Exchangeable representations of measure-valued population processes I

work with Peter Donnelly [1, 2]

A simple population model: Wright-Fisher

n population size (constant in time)

Two types $0, 1$

X_m the fraction of the population in generation m that is type 1

μ_{01}, μ_{10} mutation probabilities

$$P\{X_{m+1} = \frac{k}{n} | X_m = x\} = \binom{n}{k} (x(1-\mu_{10}) + (1-x)\mu_{01})^k ((1-x)(1-\mu_{01}) + x\mu_{10})^{n-k}$$

$$E[X_{m+1}|\mathcal{F}_m] = E[X_{m+1}|X_m] = X_m(1-\mu_{10}) + (1-X_m)\mu_{01} = X_m(1-\mu_{10}-\mu_{01}) + \mu_{01}$$

If $\mu_{01} = \mu_{10} = 0$, $E[X_{m+1}|\mathcal{F}_m] = X_m$ and

$$E[(X_{m+1} - X_m)^2|\mathcal{F}_m] = \frac{1}{n} X_m(1-X_m)$$

Correct time scale: n generations per unit time or equivalently n^2 births per unit time.

Simple Moran model

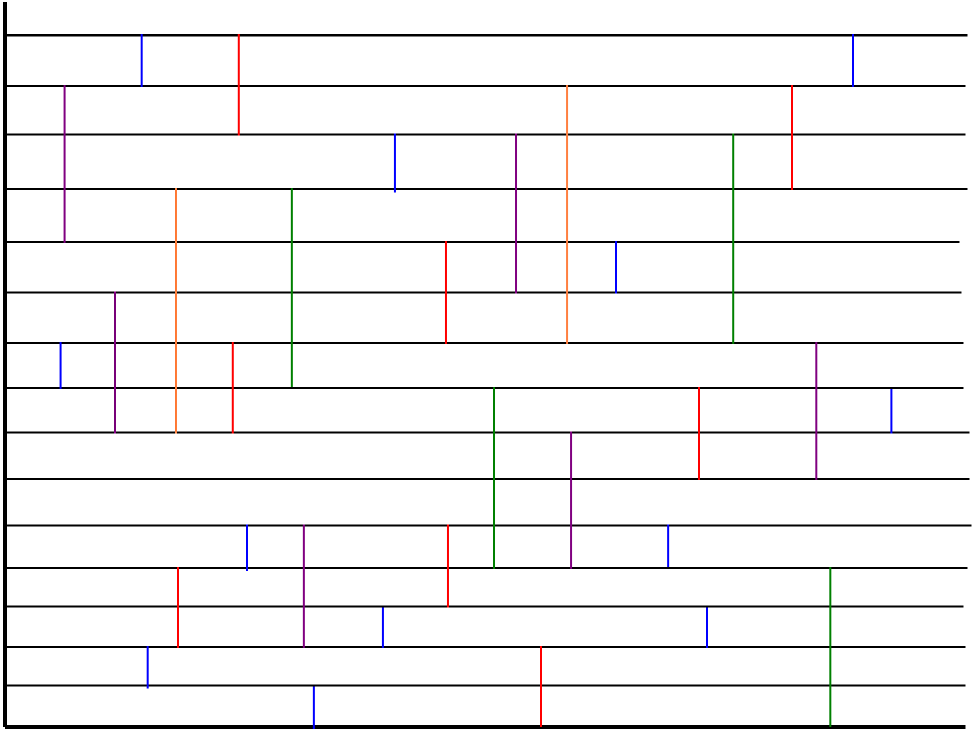
Specify a continuous time Markov model by its infinitesimal behavior determining the process as the solution of a martingale problem

