Random permutations with logarithmic cycle weights

Dirk Zeindler
(joint with Nicolas Robles)

Lancaster University

30 January 2020

Setting the scene

Permutations and their cycle structure

1 2 3 4 5 6 7 8 9 10

Permutations and their cycle structure

1	2	3	4	5	6	7	8	9	10
9	1	7	4	3	2	5	8	10	6

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1 \;\; 9$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1 \;\; 9 \;\; 10$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = \big(1 \;\; 9 \;\; 10 \;\; 6$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & {\color{red}6} & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & {\color{red}2} & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1 \;\; 9 \;\; 10 \;\; 6 \;\; 2$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = \begin{pmatrix} 1 & 9 & 10 & 6 & 2 \end{pmatrix}$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1\ \ 9\ \ 10\ \ 6\ \ 2)(3\ \ 7\ \ 5)$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \dots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1\ 9\ 10\ 6\ 2)(3\ 7\ 5)(4)$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1 \;\; 9 \;\; 10 \;\; 6 \;\; 2)(3 \;\; 7 \;\; 5)(4)(8)$$

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \dots, n\}$ and denote by S_n the symmetric group.

$$\pi = \begin{pmatrix} 1 & 9 & 10 & 6 & 2 \end{pmatrix} \begin{pmatrix} 3 & 7 & 5 \end{pmatrix} \begin{pmatrix} 4 \end{pmatrix} \begin{pmatrix} 8 \end{pmatrix}$$

Two cycles $(s_0 \dots s_{k-1})$ and $(t_0 \dots t_{m-1})$ are called disjoint if the sets $\{s_0, \dots, s_{k-1}\}$ and $\{t_0, \dots, t_{m-1}\}$ are disjoint.

Permutations and their cycle structure

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 9 & 1 & 7 & 4 & 3 & 2 & 5 & 8 & 10 & 6 \end{pmatrix}$$

Consider $\{1, \ldots, n\}$ and denote by S_n the symmetric group.

$$\pi = (1 \ \ 9 \ \ 10 \ \ 6 \ \ 2)(3 \ \ 7 \ \ 5)(4)(8)$$

All $\sigma \in S_n$ decompose into disjoint cycles: $\sigma = \sigma_1 \sigma_2 \ldots \sigma_\ell$.

Cycle-type

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.

Cycle-type

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.

We write λ_j for the length of the cycle σ_j.

Cycle-type

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.

We write λ_j for the length of the cycle σ_j.

W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$.

Cycle-type

If $\sigma \in S_n$ is given, then it can be written as

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$$

where $\sigma_1, \ldots, \sigma_\ell$ are disjoint cycles.

We write λ_j for the length of the cycle σ_j.

W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_\ell$.

The partition $\lambda = (\lambda_1, \ldots, \lambda_\ell)$ is called the *cycle-type* of σ.

Random permutations on S_n

Classical measures on S_n

Classical measures on S_n

- Uniform permutations:

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

- Ewens measure:

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

- Ewens measure:
  Let $\vartheta > 0$.

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

- Ewens measure:
  Let $\vartheta > 0$. We write $\sigma \in S_n$ as $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$ with $\sigma_1, \dots, \sigma_\ell$ disjoint cycles.

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

- Ewens measure:
  Let $\vartheta > 0$. We write $\sigma \in S_n$ as $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$ with $\sigma_1, \ldots, \sigma_\ell$ disjoint cycles. The Ewens measure is then defined as

$$\mathbb{P}_\vartheta[\sigma] := \frac{\vartheta^\ell}{h_n n!}$$

Classical measures on S_n

- Uniform permutations:

$$\mathbb{P}[\sigma] = \frac{1}{n!} \text{ for all } \sigma \in S_n.$$

Studied since 1940, many applications in probability.

- Ewens measure:
Let $\vartheta > 0$. We write $\sigma \in S_n$ as $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell$ with $\sigma_1, \dots, \sigma_\ell$ disjoint cycles. The Ewens measure is then defined as

$$\mathbb{P}_\vartheta[\sigma] := \frac{\vartheta^\ell}{\vartheta(\vartheta+1)\cdots(\vartheta+n-1)}$$

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))
- It plays a crucial role for virtual permutations since it is central and stable under the restriction $S_n \to S_{n-1}$

Ewens measure

The Ewens measure was introduced by Ewens (1972) in population genetics.
But it has various applications, for instance

- It has a connection with Kingman's coalescent process (1982).
- It has been used to model the dynamics of tumour evolution. (Barbour and Tavaré (2010))
- It appears in a Bayesian non parametric statistics stetting. (Antoniak (1974))
- It plays a crucial role for virtual permutations since it is central and stable under the restriction $S_n \to S_{n-1}$
- ....

The weighted measure

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

The weighted measure

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

- Ewens measure:

$$\mathbb{P}_\vartheta[\sigma] = \frac{\vartheta^\ell}{h_n n!}$$

The weighted measure

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

- Ewens measure:

$$\mathbb{P}_\vartheta[\sigma] = \frac{\vartheta^\ell}{h_n n!}$$

The weighted measure

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

- Ewens measure:

$$\mathbb{P}_\vartheta[\sigma] = \frac{\vartheta^\ell}{h_n n!} = \frac{1}{h_n n!} \prod_{j=1}^{\ell} \vartheta, \quad \vartheta > 0.$$

The weighted measure

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

- Ewens measure:

$$\mathbb{P}_\vartheta[\sigma] = \frac{\vartheta^\ell}{h_n n!} = \frac{1}{h_n n!} \prod_{j=1}^{\ell} \vartheta, \quad \vartheta > 0.$$

- Let $\Theta := (\theta_m)_{m \geq 1}$ non-negative weights.

