Diffusion Approximation
Effective population size estimation

Jing Liu

Department of Statistics
UoA

November 29, 2009

Outline

Return to Hyperlinks

Outline

Return to Hyperlinks

Outline

1. Introduction
   - Motivation
   - What is Effective population size
2. Question
   - Objective
   - Issue
3. Solution
   - Diffusion approximation
   - Simplex expansion

Return to Hyperlinks

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion ≈ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion
  1 Male (*Distorted Sex ratio*)
  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion

  1 Male (*Distorted Sex ratio*)

  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion
  1 Male (*Distorted Sex ratio*)
  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

Motivation

- Successfulness of a species

  Population size is a sensible indicator

- But...... Imagine a species

  6 Billion
  1 Male (*Distorted Sex ratio*)
  Clones (*limited genetic variation*)

- Is this so different from a species with only one female and one male left?

  6 Billion $\approx$ 2

- Something better is needed

N_e

- Effective population size N_e
  - Abstract and conceptual population size
  - Sex ratio, Genetic variation, Breeding structure
- Ecology, Population Genetics, Conservation

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
- $$r = cor(\quad \#\text{Allele A}, \quad \#\text{Allele B})$$
  $$\uparrow \qquad\qquad \uparrow$$
  $$\text{locus 1} \qquad \text{locus 2}$$
- Sample mean $\hat{r}^2$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
-

$$r = cor(\underset{\text{locus 1}}{\underset{\uparrow}{\#\text{Allele A}}}, \underset{\text{locus 2}}{\underset{\uparrow}{\#\text{Allele B}}})$$

- Sample mean $\hat{r}^2$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
- $$r = cor(\ \#\text{Allele A},\ \#\text{Allele B})$$
  $$\underset{\text{locus 1}}{\uparrow} \qquad \underset{\text{locus 2}}{\uparrow}$$
- Sample mean $\hat{r}^2$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
- $$r = cor(\underset{\substack{\uparrow \\ \text{locus 1}}}{\#\text{Allele A}}, \underset{\substack{\uparrow \\ \text{locus 2}}}{\#\text{Allele B}})$$
- Sample mean $\hat{r^2}$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
- $$r = cor\left(\underset{\text{locus 1}}{\underset{\uparrow}{\#\text{Allele A}}},\ \underset{\text{locus 2}}{\underset{\uparrow}{\#\text{Allele B}}}\right)$$
- Sample mean $\hat{r^2}$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

$$\text{Matching sample } r^2 \text{ with } \mathbb{E}[r^2]$$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
- $$r = cor(\underset{\text{locus 1}}{\underset{\uparrow}{\#\text{Allele A}}},\ \underset{\text{locus 2}}{\underset{\uparrow}{\#\text{Allele B}}})$$
- Sample mean $\hat{r^2}$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Estimation of N_e

- How to estimate N_e given a sample of genetic data?
- $\mathbb{E}[r^2] = f(N_e)$
-

$$r = cor\left(\underset{\text{locus 1}}{\underset{\uparrow}{\#\text{Allele A}}},\ \underset{\text{locus 2}}{\underset{\uparrow}{\#\text{Allele B}}} \right)$$

- Sample mean $\hat{r^2}$ can be obtained from data
- But exact function link between $\mathbb{E}[r^2]$ and N_e is not known
- If we know.....then ......

Matching sample r^2 with $\mathbb{E}[r^2]$

Real probelm

- Real problem is to derive $\mathbb{E}[r^2]$ analytically in terms of N_e and other parameters

Real probelm

- Real problem is to derive $\mathbb{E}[r^2]$ analytically in terms of N_e and other parameters

- r^2 can be defined in terms of Genotype frequency

Real probelm

- Real problem is to derive $\mathbb{E}[r^2]$ analytically in terms of N_e and other parameters

- r^2 can be defined in terms of Genotype frequency
- Analytic derivation of the expectation by integration is not going to be easy

