Ensemble behaviour in population processes

Phil Pollett

http://www.maths.uq.edu.au/~pkp
Motivating example

Patients in later stages of congestive heart failure.

Clinicians claimed that numbers appear to be “quasi-stationary”.
Proportions of patients

Proportions alive at $t = 110$ days ($n = 1000$ patients)

Stage

Proportion

A

B

C

D

Proportions A and B are similar, while C has the highest proportion and D has the lowest.
Proportions of patients

Proportions alive at $t = 120$ days ($n = 1000$ patients)
Proportions of patients

Proportions alive at $t = 130$ days ($n = 1000$ patients)

- Stage A: 10%
- Stage B: 20%
- Stage C: 40%
- Stage D: 20%
Proportions of patients

Proportions alive at $t = 140$ days ($n = 1000$ patients)

Stage Proportion

A 10%

B 20%

C 30%

D 40%
Proportions of patients

Proportions alive at $t = 150$ days \((n = 1000 \text{ patients})\)

- **Stage A**: 10%
- **Stage B**: 20%
- **Stage C**: 40%
- **Stage D**: 30%
Proportions of patients

Proportions alive at $t = 160$ days ($n = 1000$ patients)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10%</td>
</tr>
<tr>
<td>B</td>
<td>20%</td>
</tr>
<tr>
<td>C</td>
<td>30%</td>
</tr>
<tr>
<td>D</td>
<td>40%</td>
</tr>
</tbody>
</table>

MASCOS APWSPM08, February 2008 - Page 8
A discrete-time Markov chain with state space 
\( S = \{0, 1, 2, 3, 4\} \) and with 1-step transition matrix \( P = (p_{ij}) \) given by

\[
p_{i,i-1} = 1 - p_{ii} = r_i \quad (i = 1, \ldots, 4) \quad (r_1, \ldots, r_4 \text{ given})
\]
\[
p_{00} = 1.
\]
Their model

A discrete-time Markov chain with state space $S = \{0, 1, 2, 3, 4\}$ and with 1-step transition matrix $P = (p_{ij})$ given by

$$p_{i,i-1} = 1 - p_{ii} = r_i \quad (i = 1, \ldots, 4) \quad (r_1, \ldots, r_4 \text{ given}).$$

$p_{00} = 1$.

Comments please.
Their model

A discrete-time Markov chain with state space $S = \{0, 1, 2, 3, 4\}$ and with 1-step transition matrix $P = (p_{ij})$ given by

$$p_{i,i-1} = 1 - p_{ii} = r_i \quad (i = 1, \ldots, 4) \quad (r_1, \ldots, r_4 \text{ given}).$$

$$p_{00} = 1.$$

Comments please.

Their method of analysis involved evaluating the conditional probability $p_j(t)/(1 - p_0(t))$ $(j = 1, \ldots, 4)$, where $p_j(t) = (P^t)_{ij}$ ($i$ is the initial state).
Their model

A discrete-time Markov chain with state space \( S = \{0, 1, 2, 3, 4\} \) and with 1-step transition matrix \( P = (p_{ij}) \) given by

\[
p_{i,i-1} = 1 - p_{ii} = r_i \quad (i = 1, \ldots, 4) \quad (r_1, \ldots, r_4 \text{ given}).
\]

\[
p_{00} = 1. \]

Comments please.

Their method of analysis involved evaluating the conditional probability \( p_j(t)/(1 - p_0(t)) \) \( (j = 1, \ldots, 4) \), where \( p_j(t) = (P^t)_{ij} \) \( (i \text{ is the initial state}) \).

Correct!
Why did our clinicians propose a model for the progress of a disease in a single patient, when they were interested in the behaviour of a large group?
Why did our clinicians propose a model for the progress of a disease in a single patient, when they were interested in the behaviour of a large group?

... because the *proportion* of patients in stage $s$ at time $t$ should be approximately equal to $p_s(t)$, the *probability* that the *individual* patient is in stage $s$ at time $t$. 
Why did our clinicians propose a model for the progress of a disease in a single patient, when they were interested in the behaviour of a large group?

