Diffusion Approximations for Ecological Models

P.K. Pollett

*Department of Mathematics, The University of Queensland, Queensland
4072 Australia (pkp@maths.uq.edu.au)*

Abstract: Diffusion models are widely used in ecology, and in more general population biology contexts, for predicting population-size distributions and extinction times. They are often used because they are particularly simple to analyse and give rise to explicit formulae for most of the quantities of interest. However, whilst diffusion models are ubiquitous in the literature on population models, their use is frequently inappropriate and often leads to inaccurate predictions of critical quantities such as persistence times. This paper examines diffusion models in the context in which they most naturally arise: as approximations to discrete-state Markovian models, which themselves are often more appropriate in describing the behaviour of the populations in question, yet are difficult to analyse from both an analytical and a computational point of view. We will identify a class of Markovian models (called *asymptotically density dependent* models) that permit a diffusion approximation through a simple limiting procedure. This procedure allows us to immediately identify the most appropriate approximating diffusion and to decide whether the diffusion approximation, and hence a diffusion model, is appropriate for describing the population in question. This will be made possible through the remarkable work of Tom Kurtz and Andrew Barbour, which is frequently cited in the applied probability literature, but is apparently not widely accessible to practitioners. Their results will be presented here in a form that most easily allows their direct application to population models. We will also present results that allow one to assess the accuracy of diffusion approximations by specifying for how long and over what ranges the underlying Markovian model is faithfully approximated. We will explain why diffusion models are not generally useful for estimating extinction times, a serious shortcoming that has been identified by other authors using empirical means.

Keywords: Diffusions approximations; Ecology; stochastic modelling

1. INTRODUCTION

One of the central themes of population biology is the estimation of persistence times. Diffusion models are frequently used because their parameters can be estimated simply from very little data and they offer explicit expressions for important quantities of interest, such as the expected time to extinction. Whilst it might appear trite to suggest that practitioners should use such expressions with care, because the assumptions of the model might not always be satisfied in any given situation, there does seem to be a prevailing view that they can be applied universally. In a recent paper, Wilcox and Possingham [2001] show, by empirical means, that the formula given in Foley [1994] for the mean time to extinction of a population may not always be accurate. However, there are often simple reasons why one would not expect formulae such as this to apply in all situations. Foley's formula is nothing more than an expression for the expected time it takes Brownian motion with drift to reach state 0 (representing extinction) from a given state n_0 (representing the initial population size). We should not possibly expect this to be accurate if, in the first place, the underlying discrete-time model is inappropriate in describing the population in question, or, if the limiting procedure that gives rise to the diffusion has no physical analogue. While these matters might be construed as "mathematical scruples", as Foley puts it (p. 136), they are in fact key scientific issues: the kind of issues that lie at the very core of whether mathematical models can be useful in describing the physical world. In Foley's case, the diffusion in question (Brownian motion with drift) is realized as an approximation of the simplest discrete-time random walk with state-independent jump probabilities, and this is achieved by correctly scaling time and space, and the parameters accordingly. The for-

mal approximation procedure mirrors the way in which the Central Limit Theorem works, with the position of the random walk at any given time instant being expressed as the sum of a large number of independent increments [for a simple exposition, see Section 4.8 of Ross, 1996]. So, whether Foley's formula could possibly be accurate depends firstly on whether the population in question could be modelled as a discrete-time process that jumps up or down by 1 at each time point (independently at each time point), with jump probabilities that do not depend on the population size, and, secondly, whether there have been sufficiently many transitions for the approximation procedure to work. To be emphatic, it would be naive to expect Foley's formula to be useful otherwise.

The above discussion also throws up the important question of whether a continuous-time model might be more appropriate in any given context. There seems to be a view that, because data are collected at discrete time points, say at the end of each breeding cycle, a discrete-time model is preferred. However, populations evolve continuously in time, and it should be the internal workings of the population that determine the nature of the model. After all, it is the population that is being modelled, rather than the estimation procedure. The fact that data might be collected only periodically, does not prevent one from estimating the parameters of a continuous-time model. This is done by way of certain embedded discrete-time processes, which are indeed determined by the sampling procedure.