The weighted measure

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

- Ewens measure:

$$\mathbb{P}_\vartheta[\sigma] = \frac{\vartheta^\ell}{h_n n!} = \frac{1}{h_n n!} \prod_{j=1}^{\ell} \vartheta, \quad \vartheta > 0.$$

- Let $\Theta := (\theta_m)_{m \geq 1}$ non-negative weights. The weighted measure is then defined as

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{i=1}^{\ell} \theta_{\lambda_i}$$

The weighted measure

Typical weights $(\theta_m)_{m \geq 1}$ studied in the literature are

The weighted measure

Typical weights $(\theta_m)_{m\geq 1}$ studied in the literature are

- $\theta_m \equiv 1$: Uniform measure,

The weighted measure

Typical weights $(\theta_m)_{m\geq 1}$ studied in the literature are

- $\theta_m \equiv 1$: Uniform measure,
- $\theta_m \equiv \vartheta$: Ewens measure,

The weighted measure

Typical weights $(\theta_m)_{m\geq 1}$ studied in the literature are

- $\theta_m \equiv 1$: Uniform measure,
- $\theta_m \equiv \vartheta$: Ewens measure,
- '$\theta_m \to \vartheta$': Generalised Ewens measure.

The weighted measure

Typical weights $(\theta_m)_{m\geq 1}$ studied in the literature are

- $\theta_m \equiv 1$: Uniform measure,
- $\theta_m \equiv \vartheta$: Ewens measure,
- '$\theta_m \to \vartheta$': Generalised Ewens measure.
- $\theta_m = m^\gamma$ with $\gamma > 0$.

The weighted measure

Typical weights $(\theta_m)_{m\geq 1}$ studied in the literature are

- $\theta_m \equiv 1$: Uniform measure,
- $\theta_m \equiv \vartheta$: Ewens measure,
- '$\theta_m \to \vartheta$': Generalised Ewens measure.
- $\theta_m = m^\gamma$ with $\gamma > 0$.

We study in this talk weights of the form

$$\theta_m = \log^k(m), \quad \text{with } k \geq 1 \text{ fix.}$$

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

- C_m: the cycle counts,

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

- C_m: the cycle counts,
- ℓ_1: the length of the cycle containing 1,

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

- C_m: the cycle counts,
- ℓ_1: the length of the cycle containing 1,
- K_n: the total number of cycles ($K_n = \sum_{m=1}^{n} C_m$),

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

- C_m: the cycle counts,
- ℓ_1: the length of the cycle containing 1,
- K_n: the total number of cycles ($K_n = \sum_{m=1}^{n} C_m$),
- λ_1: the length of the longest cycle.

Interesting random variables

$$\mathbb{P}_\Theta[\sigma] := \frac{1}{h_n n!} \prod_{m=1}^{n} \theta_m^{C_m(\sigma)}.$$

We now use this measure on S_n and let $n \to \infty$.
What are interesting random variables to study?

- C_m: the cycle counts,
- ℓ_1: the length of the cycle containing 1,
- K_n: the total number of cycles ($K_n = \sum_{m=1}^{n} C_m$),
- λ_1: the length of the longest cycle.
- $(\lambda_1, \lambda_2, \ldots)$: the length of the longest cycles.

The cycle counts C_m

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i.

The cycle counts C_m

Let $\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell \in S_n$ be given with σ_i disjoint cycles and cycle lengths λ_i. We then define

$$C_m := \#\{i;\ \lambda_i = m\}.$$

The cycle counts C_m under $\theta_m = \vartheta$

We have for the uniform and Ewens measure is

The cycle counts C_m under $\theta_m = \vartheta$

We have for the uniform and Ewens measure is

Theorem (Shepp,Loyd $\vartheta = 1$, Watterson general ϑ)

We have for each $b > 0$

$$(C_1, \ldots, C_b) \xrightarrow{d} (Y_1, \ldots, Y_b)$$

with Y_m independent Poisson distributed with

$$\mathbb{E}[Y_m] = \frac{\theta_m}{m} = \frac{\vartheta}{m}.$$

The cycle counts C_m under $\theta_m = \vartheta$

We have for the uniform and Ewens measure is

Theorem (Shepp,Loyd $\vartheta = 1$, Watterson general ϑ)

We have for each $b > 0$

$$(C_1, \dots, C_b) \xrightarrow{d} (Y_1, \dots, Y_b)$$

with Y_m independent Poisson distributed with

$$\mathbb{E}[Y_m] = \frac{\theta_m}{m} = \frac{\vartheta}{m}.$$

This result can be improved further.

The cycle counts C_m under $\theta_m = \vartheta$

Let Ω be a countable set and $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ be two probability measures on Ω.

The cycle counts C_m under $\theta_m = \vartheta$

Let Ω be a countable set and $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ be two probability measures on Ω.

The total variation distance between $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ is defined as

$$\mathrm{d}_{\mathrm{TV}}(\mathbb{P}, \widetilde{\mathbb{P}}) = \frac{1}{2} \sum_{\omega \in \Omega} |\mathbb{P}(\omega) - \widetilde{\mathbb{P}}(\omega)|.$$

The cycle counts C_m under $\theta_m = \vartheta$

For instance

Theorem (Arratia and Tavaré)

Let S_n be endowed with the uniform measure. Then

$$d_{\mathrm{TV}}\big((C_1, \dots, C_b), (Y_1, \dots, Y_b)\big) \to 0$$

if and only if $b = o(n)$, where d_{TV} denotes the total variation distance.

The cycle counts C_m under $\theta_m = m^\gamma$

We have for the cycle weights $\theta_m = m^\gamma$ with $\gamma > 0$

The cycle counts C_m under $\theta_m = m^\gamma$

We have for the cycle weights $\theta_m = m^\gamma$ with $\gamma > 0$

Theorem (Ercolani and Ueltschi)

We have for each $b > 0$

$$(C_1, \dots, C_b) \xrightarrow{d} (Y_1, \dots, Y_b)$$

with Y_m independent Poisson distributed with

$$\mathbb{E}[Y_m] = \frac{\theta_m}{m} = m^{\gamma - 1}.$$

The cycle counts C_m under $\theta_m = m^\gamma$

As for the Ewens measure, we can also compute the total variation distance.

