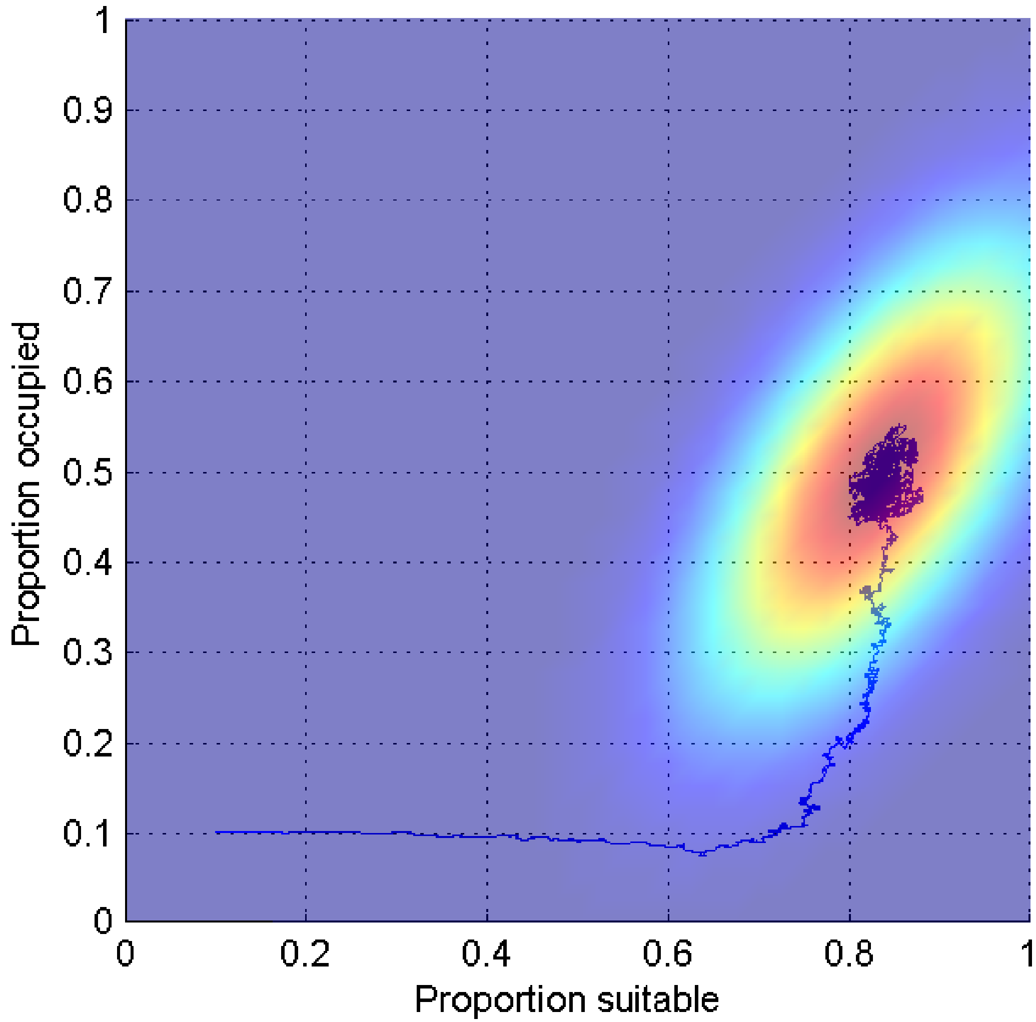
Quasi-stationary Distributions

Phil Pollett

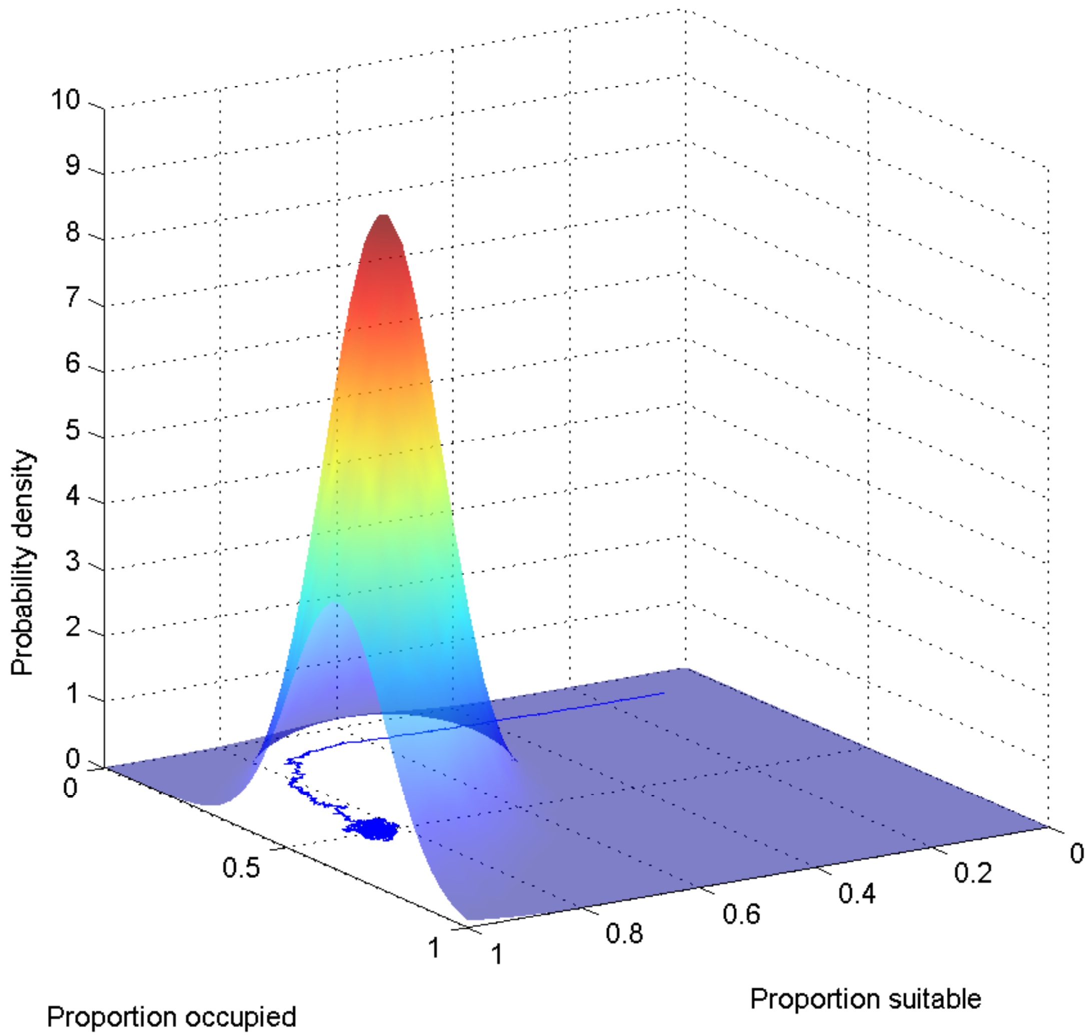
Department of Mathematics and MASCOS
University of Queensland

AUSTRALIAN RESEARCH COUNCIL
Centre of Excellence for Mathematics
and Statistics of Complex Systems

Habitat dynamics



Habitat dynamics



Habitat dynamics

A continuous-time Markov chain $\{(m(t), n(t)),\ t \geq 0\}$.

- $S = \{(m, n) : 0 \leq n \leq m \leq M\}$
- $m =$ the number of suitable patches
- $n =$ the number of occupied patches
- $M =$ the total number of patches (fixed)
- Transition rates $\{q(x, y),\ x, y \in S\}$:

$$q((m, n), (m + 1, n)) = r(M - m)$$

$$q((m, n), (m - 1, n)) = s(m - n)$$

$$q((m, n), (m - 1, n - 1)) = sn$$

$$q((m, n), (m, n + 1)) = c\frac{n}{M}(m - n)$$

$$q((m, n), (m, n - 1)) = en$$

Metapopulation network

A continuous-time Markov chain $\{(n_1(t), \ldots, n_M(t)),\ t \geq 0\}$.

- $S = \{0, \ldots, N_1\} \times \cdots \times \{0, \ldots, N_M\}$
- $n_i =$ the patch-i population size (capacity N_i)
- $M =$ the total number of patches (fixed)
- Transition rates $\{q(x, y),\ x, y \in S\}$:

$$q(n, n + e_i) = b\frac{n_i}{N_i}(N_i - n_i)$$

$$q(n, n - e_i + e_j) = \gamma_{ij}\frac{n_i}{N_j}(N_j - n_j) \qquad (j \neq i)$$

$$q(n, n - e_i) = \mu n_i.$$

Here b is the local birth rate, γ_{ij} is the rate of migration from patch i to patch j, and μ is the per-capita death rate.

An auto-catalytic reaction

Consider the reaction $A \stackrel{X}{\to} B$, where X is a catalyst.