It is also apparent that the term "diffusion approximation" is used in some of the ecology literature without specifically identifying any approximation procedure. This may help account for the sometimes-inappropriate use of formulae derived from diffusion models. However, in contrast to the above-mentioned space-time scaling procedures, there are many instances where the approximation procedure can be realized explicitly in terms of the parameters of the underlying model, and it is often then a simple task to identify conditions under which the approximation is likely to be accurate. For example, the diffusion approximation might become more accurate as the area of the habitat grows, or as the immigration rate becomes large, or, as the carrying capacity grows. In these situations it may also be possible to estimate the error in the approximation.

We shall identify a class of Markovian models that permit a diffusion approximation through an explicitly defined limiting procedure that identifies the most appropriate approximating diffusion. Explicit expressions are derived for various quantities of interest, and, in the stationary case, error bounds are given to assess the accuracy of the approximation. However, as we shall see, even when the diffusion process accurately models (say) the population size, it might be that the expected extinction times it predicts are different from those predicted by the underlying model. In these instances one is forced to deal directly with the underlying model. Since it is Markovian, one can always obtain the expected extinction times as the solution to a system of linear equations, and, if it is as simple as a birth-death process, then an explicit expression is available.

2. DENSITY DEPENDENCE

We shall restrict our attention to processes that are density dependent in a sense that might not accord with general usage of this term in the ecology literature. The basic idea is as follows. First, we identify a scale parameter N (the approximation will be achieved by letting N become large). Then, if $n(t)$ denotes the *number* in the population at time t, $n(\cdot)$ is density dependent if the rate of change of $n(t)$ depends on $n(t)$ only through the population *density* $n(t)/N$, whence the population density can usually be modelled accurately by a diffusion, one whose parameters can be written down explicitly in terms of the parameters of the underlying model. We shall make this precise shortly.

Suppose that the population in question can be modelled as a continuous-time Markov chain $(n(t), t \geq 0)$ taking values in a discrete set S with a (conservative) set of transition rates $Q = (q(m, n), m, n \in S)$, with $q(m, n)$ representing the rate of transition from state m to state n, for $m \neq n$, and $q(m) := -q(m, m) = \sum_{n \neq m} q(m, n)$ representing the total rate out of state m. For simplicity, we restrict our attention to *one-dimensional* Markov chains, and we take $n(t)$ to represent the number in the population at time t. We also restrict attention to the case when the state space S is *finite*, though all of the results presented here hold more generally. Of course the "state" of the population will usually be more complicated. Fortunately, the general procedure we shall describe carries over almost immediately to multi-dimensional processes, but further technical conditions are required to handle the infinite-state case; see Pollett [1990].

Now let $\{n_N(\cdot)\}$ be a *family* of such processes indexed by $N > 0$, and suppose that $n_N(\cdot)$ takes values in S_N, a finite subset of the integers Z, and has transition rates $Q_N = (q_N(m, n), m, n \in S_N)$. In practice, one has great freedom in identifying an index parameter. For definiteness, let us imagine N is the carrying capacity.

Definition *Suppose that there exists an interval $E \subseteq \mathbb{R}$ and a family $\{f_N, N > 0\}$ of continuous*

functions, with $f_N : E \times \mathbb{Z} \to \mathbb{R}$, such that

$$q_N(n, n+l) = N f_N \left(\frac{n}{N}, l \right), \quad l \neq 0.$$

Then, the family of Markov chains is (asymptotically) density dependent if, additionally, there exists a function $F : E \to \mathbb{R}$ such that $\{F_N\}$, given by $F_N(x) = \sum_l l f_N(x, l)$, $x \in E$, converges (pointwise) to F on E.

This definition of density dependence is more general than that introduced by Kurtz [1970], which has f_N (and hence F_N) being the same for all N. Roughly speaking, the family is density dependent if the transition rates of the corresponding "density process" $X_N(\cdot)$, defined by $X_N(t) = n_N(t)/N$, $t \geq 0$, depend on the present state n only through the density n/N, or, failing this, if this property is exhibited *asymptotically* for large N. Thus, there is a natural way to associate with this process, a
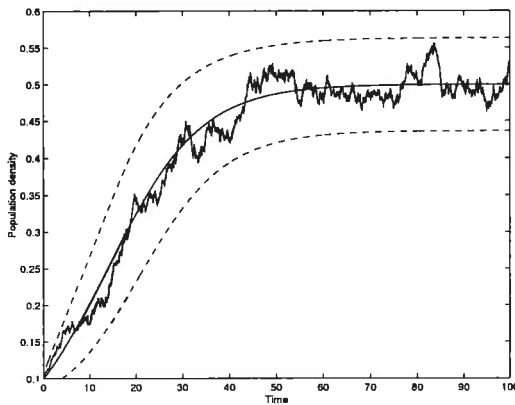

