

**STOCHASTIC POPULATION MODELS
(SPRING 2011)**

STEFAN GERITZ
DEPARTMENT OF MATHEMATICS AND STATISTICS
UNIVERSITY OF HELSINKI

9. FOKKER-PLANCK APPROXIMATION FOR SEMI-LARGE SYSTEMS

9.1. **System size.** Consider the birth-death process with probability distribution $P_n(t)$ satisfying

$$(1) \quad \frac{dP_n}{dt} = B_{n-1}P_{n-1} + D_{n+1}P_{n+1} - (B_n + D_n)P_n \quad \forall n \geq 0$$

where $B_n = D_n = 0$ for $n \leq 0$. We introduce the notion of *system size* as the total area or volume in which the population lives. Given system size Ω and population size n , the population density is n/Ω . The objective is to re-write system (1) in terms of population density (instead of number of individuals) and probability density (instead of mere probability) and see what the effect is of increasing the system size. To this end we introduce the following change of variables:

$$(2) \quad \begin{aligned} \varepsilon &= \Omega^{-1} \\ x &= \varepsilon n \\ \varepsilon p(t, x) &= P_n(t) \\ b(x)/x &= B_n/n \\ d(x)/x &= D_n/n \end{aligned}$$

Note, that $\varepsilon = \Omega^{-1}$ is the smallest unit by which the population density can change: one individual per total system area or volume. Also note that $b(x)/x = B_n/n$ and $d(x)/x = D_n/n$ are the *per capita* birth and death rates expressed as functions of the population density x (left sides) or as a function of the number of individuals n (right sides).

System (1) rewritten in terms of the new variables becomes

$$(3) \quad \varepsilon \frac{\partial p(t, x)}{\partial t} = b(x - \varepsilon)p(t, x - \varepsilon) + d(x + \varepsilon)p(t, x + \varepsilon) - (b(x) + d(x))p(t, x)$$

Taylor-expansion for small ε gives

$$(4) \quad \frac{\partial p}{\partial t} = -\frac{\partial(b-d)p}{\partial x} + \frac{\varepsilon}{2} \frac{\partial^2(b+d)p}{\partial x^2} + O(\varepsilon^2)$$

9.2. Infinitely large systems. Letting $\varepsilon \rightarrow 0$, equation (4) converges point-wise (i.e., for each fixed x) to

$$(5) \quad \frac{\partial p}{\partial t} = -\frac{\partial(b-d)p}{\partial x}$$

which is known as the *transport equation*, and which corresponds to the deterministic population model

$$(6) \quad \frac{dx}{dt} = b(x) - d(x)$$

In other words, if we increase system size while keeping the population density constant, we lose all demographic stochasticity and eventually are left with a pure deterministic population model. Any uncertainty in the value of $x(t)$ as described by the distribution $p(t, x)$ is only due to uncertainty in the initial condition $x(0)$; no other source of stochasticity remains. The difference between equations (5) and (6) is that the former describes how the probability distribution changes in time while the latter describes a single (deterministic) orbit.

9.3. Semi-large systems. Instead of taking the limit $\varepsilon \rightarrow 0$ in equation (4) we also can simply take $\varepsilon > 0$ but small and ignore the $O(\varepsilon^2)$ terms, which are even smaller. This gives us the *Fokker-Planck* approximation of the birth-death process for *semi-large systems*, i.e., for systems that are large but still have a finite area or volume:

$$(7) \quad \frac{\partial p}{\partial t} = -\frac{\partial(b-d)p}{\partial x} + \frac{\varepsilon}{2} \frac{\partial^2(b+d)p}{\partial x^2}$$

The Fokker-Planck equation (7) (also known as the *Kolmogorov forward equation*) relates to the stochastic differential equation

$$(8) \quad dx = (b(x) - d(x))dt + \sqrt{\varepsilon(b(x) + d(x))} dW \quad (\text{Ito})$$

in the same way as the transport equation (5) relates to the deterministic equation (6), i.e., both the Fokker-Planck equation and the transport equation describe how the probability distribution of x at a given time changes with time, while both the deterministic equation (6) and the stochastic equation (8) describe a single orbit or sample path across time. In the case of (8), however, the effect of demographic stochasticity is not lost as it was in (6).

