Quasi stationarity

Phil. Pollett

The University of Queensland

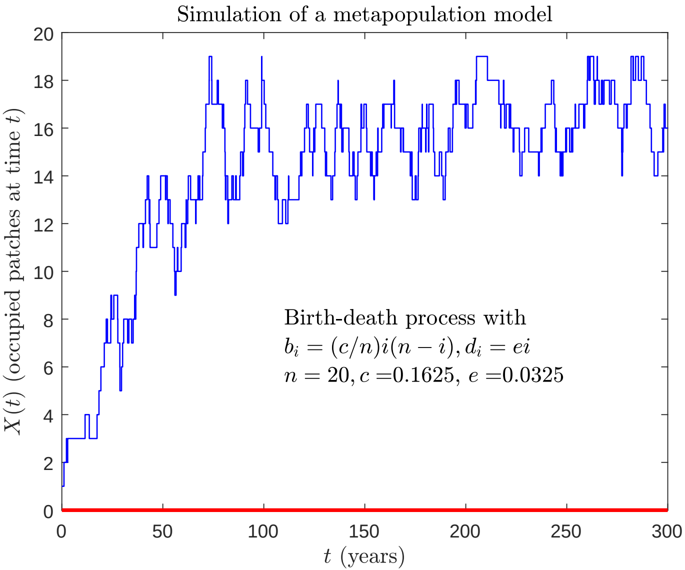
10th International Conference on Matrix-Analytic Methods in Stochastic Models

University of Tasmania, 13 February 2019

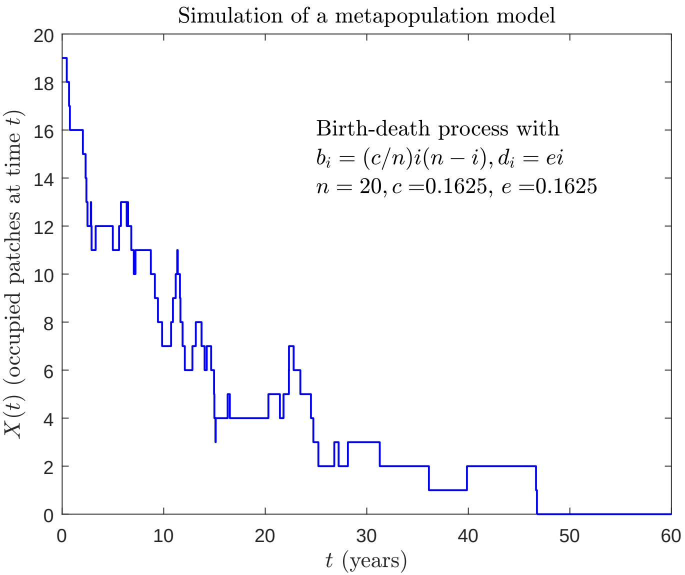
ACEMS

AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR MATHEMATICAL AND STATISTICAL FRONTIERS

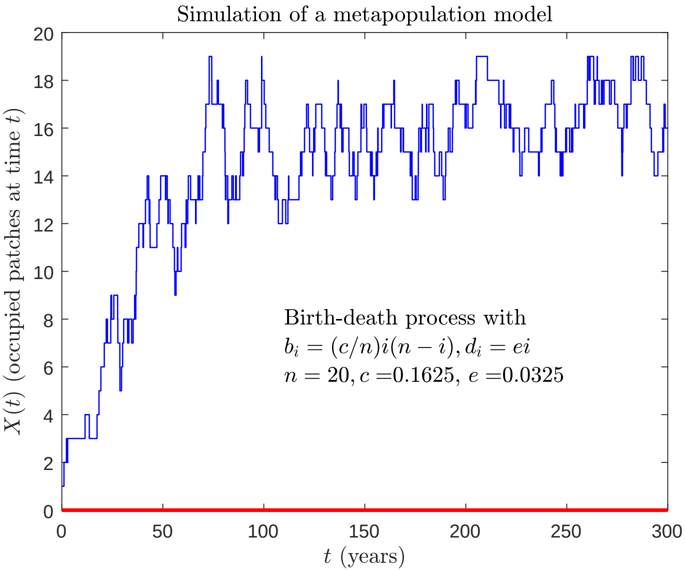
Quasi stationarity



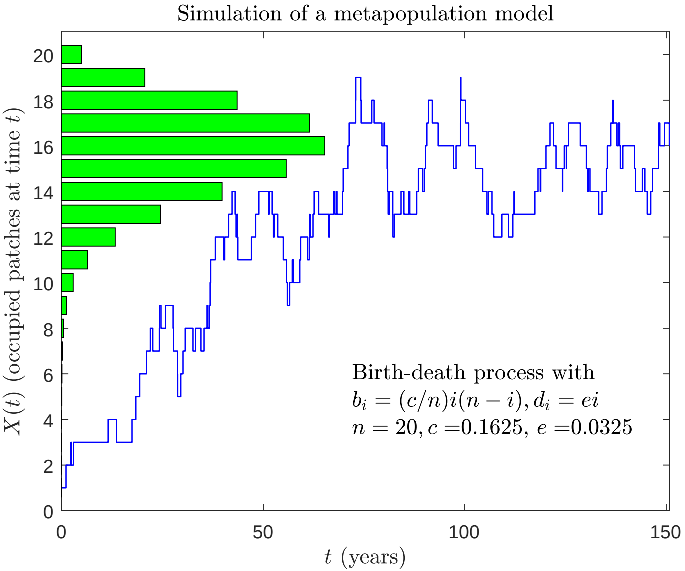
Evanescence



Quasi stationarity



A "quasi-stationary" distribution



A “quasi-stationary” distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied.