Figure 1. Simulation of the stochastic logistic model, together with its diffusion approximation ($N = 500$, $\rho = 0.5$, $\mu = 0.1$ and $X_N(0) = 0.1$).

density dependent *deterministic* process which, for large N, is "tracked" by the process. (Figure 1 illustrates this with reference to the stochastic logistic model to be described in Section 3.) The following (functional) law of large numbers establishes a deterministic approximation under appropriate conditions. It can be deduced immediately from Theorem 3.1 of Kurtz [1970].

Theorem 2.1 *Suppose that $f_N(\cdot, l)$ is bounded, for each l and N, that F is Lipschitz continuous on E and that $\{F_N\}$ converges uniformly to F on E. Then, if $\lim_{N \to \infty} X_N(0) = x_0$, the density process $X_N(\cdot)$ converges uniformly in probability on $[0, t]$ to $X(\cdot, x)$, the unique (deterministic) trajectory satisfying $X(0, x) = x$, $X(s, x) \in E$, $s \in [0, t]$, and*

$$\frac{\partial}{\partial s} X(s, x) = F(X(s, x)). \tag{1}$$

The following (functional) central limit law establishes that, for large N, the fluctuations about the deterministic path follow a Gaussian diffusion, provided that certain "second-order" conditions are satisfied. It can be deduced from Theorems 3.1 and 3.5 of Kurtz [1971]. In particular, (2) strengthens the condition that $\{F_N\}$ converges uniformly to F, to ensure that this convergence happens at the correct rate, while (3) provides the initial value of the approximating diffusion.

Theorem 2.2 *Suppose $f_N(\cdot, l)$ is bounded, that F is Lipschitz continuous and has uniformly continuous first derivative on E, and that*

$$\lim_{N \to \infty} \sup_{x \in E} \sqrt{N} |F_N(x) - F(x)| = 0. \tag{2}$$

Suppose also that the sequence $\{G_N\}$, where

$$G_N(x) = \sum_l l^2 f_N(x, l), \quad x \in E,$$

converges uniformly to G, where G is uniformly continuous on E. Let $x_0 \in E$. Then, if

$$\lim_{N \to \infty} \sqrt{N} (X_N(0) - x_0) = z, \tag{3}$$

the family of processes $\{Z_N(\cdot)\}$, defined by $Z_N(s) = \sqrt{N} (X_N(s) - X(s, x_0))$, $0 \leq s \leq t$, converges weakly in $D[0, t]$ (the space of right-continuous, left-hand limits functions on $[0, t]$) to a Gaussian diffusion $Z(\cdot)$ with initial value $Z(0) = z$ and with mean and variance given by $\mu_s := \mathrm{E}(Z(s)) = M_s z$, where $M_s = \exp(\int_0^s B_u du)$ and $B_s = F'(X(s, x_0))$, and $\mathrm{Var}(Z(s)) = \sigma_s^2$, where $\sigma_s^2 = M_s^2 \int_0^s M_u^{-2} G(X(u, x_0)) du.$

It follows that $X_N(s)$ has an approximate normal distribution with $\mathrm{Var}(X_N(s)) \simeq \sigma_s^2/N$. We would usually take $x_0 = X_N(0)$, thus giving $\mathrm{E}(X_N(s)) \simeq X(s, x_0)$. Figure 1 illustrates this with reference to the stochastic logistic model to be described in Section 3: the deterministic mean path is shown, together with plus and minus two standard deviations: $X(s, x_0) \pm 2\sigma_s/\sqrt{N}$.

