On a Functional Central Limit Theorem for Markov Population Processes

Andrew D. Barbour


Stable URL:
http://links.jstor.org/sici?sici=0001-8678%28197403%296%3A1%3C21%3AOAFC%3E2.0.CO%3B2-%23

Advances in Applied Probability is currently published by Applied Probability Trust.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/apt.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.
ON A FUNCTIONAL CENTRAL LIMIT THEOREM
FOR MARKOV POPULATION PROCESSES

ANDREW D. BARBOUR, University of Cambridge

Abstract

Let \( \{X_n(t)\} \) be a sequence of continuous time Markov population processes on an \( n \)-dimensional integer lattice, such that \( X_n \) has initial state \( N\mathbf{x}(0) \) and has a finite number of possible transitions \( J \) from any state \( X \); let the transition \( X \to X + J \) have rate \( N g^J(N^{-1}X) \), and let \( g^J(x) \) and \( x(0) \) be fixed as \( N \) varies. The rate of convergence of \( N^{-1}X_n(t) \) to a Gaussian diffusion is investigated, where \( \xi(t) \) is the deterministic approximation to \( N^{-1}X_n(t) \), and a method of deriving higher order asymptotic expansions for its distribution is justified. The methods are applied to two birth and death processes, and to the closed stochastic epidemic.

MARKOV POPULATION PROCESS; FUNCTIONAL CENTRAL LIMIT THEOREM; EPIDEMIC;
COMPETITION PROCESS; BIRTH AND DEATH PROCESS

1. Introduction

A common method of describing biological and ecological systems is to represent them as continuous parameter Markov processes, in one or more dimensions, with each coordinate typically representing a count of the present population of a particular species. The transition rates in these models are usually simple functions of the coordinates, and sometimes of the time parameter as well. We shall consider a sequence \( \{X_n(\cdot)\} \) of such processes, in which the indexing parameter \( N \) determines the magnitude of the initial state of the process, in that \( X_n(0) = N\mathbf{x}(0) \) for a fixed vector \( \mathbf{x}(0) \). In Barbour (1972) \([B]\) a heuristic argument is put forward, which leads to a method of successive approximation, in terms of large \( N \), to the process \( X_n(t) = N^{-1}X_N(T_n(t)) \), where \( T_n(t) \) is a suitably chosen function of \( t \); see also McNeil and Schach (1973). The approximation, which to the second order reduces to a small Gaussian diffusion about the deterministic path, was found effective in the example chosen even for relatively small values of \( N \). The limiting process has much to recommend it as a natural choice for approximation, and even the apparently restrictive initial condition requirements proved amenable in practice.

In this paper, we provide a rigorous justification for these ideas. The basis for the argument is to be found in Kurtz (1970) and (1971), where the first and second

Received in revised form 26 June 1973.
order approximations are justified under very weak conditions. In order to take full advantage of the method of [B], we will consider only a more restricted situation, which nonetheless includes all of the most common biological models, although much of the material can be proved in a more general setting.

Section 2 consists of a statement of Kurtz's functional central limit theorem, in a form compatible with the aims of this paper, and in Section 3 an associated rate of convergence result is proved. Section 4 contains a detailed study of the properties of integrals of the deterministic equations. The application of the method is illustrated in Section 5 with some examples.

2. Basic elements

For some fixed $T > 0$, let $x_N(\cdot), N = 1, 2, \ldots$, be a sequence of continuous parameter Markov processes on $[0, T]$, taking values in an $n$-dimensional lattice $L_N^n = N^{-1}Z^n$; let their stable, conservative $Q$-matrices be specified by the transition intensities

$$ x \to x + N^{-1}J \text{ at rate } Ng_J^x(x) $$

for $J$ in some fixed finite set $\mathcal{J} \subset Z^n \setminus \{0\}$, and let their common initial value be $x(0)$, independent of $N$: $Z$ here represents the integers. Assume that the deterministic solution to the process exists in $[0, T]$, that is:

**Assumption A.** There exists a unique function $\xi(t), 0 \leq t \leq T$, satisfying

$$ \dot{\xi}(t) = \sum_J J g^J(\xi(t)), \quad 0 \leq t \leq T; \quad \xi(0) = x(0). $$

Let $S \subset \mathbb{R}^n$ be defined by

$$ S = \{ x : x = \xi(t) \text{ for some } 0 \leq t \leq T \}, $$

and, for some fixed $\varepsilon > 0$, define

$$ S^\varepsilon = \{ x : |x - S| \leq \varepsilon \}. $$

**Assumption B.** The functions $g^J(x), J \in \mathcal{J}$, are multinomial in the coordinate variables $x_1, x_2, \ldots, x_n$ within $S^\varepsilon$.

Note that, when verifying Assumption B in applications, $\varepsilon$ may be taken as small as need be.

Let constants $L$ and $M_r, r = 2, 3, \ldots$, be defined by

$$ M_r = \sup_{x \in S^\varepsilon} \sum_J g^J(x) |J|^r, \quad r = 2, 3, \ldots, $$

$$ L = \sup_{x, y \in S^\varepsilon} \left[ \sum_J |J(g^J(x) - g^J(y))| / |x - y| \right]. $$

Let $y_N(t) = N^t \{ x_N(t) - \xi(t) \}$, and let $g^J_J(x)$ denote $\partial g^J_J(x) / \partial x_J$. 


23

**Theorem K.** Under Assumptions A and B:

(i) for any $0 < \delta \leq \varepsilon$, $0 \leq t \leq T$,

$$P\left[ \sup_{0 \leq s \leq t} \left| x_N(s) - \xi(s) \right| > \delta \right] \leq N^{-1}\delta^{-2}tM_2 \exp(2Lt);$$

(ii) given any integer $r \geq 3$, there exists a constant $C_r$ depending on $\varepsilon$, $T$, $L$ and $M_k$, $k = 2, 3, \ldots, r$, such that

$$P\left[ \sup_{0 \leq s \leq T} \left| x_N(s) - \xi(s) \right| > \varepsilon \right] \leq C_r N^{1-r};$$

(iii) $y_N \Rightarrow y$ in $D[0, T]$, where $y(t)$ is the diffusion with $y(0) = 0$, whose characteristic function $\Phi(\theta, t)$ satisfies

$$\frac{\partial \Phi}{\partial t} - \sum_j (\theta J) \sum_{j=1}^n g_j'(\xi(t)) \frac{\partial \Phi}{\partial \theta_j} + \frac{1}{2} \sum_j (\theta J)^2 g_j''(\xi(t)) \Phi = 0.$$  