The cycle counts C_m under $\theta_m = m^\gamma$

As for the Ewens measure, we can also compute the total variation distance.

Theorem (Storm and Zeindler)

We then have as $n \to \infty$

$$d_{\mathrm{TV}}\big((C_1, \ldots, C_b), (Y_1, \ldots, Y_b)\big) \to 0 \quad \textit{iff } b = o(n^{\frac{1}{1+\gamma}}),$$

where Y_m are independent Poisson distributed with $\mathbb{E}[Y_m] = \frac{\theta_m}{m} = m^{\gamma-1}$.

The cycle counts C_m under $\theta_m = \log^k(m)$

We have for the cycle weights $\theta_m = \log^k(m)$ with $k \geq 1$

The cycle counts C_m under $\theta_m = \log^k(m)$

We have for the cycle weights $\theta_m = \log^k(m)$ with $k \geq 1$

Theorem (Robles and Zeindler)

We have for each $b > 0$

$$(C_1, \dots, C_b) \xrightarrow{d} (Y_1, \dots, Y_b)$$

with Y_m independent Poisson distributed with

$$\mathbb{E}[Y_m] = \frac{\theta_m}{m} = \frac{\log^k(m)}{m}.$$

The cycle counts C_m under $\theta_m = \log^k(m)$

Theorem (Robles and Zeindler)

Suppose that $b(n) = o\big(n^c\big)$ with $0 < c < (3k+3)^{-\frac{1}{k+1}}$. We then have

$$d_{\mathrm{TV}}\big((C_1, \ldots, C_b), (Y_1, \ldots, Y_b)\big) \to 0$$

The cycle counts C_m under $\theta_m = \log^k(m)$

Theorem (Robles and Zeindler)

Suppose that $b(n) = o\big(n^c\big)$ with $0 < c < (3k+3)^{-\frac{1}{k+1}}$. We then have

$$d_{\mathrm{TV}}\big((C_1, \ldots, C_b), (Y_1, \ldots, Y_b)\big) \to 0$$

We expect that that the d_{TV} in this theorem goes to 0 if and only if $b(n) = o\left(\frac{n}{\log^k(n)}\right)$.

The cycle containing 1

We define

$$\ell_1 := \text{the length of the cycle containing } 1.$$

The cycle containing 1

We define

$$\ell_1 := \text{the length of the cycle containing 1.}$$

A simple computation gives that we have under the uniform measure on S_n

$$\mathbb{P}_n\left[\ell_1 = k\right] = \frac{1}{n} \quad \text{for all } 1 \leq k \leq n.$$

The cycle containing 1 and the uniform measure

We thus immediately get

Theorem

We have under the uniform measure

$$\frac{\ell_1}{n} \xrightarrow{d} U(0,1),$$

where $U(0,1)$ is the uniform measure on the interval $[0,1]$.

The cycle containing 1 and the Ewens measure

Furthermore

Theorem

We have under the Ewens measure with parameter $\vartheta > 0$

$$\frac{\ell_1}{n} \xrightarrow{d} Beta(1, \vartheta).$$

where $Beta(1, \vartheta)$ is the probability measure on the interval $[0, 1]$ with density $(1 - x)^{\vartheta - 1}$.

The cycle containing 1 under $\theta_m = m^\gamma$

Theorem (Ercolani, Ueltschi)

We have in the case $\theta_m = m^\gamma$

$$\frac{\ell_1}{n^{\frac{1}{1+\gamma}}} \xrightarrow{d} \Gamma\left(1+\gamma, \Gamma(1+\gamma)^{\frac{1}{1+\gamma}}\right)$$

where $\Gamma(\alpha, \beta)$ is the probability measure on $[0, \infty[$ with density $\frac{\beta^\alpha}{\Gamma(\alpha)} x^\alpha e^{-\beta x}$ and $\Gamma(s)$ the Gamma function.

The cycle containing 1 under $\theta_m = \log^k(m)$

Theorem (Robles, Zeindler)

We have in the case $\theta_m = \theta_m = \log^k(m)$

$$\frac{\ell_1}{\frac{n}{\log^k(n)}} \xrightarrow{d} Exp(1),$$

where $Exp(\lambda)$ is the probability measure on $[0, \infty[$ with density $\lambda e^{-\lambda x}$.

The cycle containing 1

$$d_{\mathrm{TV}}\big((C_1,\ldots,C_b),\,(Y_1,\ldots,Y_b)\big)\to 0$$

[1] conjectural

The cycle containing 1

$$d_{\mathrm{TV}}\Big((C_1,\dots,C_b),(Y_1,\dots,Y_b)\Big)\to 0$$

θ_m	$=\vartheta$	$=m^{\gamma}$	$=\log^k(m)$
$d_{\mathrm{TV}}\to 0$			
ℓ_1			

[1] conjectural

The cycle containing 1

$$d_{\mathrm{TV}}\big((C_1,\ldots,C_b),(Y_1,\ldots,Y_b)\big)\to 0$$

θ_m	$=\vartheta$	$=m^\gamma$	$=\log^k(m)$
$d_{\mathrm{TV}}\to 0$	$b=o(n)$		
ℓ_1	$\asymp n$		

[1] conjectural

The cycle containing 1

$$d_{\mathrm{TV}}\big((C_1,\dots,C_b),(Y_1,\dots,Y_b)\big)\to 0$$

θ_m	$=\vartheta$	$=m^\gamma$	$=\log^k(m)$
$d_{\mathrm{TV}}\to 0$	$b=o(n)$	$b=o\left(n^{\frac{1}{1+\gamma}}\right)$	
ℓ_1	$\asymp n$	$\asymp n^{\frac{1}{1+\gamma}}$	

[1]conjectural

The cycle containing 1

$$d_{\mathrm{TV}}\big((C_1,\dots,C_b),(Y_1,\dots,Y_b)\big)\to 0$$

θ_m	$=\vartheta$	$=m^\gamma$	$=\log^k(m)$
$d_{\mathrm{TV}}\to 0$	$b=o(n)$	$b=o\left(n^{\frac{1}{1+\gamma}}\right)$	$b=o\left(\frac{n}{\log^k(n)}\right)$ [1]
ℓ_1	$\asymp n$	$\asymp n^{\frac{1}{1+\gamma}}$	