$$E[f(X(t+\Delta t))|\mathcal{F}_t] = f(X(t)) + Af(X(t))\Delta t + o(\Delta t)$$

Scale mutation probabilities by n^{-1}

$$\begin{aligned} Af(x) &= \frac{\lambda}{2}n^2(x(1-n^{-1}\mu_{10}) + (1-x)n^{-1}\mu_{01})(1-x)(f(x+\frac{1}{n}) - f(x)) \\ &\qquad + \frac{\lambda}{2}n^2((1-x)(1-n^{-1}\mu_{01}) + xn^{-1}\mu_{10})x(f(x-\frac{1}{n}) - f(x)) \\ &\approx \frac{1}{4}\lambda(x(1-x)(2-n^{-1}(\mu_{10}+\mu_{01})) + (1-x)^2n^{-1}\mu_{01} + x^2n^{-1}\mu_{10})f''(x) \\ &\qquad + \frac{1}{2}\lambda(x(1-x)(\mu_{01}-\mu_{10}) + (1-x)^2\mu_{01} + x^2\mu_{10})f'(x) \\ &\approx \frac{\lambda}{2}x(1-x)f''(x) + \frac{\lambda}{2}((1-x)\mu_{01} - x\mu_{10})f'(x) \end{aligned}$$

The calculation suggests approximating the Markov chain by a diffusion process.



Two population models

Model I

$(X_1(t), \ldots, X_n(t))$ "types" of n individuals in a population at time t

At rate λ (per pair), a pair of individuals is selected at random, one is killed and replaced by a copy of the other

In between birth/death events, individuals may change type (mutate) independently of the other individuals in the population

Model II

Same as Model I except that when the pair is selected the lower numbered individual is copied and inserted at the the higher level. The individual at the top is killed.

Equivalence of models

Theorem 1 *Let X^I be a realization of model I and X^{II} be a realization of model II with $X^I(0) = X^{II}(0)$. Define*

$$Z^I(t) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i^I(t)}, \qquad Z^{II}(t) = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i^{II}(t)}.$$

If $\{X_i^{II}(0)\}$ is exchangeable, then for each $t > 0$, $\{X_i^{II}(t)\}$ is exchangeable and Z^I and Z^{II} have the same distribution.

$\{\xi_1, \ldots, \xi_n\}$ is *exchangeable* if the distribution of $\{\xi_{\sigma_1}, \ldots, \xi_{\sigma_n}\}$ is the same for every permutation $(\sigma_1, \ldots, \sigma_n)$ of $(1, \ldots, n)$.

