

UQAR – Chaire de Recherche EEC

Stochastic models in ecology

Analytical and numerical methods

Dominique Gravel

<http://www.chaire-eec.uqar.ca/>

August 28, 2013

- 1 Introduction
- 2 Objectives
- 3 Theory
- 4 Stochastic models
- 5 Useful algorithms

Most biological processes are inherently stochastic, yet much theoretical analysis involves deterministic models. Why?

- ▶ Pragmatic reason: deterministic models are relatively easier to solve
- ▶ Biological systems typically consist of a large collection of individuals experiencing the same ecological interactions (assumes that the dynamics of the average is a sufficient description of the system)

The question then is *when does stochasticity matter and why?*

Is the average enough?

Consider populations A and B with the same initial density (100 individuals in each one). If the growth rate of population A at year 1 is $\lambda_A(1) = 1$ and at year 2 is $\lambda_A(2) = 2$, what is the population size after year 2? If the growth rate for population B is $\lambda_B(1) = 1.5$ and $\lambda_B(2) = 1.5$? Why is there a difference?

Program a stochastic version of the geometric growth model (discrete time) where $\lambda(t) = \mu + \varepsilon(t)$, where $\varepsilon(t)$ is a normal random deviate of mean 0 and variance σ^2 .

Start with $N_0 = 10$ and consider $\bar{\lambda} = 1.05$. Increases the variance of $\varepsilon(t)$ by steps of 0.01 and for each level run 100 replicated simulations over 100 time steps. Plot the average population size and the number of extinctions as a function of the variance.

1. Demographic stochasticity
2. Environmental stochasticity
3. Jansen's inequality
4. Storage effect

1. Playing with random numbers
2. Introduction to common algorithms
3. Adding stochasticity to discrete time models
4. Stochastic models
5. Common probability distribution

Before the theory

Essential programming tips

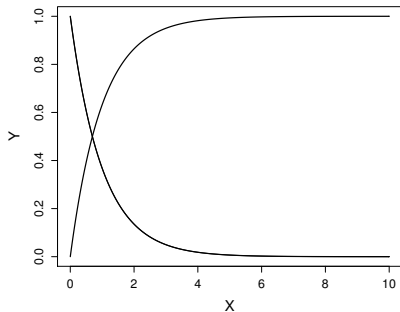
```
1 # Draw a random number between 0 and 1
2 rand = runif(n = 1, min = 0, max = 1)
3
4 # Set the seed of the random number generator
5 set.seed(1)
6 runif(n = 1, min = 0, max = 1)
7 runif(n = 1, min = 0, max = 1)
8
9 set.seed(1)
10 runif(n = 1, min = 0, max = 1)
11 runif(n = 1, min = 0, max = 1)
12
13 # Draw a random event A with probability p
14 test = 0
15 test[runif(n = 1, min = 0, max = 1) < p] = 1
16
17 # or alternatively
18 if(runif(1,0,1) < p) test = 1
19 else test = 0
```


The principle:

$$f(\bar{x}) \neq \overline{f(x)}$$

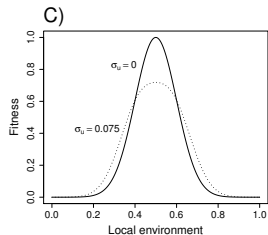
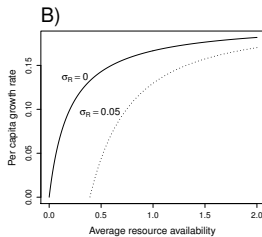
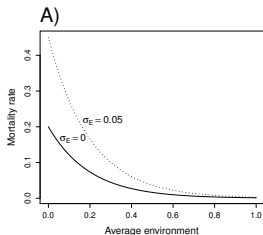
Consequences:

- ▶ $\overline{f(x)} > f(\bar{x})$ if the function is concave;
- ▶ $\overline{f(x)} < f(\bar{x})$ if the function is convex;
- ▶ $\overline{f(x)} = f(\bar{x})$ if the function linear;



Jensen's inequality

Examples

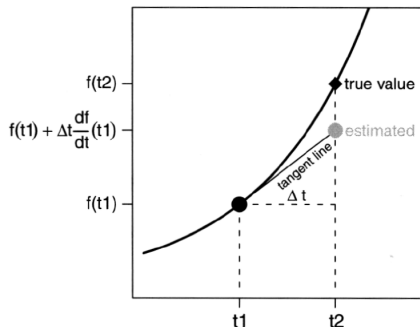


Computing the average of a non-linear function with the Taylor expansion:

$$f(x+\varepsilon) \cong f(x) + \frac{f'(x)(x-\varepsilon)}{1!} + \frac{f''(x)(x-\varepsilon)^2}{2!} + \dots$$