9.4. The meaning of $\mathbf{b}(\mathbf{x}) - \mathbf{d}(\mathbf{x})$ and $\mathbf{b}(\mathbf{x}) + \mathbf{d}(\mathbf{x})$. To understand the meaning of the $b(x) - d(x)$ and the $b(x) + d(x)$, recall from section 8.1 that the probability of a single birth event during Δt time is $B_n \Delta t + O(\Delta t)^2$. In terms of the new variables this is $\varepsilon^{-1} b(x) \Delta t + O(\Delta t)^2$. A single birth event gives a change in population density from x at time t to $x + \varepsilon$ at time $t + \Delta t$. Likewise, the probability of a single death event during Δt time is $D_n \Delta t + O(\Delta t)^2$. In terms of the new variables this is $\varepsilon^{-1} d(x) \Delta t + O(\Delta t)^2$. A single death event gives a change in population density from x at time t to $x - \varepsilon$ at time $t + \Delta t$:

event	Δx	probability in Δt time
birth	$+\varepsilon$	$\varepsilon^{-1}b(x)\Delta t + O(\Delta t)^2$
death	$-\varepsilon$	$\varepsilon^{-1}d(x)\Delta t + O(\Delta t)^2$

For the mean and variance of Δx over a time span Δt we thus find

$$(9) \quad \begin{aligned} \mathbb{E}\{\Delta x\} &= (b(x) - d(x))\Delta t + O(\Delta t)^2 \\ \text{Var}\{\Delta x\} &= \varepsilon(b(x) + d(x))\Delta t + O(\Delta t)^2 \end{aligned}$$

Division by Δt and letting $\Delta t \rightarrow 0$ gives the mean and variance of the change in x per unit of time:

$$(10) \quad \begin{aligned} \lim_{t \rightarrow 0} \frac{\mathbb{E}\{\Delta x\}}{\Delta t} &= b(x) - d(x) \\ \lim_{t \rightarrow 0} \frac{\text{Var}\{\Delta x\}}{\Delta t} &= \varepsilon(b(x) + d(x)) \end{aligned}$$

The $b(x) - d(x)$ and the $b(x) + d(x)$ are called, respectively, the deterministic drift and the demographic noise. We may use the notation

$$(11) \quad \begin{aligned} \mu(x) &:= \mathbb{E}\left\{\frac{dx}{dt}\right\} := b(x) - d(x) \\ \varepsilon\sigma^2(x) &:= \text{Var}\left\{\frac{dx}{dt}\right\} := \varepsilon(b(x) + d(x)) \end{aligned}$$

and accordingly

$$(12) \quad \partial_t p(t, x) = -\partial_x(\mu(x)p(t, x)) + \frac{\varepsilon}{2}\partial_{xx}(\sigma^2(x)p(t, x))$$

for the Fokker-Planck equation (7) and

$$(13) \quad dx = \mu(x)dt + \sqrt{\varepsilon\sigma^2(x)}dW \quad (\text{Ito})$$

for the corresponding stochastic differential equation (8).

9.5. Quasi-stationary distribution. Typically, equations (12) and (13) do not have a stationary solution for the same reason the original birth-death process of equation (1) does not have one. To study the quasi-stationary distribution in a semi-large system, we approximate the nonlinear SDE in (13) by a linear SDE centred at a deterministic equilibrium $\bar{x} > 0$ of the ODE (6), i.e., \bar{x} is a solution of

$$(14) \quad \mu(\bar{x}) = 0$$

which we assume to exist. To guarantee deterministic stability of \bar{x} we further assume that