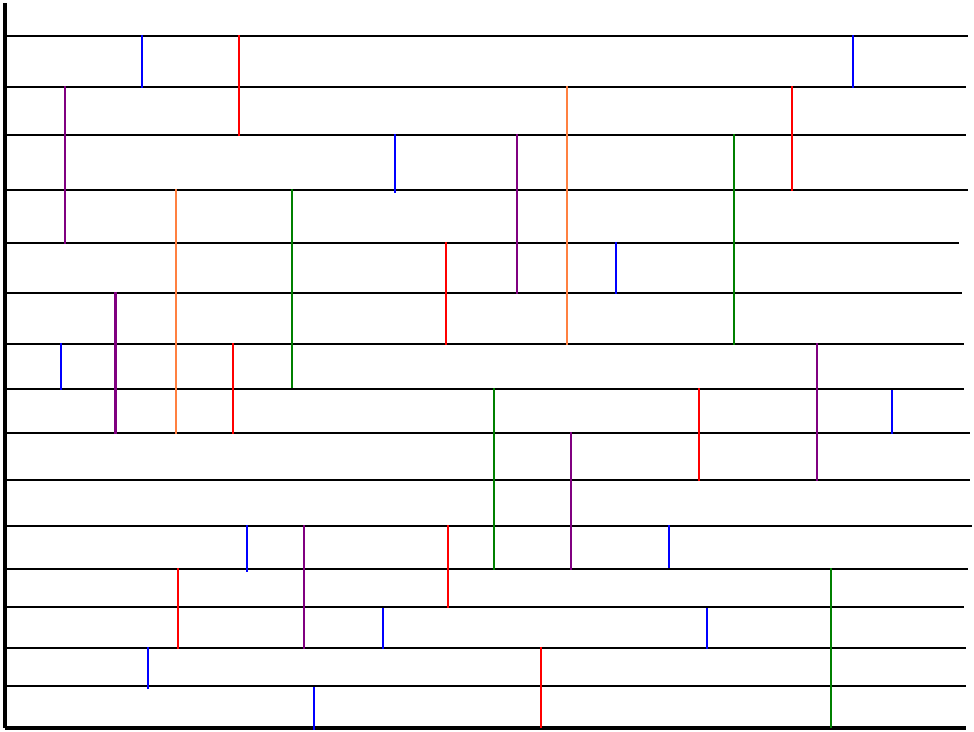
Genealogy

Consider the process looking backward in time from some fixed time T.

$a_i^T(t)$ the index of the ancestor at time $T - t$ of the ith particle alive at time T

$i \overset{t}{\sim} j$ if $a_i(t) = a_j(t)$ (an equivalence relation)

If $a_i(t) = a_j(t)$, then $a_i(s) = a_j(s)$ for all $s > t$. (The ancestral lines have coalesced.)



Coalescent

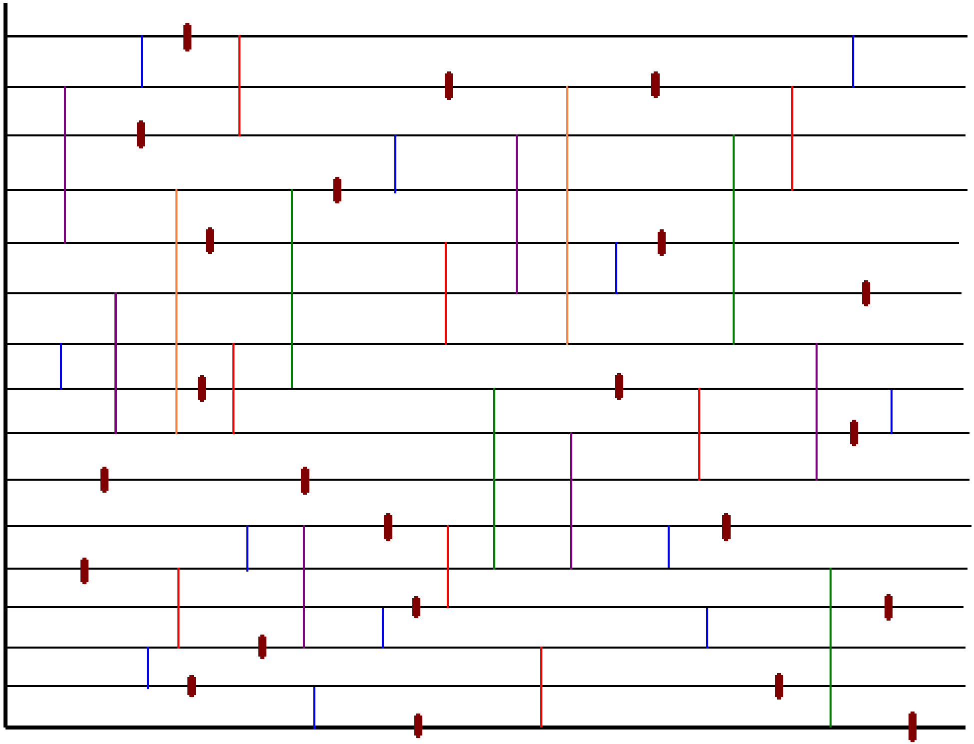
$\Gamma(t) = (\Gamma_1(t), \ldots, \Gamma_{N(t)}(t))$ collection of equivalence classes

$\Gamma(t+\Delta t) \neq \Gamma(t)$ if there is a birth event involving two distinct indices among $\{a_i(t), i = 1, \ldots, n\}$

Claim: $P\{\Gamma(t+\Delta t) \neq \Gamma(t) | \Gamma(t)\} \approx \binom{N(t)}{2} \Delta t$ (take $\lambda = 1$)

If $N(t+\Delta t) = N(t) - 1$, $\Gamma(t+\Delta t)$ is obtained from $\Gamma(t)$ by combining two of the equivalence classes, selected at random

The coalescent determines a tree. The mutations can be placed on the tree going forward in time.