**Remark.** This is a special case of Theorem (3.5) of Kurtz (1971) [K]. The developments that follow can be modified without difficulty, but at the expense of greater detail, to more general situations. For example, similar results hold if $g^J(x)$ is replaced by $g_N^J(x, t)$, where $g_N^J(x, t)$ is continuous for each $J$ as a function of $t$, uniformly in $x \in S^e$ and in $N$, and where $g_N^J(x, t)$ converges uniformly in $S^e \times [0, T]$ to $g^J(x, t)$ as $N \to \infty$. Assumption B is only made to guarantee once and for all the existence of as many derivatives of the functions $g^J$ as are necessary, and it can be weakened in the obvious way in Theorem 1. It is also possible to improve upon (ii) with an estimate of the form

$$P\left[ \sup_{0 \leq s \leq T} \left| x_N(s) - \xi(s) \right| > \varepsilon \right] \leq C \exp(-cN^4).$$

3. **Rate of convergence**

Let $h(x, t) : S^e \times [0, T] \to R$ be any solution of the equation

$$\frac{\partial h}{\partial t} + \sum_j g^J(x) \sum_k J_k \frac{\partial h}{\partial x_k} = 0$$

which has all its second partial derivatives with respect to the coordinates $x_j$, and also the mixed derivatives $\partial^2 h/\partial x_j \partial t$. Any such $h$ is an integral of the deterministic equations, and is constant along deterministic trajectories. Let $Vh = Vh(\xi(t), t)$ denote the vector of derivatives of $h(x, t)$ with respect to $x$, evaluated at $x = \xi(t)$.

**Lemma 1.** The random function $z : z(t) = y(t)$, $\forall h(\xi(t), t)$, $0 \leq t \leq T$, where $y$ is the diffusion limit specified by (2.7), is distributed in $D[0, T]$ as a Wiener process deformed by a non-random change of time scale.

**Proof.** From Equations (2.2) and (3.1),
\[
\frac{d}{dt} \left( \frac{\partial h}{\partial \xi_j} \right) = \sum_k \frac{\partial^2 h}{\partial \xi_j \partial \xi_k} \frac{d \xi_k}{dt} + \frac{\partial^2 h}{\partial \xi_j \partial t} \\
= \left[ \frac{\partial}{\partial x_j} \left( \sum_{J} g^J(x) \sum_k J_k \frac{\partial h}{\partial x_k} + \frac{\partial h}{\partial t} \right) \right]_{x = \xi(t)} - \sum_{J} \frac{\partial g^J}{\partial \xi_j} \sum_k J_k \frac{\partial h}{\partial \xi_k} \\
= -\sum_{J} \frac{\partial g^J}{\partial \xi_j} \sum_k J_k \frac{\partial h}{\partial \xi_k}.
\]

Let \( \Psi_t(\psi) \), \( 0 \leq t_0 \leq t \leq T \), denote \( E\{\exp(i \psi z(t))\} \), where \( E \) is the expectation conditional on an arbitrary given distribution for \( y(t_0) \), whose characteristic function is denoted by \( \Psi_{t_0}(\psi) \). Then

\[
\frac{d\Psi_t}{dt} = i \psi E\left\{ \exp(i \psi z(t)) y(t) \cdot \frac{d}{dt} \nabla h \right\} + \left[ \frac{d}{dt} E\{\exp(i \theta \cdot y(t))\} \right]_{\theta = \psi \nabla h},
\]

and substituting in (3.3) from (3.2) for \( d\nabla h / dt \) and from (2.7) for the second term, we obtain, after integration,

\[
\Psi_t = \Psi_{t_0} \exp \left\{ -\frac{i}{2} \psi^2 \int_{t_0}^{t} \sigma^2(u) \, du \right\},
\]

where

\[
\sigma^2(t) = \sum_{J} g^J(\xi(t)) (J \cdot \nabla h)^2.
\]

This identifies \( z \) through its finite-dimensional distributions.

Now define \( z_N(t) = y_N(t) \cdot \nabla h(\xi(t), t) \), and let \( \Psi_N(\psi, t) \) and \( \Psi(\psi, t) \) be the characteristic functions of \( z_N(t) \) and \( z(t) \) respectively, conditional on \( y_N(0) = y(0) = 0 \). Let \( \Sigma^2(t) = \int_0^t \sigma^2(u) \, du \), where \( \sigma^2(t) \) is defined by (3.5), and set

\[
H = \sup_{0 \leq t \leq T} |\nabla h(\xi(t), t)|;
\]

note that, from (3.4), \( z(t) / \Sigma(t) \) is distributed as a standard Normal random variable.

**Theorem 1.** In addition to Assumptions A and B, and assumptions about the smoothness of \( h \), suppose that

\[
\inf_{0 \leq t \leq T} \sigma^2(t) = \tau^2 > 0.
\]

Then, for any fixed \( t \), \( 0 < t \leq T \),

\[
\sup_{v} \left| P[Z_N(t) \leq v] - \Phi(v / \Sigma(t)) \right| = O(N^{-\frac{1}{2}} \log N),
\]

where the order term depends on \( h \), \( t \) and the parameters of the process, and where \( \Phi(\cdot) \) is the distribution function of a standard Normal random variable.
Proof. The error estimate is based on an estimate of the discrepancy between the characteristic functions \( \Psi_N \) and \( \Psi \). For convenience, we work with the process \( \bar{x}_N \), which is obtained from the original process \( x_N \) by stopping it when it first leaves \( S^{\pm} \); then \( \bar{x}_N(t) \in S^\pm \) almost surely for all \( N > 2 \max |J|/\varepsilon \), and yet, because of (2.6), the estimate (3.8) is unaltered by this modification of the underlying process outside \( S^{\pm} \).

Let \( I(x) \) be the function which takes the value 1 on \( S^{\pm} \) and zero otherwise, and take \( N > 2 \max |J|/\varepsilon \). Let \( k_j, j = 1, 2, \ldots \), denote constants independent of \( N \) and \( t \), but depending on \( \varepsilon, T, H \) and the parameters of the process. Let \( E \) denote expectation conditional on \( y_N(0) = 0 \), set \( Q(u) = e^u - 1 - u - \frac{1}{2}u^2 \), and let

\[
\begin{align*}
g'_k & \text{ denote } \frac{\partial g'(x)}{\partial x_k} \text{ evaluated at } x = \xi(t); \\
x & \text{ denote } \xi(t) + N^{-\frac{1}{2}}\bar{y}_N(t), \text{ or equivalently } \bar{x}_N(t); \\
\xi & \text{ denote } \xi(t);
\end{align*}
\]

and replace \( x_N \) by \( \bar{x}_N \) in earlier definitions.