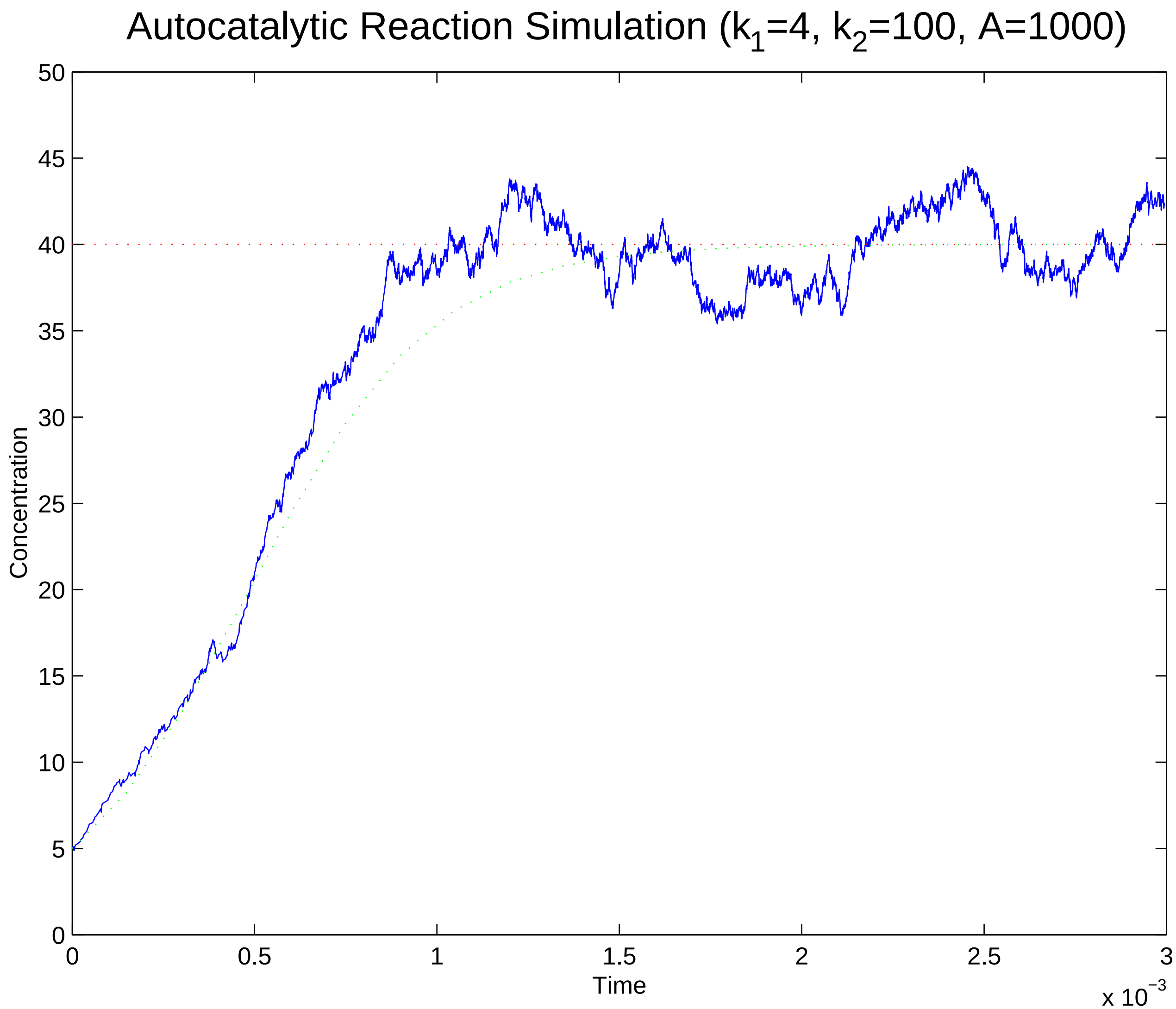
A two stage (auto-catalytic) scheme:

$$A + X \stackrel{k_1}{\to} 2X \quad \text{and} \quad 2X \stackrel{k_2}{\to} B.$$

Let $X(t) =$ number of X molecules at time t. Suppose that the concentration of A is held constant; let a be the number of molecules of A. The state space is $S = \{0, 1, 2, \ldots\}$ and the transition rates are:

$$q_{ij} = \begin{cases} k_1 a i & \text{if } j = i + 1 \\ k_2 \binom{i}{2} & \text{if } j = i - 2 \\ 0 & \text{otherwise.} \end{cases}$$

An auto-catalytic reaction



Ingredients

- A (time-homogeneous) Markov chain $(X(t),\, t \geq 0)$ in continuous time, taking values in $S = \{0, 1, 2, \ldots\}$.
- Transition rates $Q = \{q_{ij}, i, j \in S\}$: $q_{ij}\,(\geq 0)$, for $j \neq i$, is the transition rate from state i to state j and $q_{ii} = -q_i$, where $q_i = \sum_{j \neq i} q_{ij}\,(< \infty)$ is the transition rate out of state i.
- Assumptions : For simplicity, take 0 to be the sole absorbing state (that is, $q_{0j} = 0$), suppose that $C = \{1, 2, \ldots\}$ is irreducible and that we reach 0 from C with probability 1.
- Transition probabilities : $P(t) = \{p_{ij}(t), i, j \in S\}$, where $p_{ij}(t) = \Pr(X(t) = j | X(0) = i)$. State probabilities : $p(t) = \{p_j(t), j \in S\}$, where $p_j(t) = \Pr(X(t) = j)$.

Ingredients

- Initial distribution : $a = (a_j, j \in S)$ ($a_0 = 0$).
- Forward equations : the state probabilities satisfy $p'(t) = p(t)Q$, $p(0) = a$. In particular, since $q_{0j} = 0$,

$$p_j'(t) = \sum_{i \in C} p_i(t) q_{ij}, \quad j \in S, \ t > 0.$$

- Conditional state probabilities : define $m(t) = (m_j(t), j \in C)$ by

$$m_j(t) = \Pr(X(t) = j \mid X(t) \in C) = \frac{p_j(t)}{\sum_{k \in C} p_k(t)},$$

the chance of being in state j *given that the process has not reached* 0.

Quasi-stationary distributions

Question 1. Does $m(t) \to m$ as $t \to \infty$?

Question 2. Can we choose the initial distribution a in order that $m_j(t) = a_j$, $j \in C$, for all $t > 0$?