Mutation process

E type space

$\frac{\theta}{2}$ mutation rate on each level

$\mu(dz|x)$ distribution of mutant given original type

$B_i f(x) = \frac{\theta}{2} \int_E (f(\eta_i(x|z)) - f(x))\mu(dz|x_i)$

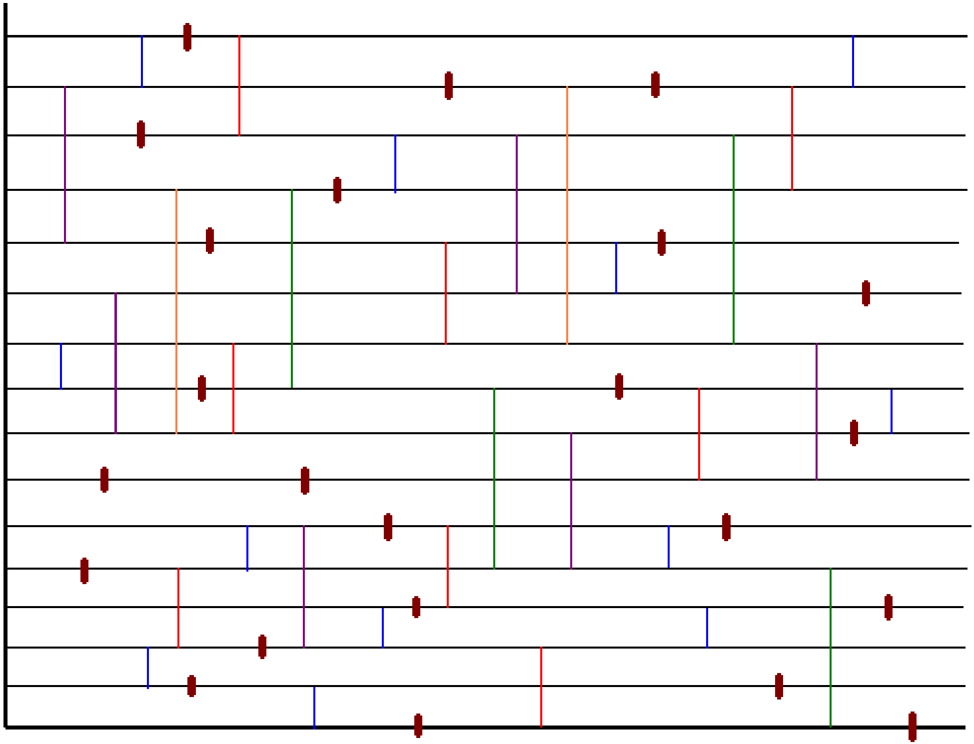
$\eta_i(x|z) = (x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$

Generator: $A^n f(x) = \sum_{i=1}^n B_i f(x) + \sum_{1 \le i < j \le n} (f(\theta_{ij}(x)) - f(x))$

$y = \theta_{ij}(x)$ is the element of E^n satisfying

$$\begin{aligned} y_k &= x_k, \quad k \le j-1 \\ y_j &= x_i \\ y_k &= x_{k-1}, \quad k > j \end{aligned}$$

that is $\theta_{ij}(x) = (x_1, \ldots, x_{j-1}, x_i, x_j, \ldots, x_{n-1})$



Infinite population limit

$f(x) = f(x_1, \ldots, x_m)$

Generator: $Af(x) = \sum_{i=1}^{m} B_i f(x) + \sum_{1\le i<j\le m}(f(\theta_{ij}(x)) - f(x))$

$$\begin{aligned} E[f(X_1(t+\Delta t), \ldots, X_m(t+\Delta t))|\mathcal{F}_t^X] &\approx f(X_1(t), \ldots, X_m(t)) \\ &\quad +Af(X_1(t), \ldots, X_m(t))\Delta t \end{aligned}$$