Define \( p_N(x, t) = P[\bar{x}_N(t) = x | \bar{x}_N(0) = x(0)] \); then the Kolmogorov forward differential equation for the process \( \bar{x}_N \) takes the form

\[
\begin{align*}
\frac{\partial p_N(x, t)}{\partial t} & = N \sum_j \{ p_N(x - JN^{-1}, t)g'_j(x - JN^{-1})I(x - JN^{-1}) \\
& \quad - Np_N(x, t) \sum_j g'_j(x)I(x) \},
\end{align*}
\]

for any state \( x \) accessible to \( \bar{x}_N \). There are only finitely many such states, since the process is confined to \( S^{\pm} \cap L_N^0 \), and so the equations have a unique solution; furthermore, multiplying them by \( \exp(\iota\theta \cdot x) \) on both sides, for any fixed \( \theta \in \mathbb{R}^n \), and summing over all accessible states \( x \),

\[
\frac{d}{dt} E \{ \exp(\iota\theta \cdot \bar{x}_N(t)) \}
\]

\[
= N \sum_j (\exp(\iota\theta \cdot JN^{-1}) - 1)E\{g'_j(\bar{x}_N(t))I(\bar{x}_N(t))\exp(\iota\theta \cdot \bar{x}_N(t))\}
\]

since all the expectations exist.

For each real \( \psi \),

\[
\frac{d\Phi_N}{dt} = i\psi E \left\{ \exp(\iota\psi z_N(t)) \cdot \frac{d}{dt} \bar{y}_N(t) \right\} + \left[ \frac{d}{dt} E\{\exp(\iota\theta \cdot \bar{y}_N(t))\} \right]_{\theta = \psi h}.
\]

Substitute from Equation (3.2) for \( d\bar{y}_N/dt \), and expand the second term using Equation (3.9), expressed in terms of \( \bar{y}_N(t) = N^{\frac{1}{2}}(\bar{x}_N(t) - \xi(t)) \). This leads, after some manipulation, to
\[
\frac{d\Psi_N}{dt} + \frac{1}{2} \psi^2(t) \psi^2 \Psi_N \\
= N \sum_j \left( \exp(i\psi J \cdot \nabla h N^{-\frac{1}{4}}) - 1 \right) E\{\exp(i\psi z_N(t))g^J(\vec{x})[I(\vec{x}) - 1]\} \\
- \frac{1}{2} \psi^2 \sum_j \left( \nabla h \cdot J \right)^2 E\{\exp(i\psi z_N(t)) \left[ g^J(\vec{x}) - g^J(\xi) \right] \} \\
+ i\psi \sum_j J \cdot \nabla h N^{\frac{1}{4}} E\{\exp(i\psi z_N(t)) \left[ g^J(\vec{x}) - g^J(\xi) - \sum_k (\vec{x}_k - \xi_k)g^J_k \right] \} \\
+ N \sum_j Q(i\psi J \cdot \nabla h N^{-\frac{1}{4}}) E\{\exp(i\psi z_N(t))g^J(\vec{x})\}.
\]

Let
\[(3.11) \quad r(\psi, t, N) = N \sum_j Q(i\psi J \cdot \nabla h N^{-\frac{1}{4}})g^J(\vec{x}_j),\]

and let
\[(3.12) \quad R(\psi, t, N) = \int_0^t r(\psi, u, N) \, du;\]

note that
\[(3.13) \quad |r(\psi, t, N)| \leq \frac{1}{2} N \sum_j |\psi J \cdot \nabla h N^{-\frac{1}{4}}| \sup_{x' \in S^*} g^J(x') \leq k_1 |\psi|^3 N^{-\frac{1}{4}},\]

and hence that
\[(3.14) \quad |R(\psi, t, N)| \leq k_1 |\psi|^3 N^{-\frac{1}{4}} t.\]

Split the last term in the R.H.S. of (3.10) by setting
\[g^J(x) = g^J(\xi) + (g^J(x) - g^J(\xi))\]

and take the part involving \(g^J(\xi)\) alone onto the L.H.S. This gives an estimate
\[(3.15) \quad \left| \frac{d\Psi_N}{dt} + \frac{1}{2} \psi^2 \sigma^2(t) \Psi_N - r(\psi, t, N) \Psi_N \right| \leq |\psi| k_2 N^{\frac{1}{2}} P[\bar{\tau}_N(t) \notin S^+] + |\psi|^2 k_3 E|\bar{\tau}_N(t) - \xi(t)|^2 + |\psi|^3 k_4 N^{-\frac{1}{2}} E|\bar{\tau}_N(t) - \xi(t)|.
\]

From Equation (2.5), it follows easily that
\[(3.16) \quad E|\bar{\tau}_N(t) - \xi(t)| \leq k_6 N^{-\frac{1}{4}},\]

\[E|\bar{\tau}_N(t) - \xi(t)|^2 \leq k_7 N^{-1} \log_e N.\]
Thus the R.H.S. of (3.15) is bounded by

\[ (3.17) \quad k_8 |\psi| \log N + \frac{k_9 |\psi|^2 N^{-\frac{1}{2}}}{k_{10}} |\psi| \left| N^{-\frac{1}{2}} + k_{20} |\psi| N^{-1} \right| \equiv K(\psi, N). \]

Now \( \Psi_N^*(\psi, t) = \exp \{ - \frac{1}{2} \psi^2 \Sigma^2(t) + R(\psi, t, N) \} \) satisfies

\[ (3.18) \quad \frac{d\Psi_N^*}{dt} + \frac{1}{2} |\psi|^2 \sigma^2(t) \Psi_N^* - r(\psi, t, N) \Psi_N^* = 0, \]

and \( \Psi_N^*(\psi, 0) = 1 = \Psi_N(\psi, 0) = \Psi(\psi, 0) \). Let

\[ (3.19) \quad M = \tau^2 / 4 k_1 ; \]

then, if \( |\psi| \leq MN^{\frac{1}{2}} \), it follows that

\[ (3.20) \quad |\Psi_N^*(\psi, t) - \Psi(\psi, t)| \leq \exp(\frac{1}{2} \psi^2 \Sigma^2(t)) |\exp(R(\psi, t, N)) - 1| \leq k_1 |\psi| |N^{-\frac{1}{2}} t \exp(-\frac{1}{2} t \tau^2 \psi^2)|. \]

Also, if \( \Delta_N(\psi, t) \equiv \Psi_N(\psi, t) - \Psi_N^*(\psi, t) \),

\[ (3.21) \quad \Delta_N(\psi, 0) = 0 \]

and

\[ (3.22) \quad \left| \frac{d\Delta_N}{dt} + (\frac{1}{2} \psi^2 \sigma^2(t) - r(\psi, t, N)) \Delta_N \right| \leq K(\psi, N). \]

Integrating (3.22), for \( |\psi| \leq MN^{\frac{1}{2}} \), gives

\[ (3.23) \quad |\Delta_N(\psi, t)| \leq K(\psi, N) \int_0^t \exp\{\frac{1}{2} \psi^2 (\Sigma^2(u) - \Sigma^2(t)) - R(\psi, u, N) + R(\psi, t, N)\} du \leq 4(\tau \psi)^{-2} K(\psi, N) (1 - \exp(-\frac{1}{2} \psi^2 \tau^2 t)). \]

But \( \Psi_N(\psi, t) \) is the characteristic function of a probability distribution with finite mean, and \( \Psi(\psi, t) \) is that of \( N(0, \Sigma^2(t)) \). Hence, applying Chung (1968), p. 208, Lemma 2, we deduce that

\[ (3.24) \quad \sup_v \left| P[z_N(t) \leq v] - \Phi(v / \Sigma(t)) \right| \leq \frac{2}{\pi} \int_0^{M / \sqrt{N}} \left( |\Delta_N(\psi, t)| + |\Psi_N^*(\psi, t) - \Psi(\psi, t)| \right) |\psi|^{-1} d\psi + 24N^{-\frac{1}{2}} (M \Sigma(t) \sqrt{(2\pi^3)})^{-1}. \]

The integral is estimated using (3.20), (3.23) and (3.17), and the fact that, for all real \( w, 1 - \exp(-w^2) \leq \min(w^2, 1) \): the theorem follows.

