Infinite-dimensional Diffusions Related to the Two-parameter Poisson-Dirichlet Distributions

Leonid Petrov

Institute for Information Transmission Problems (Moscow, Russia)

March 14, 2011

A Moran-type Population Model

N — population size;
each individual has a *type* (a number $\in [0, 1]$)

A Moran-type Population Model

N — population size;
each individual has a *type* (a number $\in [0, 1]$)

Markov dynamics in discrete time;
at each step happens one of the events:

A Moran-type Population Model

N — population size;
each individual has a *type* (a number $\in [0, 1]$)

Markov dynamics in discrete time;
at each step happens one of the events:

1. For each pair of individuals — reproduction:

$$(A, B) \longrightarrow (A, A) \text{ or } (B, B)$$

each event with probability proportional to 1

A Moran-type Population Model

N — population size;
each individual has a *type* (a number $\in [0,1]$)

Markov dynamics in discrete time;
at each step happens one of the events:

1. For each pair of individuals — reproduction:

$$(A,B) \longrightarrow (A,A) \text{ or } (B,B)$$

   each event with probability proportional to 1

2. For each individual — mutation:

$$A \longrightarrow \text{new type not present in population}$$

   with probability proportional to $\theta > 0$

Partition Representation

Population of size $N \longrightarrow$ *allele partition* $\lambda = (\lambda_1, \ldots, \lambda_\ell)$:

$$\lambda_1 + \cdots + \lambda_\ell = N$$

$$\lambda_1 \geq \cdots \geq \lambda_\ell > 0$$

$\lambda_i = \#$ of individuals with the ith most common type

Partition Representation

Population of size $N \longrightarrow$ *allele partition* $\lambda = (\lambda_1, \ldots, \lambda_\ell)$:

$$\lambda_1 + \cdots + \lambda_\ell = N$$
$$\lambda_1 \geq \cdots \geq \lambda_\ell > 0$$
$$\lambda_i = \#\text{ of individuals with the } i\text{th most common type}$$

Example.

$$(A, B, A, C, D, D, D, A, D, E, B, B, E, F, D)$$
$$\downarrow$$
$$\lambda = (5, 3, 3, 2, 1, 1)$$

Moran-type model (population size $= N$)

$\longrightarrow$ *Markov* chain $T_\theta^{(N)}$ on $\mathrm{Part}(N)$ (partitions of N) with transition probabilities:

Partition Representation

Moran-type model (population size $= N$)

$\longrightarrow$ *Markov* chain $T_\theta^{(N)}$ on $\mathrm{Part}(N)$ (partitions of N) with transition probabilities:

- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots, \lambda_j + 1, \ldots \lambda_\ell)$
  with probability $\frac{\mathbf{1}}{\mathbf{Z}}\boldsymbol{\lambda}_\mathbf{i}\boldsymbol{\lambda}_\mathbf{j}$, $i, j = 1, \ldots, \ell$, $i \neq j$;

Partition Representation

Moran-type model (population size $= N$)

$\longrightarrow$ *Markov* chain $T_\theta^{(N)}$ on $\mathrm{Part}(N)$ (partitions of N) with transition probabilities:

- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots, \lambda_j + 1, \ldots \lambda_\ell)$

  with probability $\frac{\mathbf{1}}{\mathbf{Z}}\boldsymbol{\lambda}_{\mathbf{i}}\boldsymbol{\lambda}_{\mathbf{j}}$, $i, j = 1, \ldots, \ell$, $i \neq j$;
- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_\ell)$

  with probability $\frac{\mathbf{1}}{\mathbf{Z}}\sum_{\mathbf{k}=\mathbf{1}}^{\boldsymbol{\ell}}\boldsymbol{\lambda}_{\mathbf{k}}(\boldsymbol{\lambda}_{\mathbf{k}} - \mathbf{1})$,

Partition Representation

Moran-type model (population size $= N$)

$\longrightarrow$ *Markov* chain $T_\theta^{(N)}$ on $\mathrm{Part}(N)$ (partitions of N) with transition probabilities:

- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots, \lambda_j + 1, \ldots \lambda_\ell)$

  with probability $\frac{1}{\mathbf{Z}} \boldsymbol{\lambda}_\mathbf{i} \boldsymbol{\lambda}_\mathbf{j}$, $i, j = 1, \ldots, \ell$, $i \neq j$;
- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_\ell)$

  with probability $\frac{1}{\mathbf{Z}} \sum_{\mathbf{k}=\mathbf{1}}^{\ell} \boldsymbol{\lambda}_\mathbf{k}(\boldsymbol{\lambda}_\mathbf{k} - \mathbf{1})$,
- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots \lambda_\ell, 1)$

  with probability $\frac{1}{\mathbf{Z}} \boldsymbol{\theta} \boldsymbol{\lambda}_\mathbf{i}$, $i = 1, \ldots, \ell$.

$Z = N(N - 1 + \theta)$ — normalizing constant.

Limit $N \to +\infty$

Scale time:
one step of the Nth Markov chain corresponds to
time interval $\Delta t \approx 1/N^2$

Limit $N \to +\infty$

Scale time:
one step of the Nth Markov chain corresponds to time interval $\Delta t \approx 1/N^2$

Scale space:
embed all sets $\mathrm{Part}(N)$ into the infinite-dimensional simplex

$$\overline{\nabla}_\infty = \left\{ x = (x_1, x_2, \dots) \colon x_1 \geq x_2 \geq \dots \geq 0,\ \sum_{i=1}^{\infty} x_i \leq 1 \right\}$$

as

$$\mathrm{Part}(N) \ni \lambda = (\lambda_1, \dots, \lambda_\ell) \mapsto \left(\frac{\lambda_1}{N}, \dots, \frac{\lambda_\ell}{N}, 0, 0, \dots \right) \in \overline{\nabla}_\infty.$$

Limit $N \to +\infty$

Theorem [Ethier–Kurtz 1981]

1. As $N \to +\infty$ under the above space and time scalings, the Markov chains $T_\theta^{(N)}$ on partitions converge to a continuous-time Markov process $(X_\theta(t))_{t \geq 0}$ on $\overline{\nabla}_\infty$. It has *continuous* sample paths and can start from any point of $\overline{\nabla}_\infty$ (= *infinite-dimensional diffusion*).

Theorem [Ethier–Kurtz 1981]

1. As $N \to +\infty$ under the above space and time scalings, the Markov chains $T_\theta^{(N)}$ on partitions converge to a continuous-time Markov process $(X_\theta(t))_{t \geq 0}$ on $\overline{\nabla}_\infty$.
   It has *continuous* sample paths and can start from any point of $\overline{\nabla}_\infty$ (= *infinite-dimensional diffusion*).

2. The process $X_\theta(t)$ has a *unique invariant probability distribution* on $\overline{\nabla}_\infty$ — the Poisson-Dirichlet distribution $PD(\theta)$. The process $X_\theta(t)$ is *reversible* and *ergodic* with respect to $PD(\theta)$.

Theorem [Ethier–Kurtz 1981]

1. As $N \to +\infty$ under the above space and time scalings, the Markov chains $T_\theta^{(N)}$ on partitions converge to a continuous-time Markov process $(X_\theta(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$. It has *continuous* sample paths and can start from any point of $\overline{\nabla}_\infty$ (= *infinite-dimensional diffusion*).

2. The process $X_\theta(t)$ has a *unique invariant probability distribution* on $\overline{\nabla}_\infty$ — the Poisson-Dirichlet distribution $PD(\theta)$. The process $X_\theta(t)$ is *reversible* and *ergodic* with respect to $PD(\theta)$.