Out[2]=

$$
\begin{aligned}
&\Big(-4\Big(p1+p2+\tfrac{1}{2}(p3+p4+p6+p7)+p5\Big) \\
&\quad \Big(p1+\tfrac{1}{2}(p2+p4+p6+p9)+p3+p8\Big)+ \\
&\quad 4\,p1+2\,p2+2\,p3+p4+p6\Big)^2 \Big/ \\
&\Big(4\Big(-\Big(p1+p2+\tfrac{1}{2}(p3+p4+p6+p7)+p5\Big)^2- \\
&\quad \tfrac{1}{4}(2\,p1+2\,p2+p3+p4+2\,p5+p6+p7-2) \\
&\quad (2\,p1+2\,p2+p3+p4+2\,p5+p6+p7)+p1+p2+p5\Big) \\
&\quad \Big(-\Big(p1+\tfrac{1}{2}(p2+p4+p6+p9)+p3+p8\Big)^2- \\
&\quad \tfrac{1}{4}(2\,p1+p2+2\,p3+p4+p6+2\,p8+p9-2) \\
&\quad (2\,p1+p2+2\,p3+p4+p6+2\,p8+p9)+p1+p3+p8\Big)\Big)
\end{aligned}
$$

Figure: Rsq

Diffusion

- Diffusion Approximation
  - Diffusion Model built upon assumptions from a very popular reproduction model in population genetics, namely Wright-Fisher's Model
- Differential Operator

$$\mathcal{D}iff(G(p_1, p_2, \ldots, p_9)) = g(p_1, p_2, \ldots, p_9)$$

$\uparrow$ Any function $\qquad \uparrow$ Special Deriative

- $$\mathbb{E}[g] = 0$$
- A powerful "machine" produce all kinds of expressions of $\mathbb{E}$ given a "recipe" $G(p_1, p_2, \ldots, p_9)$

Diffusion

- Diffusion Approximation
  - Diffusion Model built upon assumptions from a very popular reproduction model in population genetics, namely Wright-Fisher's Model
- Differential Operator

$$\mathcal{D}iff(G(p_1, p_2, \ldots, p_9)) = g(p_1, p_2, \ldots, p_9)$$

$\uparrow$ Any function $\qquad \uparrow$ Special Deriative

- $$\mathbb{E}[g] = 0$$
- A powerful "machine" produce all kinds of expressions of $\mathbb{E}$ given a "recipe" $G(p_1, p_2, \ldots, p_9)$

Diffusion

- Diffusion Approximation
  - Diffusion Model built upon assumptions from a very popular reproduction model in population genetics, namely Wright-Fisher's Model
- Differential Operator

$$\mathcal{D}iff\,(G(p_1, p_2, \ldots, p_9)) = \; g(p_1, p_2, \ldots, p_9)$$

$\uparrow$ Any function $\qquad\qquad$ $\uparrow$ Special Deriative

- $$\mathbb{E}[g] = 0$$
- A powerful "machine" produce all kinds of expressions of $\mathbb{E}$ given a "recipe" $G(p_1, p_2, \ldots, p_9)$

Diffusion

- Diffusion Approximation
  - Diffusion Model built upon assumptions from a very popular reproduction model in population genetics, namely Wright-Fisher's Model
- Differential Operator

$$\mathcal{D}iff\,(G(p_1, p_2, \ldots, p_9)) = \quad g(p_1, p_2, \ldots, p_9)$$

↑ Any function ↑ Special Deriative

- $$\mathbb{E}[g] = 0$$
- A powerful "machine" produce all kinds of expressions of $\mathbb{E}$ given a "recipe" $G(p_1, p_2, \ldots, p_9)$

Diffusion

- Diffusion Approximation
  - Diffusion Model built upon assumptions from a very popular reproduction model in population genetics, namely Wright-Fisher's Model
- Differential Operator

$$\mathcal{D}iff\,(G(p_1, p_2, \ldots, p_9)) = \quad g(p_1, p_2, \ldots, p_9)$$

$$\uparrow \qquad\qquad\qquad\qquad \uparrow$$

Any function — Special Deriative

- $$\mathbb{E}[g] = 0$$

- A powerful "machine" produce all kinds of expressions of $\mathbb{E}$ given a "recipe" $G(p_1, p_2, \ldots, p_9)$