**Remark.** It is easily possible to calculate an explicit formula for the upper bound of the error (3.8), in terms of the parameters of the process. It is omitted here for clarity, and because, in practice, the accuracy of the approximation seems vastly better than the bound suggests.
4. Further properties of the deterministic integrals

One of the main points of the method of [B] is that it shows how to get a more accurate asymptotic description of the process \( x_N \) than that given by the central limit theorem. The essential feature in the argument is a sequence of functions \( \phi_r(x, t), \quad r = 0, 1, \ldots, \) such that \( \sum_{j=0}^r N^{-j} \phi_j(x, t) \) is in some sense a martingale, but for an error of order \( N^{-r-1} \); the appropriate functions to take for \( \phi_0 \) are integrals of the deterministic equations, and the functions \( \phi_j, j \geq 1, \) can in principle be calculated sequentially. The discussion in [B] is heuristic, but it can be made precise on the basis of the following paragraphs.

Let \( \phi(x, t): R^+ \times R^+ \to R \) be measurable, and \( t \)-differentiable in \( t \geq 0 \) for each \( x \). Define

\[
Q_\phi(x, t) = \partial \phi(x, t)/\partial t + N \sum_j g_j'(x) \{ \phi(x + JN^{-1}t) - \phi(x, t) \}.
\]

Let \( t = \inf \{ t: t \geq 0, x_N(t) \notin S^{\delta} \} \), and set \( \bar{t} = \min(t, \tau) \). Let \( \mathcal{F}_t \) denote the \( \sigma \)-field generated by \( x_N(s), 0 \leq s \leq t \); let \( U \) denote \( S^\delta \times [0, T] \), and \( C(U) \) the set of functions \( \phi(x, t): U \to R \) which are infinitely differentiable in \( x \), and once differentiable in \( t \).

**Lemma 2.** Let

\[
Z_N(t) \equiv \phi(x_N(\bar{t}), \bar{t}) - \int_0^t Q_\phi(x_N(s), s) ds:
\]

then the family \( \{Z_N(t), \mathcal{F}_t, 0 \leq t \leq T \} \) is a martingale.

**Proof.** Because the event \( \{ \tau \leq s \} \) is \( \mathcal{F}_s \)-measurable, it is enough to show that

\[
E[Z_N(t) | \mathcal{F}_s] = Z_N(s), \quad \text{whenever} \quad x_N(s) \in S^{\delta} \cap L^0_N.
\]

This is easily verified, because \( x_N(\bar{t}) \) can only take one of a finite set of values, either by semi-group methods, similar to those in [K], or directly by using the forward differential equations.

The sequences of functions \( \{ \phi_r \} \) of [B] are constructed as follows. \( \phi_0(x, t) \) is taken to be any solution of Equation (3.1), or, equivalently, any integral of the deterministic equations

\[
\dot{x} = \sum_j Jg_j'(x).
\]

The functions \( \phi_r \) are then determined recursively from the formula

\[
\phi_r(x, t) = - \int_0^{t+r+1} \sum_{s=2}^{r+1} L_s(\phi_{r+1-s})(\eta(w), w) dw,
\]

where

\[
L_s(\phi)(x, t) = \frac{1}{r!} \sum_j g_j'(x) \left( \sum_{j_1, \ldots, j_r=1}^n \prod_{m=1}^r J_{j_m} \partial / \partial x_{j_m} \right) \phi(x, t),
\]

and

\[
\sum_j g_j'(x) = 0.
\]
and where $\eta(w) = \eta(x, t, w)$ is that deterministic trajectory which has $\eta(t) = x$: thus, $x(u) = \eta(x, t, u)$ satisfies (4.3) together with $x(t) = x$. Under Assumption B, it is always possible, for some $\varepsilon > 0$, to find functions $\phi_0(x, t)$ in $C(U)$, and only such functions will henceforth be considered. A set of $n$ such independent functions (which can be taken as a natural coordinate system for the process) is given by the components of the vector $\eta(x, t) = \eta(x, t, 0)$: see Hartman (1964), chapter V. This, and the form of (4.4), implies that the sequence can be continued for any number of terms, and that $\psi_r(x, t) = \sum_{j=0}^r N^{-j} \phi_j(x, t)$ is in $C(U)$.

Lemma 3. For any $0 \leq s \leq t$, let $t^*$ be a stopping time such that $s \leq t^* \leq t$ almost surely. Then

$$E[\psi_r(x(t^*), t^*)]_{t^*} = \psi_r(x(s), s) + O(N^{-r-1}), \tag{4.6}$$

where the order term is independent of $t^*$ and $t_s$.

Proof. The construction of $\psi_r$ was chosen precisely in such a way that $Q_{\psi_r}(x, t) = N^{-r-1}f(x, t, N)$, where, for each $(x, t) \in U, f(x, t, N)$ is of order 1 as $N \to \infty$. In fact, Expression (4.4) ensures that the equations

$$\partial \phi_j(x, t)/\partial t + \sum_{s=1}^{r+1} L_s(\phi_{r+1-s})(x, t) = 0 \tag{4.7}$$

are satisfied for $r \geq 1$, and Equation (4.7) is also satisfied for $r = 0$, when it reduces to Equation (3.1). This gives

$$Q_{\psi_r}(x, t) = \frac{\partial}{\partial t} \left( \sum_{j=0}^r N^{-j} \phi_j(x, t) \right) + N \sum_j g^f(x) \sum_{j=0}^r N^{-j} (\phi_j(x + jN^{-1}, t) - \phi_j(x, t))$$

$$= \sum_{j=0}^r N^{-j} \frac{\partial \phi_j}{\partial t}(x, t) + N \sum_{j=0}^r N^{-j+1} \sum_{s=1}^{r+1-j} N^{-s} L_s(\phi_j)(x, t) + O(N^{-r-1}), \tag{4.8}$$

where the order term is the sum of Taylor remainders in the estimates of differences $\phi_j(x + jN^{-1}, t) - \phi_j(x, t)$, and hence, since each $\phi_j$ is in $C(U)$, is uniform over $U$. But the coefficient of $N^{-j}$ in (4.8), $j = 0, 1, \cdots$, vanishes because of (4.7), and so

$$\sup_{(x, t) \in U} \left| f(x, t, N) \right| = O(1) \quad \text{as} \quad N \to \infty.$$ 

The lemma now follows from Lemma 2 and the optional stopping theorem, since the corresponding $Z$ martingale is bounded in $U$. 