3. The *generator* of $X_\theta(t)$ is explicitly computed (see below).

$X_\theta(t)$ is called the
Infinitely Many Neutral Alleles Diffusion Model (IMNA)

Scheme of proof

Approximate infinite-dimensional diffusions $X_\theta(t)$ on $\overline{\nabla}_\infty$ by finite-dimensional *Wright-Fisher* diffusions on simplices $\left\{x_1 \geq 0, \ldots, x_K \geq 0 \colon \sum_{i=1}^K x_i = 1\right\}$ of growing dimension

Scheme of proof

Approximate infinite-dimensional diffusions $X_\theta(t)$ on $\overline{\nabla}_\infty$ by finite-dimensional *Wright-Fisher* diffusions on simplices $\left\{x_1 \geq 0, \ldots, x_K \geq 0 \colon \sum_{i=1}^K x_i = 1\right\}$ of growing dimension

The Markov chains $T_\theta^{(N)}$ have the same limit as these finite-dimensional diffusions

Scheme of proof

Approximate infinite-dimensional diffusions $X_\theta(t)$ on $\overline{\nabla}_\infty$ by finite-dimensional *Wright-Fisher* diffusions on simplices $\left\{x_1 \geq 0, \dots, x_K \geq 0\colon \sum_{i=1}^{K} x_i = 1\right\}$ of growing dimension

The Markov chains $T_\theta^{(N)}$ have the same limit as these finite-dimensional diffusions

On finite-dimensional simplices the invariant distribution is the symmetric Dirichlet distribution (= "multivariate Beta distribution") with density

$$\frac{\Gamma(K\gamma)}{\Gamma(\gamma)^K} x_1^{\gamma-1} \dots x_K^{\gamma-1} dx_1 \dots dx_{K-1}, \qquad \gamma = \frac{\theta}{K-1}$$

These distributions converge to $PD(\theta)$ as $K \to +\infty$

Scheme of proof

The finite-dimensional generators are

$$\sum_{i,j=1}^{K} x_i(\delta_{ij} - x_j)\frac{\partial^2}{\partial x_i \partial x_j} - \frac{\theta}{K-1}\sum_{i=1}^{K}(Kx_i - 1)\frac{\partial}{\partial x_i}$$

Scheme of proof

The finite-dimensional generators are

$$\sum_{i,j=1}^{K} x_i(\delta_{ij} - x_j)\frac{\partial^2}{\partial x_i \partial x_j} - \frac{\theta}{K-1}\sum_{i=1}^{K}(Kx_i - 1)\frac{\partial}{\partial x_i}$$

The infinite-dimensional generator is

$$\sum_{i,j=1}^{\infty} x_i(\delta_{ij} - x_j)\frac{\partial^2}{\partial x_i \partial x_j} - \theta\sum_{i=1}^{\infty} x_i\frac{\partial}{\partial x_i}.$$

It acts on *continuous symmetric polynomials* in the coordinates $x_1, x_2, \dots$ (= polynomials in $p_r(x) := \sum_{i=1}^{\infty} x_i^r$, $r = 2, 3, \dots$).

Two-parameter generalization

Two-parameter Poisson-Dirichlet distribution [Pitman 1992], [Pitman-Yor 1997]

$\mathbf{PD}(\boldsymbol{\alpha}, \boldsymbol{\theta})$ ($0 \le \alpha < 1$, $\theta > -\alpha$)
— probability measures on the infinite-dimensional simplex $\overline{\nabla}_\infty$

$PD(\theta) \equiv PD(0, \theta)$

Two-parameter generalization

Two-parameter Poisson-Dirichlet distribution [Pitman 1992], [Pitman-Yor 1997]

$\mathbf{PD}(\boldsymbol{\alpha}, \boldsymbol{\theta})$ $(0 \leq \alpha < 1,\ \theta > -\alpha)$
— probability measures on the infinite-dimensional simplex $\overline{\nabla}_\infty$

$PD(\theta) \equiv PD(0, \theta)$

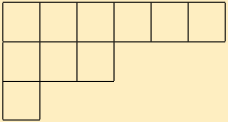
Program

1. Construct Markov chains $T^{(N)}_{\alpha,\theta}$ on $\mathrm{Part}(N)$
2. Study their limit as $N \to +\infty$
3. Thus obtain infinite-dimensional diffusions $X_{\alpha,\theta}(t)$ on $\overline{\nabla}_\infty$ preserving $PD(\alpha, \theta)$.