For $Z(t) = \lim_{m\to\infty} \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(t)}$

$$\begin{aligned} E[f(X_1(t+\Delta t), \ldots, X_m(t+\Delta t))|\mathcal{F}_t^Z] &= E[\langle f, Z(t+\Delta t)^m\rangle|\mathcal{F}_t^Z] \\ &\approx \langle f, Z(t)^m\rangle + \langle Af, Z(t)^m\rangle\Delta t \end{aligned}$$

$$\langle Af, Z(t)^m\rangle = \sum_{i=1}^{m}\langle B_i f, Z(t)^m\rangle + \sum_{1\le i<j\le m}(\langle \Phi_{ij} f, Z(t)^{m-1}\rangle - \langle f, Z(t)^m\rangle)$$

Measure-valued limit

(Fleming-Viot process)

$F(\mu) = \langle f, \mu^m \rangle$

$$\mathbb{A}F(\mu) \equiv \sum_{i=1}^{m} \langle B_i f, \mu^m \rangle + \sum_{1 \leq i < j \leq m} (\langle \Phi_{ij} f, \mu^{m-1} \rangle - \langle f, \mu^m \rangle)$$

For $f \in B(E^m)$, $\Phi_{ij} f \in B(E^{m-1})$ is the function obtained by setting the ith and jth variables equal.

Z is a solution of the martingale problem for $\mathbb{A}$.

Two types

$E = \{0, 1\}$

$f(x_1, \ldots, x_m) = \prod_{i=1}^m x_i$

$$Bf(0) = \mu_{01}(f(1) - f(0)) \qquad Bf(1) = \mu_{10}(f(0) - f(1))$$

If $Z(t) = z\delta_1 + (1 - z)\delta_0$, then

$$\begin{aligned} \langle Af, Z(t)^m \rangle &= m\mu_{01}(1 - z)z^{m-1} - m\mu_{10}z^m + \binom{m}{2}(z^{m-1} - z^m) \\ &= (\mu_{01}(1 - z) - \mu_{10}z)mz^{m-1} + \frac{1}{2}z(1 - z)m(m - 1)z^{m-2} \\ &= (\mu_{01}(1 - z) - \mu_{10}z)g'(z) + \frac{1}{2}z(1 - z)g''(z) \end{aligned}$$

for $g(z) = z^m$.

Examples

Infinitely many alleles model

$$Bf(x) = \frac{1}{2}\theta \int_E (f(y) - f(x))\nu(dy)$$

with ν nonatomic (each mutation is new).