Lemma 4. Let \( h(x,t) \in C(U) \) be any integral of (4.3), such that \( h(x(0),0) = 0 \). Then, for any \( r \geq 0 \),

\[
E \left| h(x_N(t^*),t^*) \right|^r = O(N^{-\frac{1}{2}r}) \text{ independently of } t^*,
\]

where \( t^* \) is defined as in Lemma 3.

Proof. It is enough to prove (4.9) for arbitrary even integral \( r \). Let \( \phi \) without suffix denote a generic member of \( C(U) \), not necessarily the same at each appearance. Then, if \( \psi = \phi^{k_s} \), \( L_s(\psi) = \phi^{k_s-s} \) for each integer \( s, 2 \leq s \leq k \). Hence, from (4.4), constructing the sequence \( \{\phi_s\} \) starting with \( \phi_0 = h^r \),

\[
\phi_1(x,t) = -\int_0^t L_2(h^r)(\eta(x,t,w),w)\,dw
\]

\[
= -h^{r-2}(x,t)\int_0^t \phi(\eta(x,t,w),w)\,dw,
\]

since \( h \) is an integral of (4.3). Proceeding by induction, using (4.4), it follows that, for all \( 0 \leq s \leq \frac{1}{2}r \),

\[
\phi_s(x,t) = \phi(x,t)[h(x,t)]^{r-2s}.
\]

Hence, applying Lemma 3 to \( \psi_s(x,t) = \sum_{j=0}^{\frac{1}{2}r} N^{-j}\phi_j(x,t) \), and using (4.10), we have

\[
E[h(x_N(t^*),t^*)]^r = O \left( \sum_{j=1}^{\frac{1}{2}r} N^{-j}E[h(x_N(t^*),t^*)]^{r-2j} \right), \text{ even } r \geq 2,
\]

where the order term depends on \( t^* \) only through the expectations. Equation (4.9) now follows, for all even \( r \geq 2 \), inductively from (4.11).

Remark. From Lemma 3 with \( r = 0 \), it follows also that

\[
E h(x_N(t^*),t^*) = O(N^{-1}).
\]

Taking the components of \( \eta(x,t) = \eta(x,t,0) - x(0) \) as the functions \( h \) in Lemma 4, it is easy to deduce that, for all \( r \geq 0 \),

\[
E \left| \eta(x_N(t^*),t^*,0) - x(0) \right|^r = O(N^{-\frac{1}{2}r}).
\]

Now, from the definition of \( L \) and the fact that \( x(u) = \eta(x,t,u) \) satisfies (4.3), it follows that, for any \( t \in [0,T] \) and any \( x^1, x^2 \) such that \( \eta(x^I,0,v) \in S^e \) for all \( 0 \leq v \leq t \),

\[
\left| \eta(x^1,0,t) - \eta(x^2,0,t) \right|
\]

\[
\leq |x^1 - x^2| + \left\{ \sum_{j=0}^{t} \left( \sum_{I} g^I(\eta(x^I,0,v)) - \sum_{I} g^I(\eta(x^2,0,v)) \right) dv \right\}
\]

\[
\leq |x^1 - x^2| + \int_0^t L|\eta(x^1,0,v) - \eta(x^2,0,v)|dv.
\]
Thus, from Gronwall’s inequality,
\[ |\eta(x^1, 0, t) - \eta(x^2, 0, t)| \leq |x^1 - x^2|e^{Lt}, \]
and so, taking \( \eta(x^j, t, 0) \) for \( x^j \),
\[ |x^1 - x^2| \leq |\eta(x^1, t, 0) - \eta(x^2, t, 0)|e^{Lt}, \]
provided that \( \eta(x^j, t, v) \in S^c \) for all \( 0 \leq v \leq t \). It now follows easily from (4.12) that
\[ (4.13) \quad E\left| x_N(t^*) - \xi(t^*) \right|^r = O(N^{-r}), \]
the order term depending on \( t^* \) only through \( T \).

Lemma 3 and Equation (4.13) are between them enough to justify the detailed calculations in [B], insofar as the distribution and moments under investigation belong to \( x_N(t) \); but taking \( t^* = \bar{t} \) in Equation (4.13), it follows that, for any \( \bar{r} \geq 0 \),
\[ (4.14) \quad P[\bar{t} < t] = P\left[\left| x_N(\bar{t}) - \xi(\bar{t}) \right| > \frac{1}{2}\varepsilon \right] = O(N^{-\bar{r}}), \]
so that the distribution of \( x_N(\bar{t}) \) is an accurate asymptotic approximation to that of \( x_N(t) \). Thus, although moments evaluated for \( x_N(\bar{t}) \) may not be those of \( x_N(t) \), any 
\textit{distributional} information obtained from them applies, within the limits of
Equation (4.14), to \( x_N(t) \) also.

The results of Lemma 4 can be complemented as follows. For \( h \) as in Lemma 4, let
\[ H = \sup_{(x, \bar{t}) \in U} \left| \nabla h(x, t) \right|; \]
\[ W^J = \sup_{(x, \bar{t}) \in U} \left| \sum_j \sum_k J_j J_k \partial^2 h(x, t) / \partial x_j \partial x_k \right|; \]
\[ G = \sup_{(x, \bar{t}) \in U} \sum_j g^J(x) \left| J \right|^2; \]
\[ W = \sup_{(x, \bar{t}) \in U} \sum_j g^J(x) W^J; \]
\[ c = H \max \left| J \right|. \]

\textbf{Lemma 5.} For any \( \bar{k} \geq 0 \), under the assumptions of Lemma 4,
\[ (4.16) \quad E\left\{ \exp\left( kh(x_N(t^*), t^*) \right) \right\} \leq \exp \{ C(N, k)T \}, \]
where
\[ (4.17) \quad C(N, k) = \frac{1}{2} k N^{-1} W + \frac{1}{2} k^2 N^{-1} G H^2 \exp(k N^{-1} c). \]

\textbf{Proof.} Let \( \phi(x, t) \equiv \exp(kh(x, t)) \), and, for any \( v \in [0, T] \) define \( v' = \min(v, t^*) \). From Lemma 2 and the optional stopping theorem, since \( h(x(0), 0) = 0 \),
\[ E\{\exp(\theta h(x_N(v'), v'))\} = E \int_0^{v'} Q_\phi(x_N(u), u) \, du + 1 \]
\[ \leq \int_0^{v'} E|Q_\phi(x_N(u'), u')| \, du + 1, \]
where, defining
\[ \Delta h^J(x, t) = h(x + JN^{-1}, t) - h(x, t), \]
\(Q_\phi\) is given by
\[ Q_\phi(x, t) = \exp(\theta h(x, t)) \{ k \frac{\partial h(x, t)}{\partial t} \}
+ N \sum_J g^J(x) [\exp(k \Delta h^J(x, t)) - 1]. \]

