The fast fragmentation-coalescence process

Steven Pagett

University of Bath

May 16, 2016

[1] Joint work with Tim Rogers, Andreas Kyprianou and Jason Schweinsberg

Kingman's Coalescent

The Kingman n-coalescent is a coalescent process that was developed to describe the relations between a selection of haploids in a population.

Kingman's Coalescent

The Kingman n-coalescent is a coalescent process that was developed to describe the relations between a selection of haploids in a population.

This is a Markov chain on the partitions of $\{1, \ldots, n\}$ often denoted by

$$\Pi^{(n)}(t) = \{\Pi_1^{(n)}(t), \ldots, \Pi_n^{(n)}(t)\},$$

where $\Pi_i^{(n)}(t)$ is the subset of elements that make up the i^{th} block of the partition. These are ordered by smallest element and some of them are allowed to be empty.

Kingman's Coalescent

Start with everyone in their own block i.e.

$$\Pi^{(n)}(0) = \{\{1\}, \ldots, \{n\}\},$$

then every pair of blocks merges at rate $c > 0$.

Kingman's Coalescent

Start with everyone in their own block i.e.

$$\Pi^{(n)}(0) = \{\{1\}, \ldots, \{n\}\},$$

then every pair of blocks merges at rate $c > 0$.

Hence if N_t is the number of non-empty blocks of $\Pi^{(n)}(t)$, then $N_0 = n$, and N_t is a pure death process with death rate $c\binom{N_t}{2}$.

Example

$$\{\{1\},\{2\},\{3\},\{4\},\{5\}\}$$
$$\downarrow$$
$$\{\{1,2\},\{3\},\{4\},\{5\}\}$$
$$\downarrow$$
$$\{\{1,2\},\{3\},\{4,5\}\}$$
$$\downarrow$$
$$\{\{1,2,4,5\},\{3\}\}$$
$$\downarrow$$
$$\{\{1,2,3,4,5\}\}$$

Extension to $\mathbb{N}$

This can be extended naturally to a process, Π, on partitions of $\mathbb{N}$, such that when you restrict Π to $\{1, \ldots, n\}$ you get an n-coalescent.

Extension to $\mathbb{N}$

This can be extended naturally to a process, Π, on partitions of $\mathbb{N}$, such that when you restrict Π to $\{1, \ldots, n\}$ you get an n-coalescent.

An interesting result for this extension is that although $N_0 = \infty$, we have that

$$\mathbb{P}(\inf\{t > 0 : N_t < \infty\} = 0) = 1$$

This is known as "coming down from infinity".

Adding Fragmentation

These processes can be generalised to allow fragmentation to also occur. This works by picking a block of the partition and splitting it up into smaller blocks.

Adding Fragmentation

These processes can be generalised to allow fragmentation to also occur. This works by picking a block of the partition and splitting it up into smaller blocks.

The main questions that are asked about these processes are

1. How many blocks are there?
2. How big are these blocks?

Example

$$\{\{1,2,3,4,5,6,7\}\}$$
$$\downarrow$$
$$\{\{1,3,5\},\{2,6,7\},\{4\}\}$$
$$\downarrow$$
$$\{\{1,3\},\{2,6,7\},\{4\},\{5\}\}$$
$$\downarrow$$
$$\{\{1,3,4\},\{2,6,7\},\{5\}\}$$
$$\downarrow$$
$$\{\{1,3,4\},\{2\},\{5\},\{6\},\{7\}\}$$
$$\downarrow$$
$$\{\{1,3,4,7\},\{2\},\{5\},\{6\}\}$$

Kingman + Fragmentation

(Berestycki 2004) If you have Kingman's coalescent and let the fragmentation satisfy

- Rate of fragmenting each block is finite
- It fragments a block into finitely many smaller blocks
- and if

  $$p_k = \mathbb{P}(\text{Fragment a block into } k+1 \text{ smaller blocks}),$$

  and

  $$\sum_{k=1}^{\infty} p_k \log p_k < \infty$$

then you still come down from infinity, i.e. you have finitely many blocks immediately after time 0.

Question

Is there any finite rate fragmentation mechanism which will stop Kingman's coalescent from coming down from infinity?

