Birth-Death Processes and Orthogonal Polynomials

Phil. Pollett

The University of Queensland

India Institute of Technology Bombay

26 October 2015
A metapopulation model: $X_t$ is the number of occupied patches at time $t$

$\lambda_n = (c/N)n(N - n)$ \hspace{1cm} (c = 8)

$\mu_n = en$ \hspace{1cm} (e = 2)
A metapopulation model: $X_t$ is the number of occupied patches at time $t$

$\uparrow \lambda_n = (c/N)n(N - n) \quad (c = 8)$

$\downarrow \mu_n = en \quad (e = 2)$

$m_j = \lim_{t \to \infty} \Pr(X_t = j | X_t > 0) \quad (j = 1, \ldots, 20)$
\[ p_{ij}(t) : = \Pr(X_{s+t} = j | X_s = i) \]

\[ = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \]
Birth-death processes

A *birth-death* process is a continuous-time Markov chain \((X_t, \ t \geq 0)\) taking values in \(S \cup \{−1\}\), where \(S \subseteq \{0, 1, \ldots \}\), with

\[
\Pr(X_{t+h} = n + 1 | X_t = n) = \lambda_n h + o(h)
\]

\[
\Pr(X_{t+h} = n - 1 | X_t = n) = \mu_n h + o(h)
\]

\[
\Pr(X_{t+h} = n | X_t = n) = 1 - (\lambda_n + \mu_n) h + o(h)
\]

(as \(h \to 0\)). Other transitions happen with probability \(o(h)\).
Birth-death processes

A birth-death process is a continuous-time Markov chain \((X_t, \ t \geq 0)\) taking values in \(S \cup \{-1\}\), where \(S \subseteq \{0, 1, \ldots \}\), with

\[
\Pr(X_{t+h} = n + 1|X_t = n) = \lambda_n h + o(h)
\]

\[
\Pr(X_{t+h} = n - 1|X_t = n) = \mu_n h + o(h)
\]

\[
\Pr(X_{t+h} = n|X_t = n) = 1 - (\lambda_n + \mu_n)h + o(h)
\]

(as \(h \to 0\)). Other transitions happen with probability \(o(h)\).

The birth rates \((\lambda_n, \ n \geq 0)\) and the death rates \((\mu_n, \ n \geq 0)\) are all strictly positive except perhaps \(\mu_0\), which could be 0. State \(-1\) is a “extinction state”, which can be reached if \(\mu_0 > 0\).
Birth-death processes

The \textit{birth rates} \((\lambda_n, \ n \geq 0)\) and the \textit{death rates} \((\mu_n, \ n \geq 0)\) are all strictly positive, except perhaps \(\mu_0\), which could be 0. State \(-1\) is a “extinction state”, which can be reached if \(\mu_0 > 0\).

\[ \mu_0 > 0 \]

\[ \begin{align*}
-1 & \xrightarrow{\mu_0} 0 \\
0 & \xrightarrow{\lambda_0} 1 \\
1 & \xrightarrow{\mu_1} \cdots \\
n-1 & \xrightarrow{\lambda_{n-1}} n \\
n & \xrightarrow{\mu_n} n+1 \\
n+1 & \xrightarrow{\lambda_n} \cdots \\
\end{align*} \]
Birth-death processes

The *birth rates* \((\lambda_n, \ n \geq 0)\) and the *death rates* \((\mu_n, \ n \geq 0)\) are all strictly positive, except perhaps \(\mu_0\), which could be 0. State \(-1\) is a “extinction state”, which can be reached if \(\mu_0 > 0\).

![Diagram of birth-death processes]

\[
\mu_0 > 0
\]

\[
\mu_0 = 0
\]
Explosive birth-death processes

Suppose that $\lambda_n = 2^{2n}$, $\mu_n = 2^{2n-1}$ ($n \geq 1$), and $\mu_0 = 0$, with $S = \{0, 1, \ldots \}$. Thus births are twice as likely as deaths, and so the process will have positive drift.
Explosive birth-death processes

Suppose that \( \lambda_n = 2^{2n} \), \( \mu_n = 2^{2n-1} \) (\( n \geq 1 \)), and \( \mu_0 = 0 \), with \( S = \{0, 1, \ldots \} \).