By Taylor's theorem, for \((x, t) \in U,\)
\[ |\Delta h^J(x, t)| \leq N^{-1} H|J|; \]
\[ |\Delta h^J(x, t) - N^{-1} J \cdot \nabla h(x, t)| \leq \frac{1}{2} N^{-2} W^J; \]
and so the term in braces in the R.H.S. of (4.20) can be written as
\[ k[\partial h(x, t)/\partial t + \sum_J g^J(x) J \cdot \nabla h(x, t)] \]
\[ + \frac{1}{2} kN^{-1} \theta_1 W + \frac{1}{2} k^2 N^{-1} GH^2 \theta_2 \exp(kN^{-1}c), \]
where \(\theta_1\) and \(\theta_2\) denote quantities of modulus not exceeding unity, and where the term in square brackets is zero because \(h\) satisfies Equation (3.1). Thus (4.18) can be reduced to
\[ E\{\exp(\theta h(x_N(v'), v'))\} \leq 1 + \int_0^{v'} C(N, k) E\{\exp(\theta h(x_N(u'), u'))\} \, du, \]
and by Gronwall's inequality this implies that
\[ E\{\exp(\theta h(x_N(v'), v'))\} \leq \exp\{C(N, k)v\}. \]

Taking \(v = T,\) the Lemma follows, because \(T' = t^*\) a.s.

**Corollary 1.** Under the assumptions and definitions of Lemma 5,
\[ P[|h(x_N(t), t)| \geq y] \leq 2 \exp\{-ky + TC(N,k)\}. \]

**Proof.** It is immediate from Lemma 5, taking \(t^* = \bar{t},\) that
\[ P[h(x_N(t), t) \geq y] \leq \exp\{-ky + TC(N,k)\}, \]
and the Corollary follows because \(-h\) is also a solution of (3.1), and has the same bounds (4.15) as \(h.\)
The choice of \( k \) is arbitrary, and for the purposes of Corollary 1 may be chosen depending on \( y \) and \( N \) to give a convenient estimate in (4.23). For example, if \( k \) is taken to be \( N^{\frac{1}{2}} \), Equation (4.23) implies that

\[
P[|h(x_N(t), i)| \geq y] \leq 2 \exp \left\{ -yN^{\frac{1}{2}} + \frac{1}{2} TGH^2 (1 + O(N^{-\frac{1}{2}})) \right\}.
\]

It may in practice be easier to work with this simple estimate than the better one given by the following corollary.

**Corollary 2.** Under the assumptions and definitions of Lemma 5, let \( a = \frac{1}{2} W T \) and \( b = \frac{1}{2} GH^2 T \), and let \( \tilde{y} \) denote \( y - aN^{-1} \). Then, for \( y \geq aN^{-1} \),

\[
P[|h(x_N(t), i)| \geq y] \leq 2 \exp \left\{ -N\tilde{y}u(\tilde{y})(1 + cu(\tilde{y}))/ (2 + cu(\tilde{y})) \right\}
\]

\[
\leq 2 \exp \left\{ -\frac{1}{2}N(y - aN^{-1})u(y - aN^{-1}) \right\},
\]

where \( u(w) \) is the unique non-negative root \( u \) of the equation

\[
ub(2 + cu)e^{cw} = w.
\]

**Remark.** Equation (4.26) implies that \( u(\tilde{y}) \leq \tilde{y}/2b \), with asymptotic equality as \( \tilde{y} \to 0 \): since the significant order of \( y \) turns out to be \( N^{-\frac{1}{2}} \), this estimate of \( u \) is extremely useful.

**Proof.** Equation (4.25) is deduced directly from Equation (4.23) by choosing \( k = k(y, N) \) to minimise its R.H.S.

Lemma 5 can be used to provide information about the tails of the distribution of \( x(t) - \xi(t) \) in the same way as was used after Lemma 4. It does not require the full strength of Assumption B, because it only involves the existence of continuous derivatives of \( h \) up to the second order. By Hartman (1964), Chapter V, Theorem 4.1, if the components of \( \eta(x, t) = \eta(x, t, 0) - x(0) \) are taken for \( h \), these derivatives exist and are continuous, if the functions \( g^J(x) \) are all twice continuously differentiable. Lemma 5 can in fact be further strengthened, by requiring only the existence of continuous first derivatives for \( g^J(x) \). Under this hypothesis, the second estimate in (4.21) is changed, with a corresponding modification of \( C(N, k) \), to \( N^{-1/2}f^J(N) \) where \( f^J(N) = o(1) \) as \( N \to \infty \).

Note that no use is made of Theorem K in Section 4: however, Lemma 2 corresponds closely to Proposition (2.1) of [K].

5. Examples

(a) **The linear birth and death process**

Let \( Z_N(t) \): \( 0 \leq t \leq T \) be the continuous parameter Markov process on the non-negative integers, with conservative \( Q \)-matrix specified by the transition rates

\[
q_{j, j+1} = ja, \quad j = 0, 1, 2, \ldots;
\]

\[
q_{j, j-1} = jb,
\]
and with initial state \( Z_N(0) = N x_0 \). Define the sequence of processes \( \{x_N(t): 0 \leq t \leq T\} \) with state space \( L^1_N \) by
\[
x_N(t) = N^{-1} Z_N(t), \quad 0 \leq t \leq T.
\]
Then \( x_N(0) = x_0 \), fixed independent of \( N \), and \( x_N(t) \) is a continuous parameter Markov process of the form assumed in Equation (2.1), with \( \mathcal{J} = \{-1, 1\} \) and
\[
g^+(x) = ax, \quad g^{-1}(x) = bx.
\]

The deterministic Equations (2.2) for this process are
\[
\dot{\xi}(t) = (a - b) \xi(t), \quad 0 \leq t \leq T; \quad \xi(0) = x_0,
\]
and Assumption A is satisfied with
\[
\xi(t) = x_0 \exp((a - b)t), \quad 0 \leq t \leq T.
\]

The set \( S^\varepsilon \), for any fixed \( \varepsilon > 0 \), is then the closed interval \([x_0 m_1 - \varepsilon, x_0 m_2 + \varepsilon]\), where we define
\[
m_1 = \min\{1, e^{(a-b)T}\}, \quad m_2 = \max\{1, e^{(a-b)T}\},
\]
and Assumption B is clearly satisfied for any \( \varepsilon, 0 < \varepsilon \leq x_0 m_1 \).