$$(15) \quad \mu'(\bar{x}) < 0$$

Linearisation of the SDE near \bar{x} gives

$$(16) \quad d(x - \bar{x}) = \mu'(\bar{x})(x - \bar{x})dt + \sqrt{\varepsilon\sigma^2(\bar{x})}dW$$

which is the Ornstein-Uhlenbeck process. The stationary distribution of x therefore will be approximately Gaussian

$$(17) \quad x \sim \mathcal{N}\left(\bar{x}, \frac{\varepsilon\sigma^2(\bar{x})}{2|\mu'(\bar{x})|}\right)$$

with auto-covariance

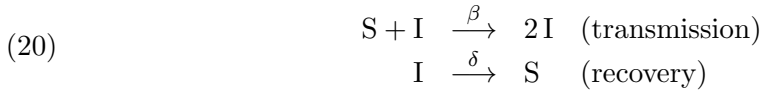
$$(18) \quad C(\tau) = \frac{\varepsilon\sigma^2(\bar{x})}{2|\mu'(\bar{x})|} e^{-|\tau\mu'(\bar{x})|}$$

and corresponding spectral density

$$(19) \quad S(\omega) = \frac{\varepsilon\sigma^2(\bar{x})}{\omega^2 + |\mu'(\bar{x})|^2}$$

The approximation improves towards smaller values of ε (i.e., larger system size) because the demographic noise in (16) becomes smaller and the population will stay closer to the deterministic equilibrium.

9.6. Example. Consider the SIS-model in which “S” denotes an uninfected (but “susceptible”) individual and “I” an infected individual, and consider the following processes:



The first process represents the transmission of an infection from an infected person to an uninfected person through direct contact. The second process represents the recovery of an infected individual but without acquiring immunity. From the point of view of the susceptible, the first represents a “death” event and the second a “birth” event. From the point of view of the infected it is the other way around. We here take the point of view of the infected.

Applying the principle of mass-action (see section 1.5) the population “birth” and “death” rates are

$$(21) \quad \begin{aligned} b(x) &= \beta x(k - x) \\ d(x) &= \delta x \end{aligned}$$

where x is the population density of infected individuals, k the total population density (which remains constant) and $k - x$ the population density of susceptible individuals. The deterministic drift and demographic noise thus are

$$(22) \quad \begin{aligned} \mu(x) &= \beta x(k - x) - \delta x \\ \sigma^2(x) &= \beta x(k - x) + \delta x \end{aligned}$$

The drift is zero for $x = 0$ or $x = k - \delta/\beta$, which are the deterministic equilibria. We assume that $k - \delta/\beta > 0$ so that we can take

$$(23) \quad \bar{x} = k - \delta/\beta$$

which is stable because $\mu'(\bar{x}) = -\beta(k - \delta/\beta) < 0$. The demographic noise at \bar{x} is $\sigma^2(\bar{x}) = 2\delta(k - \delta/\beta)$. As approximation of the quasi-stationary distribution of x we have

$$(24) \quad x \sim \mathcal{N}\left(\bar{x}, \frac{\delta\varepsilon}{\beta}\right)$$

with auto covariance

$$(25) \quad C(\tau) = \frac{\delta\varepsilon}{\beta} e^{-|\tau|(\beta k - \delta)}$$

and corresponding spectral density

$$(26) \quad S(\omega) = \frac{2\delta\varepsilon(k - \delta/\beta)}{\omega^2 + (\beta k - \delta)^2}$$

How good is the above approximation of the quasi-stationary distribution? To answer that question, we have to compare the approximation with the exact distribution, which can only be obtained from the original process for small populations as described by the system (1). We use the change of variable introduced in (2), but now in reverse, to get expressions of the B_n and the D_n , which gives

$$(27) \quad \begin{aligned} B_n &= b(\varepsilon n)/\varepsilon = \varepsilon\beta n(K - n) \\ D_n &= d(\varepsilon n)/\varepsilon = \delta n \end{aligned}$$

where

$$(28) \quad K := k/\varepsilon$$

is a positive integer. We now calculate the quasi-stationary distribution as the normalised right-eigenvector corresponding to the dominant eigenvalue of the tridiagonal matrix A given in equation (34) in section 8.6. The results are shown in the following figure for two different system sizes. It can be seen that the approximation becomes better as system size increases.

