WWU MÜNSTER

Markov renewal theory in the analysis of tries and strings

Gerold Alsmeyer

May 20, 2019

living.knowledge

Institute for Mathematical Stochastics

Overview

Janson: *Renewal theory in the analysis of tries and strings*

↪ survey of simple applications of **renewal theory** to problems on random strings and tries in iid setting (Bernoulli sources)

↪ intention: to advertise the use of renewal theory in more general settings such as Markov sources

Overview

Janson: *Renewal theory in the analysis of tries and strings*

↪ survey of simple applications of **renewal theory** to problems on random strings and tries in iid setting (Bernoulli sources)

↪ intention: to advertise the use of renewal theory in more general settings such as Markov sources

This talk: transfer of Janson's method to Markov model via Markov renewal theory, exemplified by the analysis of the expected depth of a string in a trie

Strings, tries and depth

The iid setting (Bernoulli source)

- Distributional equality
- Result and ideas for the proof
- Lattice-type

The Markov setting

- Approach
- Lattice-type

This is joint work with Philipp Godland.

Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \ldots) = \xi_1\xi_2\cdots$ of letters $\xi_i \in \{0, 1\}$

Trie: Ordered, rooted tree, that stores finite set of strings.

Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, ...) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

Example

$\Xi^{(1)} = 0011 \cdots$
$\Xi^{(2)} = 1010 \cdots$
$\Xi^{(3)} = 0111 \cdots$
$\Xi^{(4)} = 0010 \cdots$

Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \dots) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

$\Xi^{(1)}, \Xi^{(2)}, \Xi^{(3)}, \Xi^{(4)}$

Example

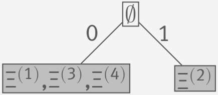
$\Xi^{(1)} = 0011 \cdots$
$\Xi^{(2)} = 1010 \cdots$
$\Xi^{(3)} = 0111 \cdots$
$\Xi^{(4)} = 0010 \cdots$

Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \ldots) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

Example

$\Xi^{(1)} = 0011\cdots$
$\Xi^{(2)} = 1010\cdots$
$\Xi^{(3)} = 0111\cdots$
$\Xi^{(4)} = 0010\cdots$

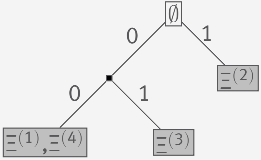


Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \ldots) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

Example

$\Xi^{(1)} = 0011\cdots$
$\Xi^{(2)} = 1010\cdots$
$\Xi^{(3)} = 0111\cdots$
$\Xi^{(4)} = 0010\cdots$

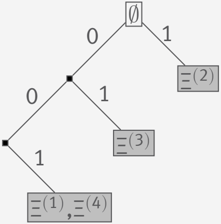


Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \ldots) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

Example

$\Xi^{(1)} = 0011 \cdots$
$\Xi^{(2)} = 1010 \cdots$
$\Xi^{(3)} = 0111 \cdots$
$\Xi^{(4)} = 0010 \cdots$

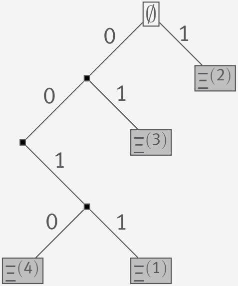


Strings and tries

String: Infinite sequence $\Xi = (\xi_1, \xi_2, \ldots) = \xi_1 \xi_2 \cdots$ of letters $\xi_i \in \{0, 1\}$
Trie: Ordered, rooted tree, that stores finite set of strings.

Example

$\Xi^{(1)} = 0011\cdots$
$\Xi^{(2)} = 1010\cdots$
$\Xi^{(3)} = 0111\cdots$
$\Xi^{(4)} = 0010\cdots$



Depth of a string in a trie

Depth of string Ξ in a trie $\stackrel{\wedge}{=}$ path length from root to the node that stores Ξ

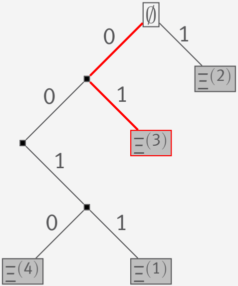
Depth of a string in a trie

Depth of string Ξ in a trie $\triangleq$ path length from root to the node that stores Ξ