A "quasi-stationary" distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied. What is the chance of there being precisely i patches occupied?

A "quasi-stationary" distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied. What is the chance of there being precisely i patches occupied?

If we were equipped with the full set of state probabilities $p_i(t) = \mathbb{P}(X(t) = i)$, $i = 0, 1, \ldots, n$, we would evaluate the *conditional probability*

$$u_i(t) = \mathbb{P}(X(t) = i \mid X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)},$$

for i in the set $C = \{1, \ldots, n\}$ of transient states.

A "quasi-stationary" distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied. What is the chance of there being precisely i patches occupied?

If we were equipped with the full set of state probabilities $p_i(t) = \mathbb{P}(X(t) = i)$, $i = 0, 1, \ldots, n$, we would evaluate the *conditional probability*

$$u_i(t) = \mathbb{P}(X(t) = i \mid X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)},$$

for i in the set $C = \{1, \ldots, n\}$ of transient states.

Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\boldsymbol{u} = (u_i,\, i \in C)$ over C such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

A "quasi-stationary" distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied. What is the chance of there being precisely i patches occupied?

If we were equipped with the full set of state probabilities $p_i(t) = \mathbb{P}(X(t) = i)$, $i = 0, 1, \ldots, n$, we would evaluate the *conditional probability*

$$u_i(t) = \mathbb{P}(X(t) = i \mid X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)},$$

for i in the set $C = \{1, \ldots, n\}$ of transient states.

Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\boldsymbol{u} = (u_i,\ i \in C)$ over C such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

Such a distribution is called a *stationary conditional distribution* or *quasi-stationary distribution* (QSD).

A "quasi-stationary" distribution

Think of an observer who at some time t is *aware of the occupancy of some patches*, yet cannot tell exactly which of n patches are occupied. What is the chance of there being precisely i patches occupied?

If we were equipped with the full set of state probabilities $p_i(t) = \mathbb{P}(X(t) = i)$, $i = 0, 1, \ldots, n$, we would evaluate the *conditional probability*

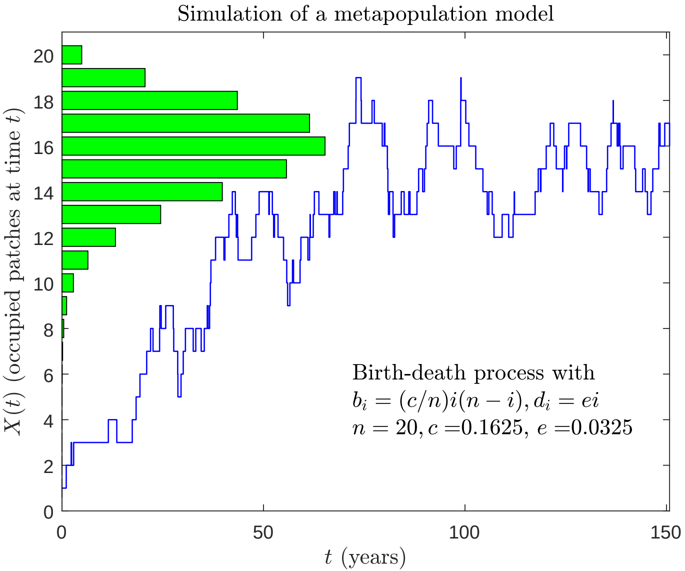
$$u_i(t) = \mathbb{P}(X(t) = i \,|\, X(t) \neq 0) = \frac{p_i(t)}{1 - p_0(t)},$$

for i in the set $C = \{1, \ldots, n\}$ of transient states.

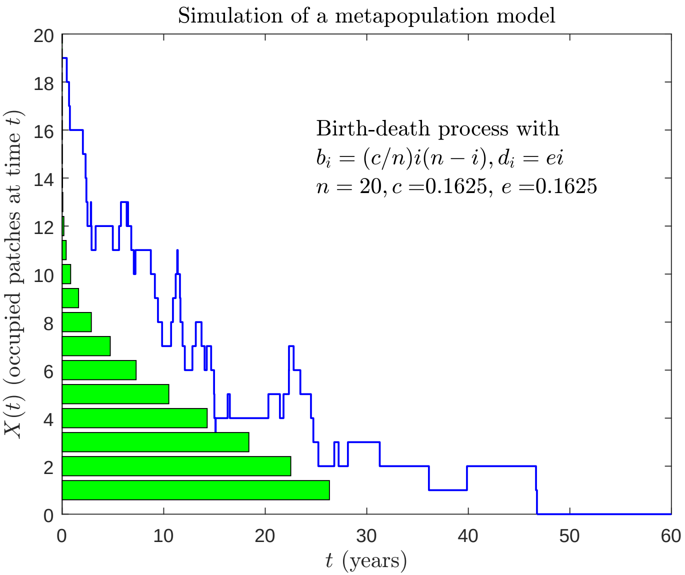
Then, in view of the behaviour observed in our simulation, it would be natural for us to seek a distribution $\boldsymbol{u} = (u_i,\, i \in C)$ over C such that if $u_i(t) = u_i$ for a particular $t > 0$, then $u_i(s) = u_i$ for all $s > t$.