... because the *proportion* of patients in stage $s$ at time $t$ should be approximately equal to $p_s(t)$, the *probability* that the *individual* patient is in stage $s$ at time $t$.

Can properties of an ensemble of individuals be deduced from a model for the behaviour of the individual?
Why did our clinicians propose a model for the progress of a disease in a single patient, when they were interested in the behaviour of a large group?

... because the *proportion* of patients in stage $s$ at time $t$ should be approximately equal to $p_s(t)$, the *probability* that the *individual* patient is in stage $s$ at time $t$.

Can properties of an ensemble of individuals be deduced from a model for the behaviour of the individual?

Further examples ....
Example  A population network, where a fixed number of individuals occupies geographically separated “patches”.

Patches may become empty, but can be recolonized through migration from other patches.

The individual spends a period of time in a given patch and might then emigrate to another patch, spend a period there, and so forth.

We could model the progress of the individual as a random walk on the patches, and thus evaluate quantities such as the probability $p_j(t)$ that the individual occupies patch $j$ at time $t$. We expect that the proportion of individuals in patch $j$ at time $t$ should be approximately equal to $p_j(t)$. 
**Example**  A variant where we allow death or external emigration from any patch.

There are two cases: (i) the *open* network, where there is external immigration to one or more patches, and (ii) the *closed* network, where all individuals eventually disappear from the network through death or external emigration.

Now individuals (perhaps arriving from outside the network) perform a random walk on the patches but then eventually leave.

The total number of individuals is now *random*, but we would expect to be able to draw similar conclusions concerning ensemble proportions.
Butterfly life cycle

Life cycle simulation

- Egg
- Larva
- Pupa
- Adult
- Death

Time (days)
Butterfly life cycle

Egg \( \sim 4 \text{ days} \)

Larva (caterpillar) \( \sim 14 \text{ days} \)

Pupa (chrysalis) \( \sim 7 \text{ days} \)

Adult (butterfly) \( \sim 14 \text{ days} \)
Butterfly life cycle

Life cycle simulation

- Egg
- Larva
- Pupa
- Adult
- Death

Time (days)
Ensemble of organisms

Life cycle simulation ($n = 7$ butterflies)
Ensemble of organisms

Life cycle simulation ($n = 7$ butterflies)
Can properties of the ensemble, be deduced from a model for the behaviour of an individual?
Can properties of the ensemble, be deduced from a model for the behaviour of an individual?

For example, suppose we have $n$ butterflies.

Our intuition tells us that, for the ensemble, the proportion of organisms in stage $s$ at time $t$ should be approximately equal to $p_s(t)$, the probability that the individual organism is in stage $s$ at time $t$. 
Can properties of the ensemble, be deduced from a model for the behaviour of an individual?

For example, suppose we have $n$ butterflies.

Our intuition tells us that, for the ensemble, the *proportion* of organisms in stage $s$ at time $t$ should be approximately equal to $p_s(t)$, the *probability* that the *individual* organism is in stage $s$ at time $t$.

*So strong is this intuition that scientists frequently model population proportions using individual-level models.*
State probabilities (individual)

Life cycle simulation ($n = 7$ butterflies)

- Egg
- Larva
- Pupa
- Adult
- Death
State probabilities (individual)

State probabilities ($t = 10$)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>0.1</td>
</tr>
<tr>
<td>Larva</td>
<td>0.6</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.2</td>
</tr>
<tr>
<td>Adult</td>
<td>0.1</td>
</tr>
<tr>
<td>Death</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**State probabilities (individual)**

**State probabilities ($t = 10$)**

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>0.1</td>
</tr>
<tr>
<td>Larva</td>
<td>0.6</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.2</td>
</tr>
<tr>
<td>Adult</td>
<td>0.1</td>
</tr>
<tr>
<td>Death</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Proportions at \( t = 10 \text{ days} \) (\( n = 7 \text{ butterflies} \))
Simulated proportions (ensemble)

Proportions at $t = 10$ days ($n = 1000$ butterflies)

- Egg: 10%
- Larva: 20%
- Pupa Stage: 30%
- Adult: 40%
- Death: 50%
Perhaps not surprising …

If the individual organisms behave independently, we can employ the Law of Large Numbers.