Too hard

- Theoretically a careful choice of $G(p_1, p_2, \ldots, p_9)$
  - $$F(\mathbb{E}[r^2]) = 0$$
- But.....

$$
\begin{aligned}
\mathcal{D}iff = & \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2} + \frac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2} \\
& + \frac{1}{2}[x(1-x)y(1-y) + D(1-2x)(1-2y) - D^2]\frac{\partial^2}{\partial D^2} \\
& + D\frac{\partial^2}{\partial x \partial y} + D(1-2x)\frac{\partial^2}{\partial x \partial D} + D(1-2y)\frac{\partial^2}{\partial y \partial D} \\
& + \frac{\theta}{4}(1-2x)\frac{\partial}{\partial x} + \frac{\theta}{4}(1-2y)\frac{\partial}{\partial y} - D(1 + \frac{\rho}{2} + \theta)\frac{\partial}{\partial D}
\end{aligned}
$$

- Finding the perfect G isn't going to be easy for such a target function and differential operator

Too hard

- Theoretically a careful choice of $G(p_1, p_2, \ldots, p_9)$
  - $$F(\mathbb{E}[r^2]) = 0$$
- But.....

$$\begin{aligned}\mathcal{D}iff &= \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2} + \frac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2} \\ &+ \frac{1}{2}[x(1-x)y(1-y) + D(1-2x)(1-2y) - D^2]\frac{\partial^2}{\partial D^2} \\ &+ D\frac{\partial^2}{\partial x \partial y} + D(1-2x)\frac{\partial^2}{\partial x \partial D} + D(1-2y)\frac{\partial^2}{\partial y \partial D} \\ &+ \frac{\theta}{4}(1-2x)\frac{\partial}{\partial x} + \frac{\theta}{4}(1-2y)\frac{\partial}{\partial y} - D(1+\frac{\rho}{2}+\theta)\frac{\partial}{\partial D}\end{aligned}$$

- Finding the perfect G isn't going to be easy for such a target function and differential operator

Too hard

- Theoretically a careful choice of $G(p_1, p_2, \ldots, p_9)$
  -
$$F(\mathbb{E}[r^2]) = 0$$
- But.....

$$\begin{aligned}\mathcal{D}iff &= \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2} + \frac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2} \\ &+ \frac{1}{2}[x(1-x)y(1-y) + D(1-2x)(1-2y) - D^2]\frac{\partial^2}{\partial D^2} \\ &+ D\frac{\partial^2}{\partial x \partial y} + D(1-2x)\frac{\partial^2}{\partial x \partial D} + D(1-2y)\frac{\partial^2}{\partial y \partial D} \\ &+ \frac{\theta}{4}(1-2x)\frac{\partial}{\partial x} + \frac{\theta}{4}(1-2y)\frac{\partial}{\partial y} - D(1+\frac{\rho}{2}+\theta)\frac{\partial}{\partial D}\end{aligned}$$

- Finding the perfect G isn't going to be easy for such a target function and differential operator

Too hard

- Theoretically a careful choice of $G(p_1, p_2, \ldots, p_9)$
    - $$F(\mathbb{E}[r^2]) = 0$$
- But.....

$$\begin{aligned}\mathcal{D}iff &= \frac{1}{2}x(1-x)\frac{\partial^2}{\partial x^2} + \frac{1}{2}y(1-y)\frac{\partial^2}{\partial y^2} \\ &+ \frac{1}{2}[x(1-x)y(1-y) + D(1-2x)(1-2y) - D^2]\frac{\partial^2}{\partial D^2} \\ &+ D\frac{\partial^2}{\partial x \partial y} + D(1-2x)\frac{\partial^2}{\partial x \partial D} + D(1-2y)\frac{\partial^2}{\partial y \partial D} \\ &+ \frac{\theta}{4}(1-2x)\frac{\partial}{\partial x} + \frac{\theta}{4}(1-2y)\frac{\partial}{\partial y} - D(1+\frac{\rho}{2}+\theta)\frac{\partial}{\partial D}\end{aligned}$$

- Finding the perfect G isn't going to be easy for such a target function and differential operator