Markov chains $T_\theta^{(N)}$ as two-step processes

Partitions = Young diagrams

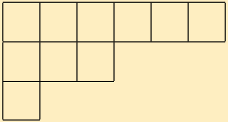
$\lambda = (6, 3, 1)$:



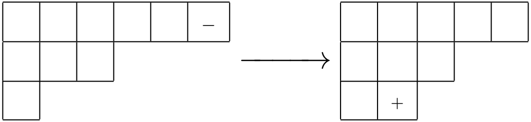
Markov chains $T_\theta^{(N)}$ as two-step processes

Partitions = Young diagrams

$\lambda = (6, 3, 1)$:



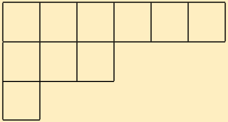
One step of the chain $T_\theta^{(N)}$ = move a box from one place to another:



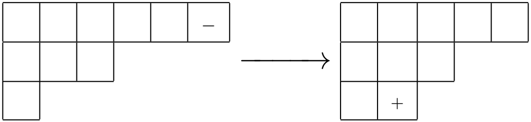
Markov chains $T_\theta^{(N)}$ as two-step processes

Partitions = Young diagrams

$\lambda = (6, 3, 1)$:



One step of the chain $T_\theta^{(N)}$ = move a box from one place to another:



move a box = delete then add

Markov chains $T_\theta^{(N)}$ as two-step processes

Delete a box

Choose any box uniformly, delete it; then rearrange

Markov chains $T_\theta^{(N)}$ as two-step processes

Delete a box

Choose any box uniformly, delete it; then rearrange

Add a box: *"Chinese restaurant"*

- Add a box next to m other boxes with probability $\dfrac{m}{N+\theta}$; then rearrange
- Or add a new row with probability $\dfrac{\theta}{N+\theta}$

Markov chains $T_\theta^{(N)}$ as two-step processes

Delete a box

Choose any box uniformly, delete it; then rearrange

Add a box: *"Chinese restaurant"*

- Add a box next to m other boxes with probability $\dfrac{m}{N+\theta}$; then rearrange
- Or add a new row with probability $\dfrac{\theta}{N+\theta}$

The Markov chain $T_\theta^{(N)}$ = delete-add process

Two-parameter Markov chains $T^{(N)}_{\alpha,\theta}$

Modified "add a box": *Two-parameter "Chinese restaurant"*

- Add a box next to m other boxes with probability $\dfrac{m-\alpha}{N+\theta}$; then rearrange
- Or add a new row with probability $\dfrac{\theta+\ell(\lambda)\cdot\alpha}{N+\theta}$

Two-parameter Markov chains $T^{(N)}_{\alpha,\theta}$

Modified "add a box": *Two-parameter "Chinese restaurant"*

- Add a box next to m other boxes with probability $\dfrac{m-\alpha}{N+\theta}$; then rearrange
- Or add a new row with probability $\dfrac{\theta + \ell(\lambda)\cdot\alpha}{N+\theta}$

Two-parameter Markov chains $T^{(N)}_{\alpha,\theta}$ (here $Z = N(N-1+\theta)$)

- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots, \lambda_j + 1, \ldots \lambda_\ell)$
  with probability $\frac{1}{Z}\boldsymbol{\lambda_i(\lambda_j - \alpha)}$, $i, j = 1, \ldots, \ell$, $i \neq j$;
- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_\ell)$
  with probability $\frac{1}{Z}\sum_{k=1}^{\ell}\boldsymbol{\lambda_k(\lambda_k - 1 - \alpha)}$,
- $(\lambda_1, \ldots, \lambda_\ell) \to (\lambda_1, \ldots, \lambda_i - 1, \ldots \lambda_\ell, 1)$
  with probability $\frac{1}{Z}\boldsymbol{(\theta + \ell\alpha)\lambda_i}$, $i = 1, \ldots, \ell$.