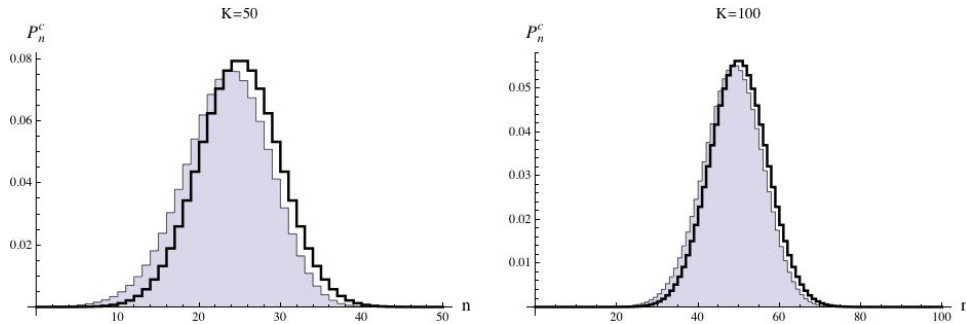


FIGURE 1. Exact stationary distribution (shaded) and the approximation (black) for $\beta = 2$, $\delta = 1$, $k = 1$, $\varepsilon = .1$ (left) and $\varepsilon = .01$ (right).

9.7. **Time till extinction.** Let $T(x)$ be the time till extinction of a population that presently has a density x , and let

$$(29) \quad R(t, x) := \text{Prob}\{T(x) > t\}$$

The probability density of $T(x)$ then is

$$(30) \quad r(t, x) := -\partial_t R(t, x)$$

Let further $k(\Delta x, x, \Delta t)$ denote the probability density of the change in population density Δx over Δt time if the present population density is x . The expectation and variance of Δx we already calculated in (9). Then,

$$(31) \quad \begin{aligned} R(t + \Delta t, x) &= \\ &= \int k(\Delta x, x, \Delta t) R(t, x + \Delta x) d\Delta x \\ &= \int k(\Delta x, x, \Delta t) \left[R(t, x) + \Delta x \partial_x R(t, x) + \frac{1}{2} \Delta x^2 \partial_{xx} R(t, x) + \dots \right] d\Delta x \\ &= R(t, x) + \mathbb{E}\{\Delta x\} \partial_x R(t, x) + \frac{1}{2} \mathbb{E}\{\Delta x^2\} \partial_{xx} R(t, x) + \dots \end{aligned}$$

where the dots represent third- and higher-order moments of Δx . These can be calculated explicitly from the table above equation (9), but they turn out to be of $O(\varepsilon^2)$ and so will be ignored as we did also in (4) to get the Fokker-Planck approximation (7) for semi-large systems. The above equation than can be re-organized as

$$(32) \quad \frac{R(t + \Delta t, x) - R(t, x)}{\Delta t} = \frac{\mathbb{E}\{\Delta x\}}{\Delta t} \partial_x R(t, x) + \frac{1}{2} \frac{\mathbb{E}\{\Delta x^2\}}{\Delta t} \partial_{xx} R(t, x)$$

The $\mathbb{E}\{\Delta x\}/\Delta t$ and $\mathbb{E}\{\Delta x^2\}/\Delta t$ we already calculated before in (9). Thus, in the limit $\Delta t \rightarrow 0$ we get the *Kolmogorov backward equation*

$$(33) \quad \partial_t R(t, x) = \mu(x) \partial_x R(t, x) + \frac{\varepsilon}{2} \sigma^2(x) \partial_{xx} R(t, x)$$

We are interested in solutions $R(t, x)$ that satisfy

$$(34) \quad \begin{aligned} R(t, 0) &= 0 \quad \text{and} \quad \partial_x R(t, x_{\max}) = 0 \quad \forall t \geq 0 \\ R(0, x) &= 1 \quad \text{and} \quad \lim_{t \rightarrow \infty} t R(t, x) = 0 \quad \forall x \in (0, x_{\max}] \end{aligned}$$

where $x_{\max} \leq \infty$ is an upper boundary to the population density.