An “equilibrium distribution” exists: \( \pi_n = (\frac{1}{2})^{n+1} \). But ...
Suppose that \( \lambda_n = 2^{2n}, \mu_n = 2^{2n-1} \) (\( n \geq 1 \)), and \( \mu_0 = 0 \), with \( S = \{0, 1, \ldots \} \).

An “equilibrium distribution” exists: \( \pi_n = \left(\frac{1}{2}\right)^{n+1} \). But …

When a jump occurs it is a birth with probability

\[
p_n = \frac{2^{2n}}{2^{2n} + 2^{2n-1}} = \frac{2}{3}.
\]

Thus births are twice as likely as deaths, and so the process will have positive drift.
Explosive birth-death processes

Birth-death simulation ($\lambda_n = 2^{2n}, \mu_n = 2^{2n-1}$) - Explosion
Explosive birth-death processes

Birth-death simulation \(\lambda_n = 2^{2n}, \mu_n = 2^{2n-1}\) - Explosion

\[ X_t \]

\[ t \]

0 1 2 3 4 5 6 7 8 9
0 50 100 150 200 250
Explosive birth-death processes

Birth-death simulation \( (\lambda_n = 2^{2n}, \mu_n = 2^{2n-1}) \) - Restart in State 0 after explosion
Explosive birth-death processes

Birth-death simulation \((\lambda_n = 2^{2n}, \mu_n = 2^{2n-1})\) - Regular boundary
The Kolmogorov differential equations

The conditions we have imposed ensure that the transition probabilities

\[ p_{ij}(t) = \Pr(X_{s+t} = j | X_s = i) \ (i, j \in S, \ s, t \geq 0) \]

do not depend on \( s \).
The Kolmogorov differential equations

The conditions we have imposed ensure that the transition probabilities
\( p_{ij}(t) = \Pr(X_{s+t} = j|X_s = i) \) \((i, j \in S, s, t \geq 0)\) do not depend on \(s\).

For any such time-homogeneous continuous-time Markov chain with (conservative) transition rate matrix \(Q = (q_{ij})\), the transition function \(P(t) = (p_{ij}(t))\) satisfies the backward equations

\[
P'(t) = QP(t) \tag{BE}
\]

but not necessarily the forward equations

\[
P'(t) = P(t)Q \tag{FE}
\]

(the derivative is taken elementwise). Note that \(Q = P'(0+)\).
The Kolmogorov differential equations

The conditions we have imposed ensure that the transition probabilities

\[ p_{ij}(t) = \Pr(X_{s+t} = j|X_s = i) \ (i, j \in S, \ s, t \geq 0) \]

do not depend on \( s \).

For any such time-homogeneous continuous-time Markov chain with (conservative) transition rate matrix \( Q = (q_{ij}) \), the transition function \( P(t) = (p_{ij}(t)) \) satisfies the backward equations

\[ P'(t) = QP(t) \quad \text{(BE)} \]

but not necessarily the forward equations

\[ P'(t) = P(t)Q \quad \text{(FE)} \]

(the derivative is taken elementwise). Note that \( Q = P'(0+) \).

There is however a minimal solution \( F(t) = (f_{ij}(t)) \) to (BE) and this satisfies (FE).
The Kolmogorov differential equations

The conditions we have imposed ensure that the \textit{transition probabilities}
\[ p_{ij}(t) = \Pr(X_{s+t} = j | X_s = i) \quad (i, j \in S, \ s, t \geq 0) \]
do not depend on \( s \).

For any such \textit{time-homogeneous} continuous-time Markov chain with (conservative) transition rate matrix \( Q = (q_{ij}) \), the \textit{transition function} \( P(t) = (p_{ij}(t)) \) satisfies the \textit{backward equations}

\[
P'(t) = QP(t) \quad \text{(BE)}
\]
but not necessarily the \textit{forward equations}

\[
P'(t) = P(t)Q \quad \text{(FE)}
\]
(the derivative is taken elementwise). Note that \( Q = P'(0+) \).

There is however a \textit{minimal} solution \( F(t) = (f_{ij}(t)) \) to (BE) and this \textit{satisfies} (FE).

Non-explosivity corresponds to \( F \) being the \textit{unique} solution to (BE). Otherwise \( F \) governs the process \textit{up to the time of the (first) explosion}.
The Kolmogorov differential equations

For birth-death processes the full range of behaviour is possible.