The 'fast' EFC process

Coalescent part: Kingman at rate c

The 'fast' EFC process

Coalescent part: Kingman at rate c
Fragmentation part: Each block fragments at rate λ; when it fragments it completely fragments into singletons.

The 'fast' EFC process

Coalescent part: Kingman at rate c
Fragmentation part: Each block fragments at rate λ; when it fragments it completely fragments into singletons.

We know how N_t behaves in finite states

$$Q^N_{i,j} = \begin{cases} c\binom{i}{2} & \text{if } j = i-1,\ i < \infty \\ \lambda i & \text{if } i < \infty,\ j = \infty \end{cases}$$

The 'fast' EFC process

Coalescent part: Kingman at rate c
Fragmentation part: Each block fragments at rate λ; when it fragments it completely fragments into singletons.

We know how N_t behaves in finite states

$$Q^N_{i,j} = \begin{cases} c\binom{i}{2} & \text{if } j = i-1,\ i < \infty \\ \lambda i & \text{if } i < \infty,\ j = \infty \end{cases}$$

How does the process behave when $N_t = \infty$? Does it come down from infinity?

Theorem

It turns out that there is a phase transition in $\theta := 2\lambda/c$.

Theorem

(i) *If* $\theta < 1$, *then* $N := (N_t : t \geq 0)$ *is a recurrent Feller process on* $\mathbb{N} \cup \{\infty\}$ *such that* ∞ *is instantaneously regular and not sticky.*

(ii) *If* $\theta \geq 1$, *then* ∞ *is an absorbing state for* N*.*

Local Time

If $\theta < 1$, then this theorem tells us that there exists a local time for N at ∞, L_t. We can also develop an excursion theory for N, and so the periods when $N_t < \infty$ can be thought of as excursions from ∞.

Local Time

If $\theta < 1$, then this theorem tells us that there exists a local time for N at ∞, L_t. We can also develop an excursion theory for N, and so the periods when $N_t < \infty$ can be thought of as excursions from ∞.

This allows us to prove a few more results about N.

Theorem

If $\theta < 1$, then

1. *N has a stationary distribution given by*

$$\rho_N(k) = \frac{1-\theta}{\Gamma(\theta)} \frac{\Gamma(k-1+\theta)}{\Gamma(k+1)}, \quad k \in \mathbb{N}.$$

2. $\lim_{t \to 0} tN_t = 2/c$.

Open Questions

1a. Subcritical Case

Excursion Theory & Local Time

We have an excursion measure $\mathbb{Q}$ when $\theta < 1$,

- The inverse local time L_t^{-1}, defined as

$$L_t^{-1} = \inf\{s > 0 : L_s > t\}$$

  This can be thought of as a measure of excursion length
- What is the distribution of the inverse local time, L_t^{-1}, which reduces to calculating Φ where

$$\mathbb{Q}\left[1 - \exp(-q\zeta)\right] = \exp(-t\Phi(q))$$

  where ζ is the length of a typical excursion.

Open Questions

1b. Subcritical Case

Stationary Distribution

- We have the stationary distribution for N.

$$\rho_N(k) = \frac{1-\theta}{\Gamma(\theta)} \frac{\Gamma(k-1+\theta)}{\Gamma(k+1)}, \quad k \in \mathbb{N}.$$

- Can we extend this to the stationary distribution of Π? This is equivalent to finding the stationary distribution of the asymptotic frequencies of the blocks, given there are k blocks.

Open Questions

2. Critical vs. Supercritical Case

- Currently all we know about the cases $\theta = 1$ and $\theta > 1$ is that $N_t = \infty$ for all t almost surely. Do any differences exist?
- In these two cases is $\rho_N(\infty) > 0$?

Open Questions

3. The First Block

Can we find the distribution of the asymptotic frequency of the first block?
The asymptotic frequency is a measure of the 'size' of a block

$$|\Pi_1(t)| := \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{j \in \Pi_1(t)\}}$$

We have the first two moments of this asymptotic frequency, but further moments are very hard to calculate at the moment.

Open Questions

4. EFC Processes

How do these results fit in the wider context of exchangeable fragmentation-coalescence processes?

- There is an entire class of coalescent processes that come down from infinity. Is there some 'measure' of how strong fragmentation needs to be to prevent this from occurring?
- Can this phase transition be found elsewhere?