[1] conjectural

The cycle containing 1

$$d_{\mathrm{TV}}\big((C_1,\dots,C_b),(Y_1,\dots,Y_b)\big)\to 0$$

θ_m	$=\vartheta$	$=m^\gamma$	$=\log^k(m)$
$d_{\mathrm{TV}}\to 0$	$b=o(n)$	$b=o\left(n^{\frac{1}{1+\gamma}}\right)$	$b=o\left(\frac{n}{\log^k(n)}\right)$ [1]
ℓ_1	$\asymp n$	$\asymp n^{\frac{1}{1+\gamma}}$	$\asymp n/\log^k(n)$

[1] conjectural

Total Number of Cycles

The Total number of cycles is defined as

$$K_n := C_1 + \cdots + C_n.$$

Total Number of Cycles

The Total number of cycles is defined as

$$K_n := C_1 + \cdots + C_n.$$

We have under the uniform and the Ewens measure.

Total Number of Cycles

The Total number of cycles is defined as

$$K_n := C_1 + \cdots + C_n.$$

We have under the uniform and the Ewens measure.

Theorem (Goncharov $\vartheta = 1$, Watterson general ϑ)

$$\frac{K_n - \vartheta \log(n)}{\sqrt{\vartheta \log(n)}} \xrightarrow{d} N(0,1)$$

Total Number of Cycles under '$\theta_m \to \vartheta$'

Theorem (Nikeghbali and Zeindler)

Consider the general Ewens measure. We then have

$$\mathbb{E}_\Theta\left[\exp\big(sK_n\big)\right] = n^{\theta(e^s-1)}\left(\frac{\Gamma(\theta)}{\Gamma(\theta e^s)} + O\left(\frac{1}{n}\right)\right)$$

with $O(.)$ uniform for bounded $s \in \mathbb{C}$.

Total Number of Cycles under '$\theta_m \to \vartheta$'

Recall, the Kolmogorov distance between two integer valued random variables X and Y is defined as

$$d_K(X, Y) := \sup_{j \in \mathbb{Z}} \left| \mathbb{P}_n [X \leq j] - \mathbb{P}_n [Y \leq j] \right|$$

Total Number of Cycles under '$\theta_m \to \vartheta$'

Recall, the Kolmogorov distance between two integer valued random variables X and Y is defined as

$$d_K(X,Y) := \sup_{j\in\mathbb{Z}} \left| \mathbb{P}_n\left[X \le j\right] - \mathbb{P}_n\left[Y \le j\right] \right|$$

Corollary

Let $P_{\theta\log(n)}$ be a Poisson distributed random variable with parameter $\theta\log(n)$. Then

$$d_K(K_n, P_{\theta\log(n)}) = O\left(\frac{1}{\sqrt{\log(n)}}\right).$$

Total Number of Cycles under $\theta_m = m^\gamma$

Theorem (Maples, Nikeghbali, Zeindler)

We have in the case $\theta_m = m^\gamma$

$$\frac{K_n - c_\gamma n^{\frac{\gamma}{1+\gamma}}}{\sqrt{d_\gamma n^{\frac{\gamma}{1+\gamma}}}} \xrightarrow{d} N(0,1)$$

Total Number of Cycles under $\theta_m = \log^k(m)$

Theorem (Robles, Zeindler)

We have for $\theta_m = \log^k(m)$

$$\frac{K_n - \frac{\log^{k+1}(n)}{k+1}}{\sqrt{\frac{\log^{k+1}(n)}{k+1}}} \xrightarrow{d} N(0,1)$$

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell.$$

We write λ_j for the length of the cycle σ_j.
W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \dots$.

Long Cycles

Let

$$\sigma = \sigma_1 \sigma_2 \cdots \sigma_\ell .$$

We write λ_j for the length of the cycle σ_j.
W.l.o.g. we can assume $\lambda_1 \geq \lambda_2 \geq \ldots$.

Theorem (Vershik and Shmidt resp. Kingman)

$$\left(\frac{\lambda_1}{n}, \frac{\lambda_2}{n}, \ldots \right) \xrightarrow{d} \mathcal{PD}(\vartheta), \qquad (n \to \infty)$$

with $\mathcal{PD}(\vartheta)$ the Poisson–Dirichlet distribution with parameter ϑ.

Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?

Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.

Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$.

Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$.
Consider a stick of length 1

Poisson–Dirichlet distribution

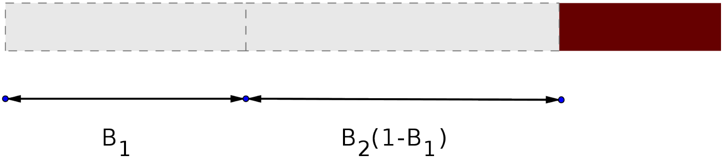
What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1,\vartheta)$.
Consider a stick of length 1

Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k \in \mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$.
Consider a stick of length 1

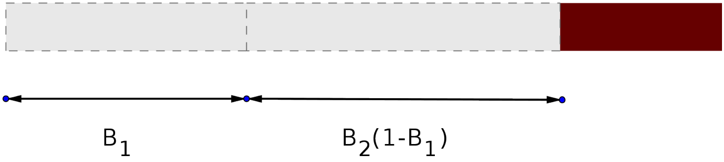
Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k\in\mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$.
Consider a stick of length 1



Poisson–Dirichlet distribution

What is the Poisson–Dirichlet distribution?
Keyword: stick breaking process with size ordering.
Let $(B_k)_{k \in \mathbb{N}}$ be iid Beta distributed with parameters $(1, \vartheta)$.
Consider a stick of length 1



Ordering the sticks obtained by this process by size then has a Poisson–Dirichlet distribution.

Long Cycles under $\theta_m = m^\gamma$

Let us now consider the case $\theta_m = m^\gamma$.