Such a distribution is called a *stationary conditional distribution* or *quasi-stationary distribution* (QSD). It might then also be a *limiting conditional distribution* (LCD) in that $u_i(t) \to u_i$ as $t \to \infty$.

The QSD for the n-patch metapopulation model



Evanescence - yes, there *is* a QSD



Quasi-stationary distributions

Consider the setting of a *non-explosive* continuous-time Markov chain $X = (X(t),\ t \geq 0)$ whose state space consists of a communicating class C and an absorbing state 0 which is accessible from C: indeed *reached with probability* 1.

Quasi-stationary distributions

Consider the setting of a *non-explosive* continuous-time Markov chain $X = (X(t),\ t \geq 0)$ whose state space consists of a communicating class C and an absorbing state 0 which is accessible from C: indeed *reached with probability* 1.

It is easy to prove that a distribution $\boldsymbol{u} = (u_i,\ i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

where $P(t) = (p_{ij}(t),\ i, j \in C)$ are the transition probabilities restricted to C.

Consider the setting of a *non-explosive* continuous-time Markov chain $X = (X(t),\ t \geq 0)$ whose state space consists of a communicating class C and an absorbing state 0 which is accessible from C: indeed *reached with probability* 1.

It is easy to prove that a distribution $\boldsymbol{u} = (u_i,\ i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

where $P(t) = (p_{ij}(t),\ i,j \in C)$ are the transition probabilities restricted to C.

On differentiating both sides near $t = 0$ we get

$$\boldsymbol{u}Q = -\mu\boldsymbol{u}, \tag{2}$$

where $Q = (q_{ij},\ i,j \in C)$ is the q-matrix of transition rates (again restricted to C).

Quasi-stationary distributions

Consider the setting of a *non-explosive* continuous-time Markov chain $X = (X(t),\ t \geq 0)$ whose state space consists of a communicating class C and an absorbing state 0 which is accessible from C: indeed *reached with probability* 1.

It is easy to prove that a distribution $\boldsymbol{u} = (u_i,\ i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

where $P(t) = (p_{ij}(t),\ i, j \in C)$ are the transition probabilities restricted to C.

On differentiating both sides near $t = 0$ we get

$$\boldsymbol{u}Q = -\mu\boldsymbol{u}, \tag{2}$$

where $Q = (q_{ij},\ i, j \in C)$ is the q-matrix of transition rates (again restricted to C).

This can be justified:

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

Quasi-stationary distributions

A distribution $\boldsymbol{u} = (u_i,\ i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

in which case

$$\boldsymbol{u}Q = -\mu\boldsymbol{u}. \tag{2}$$

The (more interesting) converse*, when a distribution $\boldsymbol{u}$ satisfying (2) also satisfies (1) (and hence is a QSD), happens *iff* $\mu = \sum_{i \in C} u_i q_{i0}$.

* Pollett, P.K., and Vere-Jones, D. (1992) A note on evanescent processes. Aust. J. Statist. 34, 531–536.

Quasi-stationary distributions

A distribution $\boldsymbol{u} = (u_i,\, i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

in which case

$$\boldsymbol{u}Q = -\mu \boldsymbol{u}. \tag{2}$$

The (more interesting) converse*, when a distribution $\boldsymbol{u}$ satisfying (2) also satisfies (1) (and hence is a QSD), happens *iff* $\mu = \sum_{i \in C} u_i q_{i0}$.

* Pollett, P.K., and Vere-Jones, D. (1992) A note on evanescent processes. Aust. J. Statist. 34, 531–536.

Note that (1) entails

$$\mathbb{P}_{\boldsymbol{u}}(X(t) \in C) = e^{-\mu t}.$$

Quasi-stationary distributions

A distribution $\boldsymbol{u} = (u_i,\, i \in C)$ is a QSD *iff*, for some $\mu > 0$,

$$\boldsymbol{u}P(t) = e^{-\mu t}\boldsymbol{u}, \qquad t > 0, \tag{1}$$

in which case

$$\boldsymbol{u}Q = -\mu\boldsymbol{u}. \tag{2}$$

The (more interesting) converse*, when a distribution $\boldsymbol{u}$ satisfying (2) also satisfies (1) (and hence is a QSD), happens *iff* $\mu = \sum_{i\in C} u_i q_{i0}$.

* Pollett, P.K., and Vere-Jones, D. (1992) A note on evanescent processes. Aust. J. Statist. 34, 531–536.

Note that (1) entails

$$\mathbb{P}_{\boldsymbol{u}}(X(t) \in C) = e^{-\mu t}.$$

There is a maximum decay rate†, and the corresponding QSD is termed "extremal".

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

Limiting conditional distributions

As for when a QSD is also a LCD, the picture is complete for irreducible (C irreducible) *finite-state* Markov chains: the QSD $\boldsymbol{u}$ is unique, and, *for all initial distributions* $\boldsymbol{w} = (w_i,\ i \in C)$, $\boldsymbol{u}$ is the LCD.

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

Limiting conditional distributions

As for when a QSD is also a LCD, the picture is complete for irreducible (C irreducible) *finite-state* Markov chains: the QSD $\boldsymbol{u}$ is unique, and, *for all initial distributions* $\boldsymbol{w} = (w_i,\ i \in C)$, $\boldsymbol{u}$ is the LCD.

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

The same can be said for *reducible* finite-state Markov chains, but we need to take into communicating class ordering an accessibility under $\boldsymbol{w}$.

Van Doorn, E.A., and Pollett, P.K. (2013) Quasi-stationary distributions for discrete-state models. European Journal of Operational Research 230, 1–14.