Here the transition rate matrix restricted to $S = \{0, 1, \ldots \}$ has the form

$$Q = \begin{pmatrix}
-(\lambda_0 + \mu_0) & \lambda_0 & 0 & 0 & 0 & \ldots \\
\mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \ldots \\
0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \ldots \\
0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}.$$
The Kolmogorov differential equations

For birth-death processes the full range of behaviour is possible.

Here the transition rate matrix restricted to $S = \{0, 1, \ldots \}$ has the form

$$Q = \begin{pmatrix} - (\lambda_0 + \mu_0) & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & - (\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & - (\lambda_2 + \mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & - (\lambda_3 + \mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$  

Returning to the example where $\lambda_n = 2^{2n}$, $\mu_n = 2^{2n-1}$ ($n \geq 1$), and $\mu_0 = 0$, we have ...
The process governed by $F$ (the “minimal process”)
A process where $P$ satisfies (BE) \textit{but not} (FE)

Birth-death simulation ($\lambda_n = 2^{2n}, \mu_n = 2^{2n-1}$) - Restart in State 0 after explosion
A process where $P$ satisfies both (BE) and (FE)
Define a sequence \((Q_n, \ n \in S)\) of polynomials by

\[
Q_0(x) = 1
\]

\[
-xQ_0(x) = -(\lambda_0 + \mu_0)Q_0(x) + \lambda_0 Q_1(x)
\]

\[
-xQ_n(x) = \mu_n Q_{n-1}(x) - (\lambda_n + \mu_n)Q_n(x) + \lambda_n Q_{n+1}(x),
\]

and a sequence of strictly positive numbers \((\pi_n, \ n \in S)\) by \(\pi_0 = 1\) and, for \(n \geq 1\),

\[
\pi_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n}.
\]
A explicit expression for $p_{ij}(t)$

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \quad t \geq 0).$$

A explicit expression for $p_{ij}(t)$

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) \, d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$

Since $p_{ij}(0) = \delta_{ij}$, it is clear that $(Q_n)$ are orthogonal with orthogonalizing measure $\psi$:

$$\int_0^\infty Q_i(x) Q_j(x) \, d\psi(x) = \frac{\delta_{ij}}{\pi_j} \quad (i, j \geq 0).$$
A explicit expression for $p_{ij}(t)$

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$

**Proof.** We use mathematical induction: on $i$ using (BE) with $j = 0$ and then on $j$ using (FE). But, showing that there is a probability measure $\psi$ with support $[0, \infty)$ whose Laplace transform is $p_{00}(t)$, that is

$$p_{00}(t) = \int_0^\infty e^{-tx} d\psi(x) \quad (\pi_0 \int_0^\infty e^{-tx} Q_0(x) Q_0(x) d\psi(x)),$$

is not completely straightforward. More on this later.
A explicit expression for $p_{ij}(t)$ - Why is it useful?

Theorem (Karlin and McGregor (1957))

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) \, d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
A explicit expression for $p_{ij}(t)$ - Why is it useful?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
A explicit expression for $p_{ij}(t)$ - Why is it useful?

Theorem (Karlin and McGregor (1957))

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
A explicit expression for $p_{ij}(t)$ - Why is it useful?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \int_{0}^{\infty} e^{-tx} Q_i(x) \pi_j Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
A explicit expression for $p_{ij}(t)$ - Why is it useful?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \int_0^\infty e^{-tx} Q_i(x) \pi_j Q_j(x) \, d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
A explicit expression for $p_{ij}(t)$ - Why is it useful?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$

This formula, together with the myriad of properties of $(Q_n)$ and $\psi$, are used to develop theory peculiar to birth-death processes.
Some properties of \((Q_n)\) and \(\psi\)

Of particular interest and importance is the “interlacing” property of the zeros \(x_{n,i}\) \((i = 1, \ldots, n)\) of \(Q_n\): they are strictly positive, simple, and they satisfy

\[
0 < x_{n+1,i} < x_{n,i} < x_{n+1,i+1}, \quad (i = 1, \ldots, n, \ n \geq 1),
\]

from which it follows that the limits \(\xi_i = \lim_{n \to \infty} x_{n,i} \ (i \geq 1)\) exist and satisfy

\[
0 \leq \xi_i \leq \xi_{i+1} < \infty.
\]
Some properties of \((Q_n)\) and \(\psi\)