Long Cycles under $\theta_m = m^\gamma$

Let us now consider the case $\theta_m = m^\gamma$.
Do have in this case also cycles of order n?

Long Cycles under $\theta_m = m^\gamma$

Let us now consider the case $\theta_m = m^\gamma$.
Do have in this case also cycles of order n? No!

Long Cycles under $\theta_m = m^\gamma$

Let us now consider the case $\theta_m = m^\gamma$.
Do have in this case also cycles of order n? No!

Long Cycles under $\theta_m = m^\gamma$

Define

$$n^* := n^{\frac{1}{1+\gamma}} \quad \text{and} \quad \ell := \gamma \log(n^*) + (\gamma - 1) \log\left(\gamma \log(n^*)\right).$$

Long Cycles under $\theta_m = m^\gamma$

Define

$$n^* := n^{\frac{1}{1+\gamma}} \quad \text{and} \quad \ell := \gamma \log(n^*) + (\gamma - 1) \log\left(\gamma \log(n^*)\right).$$

Theorem (Zeindler)

Let $K \in \mathbb{N}$ be given. We have convergence in distribution of

$$\frac{1}{n^*} \cdot \left(\widetilde{L}_1 - n^*\ell, \ldots, \widetilde{L}_K - n^*\ell\right) \xrightarrow{d} \left(-\log(E_1), \ldots, -\log\left(\sum_{j=1}^{K} E_j\right)\right).$$

as $n \to \infty$, where $(E_j)_{j=1}^{K}$ is a sequence of iid $Exp(1)$ distributed random variables.

Long Cycles under $\theta_m = m^\gamma$

There is an important difference between the cases $\theta_m = \vartheta$ and $\theta_m = m^\gamma$.

Long Cycles under $\theta_m = m^\gamma$

There is an important difference between the cases $\theta_m = \vartheta$ and $\theta_m = m^\gamma$. Let us consider

$$\frac{1}{n} \sum_{m \geq d} m C_m$$

Long Cycles under $\theta_m = m^\gamma$

There is an important difference between the cases $\theta_m = \vartheta$ and $\theta_m = m^\gamma$. Let us consider

$$\frac{1}{n} \sum_{m \geq d} m C_m$$

This can be interpreted as the fraction of indices in cycles of length at least d.

Long Cycles under $\theta_m = m^\gamma$

We have in the case $\theta_m = \vartheta$

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E} \left[\frac{1}{n} \sum_{m \geq \epsilon n} m C_m \right] = 1$$

Long Cycles under $\theta_m = m^\gamma$

We have in the case $\theta_m = \vartheta$

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E}\left[\frac{1}{n} \sum_{m \geq \epsilon n} m C_m\right] = 1$$

and in the case $\theta_m = m^\gamma$

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E}\left[\frac{1}{n} \sum_{m \geq \epsilon n^* \ell_n} m C_m\right] = 0.$$

Long Cycles under $\theta_m = m^\gamma$

We have in the case $\theta_m = \vartheta$

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E}\left[\frac{1}{n} \sum_{m \geq \epsilon n} m C_m\right] = 1$$

and in the case $\theta_m = m^\gamma$

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \mathbb{E}\left[\frac{1}{n} \sum_{m \geq \epsilon n^* \ell_n} m C_m\right] = 0.$$

Recall, we have $\mathbb{E}[\ell_1] \approx n^{\frac{1}{1+\gamma}}$.

Long Cycles under $\theta_m = m^\gamma$

We have in the case $\theta_m = \vartheta$

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{n}\sum_{m\geq\epsilon n} mC_m\right] = 1$$

and in the case $\theta_m = m^\gamma$

$$\lim_{\epsilon\to 0}\lim_{n\to\infty}\mathbb{E}\left[\frac{1}{n}\sum_{m\geq\epsilon n^*\ell_n} mC_m\right] = 0.$$

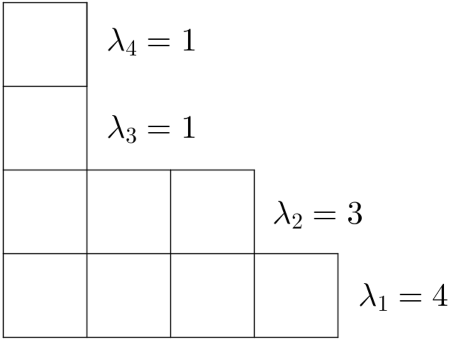
Recall, we have $\mathbb{E}\left[\ell_1\right] \approx n^{\frac{1}{1+\gamma}}$. Thus it is natural to study also the cycles the region $n^{\frac{1}{1+\gamma}}$.

Young diagram

$$\sigma = (128)(3)(4579)(6) \in S_9 \Rightarrow \lambda = (4, 3, 1, 1)$$

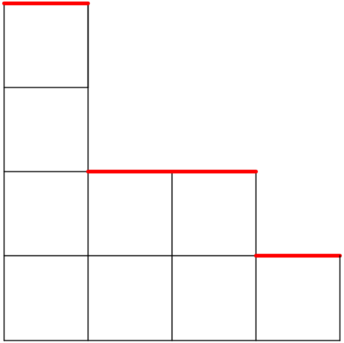
Young diagram

$$\sigma = (128)(3)(4579)(6) \in S_9 \Rightarrow \lambda = (4, 3, 1, 1)$$



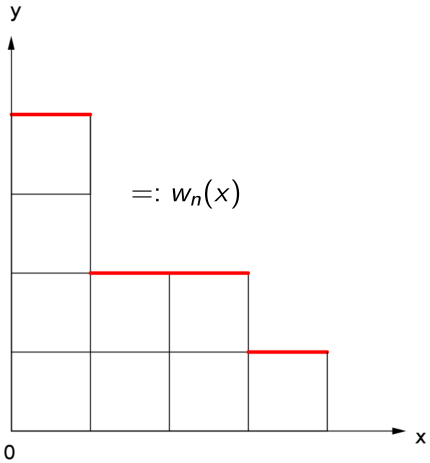
Young diagram

$$\sigma = (128)(3)(4579)(6) \in S_9 \Rightarrow \lambda = (4, 3, 1, 1)$$



Young diagram

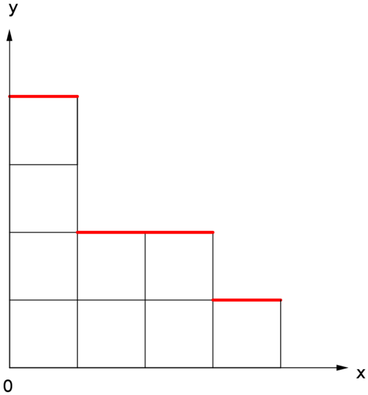
$$\sigma = (128)(3)(4579)(6) \in S_9 \Rightarrow \lambda = (4,3,1,1)$$