Equation (3.1) for this sequence of processes reduces to
\[
\frac{\partial h}{\partial t} + (a - b)x \frac{\partial h}{\partial x} = 0,
\]
and it has an integral \( h(x, t) = x e^{-(a-b)t} \); hence, applying Lemma 1, we deduce that \( W_N(t) \equiv N^{+}(N^{-1}Z_N(t)e^{-(a-b)t} - x_0) \), considered as a random element of \( D[0, T] \), converges weakly in the Skorokhod topology to the Gaussian random function \( W \) with
\[
EW(t) = 0; \quad EW(s)W(t) = \Sigma^2(s), \quad 0 \leq s \leq t \leq T;
\]
where, by Equation (3.5),
\[
\Sigma^2(s) = \int_0^s \sigma^2(u) du = \begin{cases}
x_0^{-1}(a + b)(1 - e^{-(a-b)s})(a - b)^{-1}, & a \neq b, \\
x_0^{-1}(a + b)s, & a = b.
\end{cases}
\]
Alternatively, if we define, for \( 0 \leq u \leq 1 \),
\[
W'_N(u) \equiv \sqrt{(a - b)} \left\{ x_0^{-1}(a + b)(1 - e^{-(a-b)T}) \right\}^{-\frac{1}{2}}
\times W_N(- (a - b)^{-1} \log \{ 1 - u [1 - e^{-(a-b)T})] \}),
\]
when \( a \neq b \), and
\[
W'_N(u) \equiv \{ Tx_0^{-1}(a + b) \}^{-\frac{1}{4}} W_N(uT)
\]
when \( a = b \), then \( W'_N \), considered as a random element of \( D[0,1] \), converges weakly to standard Wiener measure as \( N \to \infty \).

The rate of convergence implied by Theorem 1 for the distribution of \( W_N(t) \) is of order \( N^{-1} \log N \), and the bound, if calculated, for Equation (3.8) is found meaningful, for values of \( T \) around \( L^{-1} \), only when \( N \) exceeds 5000. This should be compared with the results obtained by regarding the \( Z_N(t) \) process as a sum of \( N \) independent processes with initial population 1, and using the Berry-Esseen theorem, giving the expected convergence rate of \( N^{-\frac{1}{2}} \), and a bound meaningful when \( N \) exceeds 100. It seems reasonable to suppose that, in other situations also, the approximation will be better than Theorem 1 suggests.

A feature of this particular example is that, for \( h(x,t) = x \exp(-(a - b)t) \), \( h(x_N(t),t) \) is actually a martingale, and not just an approximation to one. Similarly, in the notation of Section 4, the sequences \( \{\phi_s\} \) corresponding to \( \phi_0(x,t) = \{h(x,t) - x_0\}^r \) have the property that \( \phi_s(x,t) = 0 \) identically in \((x,t)\) for all \( s \geq r \), and that \( \psi_{r-1}(x_N(t),t) \) is a martingale.

In \( n \)-dimensions similar results hold. Suppose that the functions \( g^j(x) \) are linear in \( x \) in any region \( S \subseteq R^n \). Then the deterministic equations (4.3) are of the form

\[
\dot{x} = Ax + b, \quad x \in S,
\]

for some matrix \( A \) and vector \( b \). It is easily checked that, if \( \phi(x,t) \equiv x \cdot c(t) + d(t) \), where

\[
\dot{c} + cA = 0; \quad d + c \cdot b = 0;
\]

then \( Q_\phi(x,t) = 0 \) identically in \((x,t)\in S \times [0,\infty) \): and it can be shown that \( \{\phi(x_N(t),t) \in S \times [0,\infty) \} \) is a martingale.

By solving (5.1), it can be seen that the deterministic solution \( \eta(x,t,u) \), as defined following Equation (4.5), is linear in \( x \). Let \( \phi_0(x,t) \) be any integral of (5.1) which is multinomial of degree \( d \) in the coordinates \( x_j \), and construct the sequence \( \{\phi_s\} \) corresponding to it, using the formula (4.4). From Equation (4.5), and the fact that \( g^j(x) \) is linear in \( x \), \( L_2(\phi_0)(x,t) \) is of degree at most \( d - 1 \) in \( x \), for \( d \geq 2 \), and is zero if \( d = 0 \) or \( d = 1 \): hence, from Equation (4.4), \( \phi_1(x,t) \) is of degree at most \( d - 1 \) in \( x \) for \( d \geq 2 \), and is identically zero if \( d = 0 \) or \( d = 1 \). An inductive argument based on the same considerations shows that \( \phi_r(x,t) \) is of degree at most \( d - r \) for \( r \leq d - 1 \), and is identically zero if \( r \geq d \).

It is then easy to verify that \( Q_{\psi_N-1}(x,t) \) is zero identically in \((x,t)\in S \times [0,\infty) \), and it can be shown that \( \{\psi_{d-1}(x_N(t),t) \in S \times [0,\infty) \} \) is a martingale. Convenient integrals of (5.1) to take for \( \phi_0(x,t) \) are simple multinomial expressions in the coordinates of the vector \( \eta(x,t) \equiv \eta(x,t,0) - x(0) \).

The importance of the linear case is that it can often be used as a local approximation to more complicated systems. The simple form of these exact martingales is one of the advantages thereby gained.
(b) The quadratic birth and death process

The linear birth and death process was taken as an example because its
distributions can be determined by other means for the purpose of comparison. The
importance of the method is the wide class of processes to which it is applicable,
and such a simple example is unrepresentative of its power. The following example,
differing only slightly from the previous one, serves to illustrate its use for a sequence
of non-regular processes.

Let $Z_N(t): 0 \leq t \leq TN^{-1}$ be a continuous parameter Markov process on the
non-negative integers, with conservative $Q$-matrix specified by the transition rates

$$
q_{j,j+1} = j^2a, \quad j = 0,1,2,\ldots,
q_{j,j-1} = j^2b,
$$

with initial state $Z_N(0) = Nx_0$, and with $a > b$. Let

$$
x_N(t) = \frac{1}{N}x_N(tN^{-1}), \quad 0 \leq t \leq T;
$$

then $x_N(t)$ is a process of the form assumed in Equation (2.1). Proceeding as in the
previous example, we are led to conclude that, if $T < [x_0(a-b)]^{-1}$, then

$$
W_N(t) \equiv N^{-1}\{x_N(t)^{-1} - x_0^{-1} + (a-b)t\},
$$

considered as a random element of $D[0,T]$, converges weakly to the Gaussian
random function $W$ with

$$
EW(t) = 0;
$$

$$
E\{W(s)W(t)\} = (a+b)\{1 - [1 - sN^{-1}(a-b)]^2\}/(3N^3(a-b));
$$

$$
0 \leq s \leq t \leq T.
$$

However, at $t = \{x_0(a-b)\}^{-1}$, the solution of the deterministic equations

$$
\dot{\xi}(t) = (a-b)\xi^2(t); \quad \xi(0) = x_0,
$$

becomes infinite, and so, if $T \geq \{x_0(a-b)\}^{-1}$, Assumption A is violated, and the
theorems are no longer applicable.