Look at the ensemble at a fixed time $t$. Fix a stage $s$ and let

$$X_j = \begin{cases} 1 & \text{if organism } j \text{ is in stage } s \\ 0 & \text{if organism } j \text{ is in another stage.} \end{cases}$$

Clearly $X_1, X_2, \ldots$ are independent. So, $\frac{1}{n} \sum_{j=1}^{n} X_j$ (the proportion in stage $s$) converges \textit{almost surely} to $\mathbb{E}(X_1)$, being the probability that any given organism is in stage $s$. 
Individual organism

Life cycle simulation

- Egg
- Larva
- Pupa
- Adult
- Death

Time (days)
What is the probability that the organism is in stage $s$ of its life cycle at time $t$?
What is the probability that the organism is in stage $s$ of its life cycle at time $t$?

Using a simple Markov chain model, we can evaluate this for each stage $s$ and for all times $t$. 
$X(t)$ - the state of an individual at time $t \geq 0$, for example, the current stage in the individual’s life cycle.

Suppose $(X(t), t \geq 0)$ is a continuous-time Markov chain taking values in a discrete set $S$ with transition rates $(q_{ij})$:

$q_{ij}$ is the rate of transition from state $i \to j$ ($j \neq i$).

$q_i (= -q_{ii}) = \sum_{j \neq i} q_{ij}$ is the total rate out of state $i$. 
Evaluating state probabilities

$X(t)$ - the state of an individual at time $t \geq 0$, for example, the current stage in the individual’s life cycle.

Suppose $(X(t), t \geq 0)$ is a continuous-time Markov chain taking values in a discrete set $S$ with transition rates $(q_{ij})$:

$q_{ij}$ is the rate of transition from state $i \rightarrow j$ ($j \neq i$).

$q_i (= -q_{ii}) = \sum_{j \neq i} q_{ij}$ is the total rate out of state $i$.

**Example** (Butterfly life cycle) $\{4\} \rightarrow \{3\} \rightarrow \{2\} \rightarrow \{1\} \rightarrow \{0\}$

$q_4 = q_{43} = 1/4$ \hspace{1cm} ↓ Egg ($\approx 4$ days)

$q_3 = q_{32} = 1/14$ \hspace{1cm} ↓ Caterpillar ($\approx 14$ days)

$q_2 = q_{21} = 1/7$ \hspace{1cm} ↓ Chrysalis ($\approx 7$ days)

$q_1 = q_{10} = 1/14$ \hspace{1cm} ↓ Adult ($\approx 14$ days)
In matrix form

\[ Q = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
1/14 & -1/14 & 0 & 0 & 0 \\
0 & 1/7 & -1/7 & 0 & 0 \\
0 & 0 & 1/14 & -1/14 & 0 \\
0 & 0 & 0 & 1/4 & -1/4 \\
\end{pmatrix} \]
In matrix form

\[ Q = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
1/14 & -1/14 & 0 & 0 & 0 \\
0 & 1/7 & -1/7 & 0 & 0 \\
0 & 0 & 1/14 & -1/14 & 0 \\
0 & 0 & 0 & 1/4 & -1/4 \\
\end{pmatrix} \]

Why put minus the total rate on the diagonal?
Evaluating state probabilities

In matrix form

\[ Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/14 & -1/14 & 0 & 0 & 0 & 0 \\ 0 & 1/7 & -1/7 & 0 & 0 & 0 \\ 0 & 0 & 1/14 & -1/14 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & -1/4 & \end{pmatrix} \]

Why put minus the total rate on the diagonal?
For mathematical convenience \ldots the equations we must solve are then easier to write down.
The state probabilities $p(t) = (p_j(t), j \in S)$, where

$$p_j(t) = \Pr(X(t) = j),$$

can be obtained as the (unique) solution to

$$p'(t) = p(t)Q \quad \text{satisfying} \quad p(0) = \alpha,$$
Evaluating state probabilities

The state probabilities $p(t) = (p_j(t), j \in S)$, where

$$p_j(t) = \Pr(X(t) = j),$$

can be obtained as the (unique) solution to

$$p'(t) = p(t) Q \quad \text{satisfying} \quad p(0) = a,$$

where $a = (a_j, j \in S)$ is a given initial distribution.