Limiting conditional distributions

As for when a QSD is also a LCD, the picture is complete for irreducible (C irreducible) *finite-state* Markov chains: the QSD $\boldsymbol{u}$ is unique, and, *for all initial distributions* $\boldsymbol{w} = (w_i,\ i \in C)$, $\boldsymbol{u}$ is the LCD.

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

The same can be said for *reducible* finite-state Markov chains, but we need to take into communicating class ordering an accessibility under $\boldsymbol{w}$.

Van Doorn, E.A., and Pollett, P.K. (2013) Quasi-stationary distributions for discrete-state models. European Journal of Operational Research 230, 1–14.

Otherwise, the picture it is incomplete/unsatisfactory, focussing on the notion of λ-positive recurrence, which is *difficult to check* (especially from Q), and, anyway, is *not necessary* for the existence of an LCD.

The picture is also complete for specific models, such as *birth-death processes*:

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

*Kijima, M., Nair, M.G., Pollett, P.K., and van Doorn, E. (1997) Limiting conditional distributions for birth-death processes. Adv. Appl. Probab. 29, 185–204.

Conditions are given to delineate three possible cases:

(i) no QSD, and $u_i(t) \to 0$ (fixed initial state i).

(ii) a unique QSD $\boldsymbol{u}$, and $u_i(t) \to u_i$ (fixed initial state i).

(iii) a one-parameter family of QSDs, and we get convergence (again for fixed initial state) to the *extremal* QSD.

Domain of attraction problem

Let $\boldsymbol{u} = (u_i,\, i \in C)$ be a given QSD. If $\boldsymbol{u}$ is a LCD for some initial distribution $\boldsymbol{w} = (w_i,\ i \in C)$, that is

$$\lim_{t\to\infty} \mathbb{P}_{\boldsymbol{w}}(X(t) = j \,|\, X(t) \neq 0) = u_j, \quad j \in C,$$

we say that $\boldsymbol{w}$ is *in the domain of attraction of* $\boldsymbol{u}$.

Let $\boldsymbol{u} = (u_i,\, i \in C)$ be a given QSD. If $\boldsymbol{u}$ is a LCD for some initial distribution $\boldsymbol{w} = (w_i,\ i \in C)$, that is

$$\lim_{t\to\infty} \mathbb{P}_{\boldsymbol{w}}(X(t) = j \,|\, X(t) \neq 0) = u_j, \quad j \in C,$$

we say that $\boldsymbol{w}$ is *in the domain of attraction of* $\boldsymbol{u}$.

Exercise 1 : Identify the domains of attraction.

The Yaglom limit

Yaglom* was the first to identify explicitly a LCD, establishing the existence of such for the subcritical Bienaymé-Galton-Watson branching process (Heathcote, Seneta, and Vere-Jones† relaxed the condition the variance of number of offspring be finite).

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

† Heathcote, C.R., Seneta, E., and Vere-Jones, D. (1967) A refinement of two theorems in the theory of branching processes. Teor. Verojatnost. i Primenen. 12, 341–346; Theory Probab. Appl. 12, 297–301.

Theorem If the expected number m of offspring is less than 1, then

$$u_i = \lim_{n\to\infty} \mathbb{P}(X_n = i | X_n \neq 0, X_0 = 1), \qquad i \in C,$$

exists and defines a proper probability distribution $\boldsymbol{u} = (u_i,\, i \in C)$ over C.

Origins of the idea

The idea of a limiting conditional distribution goes back further than Yaglom, at least to Wright* in his discussion of gene frequencies in finite populations:

> "As time goes on, divergences in the frequencies of factors may be expected to increase more and more until at last some are either completely fixed or completely lost from the population. The distribution curve of gene frequencies should, however, approach a definite form if the genes which have been wholly fixed or lost are left out of consideration."

*Wright, S. (1931) Evolution in Mendelian populations. Genetics 16, 97–159.

Origins of the idea

The idea of a limiting conditional distribution goes back further than Yaglom, at least to Wright* in his discussion of gene frequencies in finite populations:

> "As time goes on, divergences in the frequencies of factors may be expected to increase more and more until at last some are either completely fixed or completely lost from the population. The distribution curve of gene frequencies should, however, approach a definite form if the genes which have been wholly fixed or lost are left out of consideration."

*Wright, S. (1931) Evolution in Mendelian populations. Genetics 16, 97–159.

The idea of "quasi stationarity" was crystallized by Bartlett†:

> "While presumably on the above model [for the interactions between active and passive forms of flour beetle] extinction of the population will occur after a long enough time, this may (for a deterministic 'ceiling' population not too small, but fluctuations relatively small) be so long delayed as to be negligible and an effective or quasi stationarity be established."

†Bartlett, M.S. (1957) On theoretical models for competitive and predatory biological systems. Biometrika 44, 27–42.

Origins of the idea

Bartlett* later coined the term "quasi-stationary distribution":

> "It may still happen that the time to extinction is so long that it is still of more relevance to consider the effectively ultimate distribution (called a 'quasi-stationary' distribution) of [the population process] N."

* Bartlett, M.S. (1960) Stochastic Population Models in Ecology and Epidemiology. Methuen, London. [Page 24]

Origins of the idea

Bartlett* later coined the term "quasi-stationary distribution":

> "It may still happen that the time to extinction is so long that it is still of more relevance to consider the effectively ultimate distribution (called a 'quasi-stationary' distribution) of [the population process] N."

