{n}}X_k^n$. Then

$$X_k=PX_{k-1}+M_k \qquad (15)$$

where M_k is a martingale increment. The suggestion made in [2] is to replace the martingale increment M_k in (15) by an independent Gaussian random variable G_k of mean 0 and covariance matrix $E[M_kM_k^{'}]$ whose density is denoted by ϕ_k.

That is, the signal process x_k, taking values in $\Re^3$, has dynamics

$$x_k=Px_{k-1}+v_k \qquad (16)$$

We assume that under Q the observation process is multinomial with parameters $\left(M_k,\frac{1}{4}\right)$, x_k has density ϕ_k and N_k uniformly distributed over the set $(s_1,\ldots,s_L)$. Under the 'real world' probability measure P, N_k is a Markov chain with transition matrix C, x_k has dynamics (16) and y_k has conditional probability distribution given

$$P[y_k=y_k(2)+y_k(1)+y_k(0)+y_k(u)|M_k,N_k=s_i,x_k=(x_0,x_1,x_2),N_k(u)=s_i-x_2-x_1-x_0]=\binom{M_k}{y_k(2),y_k(1),y_k(0),y_k(u)}\left(\frac{x_0}{s_i}\right)^{y_k(1)}$$

$$\bullet\left(\frac{x_2}{s_i}\right)^{y_k(2)}\left(\frac{x_1}{s_i}\right)^{y_k(1)}\left(\frac{s_i-x_2-x_1-x_0}{s_i}\right)^{M_k-y_k(2)-y_k(1)-y_k(0)}$$

P is defined in terms of Q using th G-martingale

$$\Lambda_k=\prod_{l=0}^{k}4^{M_l}\left(\frac{x_l(0)}{N_l}\right)^{y_l(0)}\left(\frac{x_l(1)}{N_l}\right)^{y_l(1)}\left(\frac{x_l(2)}{N_l}\right)^{y_l(2)}$$

$$\bullet\left(\frac{x_2}{s_i}\right)^{y_k(2)}\left(\frac{x_1}{s_i}\right)^{y_k(1)}\left(\frac{s_1-x_2-x_1-x_0}{s_i}\right)^{M_k-y_k(2)-y_k(1)-y_k(0)}$$

Nest theorem is the analog of Theorem 4.2

THEOREM 5.1: The unnormalized joint conditional probability distribution of N_k and x_k, $E_Q[I(N_k=s_i)I(x_k\in dx)\Lambda_k|y_k]:=q_k^{s_i}(x)dx$, is given recursiively as follows:

$$q_k^{s_i}(x)=2^{M_k+y_k(2)+y_k(1)}\left(\frac{x_1}{s_i}\right)^{y_k(1)}\left(\frac{x_2}{s_i}\right)^{yk(2)}\sum_{m=0}^{M_k-y_k(2)-y_k(1)}\left(\frac{x_0}{s_i}\right)^{m}$$

$$\bullet\left(\frac{s_i-x_2-x_1-x_0}{s_i}\right)^{M_k-yk(2)-yk(1)-m}\binom{M_k-yk(2)-yk(1)}{m}$$

$$\bullet\sum_{j=0}Lc_{ij}\int\phi_k(x-Pu)q_{k-1}^{s_j}(u)du$$

Acknowledgements

R.J. Elliott wishes to acknowledge support of the Natural Sciences and Engineering Research Council of Canada, grant A7964.

References

ELLIOTT R.J., AGGOUN L., MOORE J.B. (1995) *Hidden Markov Models: Estimation and Control*, Applications of Mathematics. Vol. 29 Springer-Verlag, Berlin-Heidelberg-New York.

KRICHAGINA N.V., LIPSTER R.S., and RUBINOVICH E.Y. (1985) Kalman Filter for Markov Processes. *Steklov Seminar 1984*. Ed. N. V. Krylov, R.S. Lipster and a.A. Novikov, Optimization Software Inc., New York, pp. 197-213.

SEBER G.A.F. (1982) *The Estimation of Animal Abuindance and Related Parameters*. Second Edition. Edward Arnold.

Received 13 October 1997
Accepted 15 June 1998