Infinite sites model

$E = [0, 1]^\infty$

$$Bf(x) = \frac{1}{2}\theta \int_0^1 (f((z, x)) - f(x))dz$$

Branching process

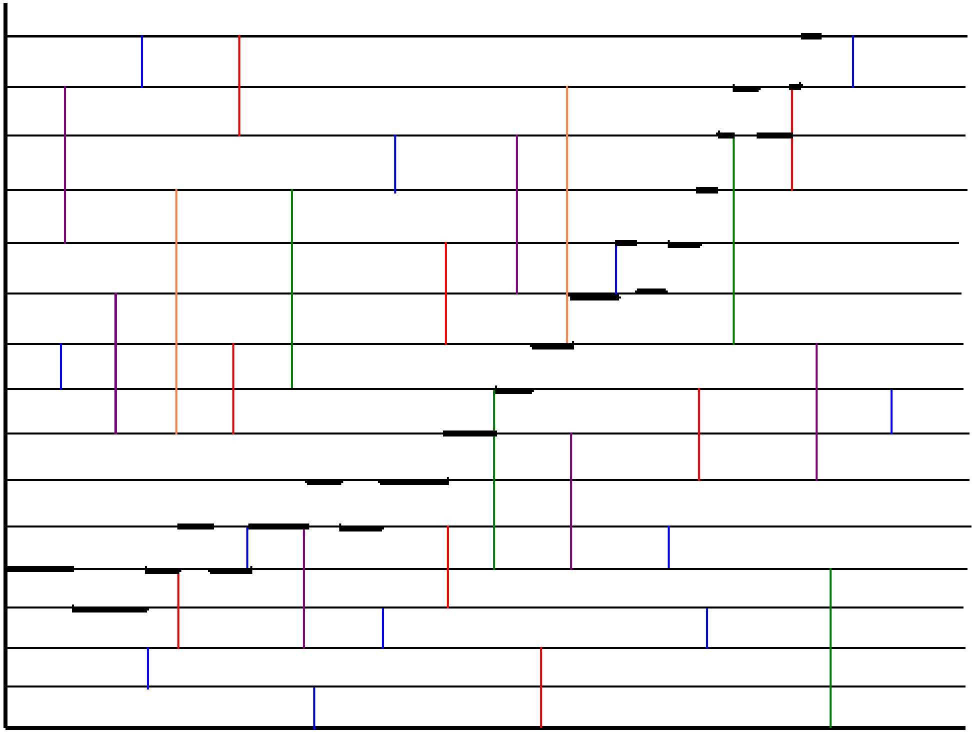
$$P^n(t) = P^n(0) + \frac{1}{n}Y_b(\int_0^t \lambda n^2 P^n(s)ds) - \frac{1}{n}Y_d(\int_0^t \lambda n^2 P^n(s)ds)$$

Assuming $P^n(0) \Rightarrow P(0)$, then $P^n \Rightarrow P$ satisfying

$$P(t) = P(0) + W_b(\int_0^t \lambda P(s)ds) - W_d(\int_0^t \lambda P(s)ds),$$

that is, a diffusion with generator

$$Gg(p) = \lambda p g''(p)$$



“Lookdown rate”

$P^n(t) = \frac{N^n(t)}{n}$

For $i < j \leq N^n(t) + 1$, the probability that i and j are in a birth event before time $t + \Delta t$ is approximately

$$\frac{\lambda n^2 P^n(t)}{\binom{N^n(t)+1}{2}} \Delta t \Rightarrow \frac{2\lambda}{P(t)} \Delta t$$

The Dawson-Watanabe process is given by $K(t) = P(t)Z(t)$

More general birth and death processes

More generally, we can consider birth and death processes satisfying

$$\begin{aligned} N_b^n(t) &= V_1(n^2\int_0^t \lambda_n(P^n(s))ds) + V_3(n^2\int_0^t \tilde{\lambda}_n(P^n(s))ds) \\ N_d^n(t) &= V_2(n^2\int_0^t \mu_n(P^n(s))ds) + V_3(n^2\int_0^t \tilde{\lambda}_n(P^n(s))ds) \\ P^n(t) &= \frac{N^n(t)}{n} = P^n(0) + \frac{1}{n}N_b^n(t) - \frac{1}{n}N_d^n(t)\ . \end{aligned}$$

If $P^n(0) \Rightarrow P(0)$ and $\lambda_n(\cdot) \to \lambda(\cdot)$, $\tilde{\lambda}_n(\cdot) \to \tilde{\lambda}(\cdot)$, and $n(\lambda_n(\cdot) - \mu_n(\cdot)) \to b(\cdot)$ uniformly on compact sets, then P^n converges to a solution of

$$P(t) = P(0) + W_1(\int_0^t \lambda(P(s))ds) - W_2(\int_0^t \lambda(P(s))ds) + \int_0^t b(P(s))ds\ . \tag{1}$$