The Poisson-Dirichlet distributions $PD(\alpha, \theta)$

Start from the empty diagram and add N boxes according to the Chinese restaurant.

$$\varnothing \xrightarrow{\mathrm{add}_{\alpha,\theta}} \square \xrightarrow{\mathrm{add}_{\alpha,\theta}} \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \xrightarrow{\mathrm{add}_{\alpha,\theta}} \ldots \xrightarrow{\mathrm{add}_{\alpha,\theta}}$$

Probability measure $M_{\alpha,\theta}^{(N)}$ on $\mathrm{Part}(N)$

The Poisson-Dirichlet distributions $PD(\alpha, \theta)$

Start from the empty diagram and add N boxes according to the Chinese restaurant.

$$\varnothing \xrightarrow{\mathrm{add}_{\alpha,\theta}} \square \xrightarrow{\mathrm{add}_{\alpha,\theta}} \begin{array}{|c|}\hline \\ \hline \\ \hline\end{array} \xrightarrow{\mathrm{add}_{\alpha,\theta}} \dots \xrightarrow{\mathrm{add}_{\alpha,\theta}} \begin{array}{l}\text{Probability} \\ \text{measure } M_{\alpha,\theta}^{(N)} \\ \text{on } \mathrm{Part}(N)\end{array}$$

$M_{\alpha,\theta}^{(N)} \longleftrightarrow$ Ewens-Pitman sampling formula:

$$M_{\alpha,\theta}^{(N)}(\lambda) = \frac{N!}{(\theta)_N} \cdot \frac{\theta(\theta+\alpha)\dots(\theta+(\ell(\lambda)-1)\alpha)}{\prod \lambda_i! \prod [\lambda : k]!} \cdot \prod_{i=1}^{\ell(\lambda)} \prod_{j=2}^{\lambda_i} (j-1-\alpha)$$

The Poisson-Dirichlet distributions $PD(\alpha, \theta)$

Start from the empty diagram and add N boxes according to the Chinese restaurant.

$$\varnothing \xrightarrow{\mathrm{add}_{\alpha,\theta}} \square \xrightarrow{\mathrm{add}_{\alpha,\theta}} \begin{array}{c}\square\\[-2pt]\square\end{array} \xrightarrow{\mathrm{add}_{\alpha,\theta}} \dots \xrightarrow{\mathrm{add}_{\alpha,\theta}}$$

Probability measure $M_{\alpha,\theta}^{(N)}$ on $\mathrm{Part}(N)$

$M_{\alpha,\theta}^{(N)} \longleftrightarrow$ Ewens-Pitman sampling formula:

$$M_{\alpha,\theta}^{(N)}(\lambda) = \frac{N!}{(\theta)_N} \cdot \frac{\theta(\theta+\alpha)\dots(\theta+(\ell(\lambda)-1)\alpha)}{\prod \lambda_i! \prod [\lambda : k]!} \cdot \prod_{i=1}^{\ell(\lambda)} \prod_{j=2}^{\lambda_i} (j-1-\alpha)$$

$PD(\alpha, \theta)$ is the limit of $M_{\alpha,\theta}^{(N)}$ as $N \to +\infty$

The processes $X_{\alpha,\theta}$ on $\overline{\nabla}_\infty$

Theorem [P.]

1. As $N \to +\infty$, under the space and time scalings, the Markov chains $T_{\alpha,\theta}^{(N)}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$.

The processes $X_{\alpha,\theta}$ on $\overline{\nabla}_\infty$

Theorem [P.]

1. As $N \to +\infty$, under the space and time scalings, the Markov chains $T^{(N)}_{\alpha,\theta}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$.
2. The Poisson-Dirichlet distribution $PD(\alpha,\theta)$ is the *unique invariant probability distribution* for $X_{\alpha,\theta}(t)$. The process is *reversible* and *ergodic* with respect to $PD(\alpha,\theta)$.

Theorem [P.]