The probability density $T(x)$ is given in (30). Integration of (33) over time t from 0 to ∞ therefore gives

$$(35) \quad -1 = \mu(x) \frac{d}{dx} \int_0^\infty R(t, x) dt + \frac{\varepsilon}{2} \sigma^2(x) \frac{d^2}{dx^2} \int_0^\infty R(t, x) dt$$

Moreover, by partial integration,

$$(36) \quad \int_0^\infty R(t, x) dt = \left[t R(t, x) \right]_{t=0}^\infty - \int_0^\infty t \partial_t R(t, x) dt = \mathbb{E}\{T(x)\}$$

because by (34) the first term is zero. Substitution of $\int R(t, x)dt$ in (35) by $\mathbb{E}\{T(x)\}$ from (36) gives

$$(37) \quad \frac{\varepsilon}{2}\sigma^2(x)\frac{d^2\mathbb{E}\{T(x)\}}{dx^2} + \mu(x)\frac{d\mathbb{E}\{T(x)\}}{dx} = -1$$

subject to the boundary conditions

$$(38) \quad \mathbb{E}\{T(0)\} = 0$$

$$\left.\frac{d}{dx}\mathbb{E}\{T(x)\}\right|_{x=x_{\max}} = 0$$

The ODE (37) with boundary condition (38) rarely can be solved explicitly, but it can be used to calculate the expected time till extinction numerically. The expected time till extinction, not for a fixed initial population density x , but for a population at the quasi-stationary distribution is obtained by taking the expectation of $T(x)$ over the distribution, i.e.,

$$(39) \quad \mathbb{E}\{T(x)\} = \int_0^{x_{\max}} T(x)p^c(x)dx$$

where $p^c(x)$ denotes the quasi-stationary distribution, which we may choose to approximate with the Gaussian distribution from equation (17).

9.8. Example continued.

The following figure gives the expected time till extinction for a population at the quasi-stationary distribution in the SIS model from section 9.6 above. Note that the approx-

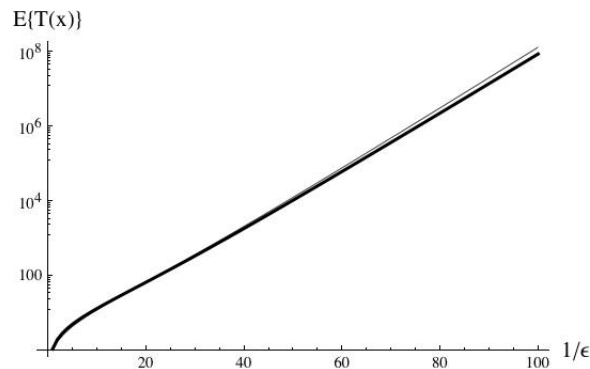


FIGURE 2. The expected time till extinction for a population at the exact quasi-stationary distribution (thin line) and the Gaussian approximation of the quasi-stationary distribution (thick line) for $\beta = 2$, $\delta = 1$ and $k = 1$.

imation gives an estimate that is systematically too low. It can be seen, however, that the expected time till extinction grows exponentially with system size. For a system size of 100, the expected time is about 10^8 times δ^{-1} , i.e., the expected duration of an infection from getting the disease to recovery. If the disease is the common cold with a recovery time of two weeks, then 10^8 amounts to 3.8 million years! The figures may

not be totally representative, but the message is that in semi-large systems, extinction becomes a negligible phenomenon on a population dynamical timescale. Still, the demographic stochasticity is significant as illustrated in figure 1 of section 9.6. This, in fact is the justification of the semi-large system approximation.