* Bartlett, M.S. (1960) Stochastic Population Models in Ecology and Epidemiology. Methuen, London. [Page 24]

He then outlined an approach to modelling quasi stationarity whereby the process is *resurrected* after extinction, and the resulting stationary distribution evaluated (extinction in finite mean time implies that this "return process" is positive recurrent).

Origins of the idea

Bartlett* later coined the term "quasi-stationary distribution":

> "It may still happen that the time to extinction is so long that it is still of more relevance to consider the effectively ultimate distribution (called a 'quasi-stationary' distribution) of [the population process] N."

* Bartlett, M.S. (1960) Stochastic Population Models in Ecology and Epidemiology. Methuen, London. [Page 24]

He then outlined an approach to modelling quasi stationarity whereby the process is *resurrected* after extinction, and the resulting stationary distribution evaluated (extinction in finite mean time implies that this "return process" is positive recurrent).

Ewens† coined term *return process*, and *pseudo-transient distribution*:

† Ewens, W.J. (1963) The diffusion equation and a pseudo-distribution in genetics. J. Royal Statist. Soc., Ser. B 25, 405–412.

† Ewens, W.J. (1964) The pseudo-transient distribution and its uses in genetics. J. Appl. Probab. 1, 141–156.

Origins of the idea

Bartlett* later coined the term “quasi-stationary distribution”:

> “It may still happen that the time to extinction is so long that it is still of more relevance to consider the effectively ultimate distribution (called a ‘quasi-stationary’ distribution) of [the population process] N.”

* Bartlett, M.S. (1960) Stochastic Population Models in Ecology and Epidemiology. Methuen, London. [Page 24]

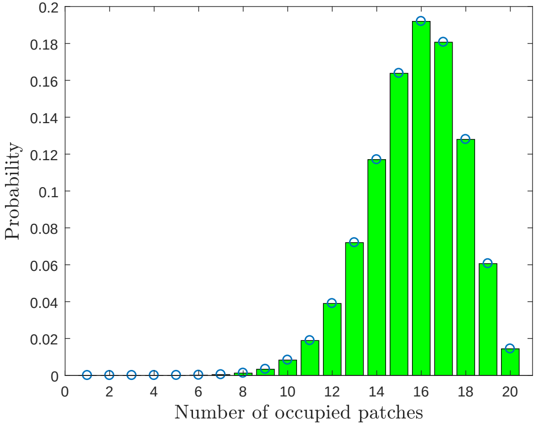
He then outlined an approach to modelling quasi stationarity whereby the process is *resurrected* after extinction, and the resulting stationary distribution evaluated (extinction in finite mean time implies that this “return process” is positive recurrent).

Darroch and Seneta† had reservations:

> “. . . we mention another objection to it, namely that $a(\alpha)$ [the pseudo-transient distribution] depends on α [the resurrection law] to such an extent that it can be made into almost any distribution over T [C here] by suitable choice of α.”

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

However, these distributions appear to be “close”



The quasi-stationary distribution (bars), and the pseudo-transient distribution when the resurrection law assigns all its mass to state 1 (circles), for the earlier 20-patch metapopulation model.

Define the *return process* $X^{\boldsymbol{\nu}}$ with state space C to have exactly the same behaviour as X while in C, but, on reaching 0, to be returned instantly to C according to a distribution $\boldsymbol{\nu} = (\nu_i,\, i \in C)$. Thus, its q-matrix $Q^{\boldsymbol{\nu}}$ will have entries

$$q_{ij}^{\boldsymbol{\nu}} = q_{ij} + q_{i0}\nu_j, \qquad i, j \in C.$$

Under extra the condition that X is absorbed in finite mean time, $X^{\boldsymbol{\nu}}$ will have a stationary distribution $\boldsymbol{\pi}^{\boldsymbol{\nu}}$.

The return process and return map

Define the *return process* $X^{\boldsymbol{\nu}}$ with state space C to have exactly the same behaviour as X while in C, but, on reaching 0, to be returned instantly to C according to a distribution $\boldsymbol{\nu} = (\nu_i,\ i \in C)$. Thus, its q-matrix $Q^{\boldsymbol{\nu}}$ will have entries

$$q_{ij}^{\boldsymbol{\nu}} = q_{ij} + q_{i0}\nu_j, \qquad i, j \in C.$$

Under extra the condition that X is absorbed in finite mean time, $X^{\boldsymbol{\nu}}$ will have a stationary distribution $\boldsymbol{\pi}^{\boldsymbol{\nu}}$.

The map $\Phi : \boldsymbol{\nu} \mapsto \boldsymbol{\pi}^{\boldsymbol{\nu}}$ is called the *return map*, and it is clear that any QSD $\boldsymbol{u}$ is is a fixed point of Φ (that is, $\boldsymbol{u} = \boldsymbol{\pi}^{\boldsymbol{u}}$), because

$$\sum_{i \in C} u_i q_{ij}^{\boldsymbol{u}} = 0 \quad \textit{iff} \quad \sum_{i \in C} u_i q_{ij} = -\mu u_j, \text{ where } \mu = \sum_{i \in C} u_i q_{i0}.$$

The return process and return map

Define the *return process* $X^{\boldsymbol{\nu}}$ with state space C to have exactly the same behaviour as X while in C, but, on reaching 0, to be returned instantly to C according to a distribution $\boldsymbol{\nu} = (\nu_i,\ i \in C)$. Thus, its q-matrix $Q^{\boldsymbol{\nu}}$ will have entries

$$q_{ij}^{\boldsymbol{\nu}} = q_{ij} + q_{i0}\nu_j, \qquad i, j \in C.$$

Under extra the condition that X is absorbed in finite mean time, $X^{\boldsymbol{\nu}}$ will have a stationary distribution $\boldsymbol{\pi}^{\boldsymbol{\nu}}$.