Simplex

- Manipulate r^2 instead of being clever with $\mathcal{D}iff$
- Apply Linear Simplex Expansion on r^2
  - Piecewise linear interpolation in high dimension
  - In every small region i of the tessellation:

$$S_i = a_{0i} + a_{1i}p_1 + a_{2i}p_2 + \cdots + a_{9i}p_9$$

Where $a_{0i}, \ldots, a_{9i}$ are found by solving a series of 10-by-10 linear system

Simplex

- Manipulate r^2 instead of being clever with $\mathcal{D}iff$
- Apply Linear Simplex Expansion on r^2
  - Piecewise linear interpolation in high dimension
  - In every small region i of the tessellation:

$$S_i = a_{0i} + a_{1i}p_1 + a_{2i}p_2 + \cdots + a_{9i}p_9$$

  Where $a_{0i}, \ldots, a_{9i}$ are found by solving a series of 10-by-10 linear system

Simplex

- Manipulate r^2 instead of being clever with $\mathcal{D}iff$
- Apply Linear Simplex Expansion on r^2
  - Piecewise linear interpolation in high dimension
  - In every small region i of the tessellation:

$$S_i = a_{0i} + a_{1i}p_1 + a_{2i}p_2 + \cdots + a_{9i}p_9$$

Where $a_{0i}, \ldots, a_{9i}$ are found by solving a series of 10-by-10 linear system

Simplex

- Manipulate r^2 instead of being clever with $\mathcal{D}iff$
- Apply Linear Simplex Expansion on r^2
  - Piecewise linear interpolation in high dimension
  - In every small region i of the tessellation:

$$S_i = a_{0i} + a_{1i}p_1 + a_{2i}p_2 + \cdots + a_{9i}p_9$$

Where $a_{0i}, \ldots, a_{9i}$ are found by solving a series of 10-by-10 linear system

Piecewise r^2

$$
\begin{aligned}
r^2 = \quad & a_{0,1}\mathbb{I}_1 + a_{1,1}p_1\mathbb{I}_1 + a_{2,1}p_2\mathbb{I}_1 + \cdots + a_{9,1}p_9\mathbb{I}_1 \\
& \vdots \\
& a_{0,i}\mathbb{I}_i + a_{1,i}p_1\mathbb{I}_i + a_{2,1}p_2\mathbb{I}_i + \cdots + a_{9,1}p_9\mathbb{I}_i \\
& \vdots \\
& a_{0,k}\mathbb{I}_k + a_{1,k}p_1\mathbb{I}_k + a_{2,k}p_2\mathbb{I}_k + \cdots + a_{9,k}p_9\mathbb{I}_k
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}[r^2] = \quad & a_{0,1}\mathbb{E}[\mathbb{I}_1] + a_{1,1}\mathbb{E}[p_1\mathbb{I}_1] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_1] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_1] \\
& \vdots \\
& a_{0,i}\mathbb{E}[\mathbb{I}_i] + a_{1,i}\mathbb{E}[p_1\mathbb{I}_i] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_i] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_i] \\
& \vdots \\
& a_{0,k}\mathbb{E}[\mathbb{I}_k] + a_{1,k}\mathbb{E}[p_1\mathbb{I}_k] + a_{2,k}\mathbb{E}[p_2\mathbb{I}_k] + \cdots + a_{9,k}\mathbb{E}[p_9\mathbb{I}_k]
\end{aligned}
$$

Piecewise r^2

$$
\begin{aligned}
r^2 = \quad & a_{0,1}\mathbb{I}_1 + a_{1,1}p_1\mathbb{I}_1 + a_{2,1}p_2\mathbb{I}_1 + \cdots + a_{9,1}p_9\mathbb{I}_1 \\
& \vdots \\
& a_{0,i}\mathbb{I}_i + a_{1,i}p_1\mathbb{I}_i + a_{2,1}p_2\mathbb{I}_i + \cdots + a_{9,1}p_9\mathbb{I}_i \\
& \vdots \\
& a_{0,k}\mathbb{I}_k + a_{1,k}p_1\mathbb{I}_k + a_{2,k}p_2\mathbb{I}_k + \cdots + a_{9,k}p_9\mathbb{I}_k \\
\mathbb{E}[r^2] = \quad & a_{0,1}\mathbb{E}[\mathbb{I}_1] + a_{1,1}\mathbb{E}[p_1\mathbb{I}_1] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_1] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_1] \\
& \vdots \\
& a_{0,i}\mathbb{E}[\mathbb{I}_i] + a_{1,i}\mathbb{E}[p_1\mathbb{I}_i] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_i] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_i] \\
& \vdots \\
& a_{0,k}\mathbb{E}[\mathbb{I}_k] + a_{1,k}\mathbb{E}[p_1\mathbb{I}_k] + a_{2,k}\mathbb{E}[p_2\mathbb{I}_k] + \cdots + a_{9,k}\mathbb{E}[p_9\mathbb{I}_k]
\end{aligned}
$$