Example

$\Xi^{(1)} = 0011\cdots$
$\Xi^{(2)} = 1010\cdots$
$\Xi^{(3)} = 0111\cdots$
$\Xi^{(4)} = 0010\cdots$

Depth of $\Xi^{(3)} = 2$



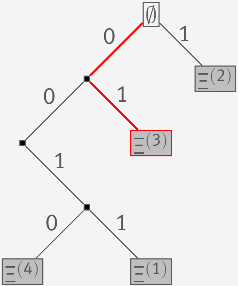
Depth of a string in a trie

Depth of string Ξ in a trie $\hat{=}$ path length from root to the node that stores Ξ

Example

$\Xi^{(1)} = 0011\cdots$
$\Xi^{(2)} = 1010\cdots$
$\Xi^{(3)} = 0111\cdots$
$\Xi^{(4)} = 0010\cdots$

Depth of $\Xi^{(3)} = 2$



Depth of a node represents retrieval cost when searching for it.

The iid setting (Bernoulli source)

The iid setting (Bernoulli source)

- $\xi_i \sim \mathrm{Bern(p)}$ and iid, $p \in (0, 1)$

The iid setting (Bernoulli source)

- $\xi_i \sim \mathrm{Bern(p)}$ and iid, $p \in (0,1)$
- $P(a_1 \cdots a_n) := \mathbb{P}((\xi_1, \dots, \xi_n) = (a_1, \dots, a_n)) = \prod_{i=1}^n P(a_i)$
- $X_i := -\log P(\xi_i)$, $i \in \mathbb{N}$
- $H := \mathbb{E}X_i = -p \log p - (1-p)\log(1-p)$, *entropy* of ξ_i

The iid setting (Bernoulli source)

- $\xi_i \sim \mathrm{Bern(p)}$ and iid, $p \in (0,1)$
- $P(a_1 \cdots a_n) := \mathbb{P}((\xi_1,\ldots,\xi_n) = (a_1,\ldots,a_n)) = \prod_{i=1}^n P(a_i)$
- $X_i := -\log P(\xi_i),\ i \in \mathbb{N}$
- $H := \mathbb{E}X_i = -p\log p - (1-p)\log(1-p)$, *entropy* of ξ_i
- $P(\xi_1 \cdots \xi_n) = \prod_{i=1}^n P(\xi_i) = \prod_{i=1}^n e^{-X_i} = e^{-S_n}$
- $\Xi^{(2)}, \Xi^{(3)}, \ldots$ are iid copies of Ξ

The iid setting (Bernoulli source)

- $\xi_i \sim \text{Bern(p)}$ and iid, $p \in (0,1)$
- $P(a_1 \cdots a_n) := \mathbb{P}((\xi_1,\ldots,\xi_n) = (a_1,\ldots,a_n)) = \prod_{i=1}^n P(a_i)$
- $X_i := -\log P(\xi_i)$, $i \in \mathbb{N}$
- $H := \mathbb{E}X_i = -p\log p - (1-p)\log(1-p)$, *entropy* of ξ_i
- $P(\xi_1 \cdots \xi_n) = \prod_{i=1}^n P(\xi_i) = \prod_{i=1}^n e^{-X_i} = e^{-S_n}$
- $\Xi^{(2)}, \Xi^{(3)}, \ldots$ are iid copies of Ξ

$(X_i)_{i\in\mathbb{N}}$ is a sequence of iid, non-negative RVs $\rightsquigarrow$ Renewal theory

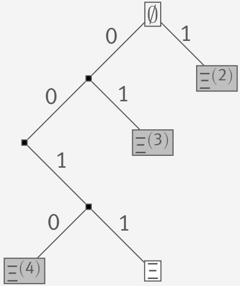
The iid setting (Bernoulli source)

$D_n :=$ depth of Ξ in trie constructed from $\Xi, \Xi^{(2)}, \ldots, \Xi^{(n)}$

The iid setting (Bernoulli source)

$D_n := $ depth of Ξ in trie constructed from $\Xi, \Xi^{(2)}, \ldots, \Xi^{(n)}$

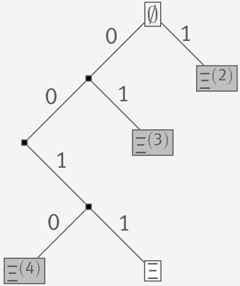
$D_n \le k$



The iid setting (Bernoulli source)