In the important special case where x_0 is chosen as an equilibrium point of (1), we can be far more precise about the approximating diffusion:

Corollary 2.1 *If x_0 satisfies $F(x_0) = 0$ then, under the conditions of Theorem 2.2, the family $\{Z_N(\cdot)\}$, defined by $Z_N(s) = \sqrt{N}(X_N(s) - x_0)$, $0 \leq s \leq t$, converges weakly in $D[0, t]$ to an Ornstein-Uhlenbeck (OU) process $Z(\cdot)$ with initial value $Z(0) = z$, and with local drift $B = F'(x_0)$ and local variance $V = G(x_0)$. In particular, $Z(s)$ is normally distributed with mean $\mu_s = e^{Bs} z$ and variance $\sigma_s^2 = \frac{V}{2B}(e^{2Bs} - 1)$.*

We conclude that, for N large, $X_N(s)$ has an approximate normal distribution with $\mathrm{Var}(X_N(s)) \simeq \sigma_s^2/N$. A "working approximation" for the mean (that is, for a fixed value of N) can be obtained by setting z equal to $\sqrt{N}\,(X_N(0) - x_0)$ (c.f. (3)):

$$\mathrm{E}(X_N(s)) \simeq x_0 + e^{Bs}(X_N(0) - x_0).$$

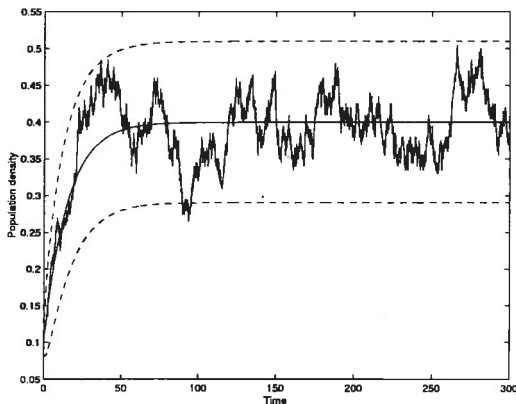The OU approximation is illustrated in Figure 2, with two standard deviations as before.

Figure 2. Simulation of the stochastic logistic model, together with its OU approximation ($N = 200$, $\rho = 0.6$, $\mu = 0.1$ and $X_N(0) = 0.1$).

In the context of population models x_0 will usually be asymptotically stable, that is $B < 0$. However, it should be emphasised that it need not be for each of the above conclusions to hold. Indeed, the OU approximation is often very accurate in describing the fluctuations about centres and unstable equilibria [see Barbour, 1976].

Although one expects the diffusion approximation detailed in Theorem 2.2 to provide a more accurate estimate of the distribution of $X_N(s)$ for s small (since it concerns random fluctuations about the deterministic trajectory, rather than a given equilibrium point), the OU approximation has the advantage that the approximate formulae for the mean and variance are explicit. Furthermore, we have at our disposal results that show how closely, for how long, and over what ranges, the density process is faithfully approximated. For example, we can obtain explicit information concerning the order of the error arising from approximating $Z_N(s)$ by a normal random variable [see Barbour, 1976]:

$$\mathrm{Pr}(Z_N(s) \ge x) = A_s(x)\left(1 + O\left(J_N^4 \frac{\log N}{\sqrt{N}}\right)\right),$$

where $A_s(x) = 1 - \Phi((x - \mu_s)/\sigma_s)$, provided J_N is of order between $(\log N)^{1/2}$ and $N^{1/8}$. This result holds uniformly over x in $[0, J_N]$ and over time intervals of the form $[s_0, \exp(\delta J_N^2)]$, for any positive s_0 and δ. Thus, the questions of "how

closely" and "over what ranges" are both answered here. Barbour also obtains results which show that, when x_0 is asymptotically stable, the time until first exit of $X_N(\cdot)$ from an interval of the form $I_N(x_0) = \{x : |x - x_0| < N^{-1/2}e_N\}$, where $e_N \to \infty$, is approximately exponentially distributed, conditional on the process leaving at, say, the left endpoint (asymptotically, left or right are equally likely). This result answers the question of "how long", because it establishes that, provided $e_N = o(N^{1/8})$, the mean time until exit is asymptotically $\sqrt{\pi V/\alpha}\,e^{\alpha e_N^2/V}/(2\alpha e_N)$, where $\alpha = -B$ (> 0). Thus, in particular, the time taken for $X_N(\cdot)$ to first leave $I_N(x_0)$ is of order $e_N^{-1}\exp(\frac{1}{2}e_N^2/\sigma_\infty^2)$, whenever $e_N = o(N^{1/8})$. Hence, it is asymptotically larger than any power of N if, for example, $e_N = O(N^{\frac{1}{8}}/\log N)$. As Barbour points out, although this may well be enough to justify treating the normal approximation around x_0 as the true equilibrium distribution, it does not yield the distribution of the time until extinction, which would require $e_N = o(N^{1/2})$.