*Customary disclaimer*: It will be convenient to restrict our attention to the case where $S$ is a *finite* set, but I note that many of the arguments presented hold more generally.
% State probabilities (butterfly life cycle)

q(1)=1/14;  q(2)=1/7;  q(3)=1/14;  q(4)=1/4;
Q=zeros(5,5);
for i=2:5
    state=i-1;  % Matlab doesn't like a 0 index
    Q(i,i-1)=q(state);  Q(i,i)=-q(state);
end
i=5;  t=10;
P=expm(Q*t);  % The solution to p'(t)=p(t)Q
p=P(i,1:5);  % with p_4(0)=1
bar(0:4,p);
Individual organism

State probabilities ($t = 10$)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>0.1</td>
</tr>
<tr>
<td>Larva</td>
<td>0.6</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.2</td>
</tr>
<tr>
<td>Adult</td>
<td>0.1</td>
</tr>
<tr>
<td>Death</td>
<td>0.0</td>
</tr>
</tbody>
</table>
The state probabilities can almost never be evaluated analytically.
The state probabilities can almost never be evaluated analytically. There are exceptions . . .

Suppose that an organism has $M$ stages of life ($M = 4$ for the butterfly), and that the expected time spent in stage $j$ is $1/q_j$ ($q_j$ is the rate of departure from stage $j$).

**Exercise** (Grimmett and Stirzaker, Exercise 6.8.31): Show that if $q_1, q_2, \ldots, q_M$ are distinct, then

\[
p_j(t) = \frac{1}{q_j} \sum_{k=j}^{M} q_k e^{-q_k t} \prod_{l=j, l \neq k}^{M} \frac{q_l}{q_l - q_k},
\]

for $j = 1, \ldots, M$, and $p_0(t) = 1 - \sum_{j=1}^{M} p_j(t)$. 

Individual organism

State probabilities ($t = 0$)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>1.0</td>
</tr>
<tr>
<td>Larva</td>
<td>0.0</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.0</td>
</tr>
<tr>
<td>Adult</td>
<td>0.0</td>
</tr>
<tr>
<td>Death</td>
<td>0.0</td>
</tr>
</tbody>
</table>
State probabilities ($t = 1$)

- Egg: 0.8
- Larva: 0.2
- Pupa: 0
- Adult: 0
- Death: 0
Individual organism

State probabilities ($t = 2$)

- Egg: 0.6
- Larva: 0.4
- Pupa Stage: 0.05
- Adult: 0.005
- Death: 0.005

MASCOS APWSPM08, February 2008 - Page 35
Individual organism

State probabilities ($t = 3$)

- Egg: 0.5
- Larva: 0.5
- Pupa Stage: 0.1
- Adult: 0.0
- Death: 0.0
Individual organism

State probabilities ($t = 5$)

- Egg
- Larva
- Pupa Stage
- Adult
- Death

Probability
Individual organism

State probabilities (t = 10)

- Egg: 0.1
- Larva: 0.6
- Pupa Stage: 0.2
- Adult: 0.1
- Death: 0.0
Individual organism

State probabilities ($t = 20$)

- Egg: 0.0
- Larva: 0.3
- Pupa Stage: 0.2
- Adult: 0.1
- Death: 0.0
Individual organism

State probabilities ($t = 30$)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>0.0</td>
</tr>
<tr>
<td>Larva</td>
<td>0.1</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.2</td>
</tr>
<tr>
<td>Adult</td>
<td>0.3</td>
</tr>
<tr>
<td>Death</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Individual organism

State probabilities ($t = 50$)