Definition. A distribution $m = (m_j, j \in C)$ satisfying $m(t) = m$ for all $t > 0$ is called a *quasi-stationary distribution* (QSD).

If $m(t) \to m$ then m is called a *limiting-conditional distribution* (LCD).

Question 3. Is the QSD unique?

Question 4. When an LCD exists, is it a QSD?

Question 5. Does the LCD depend on the initial distribution?

Origins

Yaglom proved that the LCD exists for the subcritical (Galton-Watson) branching process (in discrete time) starting from a single initial ancestor:

> A.M. Yaglom (1947) Certain limit theorems of the theory of branching processes (in Russian). *Dokl. Acad. Nauk SSSR* 56, 795–798.

The moment condition [finite variance of the number of offspring] was removed by Joffe (1967) and Heathcote, Seneta and Vere-Jones (1967).

Even earlier, Kolmogorov (1938) proved the convergence of the conditional mean number of individuals.

Origins

The finite state space case was dispensed with early on using Perron Frobenius Theory. The QSD exists uniquely and is the LCD (the same for all initial distributions):

J.N. Darroch and E. Seneta. (1965) On quasi-stationary distributions in absorbing discrete-time Markov chains. *J. Appl. Probab.* 2, 88–100.

J.N. Darroch and E. Seneta. (1967) On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probab.* 4, 192–196.

Everyone knows that ...

Digression. Suppose for the moment that S is irreducible.

- If a stationary distribution ($\pi P(t) = \pi$) exists, then it is unique.
- S is recurrent iff $mP(t) \leq m$ has a positive solution, unique up to constant multiples, which satisfies $mP(t) = m$ (m is called an invariant measure).
- $mP(t) = m$ implies $mQ = 0$. (Warning: $mQ = 0$ does not necessarily imply $mP(t) = m$.)
- $\pi Q \leq 0$ implies $\pi P(t) = \pi$ iff Q is regular (non-explosive), in which case $\pi Q = 0$ (π is called an equilibrium distribution).
- If S is recurrent, then S is positive recurrent iff the invariant measure m is finite ($\sum_j m_j < \infty$), in which case the limiting distribution is given by $\pi_i = m_i / \sum_j m_j$.

Open Problem 1

Develop a satisfactory theory of QSDs/LCDs

First major advance

Conditions for the existence of LCDs for countable-state Markov chains in discrete time.

> E. Seneta and D. Vere-Jones (1966) On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* 3, 403–434.
>
> D. Vere-Jones (1962) Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford* 13, 7–28.

The R-classification was introduced (R-recurrent, R-transient, R-null recurrent, R-positive recurrent).

First major advance

Seneta and Vere-Jones discovered that:

- QSDs may not be unique.
- Basically, R-positive recurrence (to be defined later) is sufficient for the existence of an LCD.
- R-positive recurrence is hard to check.
- R-positive recurrence is not necessary for the existence of an LCD.
- When LCDs exists, they may depend on the initial distribution (Galton-Watson branching process). Sufficient conditions were given for non-dependence.
- There are many other kinds of LCDs.

Continuous time

“Analogous” papers for the continuous-time case:

D. Vere-Jones (1969) Some limit theorems for evanescent processes. *Austral. J. Statist.* 11, 67–78.

J.F.C. Kingman (1963) The exponential decay of Markov transition probabilities. *Proc. London Math. Soc.* 13, 337–358.

The LCD can exist in the λ-transient case (continuous-time analogue of R-transient):

E. Seneta (1966) Quasi-stationary behaviour in the random walk with continuous time. *Austral. J. Statist.* 8, 92–98.

λ-classification

Following Vere-Jones (1962), Kingman (1963) proved that the limit

$$\lambda \; (= \lambda_C) \; = \lim_{t\to\infty} -\frac{1}{t} \log p_{ij}(t)$$

exists and is the same for all $i, j \in C$, where C is any irreducible class. This limit satisfies

- $0 \leq \lambda < \infty$,
- $p_{ii}(t) \leq e^{-\lambda t}$, $i \in C$, and indeed
- $p_{ij}(t) \leq M_{ij} e^{-\lambda t}$, $i, j \in C$, for suitable constants M_{ij}.

λ is called the *decay parameter* (of C).

λ-classification

The irreducible class C is said to be *λ-recurrent* if

$$\int_0^\infty e^{\lambda t} p_{ij}(t)\, dt = \infty$$

for some (and then all) $i, j \in C$. Otherwise, C is *λ-transient*.

A λ-recurrent class C is called *λ-null recurrent* if $e^{\lambda t} p_{ij}(t) \to 0$ as $t \to \infty$ for some (and then all) $i, j \in C$. It is called *λ-positive recurrent* if $e^{\lambda t} p_{ij}(t) \to m_{ij}$ (strictly positive constants) for some (and then all) $i, j \in C$.

λ-classification

For each irreducible class C there always exist λ-subinvariant quantities:

λ-subinvariant measure $m = (m_j, j \in C)$:

$$\sum_{i \in C} m_i p_{ij}(t) \le e^{-\lambda t} m_j, \quad j \in C,\ t \ge 0.$$

λ-subinvariant vector (function) $x = (x_j, j \in C)$:

$$\sum_{i \in C} p_{ji}(t) x_i \le e^{-\lambda t} x_j, \quad j \in C,\ t \ge 0.$$

These quantities are called *λ-invariant* if equality holds for some (and then all) $j \in C$.