3. BIRTH-DEATH PROCESSES

As these arise frequently in population modelling, we shall examine them briefly. We shall consider only the finite-state case with $S_N = \{0, 1, \ldots, N\}$, and birth and death rates of the form

$$a_n := q_N(n, n+1) = N\lambda(n/N) + O(1),$$
$$b_n := q_N(n, n-1) = N\mu(n/N) + O(1),$$

where $\lambda(\cdot)$ and $\mu(\cdot)$ are positive continuous functions defined on $E = [0, 1]$ with continuous first derivative. This form of rates ensures that the process is asymptotically density dependent with $F(x) = \lambda(x) - \mu(x)$ and $F_N(x) = F(x) + O(1/N)$. We also have $G(x) = \lambda(x) + \mu(x)$ and $G_N(x) = G(x) + O(1/N)$. Whilst the form of rates does not include *all* birth-death processes on S_N, it does include ones commonly encountered in population modelling, for example, those for which the transition rates are polynomials in n. Indeed, if we specialize further to this latter case, then all of the technical conditions of Theorems 2.1 and 2.2 concerning $\{f_N\}$, $\{F_N\}$, F, $\{G_N\}$ and G will be satisfied, and, provided (3) holds, there will be a valid diffusion approximation.

To illustrate this further, let us consider the *stochastic logistic model*, ubiquitous in the literature on population modelling [see for example Bartlett, 1960], but also appearing in a variety of different guises: for example, chemical kinetics [Oppenheim et al., 1977] and epidemics [Nåsell, 1999]. It is a birth-death process $n_N(\cdot)$ on $S_N = \{0, 1, \ldots, N\}$

with transition rates that are polynomials in n:

$$q_N(n, n+1) = \frac{\lambda}{N} n(N-n) = N\lambda \frac{n}{N}\left(1 - \frac{n}{N}\right),$$

$$q_N(n, n-1) = \mu n = N\mu \frac{n}{N},$$

where $\lambda, \mu > 0$. Clearly $C_N = \{1, \ldots, N\}$ is irreducible (every state in C_N can be reached by every other state in C_N), and absorption at 0 (representing the event of extinction) occurs with probability 1. We have $F(x) = \lambda x(1-x) - \mu x = \lambda x(x_0 - x)$ and $G(x) = F(x) + 2\mu x$, where $x_0 = 1 - \rho$ and $\rho = \mu/\lambda$. We will consider the most interesting case $\rho < 1$, when there is drift away from the absorbing state. Under this assumption, the system (1) has two equilibria in $E = [0, 1]$: 0 (unstable) and x_0 (asymptotically stable). Indeed, (1) has the solution

$$X(t, x) = \frac{x_0 x}{x + (x_0 - x)e^{-\lambda x_0 t}}, \quad t \geq 0, \quad (4)$$

remembering that $X(0, x) = x$. It is possible to write down the full diffusion approximation for $X_N(\cdot) = n_N(\cdot)/N$ about the deterministic path (4), but we will not pursue this further here. Suffice it to say that the standard deviation plotted in Figure 1 was evaluated by numerically integrating $\int_0^s M_u^{-2} G(X(u, x)) du$. Instead, we will consider the OU approximation about the stable equilibrium x_0. Since $F'(x) = \lambda(x_0 - 2x)$, we have local drift $B = F'(x_0) = -\lambda x_0$ and local variance $V = G(x_0) = 2\mu x_0 = 2\lambda x_0 \rho$. Thus, provided we arrange for (3) to hold, there will be a valid OU approximation. We conclude that, for N large, $X_N(t)$ has an approximate normal distribution with

$$\mathrm{E}(X_N(t)) \simeq x_0 + e^{-\alpha t}(X_N(0) - x_0),$$
$$\mathrm{Var}(X_N(t)) \simeq \rho(1 - e^{-2\alpha t})/N,$$

where $\alpha = \lambda x_0 = \lambda - \mu \; (> 0)$. This is depicted in Figure 2.

