The probability distribution of future population states

Richard Condit
15 October 2004

The goal is to describe completely the probability of all future population states, given the current population N_0, a time interval t, and descriptions of mortality and recruitment (= birth) processes. Call the desired probability distribution $p(N)$ for all N.

Survival can be modeled as a binomial process, so the probability g of observing S survivors at time t is

$$g(S) = \binom{N_0}{S} \theta^{tS} (1-\theta^t)^{N_0 - S}, \qquad \text{Eq, 1}$$

given θ as the per capita annual survival probability over the interval t (t is an arbitrary positive number, $\theta \in [0,1]$, N and S are positive integers with $S \le N$). In the binomial model, extending an annual survival probability into the future is obvious: θ^t is the survival probability after t years. A binomial model of survival is routinely used and widely accepted; it requires the assumption that all individuals are identical and have identical survival probability.

Recruitment is modeled as a Poisson process, so the probability h of observing R recruits at time t is

$$h(R) = \frac{(\pi_t N_0)^R e^{-\pi_t N_0}}{R!}, \qquad \text{Eq. 2}$$

where π_t is the per capita recruitment, or the expected number of offspring produced per individual, after t years. Unlike the binomial description of mortality, converting an annual recruitment rate π to a t-year recruitment probability, π_t, is convoluted:

$$\pi_t = \theta^t \left[\left(\frac{\pi}{\theta} + 1 \right)^t - 1 \right], \qquad \text{Eq. 3}$$

where π is the annual per capita recruitment. The unfortunate aspect is that the survival rate, θ, is involved, but this is necessary because recruits are dying throughout the interval *t*, and the goal is to define the recruitment rate as a function of N_0, the number alive at the start of the interval.

In real biological situations, the Poisson model of birth or recruitment is not straightforward to justify, at least compared to the binomial model of survival. In mammals and birds, where few offspring are produced at any one time and many survive, it makes good sense to model clutch or litter size as Poisson. But in organisms where large numbers of small seeds or larvae are produced and few recruit to a countable size, the Poisson description is not so clear. Nevertheless, it has the advantage of being simple, and I don't see a better alternative.

Accepting the binomial description, $g(S)$, of survival and the Poisson, $h(R)$, of recruitment, the goal is to write down the probability distribution $p(N)$ for each future population state. This is

$$p(N) = \sum_x h(x) g(N - x), \qquad \text{Eq. 4}$$

where the summation is over integers $x \in [0, N]$ if $N \le N_0$, or $x \in [N - N_0, N]$ if $N > N_0$. In words, the population will be *N* if there are *0* recruits and *N* survivors or if there is *1* recruit and $N - 1$ survivors or if there are *2* recruits and $N - 2$ survivors, etc.

I can't simplify the summation in Equation 4 after substituting Equations 1 and 2. But it can be calculated numerically, as in this R function (http://ctfs.si.edu/onlinepub/popbiology/ProbOfAbund.r); this produces exact probabilities. Equations 1 and 2 can also be simulated easily in R by drawing random samples on Poisson and

binomial variables (*rpois* and *rbinom*); summing the two random draws produces a random draw on $p(N)$. Both approaches indicate that $p(N)$ is symmetrical and bell-shaped, so I fit normal distributions to the results of the exact probabilities (Eq. 4).

This produced a surprisingly simple result:

$$p(N) \sim \eta\left(N_0(\pi+\theta), N_0\left(\pi+\theta-\theta^2\right)\right). \qquad \text{Eq. 5}$$

That is, $p(N)$ is normally distributed with mean equal to N_0 times recruitment plus survival (π and θ) and variance equal to N_0 times $\pi+\theta-\theta^2$. This is general, for any future time *t*, as long as π and θ indicate *t*-year recruitment and survival probabilities. I tested this over a reasonably wide range of N_0, π, and θ, and the fit is remarkably precise, even at $N = 1$. But I have not gone beyond proof by simulation; it would be interesting to see a general proof.

The neutral theory is a special case, with $\pi = 1-\theta$, in other words, where recruitment balances mortality (since mortality is $1-\theta$). In this circumstance, the mean of the probability distributions for recruitment and mortality are equal. Furthermore, it can be shown from Equation 3 that if $\pi = 1-\theta$ after one year, then $\pi_t = 1-\theta^t$ at any time in the future. In other words, given $\pi = 1-\theta$, then the mean of $p(N) = N_0$ for all times *t*; the population is neutrally stable. The variance is $N_0\left(1-\theta^{2t}\right)$, which approaches N_0 in the limit. There are caveats, though, relative to Hubbell's description of a neutral community: this formulation does not incorporate the zero-sum assumption across the community, and extinction and fixation invalidate the prediction far enough into the future.