Of particular interest and importance is the “interlacing” property of the zeros \(x_{n,i} (i = 1, \ldots, n)\) of \(Q_n\): they are strictly positive, simple, and they satisfy

\[
0 < x_{n+1,i} < x_{n,i} < x_{n+1,i+1}, \quad (i = 1, \ldots, n, \ n \geq 1),
\]

from which it follows that the limits \(\xi_i = \lim_{n \to \infty} x_{n,i} \ (i \geq 1)\) exist and satisfy

\[
0 \leq \xi_i \leq \xi_{i+1} < \infty.
\]

Interestingly, \(\xi_1 := \inf \text{supp}(\psi)\) and \(\xi_2 := \inf \{\text{supp}(\psi) \cap (\xi_1, \infty)\}\), quantities that are particularly important in the theory of \emph{quasi-stationary distributions}. 
Consider the case $\mu_0 > 0$:

$$\sum_{n=0}^{\infty} \left( \lambda_n \pi_n \right)^{-1} = \infty,$$

which ensures that the extinction state $-1$ is reached with probability 1 (and necessarily the process is non-explosive).

Let $T = \inf \{ t \geq 0 : X_t = -1 \}$ be the time to extinction. Clearly $\Pr(T > t | X_0 = i) \to 0$ as $t \to \infty$, but how fast?

Claim. \( \inf \{ a \geq 0 : \int_0^\infty e^{at} \Pr(T > t | X_0 = i) dt = \infty \} = \xi_1 \).

Consider the case \( \mu_0 > 0 \):

Suppose that \( \sum_{n=0}^{\infty} (\lambda_n \pi_n)^{-1} = \infty \), which ensures that the extinction state \(-1\) is reached with probability 1 (and necessarily the process is non-explosive).

The time to extinction

Consider the case $\mu_0 > 0$:

![Diagram showing birth-death process transition rates.]

Suppose that $\sum_{n=0}^{\infty} (\lambda_n \pi_n)^{-1} = \infty$, which ensures that the extinction state $-1$ is reached with probability 1 (and necessarily the process is non-explosive).

Let $T = \inf\{t \geq 0 : X_t = -1\}$ be the time to extinction.

---

Consider the case $\mu_0 > 0$:

$$
\begin{align*}
-1 & \xrightarrow{\mu_0} 0 \\
0 & \xrightarrow{\mu_1} 1 \\
1 & \xrightarrow{\lambda_0} \ldots \\
\ldots & \\
n-1 & \xrightarrow{\lambda_{n-1}} n \\
n & \xrightarrow{\mu_n} n+1 \\
n+1 & \xrightarrow{\lambda_n} \ldots \\
\end{align*}
$$

Suppose that $\sum_{n=0}^{\infty} (\lambda_n \pi_n)^{-1} = \infty$, which ensures that the extinction state $-1$ is reached with probability 1 (and necessarily the process is non-explosive).

Let $T = \inf\{t \geq 0 : X_t = -1\}$ be the time to extinction.

Clearly $\Pr(T > t | X_0 = i) \to 0$ as $t \to \infty$, but how fast?

---

Consider the case \( \mu_0 > 0 \):

\[
\begin{array}{cccccc}
-1 & 0 & 1 & \ldots & n-1 & n & n+1 \\
\mu_0 & \lambda_0 & \mu_1 & & \lambda_{n-1} & \mu_n & \mu_{n+1}
\end{array}
\]

Suppose that \( \sum_{n=0}^{\infty} (\lambda_n \pi_n)^{-1} = \infty \), which ensures that the extinction state \(-1\) is reached with probability 1 (and necessarily the process is non-explosive).

Let \( T = \inf\{t \geq 0 : X_t = -1\} \) be the time to extinction.

Clearly \( \Pr(T > t | X_0 = i) \to 0 \) as \( t \to \infty \), but how fast?

**Claim.** \( \inf \left\{ a \geq 0 : \int_0^{\infty} e^{at} \Pr(T > t | X_0 = i) \, dt = \infty \right\} = \xi_1. \)

Quasi-stationary distributions

A distribution \( u = (u_n, \ n \geq 0) \) is called a \textit{limiting conditional distribution} (or sometimes \textit{quasi-stationary distribution}) if \( u_{ij}(t) := \Pr(X_t = j | T > t, X_0 = i) \rightarrow u_j \) as \( t \rightarrow \infty \).