λ-classification

The irreducible class C is λ-recurrent *iff* the λ-subinvariant measure m and the λ-subinvariant vector x are unique and λ-invariant:

$$\sum_{i\in C} m_i p_{ij}(t) = e^{-\lambda t} m_j, \quad \sum_{i\in C} p_{ji}(t) x_i = e^{-\lambda t} x_j.$$

If C is *λ-recurrent*, then it is λ-positive recurrent *iff* $\sum_{k\in C} m_k x_k < \infty$, in which case

$$\lim_{t\to\infty} e^{\lambda t} p_{ij}(t) = \frac{x_i m_j}{\sum_{k\in C} m_k x_k}, \quad i, j \in C.$$

Significance of λ-positivity

$$\begin{aligned} m_{ij}(t) &:= \Pr(X(t) = j | X(t) \in C, X(0) = i) \\ &= \frac{\Pr(X(t) = j | X(0) = i)}{\Pr(X(t) \in C | X(0) = i)} \\ &= \frac{p_{ij}(t)}{\sum_{k \in C} p_{ik}(t)} = \frac{e^{\lambda t} p_{ij}(t)}{\sum_{k \in C} e^{\lambda t} p_{ik}(t)} \end{aligned}$$

So, formally, $m_{ij}(t) \to \dfrac{x_i m_j}{\sum_{k \in C} x_i m_k} = \dfrac{m_j}{\sum_{k \in C} m_k}$.

In fact, if C is λ-positive recurrent, then, for each i, $m_{ij}(t) \to m_j / \sum_{k \in C} m_k$, where $m = (m_j, j \in C)$ is the (essentially unique) λ-invariant measure (with the interpretation that the limit is 0 if $\sum_{k \in C} m_k = \infty$).

Strong ratio limit property

Since $e^{\lambda t} p_{ij}(t) \to x_i m_j / \sum_{k \in C} m_k x_k$, $i, j \in C$, when C is λ-positive recurrent, we have that

$$\frac{p_{ij}(s+t)}{p_{kl}(t)} \to e^{-\lambda s} \frac{x_i m_j}{x_k m_l}, \quad i, j, k, l \in C, \; s \geq 0.$$

This *strong ratio limit property* may hold in the λ-null recurrent case (Orey (1961), Kingman and Orey (1964), Pruit (1965), Folkman and Port (1966), Papangelou (1967) and Kersting (1974, 1976), and even the λ-transient case (Kesten (1963)).

> F. Papangelou (1968) Strong ratio limits, R-recurrence and mixing properties of discrete parameter Markov processes. *Z. Wahrscheinlich -keitstheorie und Verw. Gebiete* 8, 259–297.

Strong ratio limit property

Note that λ-positive recurrence also implies

$$\frac{\Pr(T > t + s | X(0) = i)}{\Pr(T > t | X(0) = k)} = \frac{\sum_{j \in C} p_{ij}(s + t)}{\sum_{j \in C} p_{kj}(t)} \to e^{-\lambda s} \frac{x_i}{x_k},$$

where T is the time to absorption. This again holds more generally:

> S.D. Jacka, and G.O. Roberts (1995) Weak convergence of conditioned processes on a countable state space. *J. Appl. Probab.* 32, 902–916.

Key to conditioned *process limits*.

Summary

This is all very unsatisfactory.

- λ-classification is (surprisingly) not the key to the existence of a LCD.
- The SRLP appears to be the key (but illudes us).
- When the LCD exists for a given starting state, it is a (the?) λ-invariant probability measure for P.
- The LCD may depend on the initial distribution.
- We need criteria in terms of the transition rates Q, and criteria we can check.

Birth-death processes

The LCD exists for all starting states iff $\lambda > 0$ (just like Yaglom's Theorem):

> P. Good (1968) The limiting behaviour of transient birth and death processes conditioned on survival. *J. Austral. Math. Soc. Ser. B* 8, 716–722.

Unfortunately *wrong*! Corrected in:

> E.A. van Doorn (1991) Quasi-stationary distribut -ions and convergence to quasi-stationarity of birth -death processes. *Adv. Appl. Probab.* 23, 683–700.

But, unlike for branching processes, we do not know λ explicitly. The LCD was given by van Doorn in terms of the (Karlin and McGregor) orthogonal polynomials.

Fresh start

Recall the definitions of QSD and LCD in terms of the conditional state probabilities $m(t) = (m_j(t), j \in C)$:

$$m_j(t) = \Pr(X(t) = j \,|\, X(t) \in C) = \frac{\Pr(X(t) = j)}{\Pr(X(t) \in C)} = \frac{p_j(t)}{\sum_{k \in C} p_k(t)}$$

the chance of being in state j given that the process has not reached 0.

Definition. A distribution $m = (m_j, j \in C)$ satisfying $m(t) = m$ for all $t > 0$ is called a *quasi-stationary distribution* (QSD). If $m(t) \to m$ then m is called a *limiting-conditional distribution* (LCD).