Computing it over several measurements:

$$\overline{f(x)} \cong f(\overline{x}) + \frac{f''(\overline{x})\sigma^2}{2}$$



Solution to the geometric growth model

$$N_{t+n} = N_0 \lambda^n$$

$$\log(N_{t+n}) = \log(N_0) + n \log(\lambda)$$

Estimating the average population size:

$$\overline{\log(N_{t+n})} = \log(N_0) + n \log(\bar{\lambda}) - \frac{n \sigma_{\lambda}^2}{2 \lambda^2}$$

Exercise: Calculate the maximal variance that a population of 100 individuals can stand if it has a growth rate of 1.025

Long-term average trajectory:

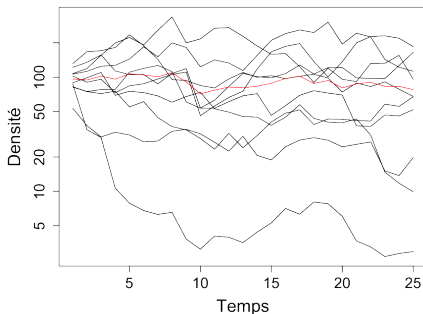
$$\overline{\log(N_{t+n})} = \log(N_0) + n\log(\bar{\lambda}) - \frac{n\sigma_{\lambda}^2}{2\bar{\lambda}^2}$$

To solve the problem, we consider there is no change on average in population size, and therefore:

$$n\log(\bar{\lambda}) = \frac{n\sigma_{\lambda}^2}{2\bar{\lambda}^2}$$

Isolating the variance term:

$$\sigma_{\lambda}^2 = 2\bar{\lambda}^2 \log(\bar{\lambda}) \sigma_{\lambda}^2 = 0.0519$$



Consider the problem of coexistence underlined by Armstrong and McGehee (1980). The community is made of two consumers (predators, P_1 and P_2) competing for a single resource (the prey, N). The dynamics are given by the following system of equations:

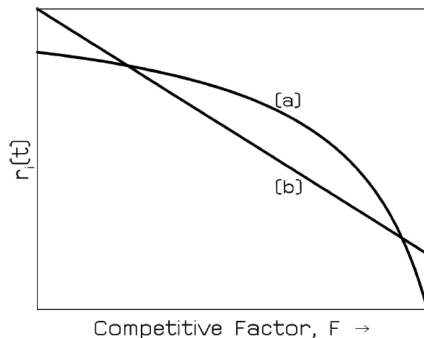
$$\frac{dN}{dt} = N(1 - N) - \frac{a_1 P_1 N}{1 + b_1 N} - \frac{a_2 P_2 N}{1 + b_2 N}$$

$$\frac{dP_1}{dt} = \frac{a_1 P_1 N}{1 + b_1 N} - m_1 P_1$$

$$\frac{dP_2}{dt} = \frac{a_2 P_2 N}{1 + b_2 N} - m_2 P_2$$

- ▶ What is the equilibrium N in presence of only one predator?
- ▶ If we consider $a_1 = 1$, $a_2 = 1$, $b_1 = 10$, $b_2 = 0$, $m_1 = 0.05$ and $m_2 = 0.2$, which species should win competition?
- ▶ Now run simulations of the model. Which species wins competition? Why?

Now we have an explanation for the coexistence observed in the Vandermeer (2006) paper. The mechanism is the *relative non-linearity*.



$$\bar{r}_i \cong b_i(k_i - k_S) - b_i(\tau_i - \tau_S)V(F^{-i})$$

Consider a simple discrete time birth-death model:

$$N_{t+\Delta t} = N_t + [B(N) - D(N)]\Delta t + \Delta Z$$

Because additive terms are statistically independent, the ensemble average yields:

$$\langle N_{t+\Delta t} \rangle = \langle N_t \rangle + [\langle B(N) \rangle - \langle D(N) \rangle]\Delta t + \langle \Delta Z \rangle$$