A distribution $u = (u_n, \ n \geq 0)$ is called a limiting conditional distribution (or sometimes quasi-stationary distribution) if $u_{ij}(t) := \Pr(X_t = j | T > t, X_0 = i) \rightarrow u_j$ as $t \rightarrow \infty$.

**Theorem**

If $\xi_1 > 0$ then $u_{ij}(t) \rightarrow u_j := \mu_0^{-1} \xi_1 \pi_j Q_j(\xi_1)$. If $\xi_1 = 0$ then $u_j(t) \rightarrow 0$.

---


Quasi-stationary distributions

A distribution \( u = (u_n, \ n \geq 0) \) is called a *limiting conditional distribution* (or sometimes *quasi-stationary distribution*) if \( u_{ij}(t) := \Pr(X_t = j| T > t, X_0 = i) \to u_j \) as \( t \to \infty \).

**Theorem**

If \( \xi_1 > 0 \) then \( u_{ij}(t) \to u_j := \mu_0^{-1}\xi_1 \pi_j Q_j(\xi_1) \). If \( \xi_1 = 0 \) then \( u_j(t) \to 0 \).

Again, how fast?

---


Quasi-stationary distributions

A distribution $u = (u_n, n \geq 0)$ is called a limiting conditional distribution (or sometimes quasi-stationary distribution) if $u_{ij}(t) := \Pr(X_t = j| T > t, X_0 = i) \to u_j$ as $t \to \infty$.

**Theorem**

If $\xi_1 > 0$ then $u_{ij}(t) \to u := \mu_0^{-1}\xi_1 \pi_j Q_j(\xi_1)$. If $\xi_1 = 0$ then $u_j(t) \to 0$.

Again, how fast?

**Claim.** $\inf \left\{ a \geq 0 : \int_0^\infty e^{at} |u_{ij}(t) - u_j| \, dt = \infty \right\} = \xi_2 - \xi_1$ (same for all $i, j \in S$).

---


**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^{\infty} e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$

**Proof.** We use mathematical induction: on $i$ using (BE) with $j = 0$ and then on $j$ using (FE). But, showing that there is a probability measure $\psi$ with support $[0, \infty)$ whose Laplace transform is $p_{00}(t)$, that is

$$p_{00}(t) = \int_0^{\infty} e^{-tx} d\psi(x) \left( = \pi_0 \int_0^{\infty} e^{-tx} Q_0(x) Q_0(x) d\psi(x) \right),$$

is not completely straightforward. More on this later.
Why does this work?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$
Why does this work?

**Theorem (Karlin and McGregor (1957))**

Let $P(t) = (p_{ij}(t))$ be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure $\psi$ with support $[0, \infty)$ such that

$$p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0).$$

Why is there a $\psi$ whose Laplace transform is $p_{00}(t)$: $p_{00}(t) = \int_0^\infty e^{-tx} d\psi(x)$?
Why does this work?

Theorem (Karlin and McGregor (1957))

Let \( P(t) = (p_{ij}(t)) \) be any transition function that satisfies both the backward and the forward equations (for example the minimal one). Then, there is a probability measure \( \psi \) with support \([0, \infty)\) such that

\[
p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) \, d\psi(x) \quad (i, j \geq 0, \ t \geq 0).
\]

Why is there a \( \psi \) whose Laplace transform is \( p_{00}(t) \): \( p_{00}(t) = \int_0^\infty e^{-tx} \, d\psi(x) \)?

**Answer.** Weak symmetry: \( \pi_i q_{ij} = \pi_j q_{ji} \) (\( \pi_i \lambda_i = \pi_{i+1} \mu_{i+1} \))
Let \((X_t, \ t \geq 0)\) be a continuous-time Markov chain taking values in \(S = \{0, 1, \ldots, N\}\) with (conservative) transition rate matrix \(Q\). So, there is collection \(\pi = (\pi_j, \ j \in S)\) of strictly positive numbers such that \(\pi Q = 0\), that is

\[
\sum_{i \in S} \pi_i q_{ij} = \pi_j \sum_{i \in S} q_{ji} \quad (j \in S).
\]
Finite state Markov chains - some linear algebra

Let \((X_t, \ t \geq 0)\) be a continuous-time Markov chain taking values in \(S = \{0, 1, \ldots, N\}\) with (conservative) transition rate matrix \(Q\). So, there is collection \(\pi = (\pi_j, \ j \in S)\) of strictly positive numbers such that \(\pi Q = 0\), that is

\[
\sum_{i \in S} \pi_i q_{ij} = \pi_j \sum_{i \in S} q_{ji} \quad (j \in S).
\]

Suppose that \(Q\) is weakly symmetric with respect to \(\pi\): \(\pi_i q_{ij} = \pi_j q_{ji}\). 