The map $\Phi : \boldsymbol{\nu} \mapsto \boldsymbol{\pi}^{\boldsymbol{\nu}}$ is called the *return map*, and it is clear that any QSD $\boldsymbol{u}$ is is a fixed point of Φ (that is, $\boldsymbol{u} = \boldsymbol{\pi}^{\boldsymbol{u}}$), because

$$\sum_{i \in C} u_i q_{ij}^{\boldsymbol{u}} = 0 \quad \textit{iff} \quad \sum_{i \in C} u_i q_{ij} = -\mu u_j, \text{ where } \mu = \sum_{i \in C} u_i q_{i0}.$$

Under mild conditions the return map is contractive, and iteration leads us to the extremal QSD*.

* Ferrari, P.A., Kesten, H., Martínez, S., and Picco, P. (1995) Existence of quasistationary distributions: a renewal dynamical approach. Ann. Probab. 23, 501-521.

The QSD and π^{ν} can be close

Write

$$d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{v}) := \sup_{A \subseteq C} |\boldsymbol{u}\{A\} - \boldsymbol{v}\{A\}| = \frac{1}{2} \sum_{k \in C} |u_k - b_k|$$

for *total variation distance* between two probability measures, $\boldsymbol{u} = (u_i,\, i \in C)$ and $\boldsymbol{v} = (v_i,\, i \in C)$, on C.

Under mild conditions, X has a unique QSD $\boldsymbol{u}$, and, *for any* probability measure $\boldsymbol{\nu}$ on C, we have $d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{\pi}^{\nu}) \leq B$. The bound is expressed solely in terms of hitting probabilities and expected hitting times of X.

Barbour, A.D., and Pollett, P.K. (2010) Total variation approximation for quasi-equilibrium distributions. J. Appl. Probab. 47, 934–946.

We argue that the bound is expected to be small *when the process spends a long time in quasi equilibrium*.

The QSD and π^{ν} can be close

Write

$$d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{v}) := \sup_{A \subseteq C} |\boldsymbol{u}\{A\} - \boldsymbol{v}\{A\}| = \frac{1}{2} \sum_{k \in C} |u_k - b_k|$$

for *total variation distance* between two probability measures, $\boldsymbol{u} = (u_i,\ i \in C)$ and $\boldsymbol{v} = (v_i,\ i \in C)$, on C.

Under mild conditions, X has a unique QSD $\boldsymbol{u}$, and, *for any* probability measure $\boldsymbol{\nu}$ on C, we have $d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{\pi}^{\nu}) \leq B$. The bound is expressed solely in terms of hitting probabilities and expected hitting times of X.

Barbour, A.D., and Pollett, P.K. (2010) Total variation approximation for quasi-equilibrium distributions. J. Appl. Probab. 47, 934–946.

We argue that the bound is expected to be small *when the process spends a long time in quasi equilibrium*.

For example, in the n-patch metapopulation model, the bound is *geometrically small* in n.

The QSD and π^ν can be close - precise statement

Write

$$\tau_A \; := \; \inf\{t > 0 : X(t) \in A, \, X(s) \notin A \text{ for some } s < t\}$$

for the first entrance time of a set A.

Theorem Suppose that there exist $s \in C$, $p > 0$, and $T < \infty$, such that

(i) $\mathbb{P}_k(X \text{ hits } s \text{ before } 0) \geq p$;

(ii) $\mathbb{E}_k(\tau_{\{s,0\}}) \leq T$,

uniformly for all $k \in C$, and suppose that $2UT/p < 1$, where

$$U = \sum_{k \in C} \frac{q_{k0}}{q_k \mathbb{E}_k(\tau_{\{k,0\}})} \qquad (q_k = -q_{kk}).$$

Then, X has a unique QSD $\boldsymbol{u}$, and, for any probability measure $\boldsymbol{\nu}$ on C,

$$d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{\pi}^{\boldsymbol{\nu}}) \leq 2UT/p.$$

ACEMS

The QSD and $\boldsymbol{\pi}^{\boldsymbol{\nu}}$ can be close - precise statement

Write

$$\tau_A \;:=\; \inf\{t > 0 : X(t) \in A,\ X(s) \notin A \text{ for some } s < t\}$$

for the first entrance time of a set A.

Theorem Suppose that there exist $s \in C$, $p > 0$, and $T < \infty$, such that

(i) $\mathbb{P}_k(X \text{ hits } s \text{ before } 0) \geq p$;

(ii) $\mathbb{E}_k(\tau_{\{s,0\}}) \leq T$,

uniformly for all $k \in C$, and suppose that $2UT/p < 1$, where

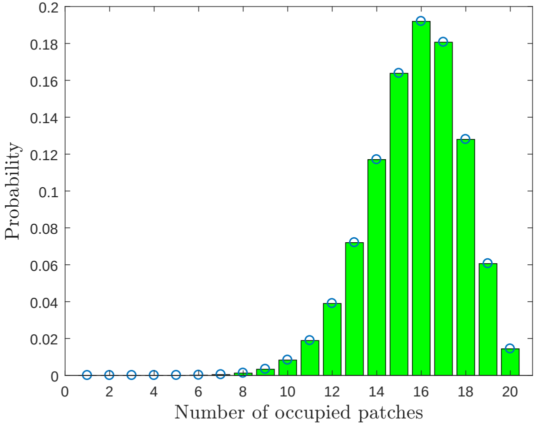
$$U = \sum_{k \in C} \frac{q_{k0}}{q_k \mathbb{E}_k(\tau_{\{k,0\}})} \qquad (q_k = -q_{kk}).$$