(c) The closed epidemic

Let $X_N(t) \equiv (X^{(1)}_N(t), X^{(2)}_N(t))$: $0 \leq t \leq T$ be a continuous parameter bivariate
Markov process on $Z^+ \times Z^+$, with conservative $Q$-matrix specified by the transition rates

$$
(j,k) \rightarrow (j-1,k+1) \quad \text{at rate } N^{-1}\alpha j k,
$$

$$
(j,k) \rightarrow (j,k-1) \quad \text{at rate } \beta k,
$$
and with initial state \((X^{(1)}_N(t), X^{(2)}_N(t)) = (N, Nh)\), where \(\alpha, \beta\) and \(h\) are fixed independent of \(N\). This process is commonly used to model the spread of an infection in a closed population, in which the number of susceptibles is denoted by \(X^{(1)}\) and the number of infectious persons by \(X^{(2)}\), the remainder of the population being presumed immune or dead. The first transition then represents the spread of the infection by homogeneous mixing of infectious and susceptible people, and the second transition represents the rate at which infectious people are removed from the mixing population. The dependence of the rates of these transitions upon \(N\), which is proportional to the total population size, is here chosen in its most natural form: however, the initial number of infectious persons is not chosen, as would be natural, to be unity, since this necessitates a further special approximation during the initial phase of the process. For two generalisations of this model, see Downton (1968) and Ridler-Rowe (1964).

Let \(x_N(t) = N^{-1}X_N(t), 0 \leq t \leq T\): then the sequence \(\{x_N(t)\}\) of processes is of the form assumed in (2.1), with \(\mathcal{F} = \{(−1, 1); (0, −1)\}\) and

\[
g^{−1, 1}(x) = \alpha x^{(1)}x^{(2)}, \quad g^{0, −1}(x) = \beta x^{(2)}.
\]

The deterministic equations (2.2) for this process are

\[
\begin{align*}
\frac{d\xi^{(1)}}{dt} &= -\alpha \xi^{(1)}\xi^{(2)}, \\
\frac{d\xi^{(2)}}{dt} &= (\alpha \xi^{(1)} - \beta)\xi^{(2)},
\end{align*}
\]

(5.2)

to be satisfied in some finite interval \(0 \leq t \leq T\) with the initial condition \(\xi(0) = (1, h)\). They have a unique solution \(\xi(t), 0 \leq t \leq T\), for any finite \(T \geq 0\), given implicitly by

\[
\xi^{(1)}(t) + \xi^{(2)}(t) - \eta \log \xi^{(1)}(t) - 1 - h = 0,
\]

(5.3)

\[
\int_{\log \xi^{(1)}(t)}^{0} (1 + h + \eta u - e^u)^{-1} du - \alpha t = 0,
\]

(5.4)

where \(\eta = \beta / \alpha\) is the threshold parameter: thus Assumption A is satisfied. Assumption B is also satisfied with

\[
0 < \varepsilon \leq \min(he^{-\beta T}, e^{-(1+h)/\eta}).
\]

The expressions on the L.H.S. of Equations (5.3) and (5.4) correspond to two independent integrals of the linearised martingale equation (3.1) for the process: they are

\[
h(x, t) = x^{(1)} + x^{(2)} - \eta \log x^{(1)} - 1 - h,
\]

(5.5)

\[
h^*(x, t) = \int_{\log x^{(1)}}^{0} (x^{(1)} + x^{(2)} - \eta \log x^{(1)} + \eta u - e^u)^{-1} du - \alpha t.
\]
With a view to Lemma 1, we define a new pair of coordinate random variables for the process,

\[ z_N(t) = (1 - \eta / \xi^{(1)}(t))y_N^{(1)}(t) + y_N^{(2)}(t) \]

(5.6)

\[ w_N(t) = \{-[\xi^{(1)}(t)]^{-1} - (\eta[\xi^{(1)}(t)]^{-1} - 1)I(t)\}y_N^{(1)}(t) - I(t)y_N^{(2)}(t), \]

where

(5.7)

\[ I(t) = \int_0^{\log \xi^{(1)}(t)} (1 + h + \eta u - e^u)^{-2} du. \]

We then deduce that, as \( N \to \infty \), \( (z_N, w_N) \) converges weakly in \( \{D[0, T]\}^2 \) to the continuous bivariate Gaussian Markov random function \( (z, w) \) specified by

\[ E^*\{z(t)\} = z_0; \quad E^*\{w(t)\} = w_0; \]

\[ E^*\{(z(t) - z_0)^2\} = [x^{-1}\eta^2 - \eta \log x]_{x = \xi^{(1)}(t_0)}, \]

(5.8)

\[ E^*\{(w(t) - w_0)^2\} = \alpha \int_{t_0}^t \left[ \xi^{(2)}(u) \right]^{-1} - 2\eta I + \eta I^2 \xi^{(2)}[\xi + \xi^{(1)}] du, \]

\[ E^*\{z(t)w(t) - z_0w_0\} = \beta \int_{t_0}^t \left[ \xi^{(1)}(u) \right]^{-1} \left\{ 1 - I\xi^{(2)}[\xi + \xi^{(1)}] \right\} du, \]

for any \( 0 \leq t_0 \leq t \leq T \) and any \( (z_0, w_0) \in \mathbb{R}^2 \), where \( E^* \) denotes the expectation conditional on \( (z(t_0), w(t_0)) = (z_0, w_0) \), and where the argument \( u \) in the functions \( \xi \) and \( I \) within the integrals has been omitted.

As an application of this result we deduce that, as \( N \to \infty \),

\[ P \left[ \max_{0 \leq t \leq T} |z_N(t)| > x \right] \to m(x, T) = P \left[ \max_{0 \leq t \leq \xi^{(1)}(T)} |B(t)| > x \right], \]

where \( B(t) \) is standard Brownian motion and where

\[ \Sigma^2(T) = \eta^2[\xi^{(1)}(T)]^{-1} - \eta \log \xi^{(1)}(T). \]

This may be rewritten, in terms of the original epidemic process \( X_N \), as

\[ P \left[ \max_{0 \leq t \leq T} \left| h(N^{-1}X_N(t), t) - h(\xi(t), t) \right| > x \right] \to m(x, T) \text{ as } N \to \infty, \]

and may hence be interpreted to give a confidence band around the deterministic trajectory of the epidemic, which the \( X_N \) process will leave before time \( T \) with probability approximating \( m(x, T) \). Such information could provide a useful guide to health authorities who, faced with an incipient epidemic, need to predict what emergency resources will be required.


Acknowledgments

The author expresses his thanks to D. G. Kendall, P. Billingsley and G. K. Eagleson for their invaluable stimulus and encouragement in this work.

References