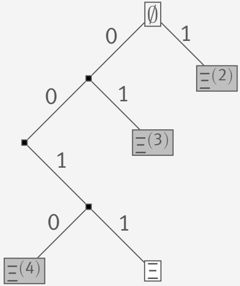
$D_n := $ depth of Ξ in trie constructed from $\Xi, \Xi^{(2)}, \ldots, \Xi^{(n)}$

$D_n \le k \Leftrightarrow$ no other string than Ξ starts with $\xi_1 \cdots \xi_k$



The iid setting (Bernoulli source)

$D_n :=$ depth of Ξ in trie constructed from $\Xi, \Xi^{(2)}, \dots, \Xi^{(n)}$



$D_n \le k \Leftrightarrow$ no other string than Ξ starts with $\xi_1 \cdots \xi_k$

$$\mathbb{P}(D_n \le k|\Xi) = (1 - P(\xi_1 \cdots \xi_k))^{n-1} = (1 - e^{-S_k})^{n-1} = \left(1 - \frac{e^{\log n - S_k}}{n}\right)^{n-1}$$

Distributional equality

For $n \geq 2$, choose starting variables $X_0^{(n)}$, independent of the rest, such that

$$\mathbb{P}(X_0^{(n)} > x) = \left(1 - \frac{e^x}{n}\right)_+^{n-1}.$$

Distributional equality

For $n \geq 2$, choose starting variables $X_0^{(n)}$, independent of the rest, such that

$$\mathbb{P}(X_0^{(n)} > x) = \left(1 - \frac{e^x}{n}\right)_+^{n-1}.$$

Then

$$\begin{aligned}\mathbb{P}(D_n \leq k|\Xi) = \left(1 - \frac{e^{\log n - S_k}}{n}\right)^{n-1} &= \mathbb{P}(X_0^{(n)} + S_k > \log n|\Xi) \\ &= \mathbb{P}(\nu(X_0^{(n)}, \log n) \leq k|\Xi),\end{aligned}$$

where $\nu(x,t) := \inf\{k \geq 0 | x + S_k > t\}$ is the *first passage time* of the random walk $(S_n)_{n\geq 0}$ beyond level t.

Distributional equality

Renewal theory deals a lot with $v(0,t)$ and $v(x,t)$.

Lemma (Janson)

$$D_n \stackrel{d}{=} v(X_0^{(n)}, \log n)$$

Note that

$$\lim_{n\to\infty} \mathbb{P}(X_0^{(n)} > x) = e^{-e^x}$$

for all $x \in \mathbb{R}$ (reflected Gumbel distribution)

Result for expected depth

Theorem (e.g. Janson)

- If $\frac{\log p}{\log(1-p)} \notin \mathbb{Q}$, then $\mathbb{E}v(0,t) = \frac{t}{H} + \frac{\mathbb{E}X_1^2}{2H^2} + o(1)$.
- If $\frac{\log p}{\log(1-p)} \in \mathbb{Q}$, then $\mathbb{E}v(0,t) = \frac{t}{H} + \frac{d}{H}\left(\frac{1}{2} - \left\{\frac{t}{d}\right\}\right) + \frac{\mathbb{E}X_1^2}{2H^2} + o(1)$, for some $d > 0$.
- If $\frac{\log p}{\log(1-p)} \notin \mathbb{Q}$, then $\mathbb{E}D_n = \frac{\log n}{H} + \frac{\mathbb{E}X_1^2}{2H^2} + \frac{\gamma}{H} + o(1)$. ($\gamma$ = Euler-Mascheroni constant)
- If $\frac{\log p}{\log(1-p)} \in \mathbb{Q}$, then $\mathbb{E}D_n = \frac{\log n}{H} + \psi(\log n) + \frac{\mathbb{E}X_1^2}{2H^2} + \frac{\gamma}{H} + o(1)$, with ψ continuous and periodic.

Idea of proof

Idea: Use existing results for $v(0,t)$ together with

$$v(X_0^{(t)},t) = v(0,t-X_0^{(t)})\mathbf{1}_{\{t-X_0^{(t)}\geq 0\}}$$

and conditioning on $X_0^{(t)}$.

Idea of proof

Idea: Use existing results for $v(0,t)$ together with

$$v(X_0^{(t)},t) = v(0,t-X_0^{(t)})\mathbf{1}_{\{t-X_0^{(t)}\ge 0\}}$$

and conditioning on $X_0^{(t)}$.