Then, X has a unique QSD $\boldsymbol{u}$, and, for any probability measure $\boldsymbol{\nu}$ on C,

$$d_{\mathrm{TV}}(\boldsymbol{u}, \boldsymbol{\pi}^{\boldsymbol{\nu}}) \leq 2UT/p.$$

For n-patch metapopulation model we may take $s := \lfloor n(1 - e/c) \rfloor$.

The QSD and π^ν are close for *any* ν under the quasi stationary regime



The quasi-stationary distribution (bars), and the pseudo-transient distribution when the resurrection law assigns all its mass to state 1 (circles), for the earlier 20-patch metapopulation model.

The return map

The return map thus provides a useful alternative to modelling quasi stationarity.

The return map

The return map thus provides a useful alternative to modelling quasi stationarity.

It is usually *simpler to evaluate* than the QSD. For example, for absorbing birth-death processes on $C = \{1, 2, \ldots, N\}$ ($N \leq \infty$), with birth rates (b_i, $i \in C$) and death rates (d_i, $i \in C$), $\boldsymbol{\nu} \mapsto \boldsymbol{\pi}^{\boldsymbol{\nu}}$ is given by

$$\pi_i^{\boldsymbol{\nu}} = \pi_1^{\boldsymbol{\nu}} \frac{d_1}{d_i} \sum_{j=1}^{i} \prod_{k=j}^{i-1} \frac{b_k}{d_k} a_j^{\boldsymbol{\nu}}, \qquad i = 1, 2, \ldots, N,$$

where $a_j^{\boldsymbol{\nu}} = \sum_{l=j}^{N} \nu_l$.

Clancy, D., and Pollett, P.K. (2003) A note on quasi-stationary distributions of birth-death processes and the SIS logistic epidemic. J. Appl. Probab. 40, 821–825.

The return map thus provides a useful alternative to modelling quasi stationarity.

It is usually *simpler to evaluate* than the QSD. For example, for absorbing birth-death processes on $C = \{1, 2, \ldots, N\}$ $(N \leq \infty)$, with birth rates $(b_i,\ i \in C)$ and death rates $(d_i,\ i \in C)$, $\boldsymbol{\nu} \mapsto \boldsymbol{\pi}^{\boldsymbol{\nu}}$ is given by

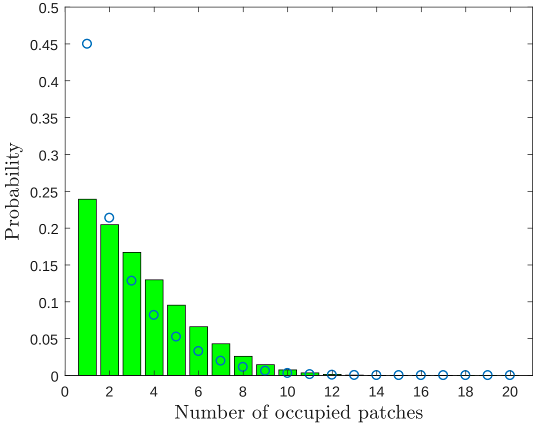
$$\pi_i^{\boldsymbol{\nu}} = \pi_1^{\boldsymbol{\nu}} \frac{d_1}{d_i} \sum_{j=1}^{i} \prod_{k=j}^{i-1} \frac{b_k}{d_k} a_j^{\boldsymbol{\nu}}, \qquad i = 1, 2, \ldots, N,$$

where $a_j^{\boldsymbol{\nu}} = \sum_{l=j}^{N} \nu_l$.

Clancy, D., and Pollett, P.K. (2003) A note on quasi-stationary distributions of birth-death processes and the SIS logistic epidemic. J. Appl. Probab. 40, 821–825.

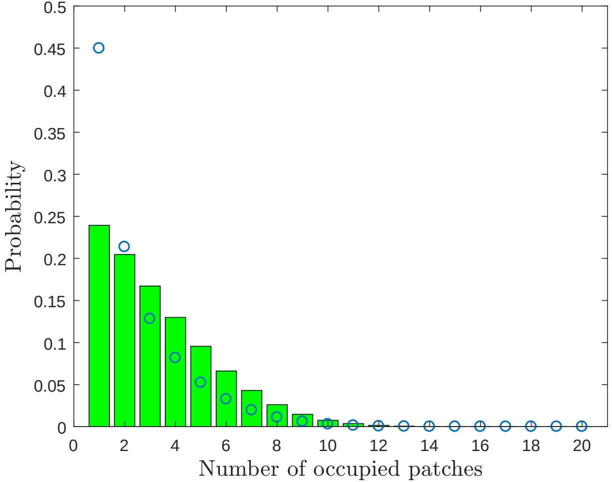
One can approximate the extremal QSD to any level accuracy by iterating the return map.