Limiting lookdown rate

$$\frac{N_b^n(\cdot)}{n^2} \Rightarrow \int_0^{\cdot} (\lambda(P(s)) + \tilde{\lambda}(P(s)))ds \ ,$$

provided the solution of (1) does not blow-up in finite time. In this case, P is a diffusion with generator

$$Gf(z) = \lambda(z)f''(z) + b(z)f'(z)$$

(see [4], Theorem 6.5.4) and

$$\frac{n^2\lambda(P^n(t)) + \tilde{\lambda}(P^n(t))}{\binom{N^n(t)+1}{2}}\Delta t \Rightarrow \frac{2(\lambda(P(t)) + \tilde{\lambda}(P(t)))}{P(t)^2}\Delta t$$

Conditioning the branching model

Etheridge and March (1991) [3]:

Theorem 2 *The neutral Dawson-Watanabe process conditioned to have total mass identically 1 for all* $t \geq 0$ *is a Fleming-Viot process.*

Type distribution at the extinction time

Let $\tau = \inf\{t : P(t) = 0\}$. If P is a continuous state branching process, then

$$\int_0^\tau \frac{1}{P(s)} ds = \infty$$

Tribe (1992) [7]

$$\lim_{t \to \tau-} Z(t) = \delta_{\xi_0}, \qquad \xi_0 = X_1(\tau)$$

Theorem 3 *Let τ be an $\{\mathcal{F}_t^P\}$-stopping time. Suppose*

$$\int_0^\tau \frac{\lambda(P(s))}{P(s)^2} ds = \infty$$

on $\{\tau < \infty\}$. Then on $\{\tau < \infty\}$,

$$\lim_{t \to \tau-} Z(t) = \delta_{X_1(\tau-)} \; .$$

Conditioning on nonextincition

Assuming $Gf(v) = avf''(v) - bvf'(v)$ $(b \geq 0)$, conditioning P on nonextinction (cf. Evans and Perkins (1990) [6]) is equivalent to replacing P with a process $\hat{P}$ with generator

$$\hat{G}f(v) = avf''(v) + (2a - bv)f'(v). \tag{2}$$

If $\hat{P}(0) > 0$, then $\hat{P}$ never hits zero, but

$$\int_0^\infty \frac{c}{\hat{P}(s)} ds = \infty.$$

It follows that eventually all particles trace their ancestory back to the bottom-level particle. In particular, the bottom-level particle in our construction is the "immortal particle" of Evans (1993) [5].

Martingale problem

E state space (a complete, separable metric space)

A generator (a linear operator with domain and range in $B(E)$

$\mu \in \mathcal{P}(E)$

X is a solution of the martingale problem for (A, μ) if and only if $\mu = PX(0)^{-1}$ and there exists a filtration $\{\mathcal{F}_t\}$ such that

$$f(X(t)) - \int_0^t Af(X(s))ds$$

is an $\{\mathcal{F}_t\}$-martingale for each $f \in \mathcal{D}(A)$

Examples of generators

Standard Brownian motion ($E = \mathbb{R}^d$)

$$Af = \frac{1}{2}\Delta f, \quad \mathcal{D}(A) = C_c^2(\mathbb{R}^d)$$

Poisson process ($E = \{0, 1, 2 \ldots\}$, $\mathcal{D}(A) = B(E)$)

$$Af(k) = \lambda(f(k+1) - f(k))$$

Pure jump process (E arbitrary)

$$Af(x) = \lambda(x) \int_E (f(y) - f(x))\mu(x, dy)$$

Diffusion ($E = \mathbb{R}^d$)

$$Af(x) = \frac{1}{2}\sum_{i,j} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} f(x) + \sum_i b_i(x)\frac{\partial}{\partial x_i} f(x), \quad \mathcal{D}(A) = C_c^2(\mathbb{R}^d)$$

Uniqueness and the Markov property

Theorem 4 *If any two solutions of the martingale problem for A satisfying $PX_1(0)^{-1} = PX_2(0)^{-1}$ also satisfy $PX_1(t)^{-1} = PX_2(t)^{-1}$ for all $t \geq 0$, then the f.d.d. of a solution X are uniquely determined by $PX(0)^{-1}$*

If X is a solution of the MGP for A and $Y_a(t) = X(a+t)$, then Y_a is a solution of the MGP for A.