- Egg
- Larva
- Pupa Stage
- Adult
- Death

<table>
<thead>
<tr>
<th>Stage</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>0</td>
</tr>
<tr>
<td>Larva</td>
<td>0.1</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>0.2</td>
</tr>
<tr>
<td>Adult</td>
<td>0.3</td>
</tr>
<tr>
<td>Death</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Individual organism

State probabilities ($t = 100$)

- Egg: 0
- Larva: 0
- Pupa Stage: 0
- Adult: 0
- Death: 1
Ensemble of organisms
Suppose that at time $t = 0$ the individuals are assigned to the states according to some rule and then each moves independently in $S$ as a Markov chain governed by $Q$.

The key assumption here is *independence*: individuals do not affect one another.
Suppose that at time $t = 0$ the individuals are assigned to the states according to some rule and then each moves independently in $S$ as a Markov chain governed by $Q$.

The key assumption here is *independence*: individuals do not affect one another.

We record only the *number* of individuals in the various states, rather than their positions.

Let $N_j(t)$ be the number of individuals in state $j$ at time $t$, and let $\mathbf{N} = (N_j, j \in S)$. The process $(\mathbf{N}(t), t \geq 0)$ is also a continuous-time Markov chain.
The ensemble model

Suppose that at time $t = 0$ the individuals are assigned to the states according to some rule and then each moves independently in $S$ as a Markov chain governed by $Q$.

The key assumption here is *independence*: individuals do not affect one another.

We record only the *number* of individuals in the various states, rather than their positions.

Let $N_j(t)$ be the number of individuals in state $j$ at time $t$, and let $N = (N_j, j \in S)$. The process $(N(t), t \geq 0)$ is also a continuous-time Markov chain.
The ensemble model

Suppose that at time $t = 0$ the individuals are assigned to the states according to some rule and then each moves independently in $\mathcal{S}$ as a Markov chain governed by $Q$.

The key assumption here is *independence*: individuals do not affect one another.

We record only the *number* of individuals in the various states, rather than their positions.

Let $N_j(t)$ be the number of individuals in state $j$ at time $t$, and let $\mathbf{N} = (N_j, j \in \mathcal{S})$. The process $(\mathbf{N}(t), t \geq 0)$ is also a continuous-time Markov chain.
Ensemble of organisms

Life cycle simulation ($n = 7$ butterflies)
Life cycle simulation \( (n = 7 \text{ butterflies}) \)

- **Egg**
- **Larva**
- **Pupa**
- **Adult**
- **Death**

Numbers at \( t = 0 \text{ days} \) \( (n = 7 \text{ butterflies}) \)
Ensemble state description

Life cycle simulation \((n = 7 \text{ butterflies})\)

Numbers at \(t = 1 \text{ days} \quad (n = 7 \text{ butterflies})\)
Life cycle simulation \((n = 7 \text{ butterflies})\)

Numbers at \(t = 2 \text{ days} \) \((n = 7 \text{ butterflies})\)

- **Egg**
- **Larva**
- **Pupa**
- **Adult**
- **Death**
Life cycle simulation \((n = 7\text{ butterflies})\)

Numbers at \(t = 3\) days \(\left(n = 7\text{ butterflies}\right)\)
Ensemble state description

Life cycle simulation \( (n = 7 \text{ butterflies}) \)

- Egg
- Larva
- Pupa
- Adult
- Death

Numbers at \( t = 4 \text{ days} \) \( (n = 7 \text{ butterflies}) \)

- Egg
- Larva
- Pupa Stage
- Adult
- Death

Numbers at \( t = 4 \text{ days} \) show a significant increase in the Larva stage compared to other stages.
Life cycle simulation ($n = 7$ butterflies)

- Egg
- Larva
- Pupa
- Adult
- Death

Numbers at $t = 5$ days ($n = 7$ butterflies)
Life cycle simulation ($n = 7$ butterflies)

- **Egg**
- **Larva**
- **Pupa**
- **Adult**
- **Death**

Numbers at $t = 6$ days ($n = 7$ butterflies)