Characterizing QSDs

Since a is the initial distribution (with $a_0 = 0$),

$$p_j(t) = \Pr(X(t) = j) = \sum_{i \in C} a_i p_{ij}(t), \quad j \in C,\ t > 0,$$

where $p_{ij}(t) = \Pr(X(t) = j | X(0) = i)$. Therefore, if m is a QSD, then

$$\sum_{i \in C} m_i p_{ij}(t) = p_j(t) = g(t) m_j, \quad j \in C,\ t > 0,$$

where $g(t) = \sum_{k \in C} p_k(t)$. It is easy to show that g satisfies: $g(s+t) = g(s)g(t)$, $s, t \geq 0$, and $0 < g(t) < 1$. Thus, $g(t) = e^{-\mu t}$, for some $\mu > 0$. The converse is also clearly true.

QSDs and μ-invariant measures

We have proved the following simple result:

Proposition. A probability distribution $m = (m_j, j \in C)$ is a QSD iff, for some $\mu > 0$, m is a μ-invariant measure, that is

$$\sum_{i \in C} m_i p_{ij}(t) = e^{-\mu t} m_j, \quad j \in C,\ t \geq 0. \tag{1}$$

Note that for (1) it is necessary that $\mu \leq \lambda$, where recall that λ is the decay parameter of C (Vere-Jones (1969)).

But, can we determine QSDs m from Q?

Characterization in terms of Q

Rewrite (1) as

$$\sum_{i \in C:\, i \neq j} m_i p_{ij}(t) = \big((1 - p_{jj}(t)) - (1 - e^{-\mu t})\big)\, m_j$$

and use the fact that q_{ij} is the right-hand derivative of $p_{ij}(\cdot)$ near 0. On dividing by t and letting $t \downarrow 0$, we get (formally)

$$\sum_{i \in C:\, i \neq j} m_i q_{ij} = (q_j - \mu) m_j, \quad j \in C,$$

or, equivalently,

$$\sum_{i \in C} m_i q_{ij} = -\mu m_j, \quad j \in C.$$

Characterization in terms of Q

Accordingly, we shall say that m is a μ-invariant measure for Q whenever

$$\sum_{i \in C} m_i q_{ij} = -\mu m_j, \quad j \in C.$$

Theorem. If m is a μ-invariant measure for P ($\mu > 0$), then m is a μ-invariant measure for Q.

R.L. Tweedie (1974) Some ergodic properties of the Feller minimal process. *Quart. J. Math. Oxford* 25, 485–495.

Corollary. If m is a QSD then, for some $\mu > 0$, m is a μ-invariant measure for Q.

Characterization in terms of Q

Is the converse true? Suppose that, for some $\mu > 0$, m is a μ-invariant measure for Q, that is

$$\sum_{i \in C} m_i q_{ij} = -\mu m_j, \quad j \in C.$$

Is m a μ-invariant measure for P? (So that if m is a probability measure, then m is a QSD).

Sum this equation over $j \in C$: we get (formally), in the case when m is a probability measure,

$$\sum_{i \in C} m_i q_{i0} = -\sum_{i \in C} m_i \sum_{j \in C} q_{ij} = -\sum_{j \in C} \sum_{i \in C} m_i q_{ij} = \mu \sum_{j \in C} m_j = \mu.$$

Characterization in terms of Q

Theorem. Let $m = (m_j, j \in C)$ be a probability distribution over C and suppose that m is a μ-invariant measure for Q. Then, $\mu \leq \sum_{j \in C} m_j q_{j0}$ with equality iff m is a QSD.

> P.K. Pollett (1995) The determination of quasi-stationary distributions directly from the transition rates of an absorbing Markov chain. *Math. Computer Modelling* 22, 279–287.

So, in order to determine QSDs we must solve

$$\sum_{i \in C} m_i q_{ij} = -\left(\sum_{k \in C} m_k q_{k0} \right) m_j, \quad j \in C.$$

Example: catastrophe process

Suppose that

$$\begin{aligned} q_{i,i+1} &= a\rho i, & i &\geq 0, \\ q_{i,i} &= -\rho i, & i &\geq 0, \\ q_{i,i-k} &= \rho i b_k, & i \geq 2,\ k &= 1, 2 \dots i-1, \\ q_{i,0} &= \rho i \textstyle\sum_{k=i}^{\infty} b_k, & i &\geq 1, \end{aligned}$$

where $\rho, a > 0$, $b_i > 0$ for at least one $i \geq 1$ and $a + \sum_{i=1}^{\infty} b_i = 1$. Jumps occur at a constant per-capita rate ρ and, at a jump time, a birth occurs with probability a, or otherwise a catastrophe occurs, the size of which is determined by the probabilities b_i, $i \geq 1$. Clearly, 0 is an absorbing state and $C = \{1, 2, \dots\}$ is an irreducible class.

Does Q admit a QSD?