Also, since x_0 is asymptotically stable, we may deduce if $e_N = o(N^{1/8})$, the time until first exit of $X_N(\cdot)$ from an interval of the form $I_N(x_0)$ has an approximate exponential distribution, conditional on the process leaving at, say, the left endpoint, and the mean time until exit is asymptotically $\sqrt{\pi \rho/2}\, e_N^{-1} \exp\left(\frac{1}{2}e_N^2/\rho\right)/\alpha$. As already indicated, we cannot use this result to estimate the time to extinction, for this would require $e_N = o(N^{1/2})$. Indeed, it is strongly suggestive that the OU approximation will not accurately predict the expected time to extinction. Since we have at our disposal an explicit expression for this quantity, we can test the hypothesis.

For a general Markov chain with transition rates $Q = (q(m, n), m, n \in S)$, whose state space S (possibly infinite) includes a subset A which is reached with probability 1, the time τ_i it takes to reach A starting in state i is the minimal (and then unique) non-negative solution to $\sum_{i \in S} q(i, j)\tau_j + 1 = 0$, $i \notin A$, with $\tau_i = 0$ for $i \in A$. This result can be found in any text on Markov chains, no matter how ancient [for a recent exposition see Norris, 1997], but it is apparently not widely used by biologists; in their paper "Four facts every conservation biologist should know about persistence", Mangel and Tier [1994] implore their readers to use it: Fact 2 "There is a simple and direct method for the computation of persistence times that virtually all biologists can use". It is often then a matter of simple arithmetic to evaluate the expected hitting times in any given situation. Failing that, a host of numerical methods exist (we are merely solving a system of linear equations).

In the present context of birth-death processes on S_N with birth rates (a_n) and death rates (b_n), the time $\tau_i(N)$ it takes to reach 0 starting in state i is given by $\tau_i(N) = \sum_{j=1}^{i} (b_j \pi_j)^{-1} \sum_{k=j}^{N} \pi_k$, with $\tau_0(N) = 0$, where the potential coefficients (π_j) are given by $\pi_1 = 1$ and $\pi_j = \prod_{k=2}^{j}(a_{k-1}/b_k)$ for $j \geq 2$. (This formula is valid in the infinite state case, replacing N by ∞.) For the logistic model,

$$\tau_i(N) = \frac{1}{\mu} \sum_{j=1}^{i} \sum_{k=0}^{N-j} \frac{1}{j+k} \prod_{l=0}^{k-1}\left(\frac{N-j-l}{N\rho}\right).$$

Whilst this admits further simplification, the form given here reflects the algorithm used to evaluate $\tau_i(N)$, the product being evaluated recursively, and the sums evaluated in such a way as to minimize round-off error.

Estimating expected extinction times by way of the OU approximation is also a simple task. For an OU process $(Z(t), t \geq 0)$ with local drift B and local variance V, let $T_a = \inf\{t : Z(t) = a\}$ and $v(z) = \mathrm{E}(T_a | Z(0) = z)$. Then, v is the unique solution to the linear ODE

$$\frac{1}{2}V\frac{d^2 v}{dz^2} + Bz\frac{dv}{dz} + 1 = 0,$$

with $v(a) = 0$ (an absorbing barrier at a) and $v'(b) = 0$ (a reflecting barrier at b) [see Section 15.3 of Karlin and Taylor, 1981]. For the logistic model, we have $V = 2\alpha\rho$ and $B = -\alpha$. The DE was solved by first converting it to two first-order ODEs in the usual way, and then using the Matlab (Version 6) command bvp4c, which implements a three-stage Lobatto IIIa formula to solve two-point boundary value problems. Note that the time it takes the density process $X_N(\cdot)$ to hit 0 is approximated by the time it takes the diffusion to reach $a = -x_0\sqrt{N}$ starting at the point $z = \sqrt{N}(X_N(0) - x_0)$. We took $b = z + 4(z - a)$.