Metaphor

- Given a right "recipe", $\mathcal{D}iff$ can be considered as a powerful "machine" capable of producing anything

Metaphor

- Given a right "recipe", $\mathcal{D}iff$ can be considered as a powerful "machine" capable of producing anything

- $\mathbb{E}[r^2]$ can be considered as a complicated object I want to make using this "machine", such as a nice ....

Metaphor

- Given a right "recipe", $\mathcal{D}iff$ can be considered as a powerful "machine" capable of producing anything

- $\mathbb{E}[r^2]$ can be considered as a complicated object I want to make using this "machine", such as a nice ....

Figure: I don't have a "Recipe"

Lego

- Simplex expansion works like Lego

Lego

- Simplex expansion works like Lego

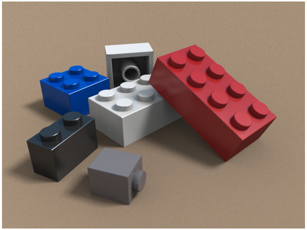

Figure: Lego

- Instead of making an actual car, I am making a Lego car

Lego

- Simplex expansion works like Lego

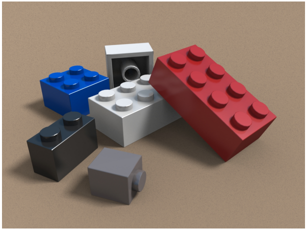

Figure: Lego

- Instead of making an actual car, I am making a Lego car

Figure: Lego car

Zoom

Lego in Diff

- Lego bricks represents those simple expectations

$$\mathbb{E}[r^2] = \begin{array}{l} a_{0,1}\mathbb{E}[\mathbb{I}_1] + a_{1,1}\mathbb{E}[p_1\mathbb{I}_1] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_1] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_1] \\ \vdots \\ a_{0,i}\mathbb{E}[\mathbb{I}_i] + a_{1,i}\mathbb{E}[p_1\mathbb{I}_i] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_i] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_i] \\ \vdots \qquad \vdots \\ a_{0,k}\mathbb{E}[\mathbb{I}_k] + a_{1,k}\mathbb{E}[p_1\mathbb{I}_k] + a_{2,k}\mathbb{E}[p_2\mathbb{I}_k] + \cdots + a_{9,k}\mathbb{E}[p_9\mathbb{I}_k] \end{array}$$

- How to solve Lego bricks using $\mathcal{D}iff$?

Lego in Diff

- Lego bricks represents those simple expectations

$$
\begin{aligned}
\mathbb{E}[r^2] = \quad & a_{0,1}\mathbb{E}[\mathbb{I}_1] + a_{1,1}\mathbb{E}[p_1\mathbb{I}_1] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_1] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_1] \\
& \vdots \\
& a_{0,i}\mathbb{E}[\mathbb{I}_i] + a_{1,i}\mathbb{E}[p_1\mathbb{I}_i] + a_{2,1}\mathbb{E}[p_2\mathbb{I}_i] + \cdots + a_{9,1}\mathbb{E}[p_9\mathbb{I}_i] \\
& \vdots \qquad \vdots \\
& a_{0,k}\mathbb{E}[\mathbb{I}_k] + a_{1,k}\mathbb{E}[p_1\mathbb{I}_k] + a_{2,k}\mathbb{E}[p_2\mathbb{I}_k] + \cdots + a_{9,k}\mathbb{E}[p_9\mathbb{I}_k]
\end{aligned}
$$

- How to solve Lego bricks using $\mathcal{D}iff$?