- **Egg**
- **Larva**
- **Pupa Stage**
- **Adult**
- **Death**
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 7$ days ($n = 7$ butterflies)
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 8$ days ($n = 7$ butterflies)
Ensemble state description

Life cycle simulation \((n = 7 \text{ butterflies})\)

Numbers at \(t = 9 \text{ days} \) \((n = 7 \text{ butterflies})\)
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 10$ days ($n = 7$ butterflies)
Life cycle simulation \( (n = 7 \text{ butterflies}) \)

Numbers at \( t = 11 \text{ days} \) \( (n = 7 \text{ butterflies}) \)
Life cycle simulation ($n = 7$ butterflies)

- Egg
- Larva
- Pupa
- Adult
- Death

Numbers at $t = 12$ days ($n = 7$ butterflies)

- Egg
- Larva
- Pupa
- Adult
- Death
Ensemble state description

Life cycle simulation ($n = 7$ butterflies)

Time (days)

Egg
Larva
Pupa
Adult
Death

Numbers at $t = 13$ days ($n = 7$ butterflies)

Number of butterflies

Egg
Larva
Pupa
Adult
Death

MASCOS APWSPM08, February 2008 - Page 62
Life cycle simulation \((n = 7\) butterflies\)

- Egg
- Larva
- Pupa
- Adult
- Death

Numbers at \(t = 14\) days \((n = 7\) butterflies\):

- Egg: 1
- Larva: 5
- Pupa: 2
- Adult: 0
- Death: 0
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 15$ days ($n = 7$ butterflies)
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 16$ days ($n = 7$ butterflies)
Life cycle simulation ($n = 7$ butterflies)

- **Egg**
- **Larva**
- **Pupa**
- **Adult**
- **Death**

Numbers at $t = 17$ days ($n = 7$ butterflies)

- **Egg**
- **Larva**
- **Pupa**
- **Adult**
- **Death**
Life cycle simulation \( (n = 7 \text{ butterflies}) \)

Numbers at \( t = 18 \) days \( (n = 7 \text{ butterflies}) \)
Life cycle simulation ($n = 7$ butterflies)

Numbers at $t = 19$ days ($n = 7$ butterflies)
Life cycle simulation \((n = 7\) butterflies\)}

Numbers at \(t = 20\) days \((n = 7\) butterflies\)}
The closed ensemble. We suppose that there is a fixed number $n$ of individuals, each moving according to $Q$. The process takes values in

$$E = \{ n \in \{0, \ldots, n\}^S : \sum_{j \in S} n_j = n \},$$

and its transition rates $Q_E = (q(n, m), n, m \in E)$ are given by

$$q(n, n + e_j - e_i) = n_i q_{ij},$$

for all states $j \neq i$ in $S$, where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is the unit vector with a 1 as its $j$-th entry (this transition corresponds to a single individual moving from state $i$ to state $j$).
Numbers at $t = 10$ days ($n = 7$ butterflies)
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 7$ butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>60%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>20%</td>
</tr>
<tr>
<td>Death</td>
<td>10%</td>
</tr>
</tbody>
</table>
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 15$ butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>50%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>20%</td>
</tr>
<tr>
<td>Death</td>
<td>10%</td>
</tr>
</tbody>
</table>
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 20$ butterflies)
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 50$ butterflies)

- Egg: 10%
- Larva: 20%
- Pupa: 30%
- Adult: 40%
- Death: 50%
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 100$ butterflies)
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 200$ butterflies)

- Egg: 10%
- Larva: 50%
- Pupa Stage: 20%
- Adult: 10%
- Death: 70%
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 500$ butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>20%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>40%</td>
</tr>
<tr>
<td>Death</td>
<td>50%</td>
</tr>
</tbody>
</table>
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 1000$ butterflies)

- Egg: 10%
- Larva: 20%
- Pupa Stage: 30%
- Adult: 40%
- Death: 50%

Stage Proportion
Ensemble proportions (simulation)

Proportions at \( t = 10 \) days \((n = 2000\) butterflies\))