Figure 3 compares the OU approximation for the expected extinction time with the exact values, for all values of the starting point $X_N(0) = x$. Notice
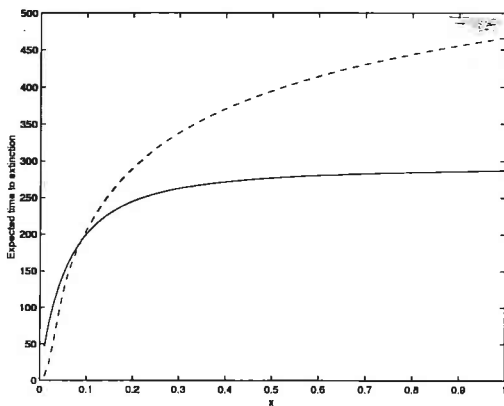

Figure 3. OU approximation (dashed) for the expected extinction time, and exact values (solid) ($N = 100$, $\rho = 0.9$, $\mu = 0.1$ and $X_N(0) = x$).

that the approximation is generally poor, but is best when x is near the equilibrium point ($x_0 = 0.1$).

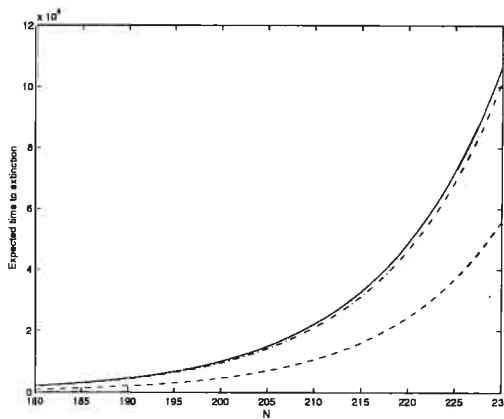Whilst the formula for $\tau_i(N)$ does not pose any significant numerical problems, several authors

Figure 4. OU approximation (dashed) for the expected extinction time, exact values (solid) and the asymptotic expansion (dash-dot) ($\rho = 0.65$, $\mu = 0.1$ and $X_N(0) = 0.05$).

have derived asymptotic expansions [for example Kryscio and Lefèvre, 1989]. Our formula is at variance with their formula (2.4), but shows a slight improvement. By evaluating the factorials as gamma integrals, and using Cauchy's method to estimate these integrals, we obtain

$$\tau_i(N) \sim \frac{\rho(1 - \rho^i)}{\mu(1 - \rho)^2} \left(\frac{e^{-(1-\rho)}}{\rho} \right)^N \sqrt{\frac{2\pi}{N}} .$$

Figure 4 compares the OU approximation with the exact values and the asymptotic expansion for various values of N.

For an alternative approach to estimating extinction times, see Nåsell [1999], who approximates the time to extinction from the "quasi-stationary state".

4. ACKNOWLEDGEMENTS

5. REFERENCES

Barbour, A., Quasi-stationary distributions in Markov population processes, *Advances in Applied Probability*, 8, 296–314, 1976.

Bartlett, M., *Stochastic Population Models in Ecology and Epidemiology*, Methuen, London, 1960.

Foley, P., Predicting extinction times from environmental stochasticity and carrying capacity, *Conservation Biology*, 8, 124–137, 1994.

Karlin, S., and H. Taylor, *A Second Course in Stochastic Processes*, Academic Press, London, 1981.

Kryscio, R., and C. Lefèvre, On the extinction of the S-I-S stochastic logistic epidemic, *Journal of Applied Probability*, 27, 685–694, 1989.

Kurtz, T., Solutions of ordinary differential equations as limits of pure jump Markov processes, *Journal of Applied Probability*, 7, 49–58, 1970.

Kurtz, T., Limit theorems for sequences of jump Markov processes approximating ordinary differential processes, *Journal of Applied Probability*, 8, 344–356, 1971.

Mangel, M., and C. Tier, Four facts every conservation biologist should know about persistence, *Ecology*, 75, 607–614, 1994.

Nåsell, I., On the quasi-stationary distribution of the stochastic logistic epidemic, *Mathematical Biosciences*, 156, 21–40, 1999.

Norris, J., *Markov Chains*, Cambridge University Press, Cambridge, 1997.

Oppenheim, I., K. Schuler, and G. Weiss, Stochastic theory of nonlinear rate processes with multiple stationary states, *Physica*, 88A, 191–214, 1977.

Pollett, P., On a model for interference between searching insect parasites, *Journal of the Australian Mathematical Society, Series B*, 31, 133–150, 1990.

Ross, S., *Stochastic Processes*, 2nd edn, Wiley, New York, 1996.

Wilcox, C., and H. Possingham, Do life history traits affect the accuracy of diffusion approximations for mean time to extinction? Submitted for publication, 2001.