Finite state Markov chains - some linear algebra

Let \((X_t, \; t \geq 0)\) be a continuous-time Markov chain taking values in \(S = \{0, 1, \ldots, N\}\) with (conservative) transition rate matrix \(Q\). So, there is collection \(\pi = (\pi_j, \; j \in S)\) of strictly positive numbers such that \(\pi Q = 0\), that is

\[
\sum_{i \in S} \pi_i q_{ij} = \pi_j \sum_{i \in S} q_{ji} \quad (j \in S).
\]

Suppose that \(Q\) is weakly symmetric with respect to \(\pi\): \(\pi_i q_{ij} = \pi_j q_{ji}\).

Let \(A\) be the symmetric matrix with entries \(a_{ij} = \sqrt{\pi_i q_{ij}}/\sqrt{\pi_j}\). It is orthogonally similar to a diagonal matrix \(D = \text{diag}\{d_0, d_1, \ldots, d_N\}: A = MDM^\top\ldots\), et cetera, \ldots
Finite state Markov chains - some linear algebra

Let \((X_t, \ t \geq 0)\) be a continuous-time Markov chain taking values in \(S = \{0, 1, \ldots, N\}\) with (conservative) transition rate matrix \(Q\). So, there is collection \(\pi = (\pi_j, \ j \in S)\) of strictly positive numbers such that \(\pi Q = 0\), that is

\[
\sum_{i \in S} \pi_i q_{ij} = \pi_j \sum_{i \in S} q_{ji} \quad (j \in S).
\]

Suppose that \(Q\) is weakly symmetric with respect to \(\pi\): \(\pi_i q_{ij} = \pi_j q_{ji}\).

Let \(A\) be the symmetric matrix with entries \(a_{ij} = \sqrt{\pi_i q_{ij}} / \sqrt{\pi_j}\). It is orthogonally similar to a diagonal matrix \(D = \text{diag}\{d_0, d_1, \ldots, d_N\}\): \(A = MDM^\top\), et cetera, ... leading to the spectral solution of \(P'(t) = QP(t)\) (BE):

\[
p_{ij}(t) = \pi_j \sum_{k=0}^N e^{dk t} Q_{ij}^{(k)} Q_{ij}^{(k)}, \quad \text{where } Q_{ij}^{(k)} = \frac{M_{ik}}{\sqrt{\pi_i}}.
\]
General symmetric Markov chains - some functional analysis

Let \( \pi = (\pi_j, \ j \in S) \) be a collection of strictly positive numbers and suppose that \( P \) is weakly symmetric with respect to \( \pi \): \( \pi_i p_{ij}(t) = \pi_j p_{ji}(t) \ (i, j \in S) \).

General symmetric Markov chains - some functional analysis

Let \( \pi = (\pi_j, j \in S) \) be a collection of strictly positive numbers and suppose that \( P \) is weakly symmetric with respect to \( \pi \): \( \pi_i p_{ij}(t) = \pi_j p_{ji}(t) \) \((i, j \in S)\).

Define \( T_t : \ell_2 \to \ell_2 \) by

\[
(T_t x)_j = \sum_{i \in S} x_i \left(\frac{\pi_i}{\pi_j}\right)^{1/2} p_{ij}(t) \quad (i \in S, \ x \in \ell_2).
\]

Let $\pi = (\pi_j, j \in S)$ be a collection of strictly positive numbers and suppose that $P$ is weakly symmetric with respect to $\pi$: $\pi_i p_{ij}(t) = \pi_j p_{ji}(t)$ ($i, j \in S$).

Define $T_t : \ell_2 \to \ell_2$ by

$$(T_t x)_j = \sum_{i \in S} x_i \left( \frac{\pi_i}{\pi_j} \right)^{1/2} p_{ij}(t) \quad (i \in S, \ x \in \ell_2).$$

Then $(T_t, \ t \geq 0)$ is a semigroup which is self adjointing $\langle T_t x, y \rangle = \langle x, T_t y \rangle$.