The QSD and π^{ν} with $\nu = \delta_{\{1\}}$ are *not close* when X is evanescent

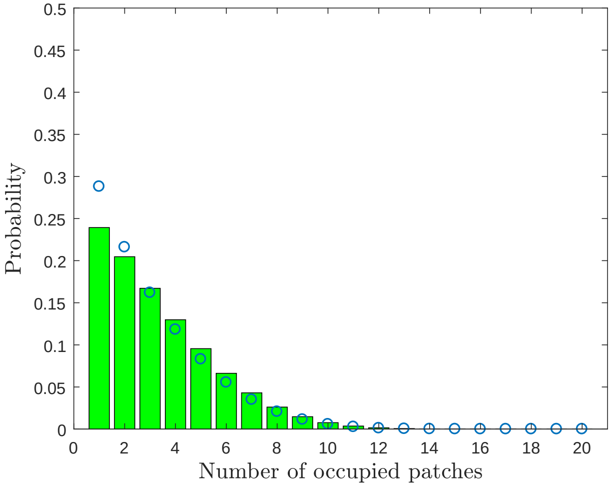


The quasi-stationary distribution (bars), and the pseudo-transient distribution when the resurrection law assigns all its mass to state 1 (circles), for the earlier 20-patch metapopulation model.

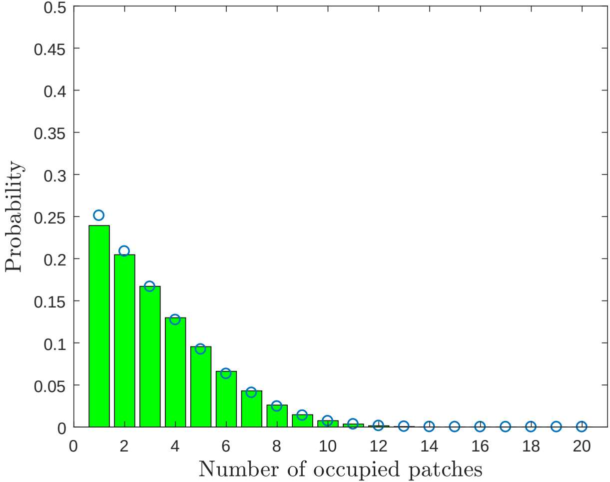
The QSD and π^ν with $\nu = \delta_{\{1\}}$ (evanescent regime)



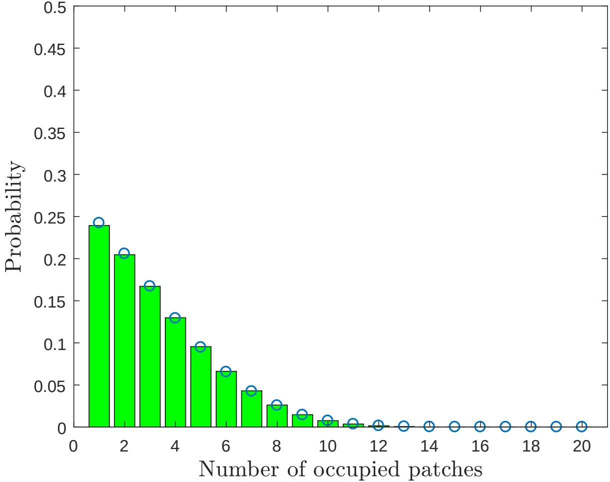
The QSD and π^{ν} under 2 iterations of the return map (evanescent regime)



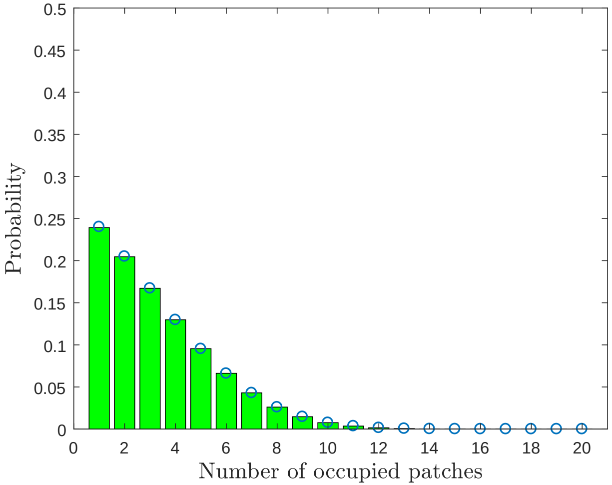
The QSD and π^ν under 2 iterations of the return map (evanescent regime)



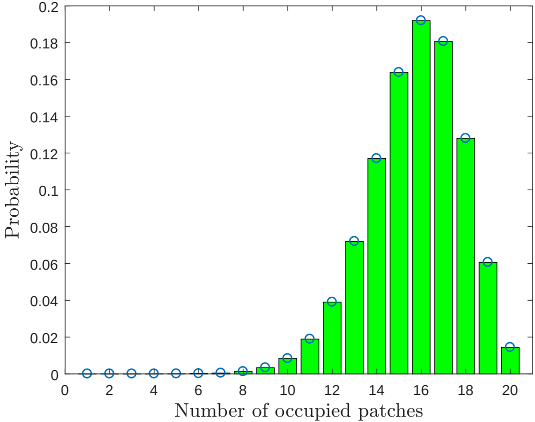
The QSD and π^ν under 4 iterations of the return map (evanescent regime)



The QSD and π^ν under 5 iterations of the return map (evanescent regime)



The QSD and π^ν are close for *any* ν under the quasi stationary regime



The quasi-stationary distribution (bars), and the pseudo-transient distribution (circles), for the earlier 20-patch metapopulation model.