Because we are interested by the equilibrium and on average $\langle \Delta Z \rangle = 0$,

$$\langle N_{t+\Delta t} \rangle = \langle N_t \rangle$$

And therefore we are left with:

$$\langle B(N) \rangle = \langle D(N) \rangle$$

So if we are now more specific about B and D, using the logistic growth model, we have:

$$\begin{aligned}\langle (a_1 N - b_1 N^2) \rangle &= \langle (a_2 N - b_2 N^2) \rangle \\ a_1 \langle N \rangle - b_1 \langle N^2 \rangle &= a_2 \langle N \rangle - b_2 \langle N^2 \rangle\end{aligned}$$

Collecting terms:

$$(a_1 - a_2) \langle N \rangle = (b_1 + b_2) \langle N^2 \rangle$$

If we now define the mean, $m = \langle N \rangle$, and variance, $V(N) = \langle N^2 \rangle - m^2$, we have:

$$\begin{aligned}(a_1 - a_2)m &= (b_1 + b_2)(V(N) + m^2) \\ m &= \frac{a_1 - a_2}{b_1 + b_2} - \frac{V(N)}{m} \\ m &\cong \frac{r}{s} - \frac{s}{r} V(N)\end{aligned}$$

where $r = a_1 - a_2$ and $s = b_1 + b_2$. The approximation assumes a small variance-to-mean ratio. The stochastic mean collapses to the deterministic expectation as $V(N)$ tends to 0.

Birth and death are discrete events inherently variable, independently of environmental stochasticity. If we consider the population growth rate λ as an average of individual probabilities of birth and death, ω , we have:

$$\lambda = \bar{\omega}$$

$$\lambda = \frac{1}{N} \sum \omega_i = \mu_{\omega} + \frac{1}{N} \sum \delta_i$$

where μ is the population average and δ_i are individual differences to this average.

Now taking the variance of the population growth rate we obtain:

$$\sigma_{\lambda}^2 = \text{Var}[\mu_{\omega}] + \frac{1}{N^2} \sum \text{Var}[\delta_i]$$

$$\sigma_{\lambda}^2 = \sigma_e^2 + \frac{\sigma_d^2}{N}$$

In ecological terms: the impact of demographic stochasticity on the variability of the population growth rate shrinks asymptotically to 0 with large population size.

The lottery model:

$$N_{1,t+1} = (1 - d_1)N_{1,t} + (d_1 N_{1,t} + d_2 N_{2,t}) \frac{f_1(t)N_{1,t}}{f_1(t)N_{1,t} + f_1(t)N_{1,t}}$$

where d_i is the probability of mortality (which we will assume equal for both species for simplicity) and $f_i(t)$ is the fecundity. The dependance of t means that it could vary over time.

After some re-arrangement of this equation, we obtain the per capita growth rate of species 1 when an invader:

$$\frac{N_{1,t+1}}{N_{1,t}} = 1 + d \left[\frac{f_1(t)}{f_2(t)} - 1 \right]$$

Reciprocal invasibility is impossible in absence of any variability (the per capita growth rate of the species with the smallest average f will be lower than 1).

Using the formula for long-term averaging we obtain:

$$\frac{\overline{N_{1,t+1}}}{\overline{N_{1,t}}} = 1 + d\left[\frac{\overline{f_1}}{\overline{f_2}} - 1\right] + d\left[\frac{\overline{f_1}}{\overline{f_2}^3}\right]\sigma_{f_2}^2$$

The second derivative with respect to f_2 is positive, meaning that any variability in fecundity of species 2 will increase the growth rate of species 1 when at low abundance.

Ingredients for the storage effect:

- ▶ Differential response to the environment
- ▶ Covariance between environment and competition
- ▶ Buffered population growth

But the problem is that at low abundance, demographic stochasticity might play a role and kick-out the invader of the system, even if its long term average growth rate is positive.

With a fair amount of algebra, Cohen and Lewontin (1969) showed that the extinction risk for geometric growth is simply given by the integral of the normal distribution $G(\mu_x, \sigma_x^2/T)$ of the growth rate x with mean μ_G and variance σ_G^2/T :

$$Pr(N_t < N_0) = \int G(x; \mu_G; \frac{\sigma_G}{\sqrt{T}}) dx$$

Interpretation: The probability of extinction

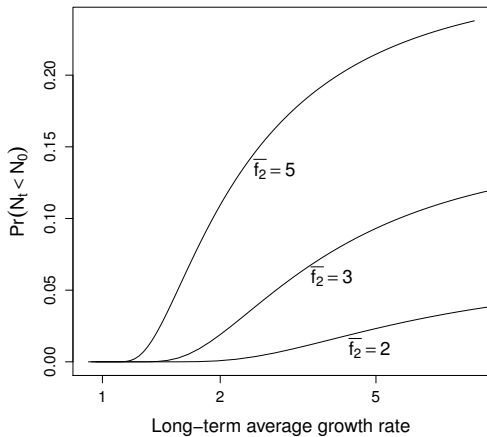
- ▶ Increases as the average per capita growth rate tends to zero
- ▶ Is a saturating positive function of the variance in the per capita growth rate

To understand the joint effects of stochasticity on coexistence, we now have to balance two potentially opposed forces:

- ▶ Stochasticity could promote the long term average population growth rate (- far from a rule though -)
- ▶ Stochasticity could increase the variance in population size, promote the drift and even random extinctions despite a positive growth rate.