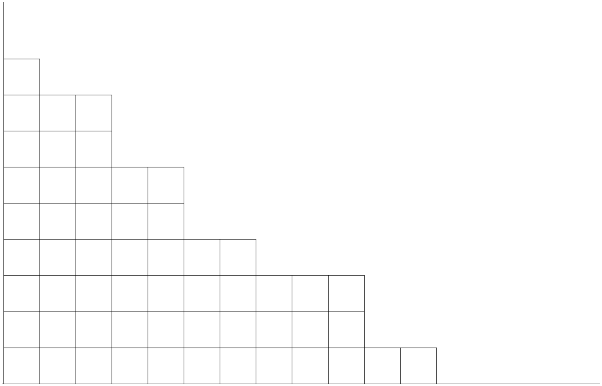
Formula for $w_n(x)$

We then obtain

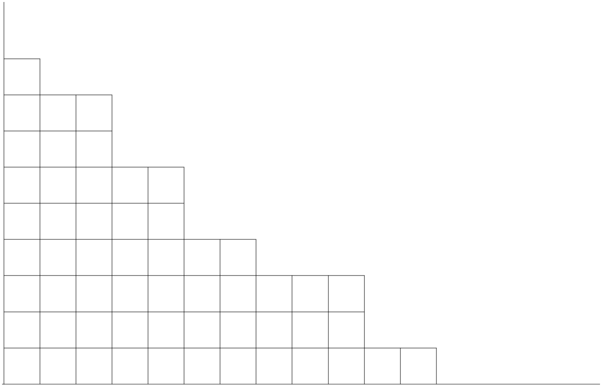
$$w_n(x) = \sum_{m \geq x}^{n} C_m$$



Limit shape

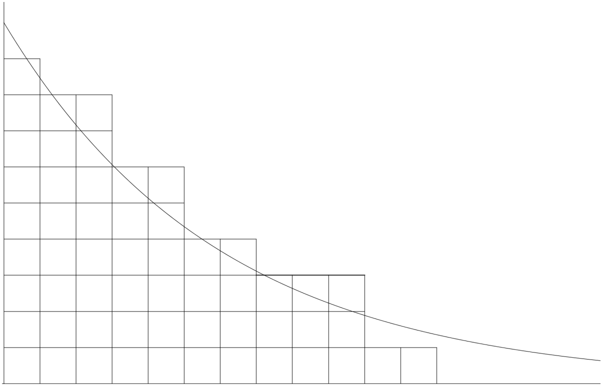


Limit shape



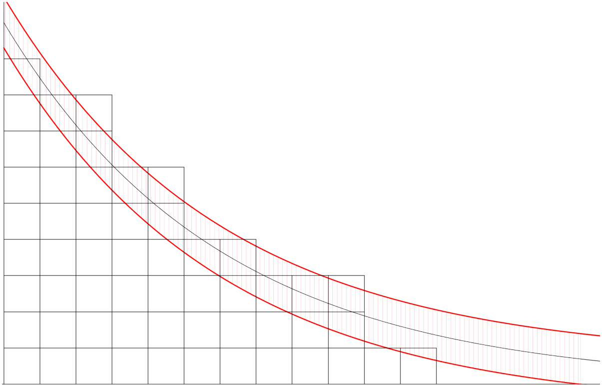
What happens if we take a σ randomly?

Limit shape



What happens if we take a σ randomly?

Limit shape

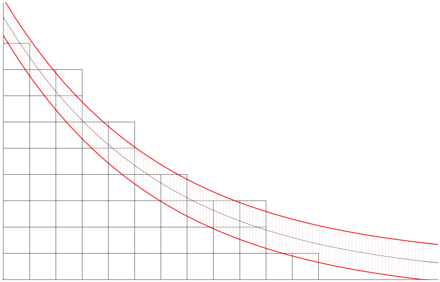


What happens if we take a σ randomly?

The *limit shape* is a function $\omega(\cdot)$ s.t. for all $\epsilon, \delta > 0$

$$\lim_{n \to +\infty} \mathbb{P}_n \left[\sup_{x \geq \delta} |\tilde{w}_n(x) - \omega(x)| \leq \epsilon \right] = 1$$

$\tilde{w}_n$ is an appropriate rescaling of w_n.



Limit shape under $\theta_m = m^\gamma$

Recall, we have $\ell_1 \asymp n^{\frac{1}{1+\gamma}}$.