Example: catastrophe process

On substituting the transition rates into the equations $\sum_{i \in C} m_i q_{ij} = -\mu m_j$, $j \in C$, we get:

$$-(\rho - \mu) m_1 + \sum_{k=2}^{\infty} k \rho b_{k-1} m_k = 0,$$

and, for $j \geq 2$,

$$(j-1)\rho a m_{j-1} - (j\rho - \mu) m_j + \sum_{k=j+1}^{\infty} k \rho b_{k-j} m_k = 0.$$

Example: catastrophe process

If we try a solution of the form $m_j = t^j$, the first equation tells us that $\mu = -\rho(f'(t) - 1)$, where

$$f(s) = a + \sum_{i \in C} b_i s^{i+1}, \qquad |s| \leq 1,$$

and, on substituting both of *these* quantities in the second equation, we find that $f(t) = t$. This latter equation has a unique solution σ on $[0, 1]$. Thus, by setting $t = \sigma$ we obtain a positive μ-invariant measure $m = (m_j, j \in C)$ for Q, which satisfies $\sum_{j \in C} m_j = 1$ whenever $\sigma < 1$.

The condition $\sigma < 1$ is satisfied only in the subcritical case, that is, when (the drift) $D = a - \sum_{i \in C} i b_i < 0$; this also guarantees that absorption occurs with probability 1.

Example: catastrophe process

Further, it is easy to show that $\sum_{i \in C} m_i q_{i0} = \mu$:

$$
\begin{aligned}
\sum_{i \in C} m_i q_{i0} &= \sum_{i=1}^{\infty} (1-\sigma)\sigma^{i-1} \rho i \sum_{k=i}^{\infty} b_k \\
&= \rho \sum_{k=1}^{\infty} b_k \sum_{i=1}^{k} (1-\sigma) i \sigma^{i-1} \\
&\vdots \\
&= \rho(1 - f'(\sigma)) = \mu.
\end{aligned}
$$

Proposition. (Pakes (1987)) The subcritical birth-death and catastrophe process has a *geometric* QSD $m = (m_j,\ j \in C)$. This is given by $m_j = (1-\sigma)\sigma^{j-1}$, $j \in C$, where σ is the unique solution to $f(t) = t$ on the interval $[0, 1]$.

Characterization in terms of Q

There are more general necessary and sufficient conditions.

P.K. Pollett (1986) On the equivalence of μ-invariant measures for the minimal process and its q-matrix. *Stochastic Process. Appl.* 22, 203–221.

Theorem. A μ-invariant measure m for Q is μ-invariant for P *iff* the equations

$$\sum_{i \in C} y_i q_{ij} = \nu y_j, \quad 0 \leq y_j \leq m_j, \; j \in C,$$

have no non-trivial solution for some (and then for all) $\nu > -\mu$.

Characterization in terms of Q

Because of the similarity with Reuter's (1957) condition for there to be a unique solution to the forward equations, we came up with the following (Hart and Pollett (1996)):

Corollary. (*The Reuter FE Condition*) If the equations

$$\sum_{i \in C} y_i q_{ij} = \nu y_j, \quad j \in C,$$

have no non-trivial, non-negative solution such that $\sum_{j \in C} y_j < \infty$, for some (and then for all) $\nu > 0$, then *all* μ-invariant probability measures for Q are μ-invariant measures (and hence QSDs) for P.

Birth-death processes

We are given strictly positive birth rates $(\lambda_j, j \in C)$ ($\lambda_0 = 0$ since 0 is absorbing) and death rates $(\mu_j, j \in C)$. We have assumed that absorption occurs with probability 1, that is,

$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i \pi_i} = \infty,$$

where $\pi_1 = 1$ and, for $j \geq 2$,

$$\pi_j = \prod_{i=2}^{j} \frac{\lambda_{i-1}}{\mu_i}.$$

Birth-death processes

Erik van Doorn proved the following characterization, extending early work of Cavender (1978), in terms of

$$D := \sum_{i=1}^{\infty} \frac{1}{\lambda_i \pi_i} \sum_{j=i+1}^{\infty} \pi_j .$$

Theorem. If $D = \infty$, then either $\lambda = 0$ and there is no QSD, or $\lambda > 0$ and there is a *one-parameter family* of QSDs, being the normalized μ-invariant measures $(0 < \mu \leq \lambda)$. If $D < \infty$, then $\lambda > 0$ and there is *exactly one* QSD, being the normalized λ-invariant measure.

E.A. van Doorn (1991) Quasi-stationary distribut
-ions and convergence to quasi-stationarity of birth
-death processes. *Adv. Appl. Probab.* 23, 683–700.

Existence of QSDs

Since $D = \infty$ (for birth-death processes) is *arithmetically equivalent* to the Reuter FE condition, we have the following conjecture for our absorbing Markov chain:

Conjecture. If the Reuter FE Condition holds, then either $\lambda = 0$ and there is no QSD, or $\lambda > 0$ and there is a one-parameter family of QSDs. If the Reuter FE Condition fails, then $\lambda > 0$ and there is *exactly one* QSD.

Conjecture. If $\lim_{i\to\infty} \mathrm{E}(T|X(0) = i) = \infty$, then either $\lambda = 0$ and there is no QSD, or $\lambda > 0$ and there is a one-parameter family of QSDs. If $\lim_{i\to\infty} \mathrm{E}(T|X(0) = i) < \infty$, then $\lambda > 0$ and there is *exactly one* QSD.

Conjecture. The above conjectures are the same!