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>20%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>40%</td>
</tr>
<tr>
<td>Death</td>
<td>50%</td>
</tr>
</tbody>
</table>

Note: The graph shows the proportions of butterflies at each life stage at the end of the 10th day.
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 3000$ butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>20%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>40%</td>
</tr>
<tr>
<td>Death</td>
<td>50%</td>
</tr>
</tbody>
</table>

Stage proportions at $t = 10$ days for 3000 butterflies.
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 5000$ butterflies)

- Egg: 10%
- Larva: 50%
- Pupa Stage: 20%
- Adult: 10%
- Death: 5%
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 10000$ butterflies)

- **Egg**: 10%
- **Larva**: 20%
- **Pupa Stage**: 30%
- **Adult**: 40%
- **Death**: 50%

Stage Proportion
Ensemble proportions (simulation)

Proportions at \( t = 10 \) days (\( n = 20000 \) butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>20%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>40%</td>
</tr>
<tr>
<td>Death</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td>70%</td>
</tr>
</tbody>
</table>
Ensemble proportions (simulation)

Proportions at $t = 10$ days ($n = 50000$ butterflies)

<table>
<thead>
<tr>
<th>Stage</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>10%</td>
</tr>
<tr>
<td>Larva</td>
<td>20%</td>
</tr>
<tr>
<td>Pupa Stage</td>
<td>30%</td>
</tr>
<tr>
<td>Adult</td>
<td>40%</td>
</tr>
<tr>
<td>Death</td>
<td>50%</td>
</tr>
<tr>
<td>Death</td>
<td>60%</td>
</tr>
</tbody>
</table>

Stage proportions at $t = 10$ days for $n = 50000$ butterflies.
Let $X^{(n)}(t) = \frac{N(t)}{n}$, where $n$ is the number of individuals, so that $X^{(n)}_j(t)$ is the proportion if individuals in state $j$. 

**Theorem 1.** If ..., then, for all ..., and for every ..., as ..., where is the unique solution to ..., namely ..., where is the matrix exponential.
Convergence of ensemble proportions

Let $X^{(n)}(t) = N(t)/n$, where $n$ is the number of individuals, so that $X_j^{(n)}(t)$ is the proportion if individuals in state $j$.

**Theorem 1.** If $X^{(n)}(0) \to a$ as $n \to \infty$, then, for all $u > 0$, and for every $\epsilon > 0$,

$$\Pr \left( \sup_{0 \leq t \leq u} \left| X^{(n)}(t) - p(t) \right| > \epsilon \right) \to 0 \quad \text{as } n \to \infty,$$

where $p(t) = (p_j(t), j \in S)$ is the unique solution to $p'(t) = p(t) Q$ satisfying $p(0) = a$, namely $p(t) = a \exp(tQ)$, where $\exp(\cdot)$ is the matrix exponential.
Convergence of ensemble proportions

2-Stage life cycle (n = 100 organisms)
Convergence of ensemble proportions

2-Stage life cycle ($n = 200$ organisms)
Convergence of ensemble proportions

2-Stage life cycle \((n = 500 \text{ organisms})\)

- Proportion in Stage A
- Proportion in Stage B

- Diagram showing the relationship between the proportion in Stage A and the proportion in Stage B for a 2-stage life cycle with 500 organisms.
Convergence of ensemble proportions

2-Stage life cycle ($n = 1000$ organisms)

Proportion in Stage A vs. Proportion in Stage B

- Blue line: Ensemble proportions
- Red line: Theoretical distribution

Figure showing the convergence of ensemble proportions in a 2-stage life cycle with $n = 1000$ organisms.
Convergence of ensemble proportions

2-Stage life cycle \((n = 2000\text{ organisms})\)

Proportion in Stage A vs Proportion in Stage B
Convergence of ensemble proportions

2-Stage life cycle (\(n=5000\) organisms)
Convergence of ensemble proportions

2-Stage life cycle ($n=10000$ organisms)
Convergence of ensemble proportions

Proportions at $t = 10$ days ($n = 100$ butterflies)
Convergence of ensemble proportions