1. As $N \to +\infty$, under the space and time scalings, the Markov chains $T^{(N)}_{\alpha,\theta}$ converge to an *infinite-dimensional diffusion process* $(X_{\alpha,\theta}(t))_{t\geq 0}$ on $\overline{\nabla}_\infty$.
2. The Poisson-Dirichlet distribution $PD(\alpha, \theta)$ is the *unique invariant probability distribution* for $X_{\alpha,\theta}(t)$. The process is *reversible* and *ergodic* with respect to $PD(\alpha, \theta)$.
3. The *generator* of $X_{\alpha,\theta}$ is explicitly computed:

$$\sum_{i,j=1}^{\infty} x_i(\delta_{ij} - x_j)\frac{\partial^2}{\partial x_i \partial x_j} - \sum_{i=1}^{\infty}(\theta x_i + \alpha)\frac{\partial}{\partial x_i}.$$

It acts on *continuous symmetric polynomials* in the coordinates $x_1, x_2, \ldots$.

Scheme of proof

No finite-dimensional approximating diffusions!

Scheme of proof

No finite-dimensional approximating diffusions!

1. The transition operators of the Markov chains $T^{(N)}_{\alpha,\theta}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathrm{Part}(N)$.

Scheme of proof

No finite-dimensional approximating diffusions!

1. The transition operators of the Markov chains $T_{\alpha,\theta}^{(N)}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathrm{Part}(N)$.
2. Write the operators $T_{\theta}^{(N)}$ in a suitable basis (*monomial symmetric functions*).

Scheme of proof

No finite-dimensional approximating diffusions!

1. The transition operators of the Markov chains $T^{(N)}_{\alpha,\theta}$ act on symmetric functions in the coordinates $\lambda_1, \ldots, \lambda_\ell$ of a partition $\lambda \in \mathrm{Part}(N)$.
2. Write the operators $T^{(N)}_{\theta}$ in a suitable basis (*monomial symmetric functions*).
3. Pass to $N \to +\infty$ limit of generators (this is done in a *purely algebraic* way)

Scheme of proof

No finite-dimensional approximating diffusions!

1. The transition operators of the Markov chains $T^{(N)}_{\alpha,\theta}$ act on symmetric functions in the coordinates $\lambda_1, \dots, \lambda_\ell$ of a partition $\lambda \in \mathrm{Part}(N)$.
2. Write the operators $T^{(N)}_{\theta}$ in a suitable basis (*monomial symmetric functions*).
3. Pass to $N \to +\infty$ limit of generators (this is done in a *purely algebraic* way)
4. Use general technique of Trotter-Kurtz to deduce convergence of the processes

Thank you for your attention

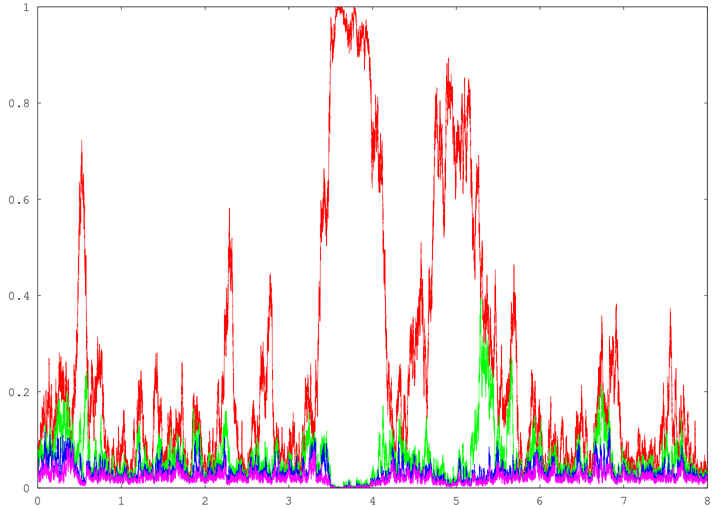


Figure: $x_1(t) \geq x_2(t) \geq x_3(t) \geq x_4(t)$