Limit shape under $\theta_m = m^\gamma$

Recall, we have $\ell_1 \asymp n^{\frac{1}{1+\gamma}}$.
Thus it is natural to use the scale w_n as

$$\widetilde{w}_n(x) := n^{-\frac{\gamma}{1+\gamma}} w_n\left(x n^{\frac{1}{1+\gamma}}\right)$$

Limit shape under $\theta_m = m^\gamma$

Recall, we have $\ell_1 \asymp n^{\frac{1}{1+\gamma}}$.
Thus it is natural to use the scale w_n as

$$\widetilde{w}_n(x) := n^{-\frac{\gamma}{1+\gamma}} w_n\left(xn^{\frac{1}{1+\gamma}}\right)$$

Proposition (Erlihson, Granovsky (2008))

The limit shape exist and is given by

$$\omega(x) := \frac{\Gamma(\gamma, x)}{\Gamma(\gamma+1)},$$

where $\Gamma(\,\cdot\,,\,\cdot\,)$ *is the incomplete* Γ*-function.*

Limit shape under $\theta_m = m^\gamma$

Recall, we have $\ell_1 \asymp n^{\frac{1}{1+\gamma}}$.
Thus it is natural to use the scale w_n as

$$\widetilde{w}_n(x) := n^{-\frac{\gamma}{1+\gamma}} w_n\left(xn^{\frac{1}{1+\gamma}}\right)$$

Proposition (Erlihson, Granovsky (2008))

The limit shape exist and is given by

$$\omega(x) := \frac{\Gamma(\gamma, x)}{\Gamma(\gamma+1)},$$

where $\Gamma(\cdot, \cdot)$ is the incomplete Γ-function.
For $\gamma = 1$, we have $\omega(x) = e^{-x}$

Limit shape under $\theta_m = \log^k(m)$

Limit shape under $\theta_m = \log^k(m)$

Recall, we have $\ell_1 \asymp \frac{n}{\log^k(n)}$.

Limit shape under $\theta_m = \log^k(m)$

Recall, we have $\ell_1 \asymp \frac{n}{\log^k(n)}$.
Thus it is natural to use the scale w_n as

$$\widetilde{w}_n(x) := \log^{-k}(n) w_n \left(x \frac{n}{\log^k(n)} \right)$$

Limit shape under $\theta_m = \log^k(m)$

Recall, we have $\ell_1 \asymp \frac{n}{\log^k(n)}$.
Thus it is natural to use the scale w_n as

$$\widetilde{w}_n(x) := \log^{-k}(n) w_n \left(x \frac{n}{\log^k(n)} \right)$$

Theorem (Robles, Zeindler (2018))

Let $k \geq 3$.

Limit shape under $\theta_m = \log^k(m)$

Recall, we have $\ell_1 \asymp \frac{n}{\log^k(n)}$.
Thus it is natural to use the scale w_n as

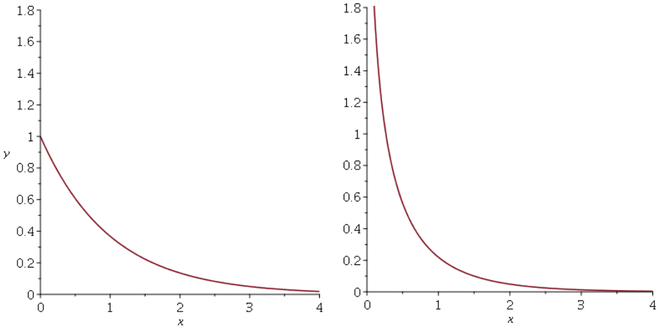
$$\widetilde{w}_n(x) := \log^{-k}(n) w_n \left(x \frac{n}{\log^k(n)} \right)$$

Theorem (Robles, Zeindler (2018))

Let $k \geq 3$. Then the limit shape exists and is given by

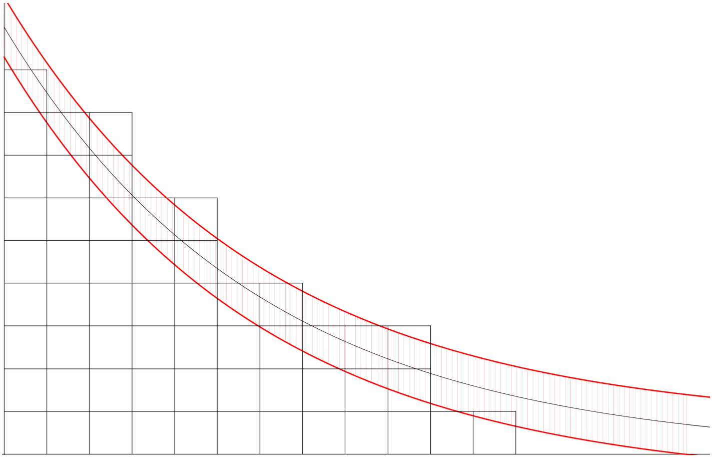
$$w_\infty(x) = \int_x^\infty u^{-1} e^{-u} \, du. \tag{1}$$