General symmetric Markov chains - some functional analysis

Let \( \pi = (\pi_j, j \in S) \) be a collection of strictly positive numbers and suppose that \( P \) is weakly symmetric with respect to \( \pi \): \( \pi_i p_{ij}(t) = \pi_j p_{ji}(t) \) \((i,j \in S)\).

Define \( T_t : \ell_2 \rightarrow \ell_2 \) by

\[
(T_t x)_j = \sum_{i \in S} x_i (\pi_i / \pi_j)^{1/2} p_{ij}(t) \quad (i \in S, \ x \in \ell_2).
\]

Then \((T_t, \ t \geq 0)\) is a semigroup which is self adjointing \( \langle T_t x, y \rangle = \langle x, T_t y \rangle \).

Kendall used a result of Riesz and Sz.-Nagy on the spectral representation of self-adjoint semigroups to show that there is a finite signed measure \( \gamma_{ij} \) with support \([0, \infty)\) such that

\[
p_{ij}(t) = (\pi_j / \pi_i)^{1/2} \int_{[0,\infty)} e^{-tx} d\gamma_{ij}(x).
\]

Furthermore, \( \gamma_{ii} \) is a probability measure.

In can be seen from the definition of the birth-death polynomials \( \mathcal{Q} = (\mathcal{Q}_n, \ n \in S) \),

\[
\mathcal{Q}_0(x) = 1 \\
-x\mathcal{Q}_0(x) = -(\lambda_0 + \mu_0)\mathcal{Q}_0(x) + \lambda_0\mathcal{Q}_1(x) \\
-x\mathcal{Q}_n(x) = \mu_n\mathcal{Q}_{n-1}(x) - (\lambda_n + \mu_n)\mathcal{Q}_n(x) + \lambda_n\mathcal{Q}_{n+1}(x),
\]

and the form of transition rate matrix restricted to \( S = \{0, 1, \ldots\} \),

\[
\mathcal{Q} = \begin{pmatrix}
-(\lambda_0 + \mu_0) & \lambda_0 & 0 & 0 & 0 & \cdots \\
\mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \cdots \\
0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \cdots \\
0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix},
\]

that \( \mathcal{Q} = \mathcal{Q}(x) \) as a column vector satisfies \( \mathcal{Q}\mathcal{Q} = -x\mathcal{Q} \) (\( \mathcal{Q}(x) \) is an \( x \)-invariant vector for \( \mathcal{Q} \)), and \( \mathcal{R} = \mathcal{R}(x) \), where \( \mathcal{R}_j(x) = \pi_j\mathcal{Q}_j(x) \), as a row vector satisfies \( \mathcal{R}\mathcal{Q} = -x\mathcal{R} \) (\( \mathcal{R}(x) \) is an \( x \)-invariant measure for \( \mathcal{Q} \)).
General symmetric Markov chains - speculation

One might speculate that

\[ p_{ij}(t) = \pi_j \int_0^{\infty} e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0) \]

holds more generally under weak symmetry \((\pi_i q_{ij} = \pi_j q_{ji})\) for a function system \(Q = (Q_n, \ n \in S)\) (necessarily orthogonal with respect to \(\psi\)) satisfying \(QQ = -xQ\).
General symmetric Markov chains - speculation

One might speculate that

\[ p_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0) \]

holds more generally under weak symmetry \((\pi_i q_{ij} = \pi_j q_{ji})\) for a function system \(Q = (Q_n, \ n \in S)\) (necessarily orthogonal with respect to \(\psi\)) satisfying \(QQ = -xQ\).

It might perhaps be too much to expect that

\[ p_{ij}(t) = \int_0^\infty e^{-tx} Q_i(x) R_j(x) d\psi(x) \quad (i, j \geq 0, \ t \geq 0) \]

holds with just \(\pi Q = 0\) for function systems \(Q = (Q_n, \ n \in S)\) and \(R = (R_n, \ n \in S)\) satisfying \(QQ = -xQ\) and \(RQ = -xR\), and, of necessity,

\[ \int_0^\infty Q_i(x) R_j(x) d\psi(x) = \delta_{ij} \quad (i, j \geq 0). \]