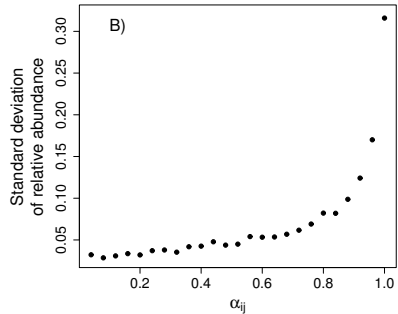
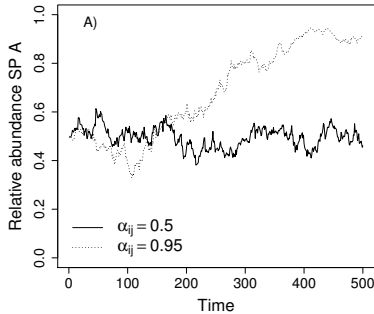
Coexistence in a variable world

Extinction versus storage effect



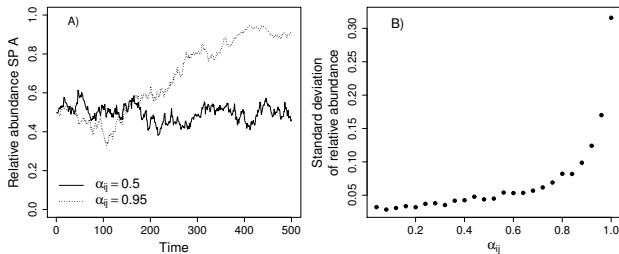
Coexistence in a variable world

Drift versus stability



Simulating stochastic models

Project A: additive noise



Ecological drift in a stable two species discrete version of the Lotka-Volterra model of competition. Environmental variability is introduced by independent normal random deviates to the carrying capacity of each species (mean of 0, standard deviation of 0.2). (a) Time series of species A relative abundance for weak ($a_{ji} = 1$ and $a_{jj} = 0.5$) and strong ($a_{ji} = 1$ and $a_{jj} = 0.95$) interactions. (b) Variability in species A relative abundance after 1000 time steps, based on 1000 replicated runs and as a function of the strength of interspecific competition. Maximal variability occurs when there is systematix fixation (one species dominates and the other goes extinct). Parameters: $r_1 = r_2 = 1.24$, $K_1 = K_2 = 1$.

Program an individual based lottery model. Consider a finite forest stand of 1 hectare, which holds approximately 250 adult trees. Recruitment follow a lottery: when an individual dies, with probability d , the identity of the recruited species is determined from a random draw among the seeds falling into the gap created by the death of the adult tree. Consider (for now) global dispersal within this forest patch. Fecundity could vary among species and years.

- ▶ Start with only 2 species and now environmental variation
- ▶ Add random variation in both species fecundity.
- ▶ Vary the death probability instead of fecundity
- ▶ Now consider the possibility of 10 species. Plot the expected species richness as a function of the variance in fecundity.

Draw 10 random numbers from a uniform distribution.

- ▶ calculate the mean
- ▶ calculate the variance
- ▶ find the largest and smallest values, without using min and max functions
- ▶ sort them
- ▶ find the median
- ▶ mix them again
- ▶ subsample 5 of them

<http://www.sorting-algorithms.com/>

The probabilities of observing events A, B and C are 0.2, 0.1 and 0.7, respectively. Program a function to draw one of these states randomly.

If probability of observing A is 0.3, observing B is 0.5 and the two are independent, what is the probability of observing neither A nor B? observing both A and B?

Write a function drawing the event B if the probability of observing is 0.2 when A is observed and 0.8 when A is not observed.

The probabilities of observing events A, B and C are 0.2, 0.1 and 0.7, respectively. Program a function to draw one of these states randomly.

Write a function to compute random numbers from a normal distribution based on a uniform random number generator

Idem but for an exponential random number generator