2-Stage life cycle \((n = 100 \text{ organisms})\)
Theorem 2. In the setup of Theorem 1, let

$$Z^{(n)}(t) = \sqrt{n}(X^{(n)}(t) - p(t)).$$

If $Z^{(n)}(0) \to z$ as $n \to \infty$, then $(Z^{(n)}(t))$ converges weakly in $D[0, t]$ (the space of right-continuous, left-hand limits functions on $[0, t]$) to a Gaussian diffusion $(Z(t))$ with initial value $Z(0) = z$ and with mean and covariance given by

$$\mu_s := \mathbb{E}(Z(s)) = e^{sQ^\top} z$$

and

$$V_s := \text{Cov}(Z(s)) = e^{sQ^\top} \left( \int_0^s e^{-uQ^\top} G(p(u))e^{-uQ} du \right) e^{sQ},$$
Theorem 2 (continued).

... where the matrix $G(x)$ has entries

$$G_{kk}(x) = x_k q_k + \sum_{i \neq k} x_i q_{ik} \quad \text{and} \quad G_{kl}(x) = -(x_l q_{lk} + x_k q_{kl}).$$
Theorem 2 (continued).

... where the matrix $G(x)$ has entries

$$G_{kk}(x) = x_kq_k + \sum_{i \neq k} x_iq_{ik} \quad \text{and} \quad G_{kl}(x) = -(x_lq_{lk} + x_kq_{kl}).$$

Theorem 2 has many implications. One immediate one is that the population proportions $X^{(n)}(t)$ have an approximate multivariate Gaussian (normal) distribution with known mean vector and covariance matrix.

This helps explain the observed fluctuations (now seen to be of order $1/\sqrt{n}$) of $X^{(n)}(t)$ about $p(t)$.
Proportions at $t = 10$ days ($n = 100$ butterflies)

- Egg: 10%
- Larva: 20%
- Pupa Stage: 30%
- Adult: 40%
- Death: 50%
Convergence of scaled fluctuations

2-Stage life cycle ($n = 100$ organisms)

Scaled fluctuations (Stage A)

Time

MASCOS APWSPM08, February 2008 - Page 99
Convergence of scaled fluctuations

2-Stage life cycle ($n=200$ organisms)

Scaled fluctuations (Stage A)

Time

MASCOS APWSPM08, February 2008 - Page 100
Convergence of scaled fluctuations

2-Stage life cycle \((n = 500\) organisms)
Convergence of scaled fluctuations

2-Stage life cycle ($n = 1000$ organisms)
Convergence of scaled fluctuations

2-Stage life cycle ($n = 2000$ organisms)
Convergence of scaled fluctuations

2-Stage life cycle ($n = 5000$ organisms)
Convergence of scaled fluctuations

2-Stage life cycle (n = 5000 organisms)

- Open ensembles

- Open ensembles
- Stationary behaviour
Further details


- Open ensembles
- Stationary behaviour
- Quasi-stationary behaviour

- Open ensembles
- Stationary behaviour
- Quasi-stationary behaviour
  - Quasi-stationary distributions (QSDs) for *reducible* Markov chains
  - QSDs for ensemble processes
In our general setup (with $C$ being the set of transient states and $\alpha$ being the decay parameter) …

**Theorem 3.** Let $\pi = (\pi_j, j \in C)$ be the QSD of the individual process. If the initial numbers $N_j(0), j \in C$, are chosen independently with $N_j(0)$ having a Poisson distribution with mean $\pi_j$, then, for all $t > 0$, $N_j(t), j \in C$, are independent with $N_j(t)$ having a Poisson distribution with mean $\pi_j e^{-\alpha t}$.

*For aficionados.* This result holds in much greater generality; $C$ need not be finite, $Q$ could be explosive, $\pi = (\pi_j, j \in C)$ could be any $\alpha$-subinvariant measure and, more remarkably still, $\pi$ need not be finite (we could have $\sum_{j \in C} \pi_j = \infty$).