Simplex and Lego

- Simplex expansion acts like a set of instruction, assembling and dismantling object, “Recipe” become irrelevant

Simplex and Lego

- Simplex expansion acts like a set of instruction, assembling and dismantling object, "Recipe" become irrelevant
-

$$
\begin{array}{lll}
\mathcal{D}iff[G_1] & \rightarrow & a^{g1}_{0,1}\mathbb{E}[\mathbb{I}_1] + \cdots + a^{g1}_{9,1}\mathbb{E}[p_9\mathbb{I}_1] + \cdots\cdots + a^{g1}_{0,k}\mathbb{E}[\mathbb{I}_k] + \cdots + a^{g1}_{9,k}\mathbb{E}[p_9\mathbb{I}_k] = 0 \\
\mathcal{D}iff[G_2] & \rightarrow & a^{g2}_{0,1}\mathbb{E}[\mathbb{I}_1] + \cdots + a^{g2}_{9,1}\mathbb{E}[p_9\mathbb{I}_1] + \cdots\cdots + a^{g2}_{0,k}\mathbb{E}[\mathbb{I}_k] + \cdots + a^{g2}_{9,k}\mathbb{E}[p_9\mathbb{I}_k] = 0 \\
 & \vdots & \vdots \qquad \vdots \\
\mathcal{D}iff[G_{10k}] & \rightarrow & a^{g10k}_{0,1}\mathbb{E}[\mathbb{I}_1] + \cdots + a^{g10k}_{9,1}\mathbb{E}[p_9\mathbb{I}_1] + \cdots\cdots + a^{g10k}_{0,k}\mathbb{E}[\mathbb{I}_k] + \cdots + a^{g10k}_{9,k}\mathbb{E}[p_9\mathbb{I}_k] = 0
\end{array}
$$

Expression

- Final expression is in the form of a Padé approximation
  - Closely related to Taylor's expansion, but often superior when expanding a rational fucntion
  - For k number of simplex division:

$$\mathbb{E}[r^2] = \frac{\alpha_0 + \alpha_1 N_e c + \alpha_2 (N_e c)^2 + \cdots + \alpha_{10k} (N_e c)^{10K}}{\beta_0 + \beta_1 N_e c + \beta_2 (N_e c)^2 + \cdots + \beta_{10k} (N_e c)^{10K}}$$

Where α and β are coefficients and c is the recombination rate

Expression

- Final expression is in the form of a Padé approximation
  - Closely related to Taylor's expansion, but often superior when expanding a rational fucntion
  - For k number of simplex division:

$$\mathbb{E}[r^2] = \frac{\alpha_0 + \alpha_1 N_e c + \alpha_2 (N_e c)^2 + \cdots + \alpha_{10k} (N_e c)^{10K}}{\beta_0 + \beta_1 N_e c + \beta_2 (N_e c)^2 + \cdots + \beta_{10k} (N_e c)^{10K}}$$

Where α and β are coefficients and c is the recombination rate

Expression

- Final expression is in the form of a Padé approximation
  - Closely related to Taylor's expansion, but often superior when expanding a rational fucntion
  - For k number of simplex division:

$$\mathbb{E}[r^2] = \frac{\alpha_0 + \alpha_1 N_e c + \alpha_2 (N_e c)^2 + \cdots + \alpha_{10k} (N_e c)^{10K}}{\beta_0 + \beta_1 N_e c + \beta_2 (N_e c)^2 + \cdots + \beta_{10k} (N_e c)^{10K}}$$

Where α and β are coefficients and c is the recombination rate

Future work

- Numerically stable algorithm
- Completely Analytic derivation of the final expression by considering the limit as $k \longrightarrow \infty$

Acknowledgment

- My Supervisor : Rachel Fewster
- New Zealand Institute Mathematics & its Applications
- IBS AR

Acknowledgment

- My Supervisor : Rachel Fewster
- New Zealand Institute Mathematics & its Applications
- IBS AR

Acknowledgment

- My Supervisor : Rachel Fewster
- New Zealand Institute Mathematics & its Applications
- IBS AR

Acknowledgment

You for listening!