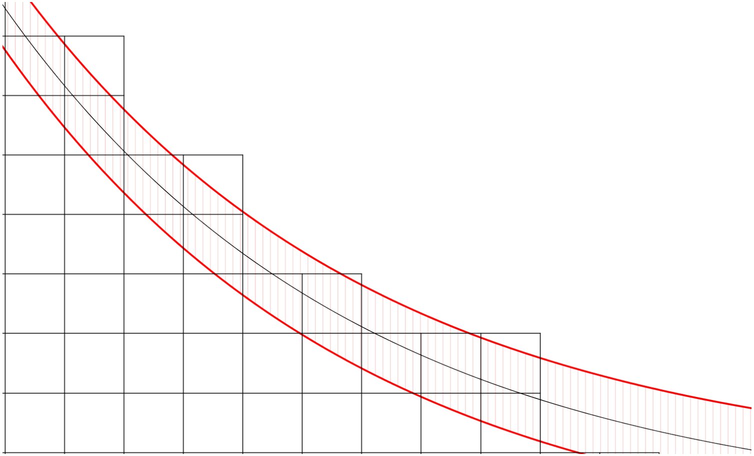
Limit shapes

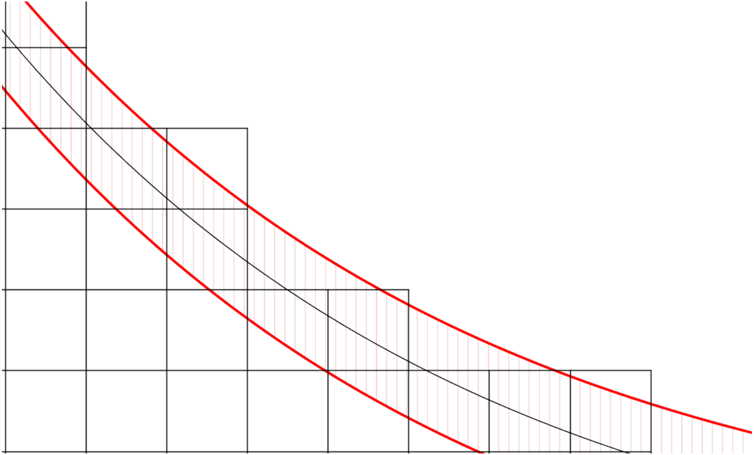


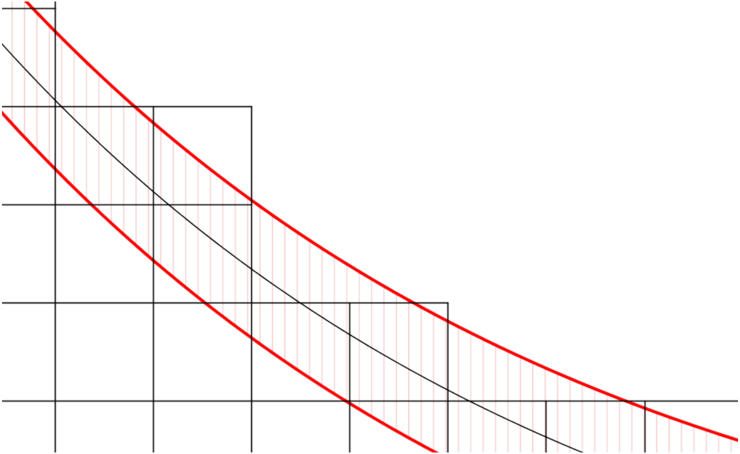
(a) $\theta_m = m^\gamma$ (b) $\theta_m = \log^k(m)$

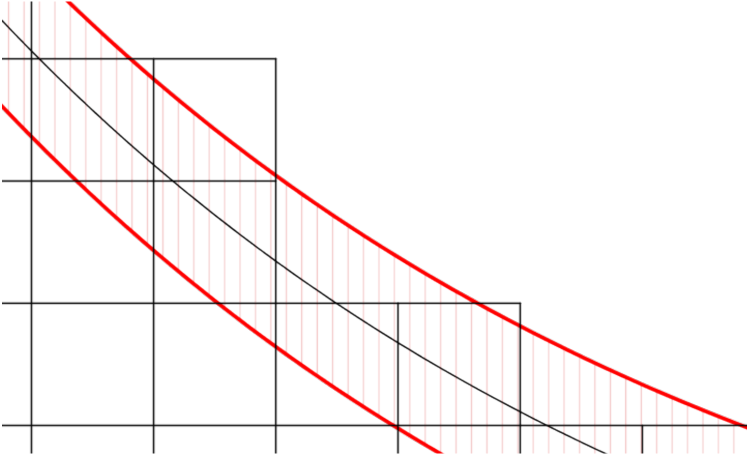
Figure: Plots of limit shapes

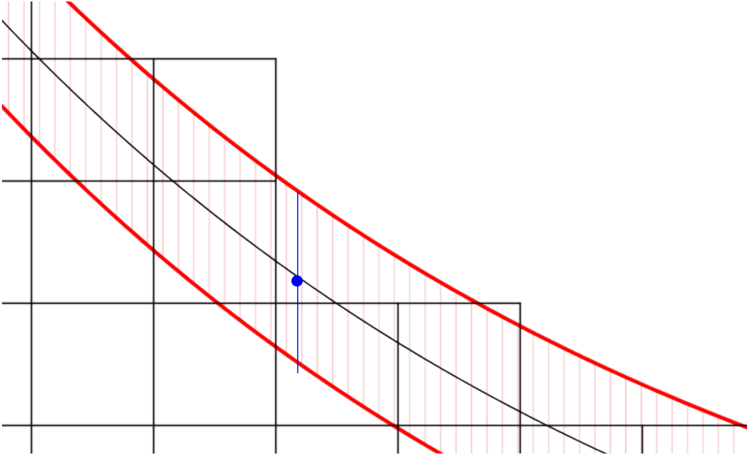












Fluctuations under $\theta_m = m^\gamma$

We define

$$\widehat{w}_n(x) := \frac{w_n\left(x\, n^{\frac{1}{1+\gamma}}\right) - n^{\frac{\gamma}{1+\gamma}}\omega(x)}{\sqrt{n^{\frac{\gamma}{1+\gamma}}}}$$

Fluctuations under $\theta_m = m^\gamma$

We define

$$\widehat{w}_n(x) := \frac{w_n\left(x\, n^{\frac{1}{1+\gamma}}\right) - n^{\frac{\gamma}{1+\gamma}}\omega(x)}{\sqrt{n^{\frac{\gamma}{1+\gamma}}}}$$

Theorem (Erlihson, Granovsky (2008))

We have as $n \to \infty$

$$\widehat{w}_n(x) \to \mathcal{N}\left(0, \omega(x) - \frac{\Gamma(\gamma+1, x)^2}{2\Gamma(\gamma+1)\Gamma(\gamma+2)}\right).$$

Fluctuations under $\theta_m = \log^k(m)$

We define

$$\widehat{w}_n(x) := \frac{w_n\left(x \frac{n}{\log^k(n)}\right) - \log^k(n)\,\omega(x)}{\sqrt{\log^k(n)}}$$

Fluctuations under $\theta_m = \log^k(m)$

We define

$$\widehat{w}_n(x) := \frac{w_n\left(x\frac{n}{\log^k(n)}\right) - \log^k(n)\,\omega(x)}{\sqrt{\log^k(n)}}$$

Theorem (Robles, Zeindler (2018))

We have for $k \geq 3$ and as $n \to \infty$

$$\widehat{w}_n(x) \to \mathcal{N}\left(0, \omega(x) + e^{-2x}\right).$$

Thank you!