Theorem 5 *If the conclusion of the above theorem holds, then any solution of the martingale problem for A is a Markov process.*

If uniqueness holds, then A is called a *generator* for the Markov process.

Exchangeability and de Finetti's theorem

$\xi_1, \xi_2, \ldots$ is *exchangeable* if

$$P\{\xi_1 \in \Gamma_1, \ldots, \xi_m \in \Gamma_m\} = P\{\xi_1 \in \Gamma_{s_1}, \ldots, \xi_m \in \Gamma_{s_m}\}$$

$(s_1, \ldots, s_m)$ any permutation of $(1, \ldots, m)$.

Theorem 6 *(de Finetti) Let $\xi_1, \xi_2, \ldots$ be exchangeable. Then there exists a random probability measure Φ such that for every bounded, measurable g,*

$$\lim_{N\to\infty} \frac{g(\xi_1) + \cdots + g(\xi_N)}{N} = \int g(x)\Phi(dx) \quad a.s.$$

almost surely, so $\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^N \delta_{\xi_i} = \Phi$.

In addition

$$E[\prod_{i=1}^m g_i(\xi_i)|\Phi] = \prod_{i=1}^m \langle \Phi, g_i\rangle = \prod_{i=1}^m \int g_i d\Phi.$$

Basic convergence lemma

Lemma 7 *For* $n = 1, 2, \ldots$*, let* $\{\xi_1^n, \ldots, \xi_{N_n}^n\}$ *be exchangeable in* S *(allowing)* $N_n = \infty$*.) Let* Ξ^n *be the empirical measure,*

$$\Xi^n = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{\xi_i^n}.$$

Assume $N_n \to \infty$*, and for each* $m = 1, 2, \ldots$*,* $\{\xi_1^n, \ldots, \xi_m^n\} \Rightarrow \{\xi_1, \ldots, \xi_m\}$ *in* S^m*.*

Then $\{\xi_i\}$ *is exchangeable and setting* $\xi_i^n = s_0 \in S$ *for* $i > N_n$*,* $\{\Xi^n, \xi_1^n, \xi_2^n \ldots\} \Rightarrow \{\Xi, \xi_1, \xi_2, \ldots\}$ *in* $\mathcal{P}(S) \times S^\infty$*, where* Ξ *is the deFinetti measure for* $\{\xi_i\}$*.*

If for each m*,* $\{\xi_1^n, \ldots, \xi_m^n\} \to \{\xi_1, \ldots, \xi_m\}$ *in probability in* S^m*, then* $\Xi^n \to \Xi$ *in probability in* $\mathcal{P}(S)$*.*

References

[1] Peter Donnelly and Thomas G. Kurtz. A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.*, 24(2):698–742, 1996.

[2] Peter Donnelly and Thomas G. Kurtz. Particle representations for measure-valued population models. *Ann. Probab.*, 27(1):166–205, 1999.

[3] Alison Etheridge and Peter March. A note on superprocesses. *Probab. Theory Related Fields*, 89(2):141–147, 1991.

[4] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[5] Steven N. Evans. Two representations of a conditioned superprocess. *Proc. Roy. Soc. Edinburgh Sect. A*, 123(5):959–971, 1993.

[6] Steven N. Evans and Edwin Perkins. Measure-valued Markov branching processes conditioned on nonextinction. *Israel J. Math.*, 71(3):329–337, 1990.

[7] Roger Tribe. The behavior of superprocesses near extinction. *Ann. Probab.*, 20(1):286–311, 1992.

Abstract

"Lookdown" constructions of measure-valued processes will be introduced. Emphasis will be on Fleming-Viot processes that arise in population genetics.