Existence of QSDs and LCDs

Suppose that $\mathrm{E}(T|X(0)=i)<\infty$. The existence of a QSD is guaranteed under *asymptotic remoteness* (AR): $\lim_{i\to\infty}\Pr(T>t|X(0)=i)=1$.

> P. Ferrari, H. Kesten, S. Martínez, and P. Picco (1995) Existence of quasi-stationary distributions. A renewal dynamic approach. *Ann. Probab.* 23, 501–521.

Theorem. Under AR a QSD exists iff $\lambda>0$. If Q is bounded ($\sup_i q_i<\infty$), the LCD exists.

(Note the absence of λ-classification in the latter.)

But (Pakes (1995)), *AR can be arbitrarily badly violated*: even $\lim_{i\to\infty}\Pr(T>t|X(0)=i)=0$.

Killer blow

Recall the Good (1968)–Van Doorn (1991) result for birth-death process and Yaglom's theorem for branching processes: the LCD exists iff $\lambda > 0$. And, for Q bounded AR is sufficient (Ferrari, et al. (1995)).

> H. Kesten (1995) A ratio limit theorem for (sub) Markov chains on $\{1, 2, \ldots\}$ with bounded jumps. *Adv. Appl. Probab.* 27, 652–691.

Suppose Q is bounded ($\sup_i q_i < \infty$). If the chain has *bounded jumps* and satisfies a *uniform irreducibility condition*, then there is at most one QSD, and the LCD exists iff $\lambda > 0$.

Kesten also provided an example (25 journal pages) for which a QSD exists, but the LCD does not exist.

Open problems

- Obtain necessary and sufficient conditions in terms of Q for the SRLP to hold and for the LCD to exist. (Kesten's result and example suggest that this might be difficult.)
- Does Kesten's result hold when $\sup_i q_i = \infty$.
- Obtain workable sufficient conditions in terms of Q.
- Solve the domain of attraction problem. Whilst the answer is known for branching processes, and several examples, it is not known for birth-death processes.
- Obtain necessary and sufficient conditions for $\lambda > 0$ for various models (solved recently for birth-death processes by Hanjun Zhang).
- Numerical methods: truncation procedures, and the GTH algorithm for dominant eigensolutions.

Barlett's idea

When the process hits 0, send it back:

> M.S. Bartlett (1960) *Stochastic Population Models in Ecology and Epidemiology*, Methuen, London.

In continuous time, we send it back instantaneously. Let $\nu = (\nu_j, j \in C)$ be a probability measure on C, and define Q^ν by $q^\nu_{ij} = q_{ij} + q_{i0}\nu_j$, $i, j \in C$. This is a stable and conservative q-matrix over Q. Note, in particular, that $\sum_{j \in C} q^\nu_{ij} = \sum_{j \in C} q_{ij} + q_{i0} = 0$. Indeed, the Q^ν process is (irreducible and) recurrent, since originally absorption occurred with probability 1; if $\mathrm{E}(T | X(0) = i) < \infty$, then it is positive recurrent. Let π be its equilibrium distribution ($\pi Q^\nu = 0$). Barlett's idea was to use π to model the long-term behaviour of the original process.

Barlett’s idea

Define a map Φ as follows. If ν is the “resurrection measure”, and π is the equilibrium distribution of Q^ν, let $\pi = \Phi(\nu)$.

Observation. Any QSD m satisfying $\mathrm{E}_m(T) < \infty$ is a fixed point of this map ($m = \Phi(m)$ means $mQ = -\mu m$, where $\mu = \sum_{i \in C} m_i q_{i0}$).

This was exploited by Ferrari, et al. (1995), and also by Clancy and Pollett (2003). We exhibited the map explicitly for birth-death processes and showed that it preserves likelihood-ratio ordering (and hence stochastic ordering). So, for example, if $\nu^{(1)} \prec^{LR} m \prec^{LR} \nu^{(2)}$, then $\Phi^n(\nu^{(1)}) \prec^{LR} m \prec^{LR} \Phi^n(\nu^{(2)})$. So, bounds on m can be obtained.

When absorption is not certain

If $\alpha_i := \lim_{t\to\infty} p_{i0}(t) < 1$, we employ an h-transform: define transition probabilities $\bar{P}(t) = \{\bar{p}_{ij}(t), i, j \in S\}$ by $\bar{p}_{ij}(t) = p_{ij}(t)\alpha_j/\alpha_i$, and corresponding transition rates $\bar{Q} = \{\bar{q}_{ij}, i, j \in S\}$ by $\bar{q}_{ij} = q_{ij}\alpha_j/\alpha_i$, Then, in an obvious notation, $\bar{\mathcal{P}}(A) = 1$ and $\mathcal{P}(\,\cdot\,|A) = \bar{\mathcal{P}}(\cdot)$, where A is the event $\{$absorption eventually occurs$\}$. This result can be traced back to Waugh (1958).

Now just reinterpret any given result for $\bar{P}$. For example, if C is λ-positive recurrent and $\sum_{i\in C} m_i\alpha_i < \infty$, then

$$\lim_{t\to\infty} \Pr(X(t) = j \,|\, X(t) \in C, A) = \frac{m_j\alpha_j}{\sum_{k\in S} m_k\alpha_k},$$

where m is the essentially unique λ-invariant measure for P.