Additional requirements: $(X_0^{(t)})$ is tight and uniformly integrable.

Lattice-type

$$X_i = -\log P(\xi_i) = \begin{cases} -\log p, & \xi_i = 1 \\ -\log(1-p), & \xi_i = 0 \end{cases}$$

Lattice-type

$$X_i = -\log P(\xi_i) = \begin{cases} -\log p, & \xi_i = 1 \\ -\log(1-p), & \xi_i = 0 \end{cases}$$

Whether $\frac{\log p}{\log(1-p)} \in \mathbb{Q}$ or $\notin \mathbb{Q}$ depends on whether $(S_n)_n$ and its renewal measure $\mathbb{U} = \sum_n \mathbb{P}(S_n \in \cdot)$ are

- supported on $d\mathbb{Z}$ for some $d > 0$ (*arithmetic*).

  or

- not supported on $d\mathbb{Z}$ for any $d > 0$ (*non-arithmetic*).

Needs to be considered in asymptotic analysis (notorious example: Blackwell's renewal theorem).

Lattice-type

$$X_i = -\log P(\xi_i) = \begin{cases} -\log p, & \xi_i = 1 \\ -\log(1-p), & \xi_i = 0 \end{cases}$$

Whether $\frac{\log p}{\log(1-p)} \in \mathbb{Q}$ or $\notin \mathbb{Q}$ depends on whether $(S_n)_n$ and its renewal measure $\mathbb{U} = \sum_n \mathbb{P}(S_n \in \cdot)$ are

- supported on $d\mathbb{Z}$ for some $d > 0$ (*arithmetic*).

  or
- not supported on $d\mathbb{Z}$ for any $d > 0$ (*non-arithmetic*).

Needs to be considered in asymptotic analysis (notorious example: Blackwell's renewal theorem).

More complicated in the *Markov setting*!

The Markov setting

Goal: Allow Ξ to be a Markov chain and obtain similar theorem:

The Markov setting

Goal: Allow Ξ to be a Markov chain and obtain similar theorem:

- $\Xi = (\xi_n)_{n\geq 0}$ time-homogeneous MC on $\{0, 1\}$ with transition matrix $(p_{i,j})$, initial value ξ_0 and $p_{i,j} \in (0, 1)$ for all i, j
- irreducible (thus positive recurrent) with unique stationary distribution π
- $\mathbb{P}_i := \mathbb{P}(\cdot|\xi_0 = i)$ for $i = 0, 1$ and $P_i(a_1 \cdots a_n) := p_{i,a_1} p_{a_1,a_2} \cdots p_{a_{n-1},a_n}$
- $X_n := -\log p_{\xi_{n-1},\xi_n}$, $n \in \mathbb{N}$
- $P_i(\xi_1 \cdots \xi_n) = e^{-S_n}$ $\mathbb{P}_i$-a.s.
- under each $\mathbb{P}_i$ the MCs $\Xi^{(2)}, \Xi^{(3)}, \ldots$ are iid copies of Ξ (with same initial state)

Markov random walk

With $X_0 := 0$, $(\xi_n, X_n)_{n\geq 0}$ is a *Markov modulated sequence*, i.e. a time-homogeneous MC on $\{0, 1\} \times \mathbb{R}$ with the specific property

$$\mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot | \xi_n, X_n) = \mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot | \xi_n).$$

Markov random walk

With $X_0 := 0$, $(\xi_n, X_n)_{n\geq 0}$ is a *Markov modulated sequence*, i.e. a time-homogeneous MC on $\{0, 1\} \times \mathbb{R}$ with the specific property

$$\mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot | \xi_n, X_n) = \mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot | \xi_n).$$

The corresponding $(\xi_n, S_n)_{n\geq 0}$ is called *Markov random walk*. As the X_n are positive, $(\xi_n, S_n)_{n\geq 0}$ is called *Markov renewal process*.

Markov random walk

With $X_0 := 0$, $(\xi_n, X_n)_{n\geq 0}$ is a *Markov modulated sequence*, i.e. a time-homogeneous MC on $\{0,1\} \times \mathbb{R}$ with the specific property

$$\mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot\,|\xi_n, X_n) = \mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot\,|\xi_n).$$

The corresponding $(\xi_n, S_n)_{n\geq 0}$ is called *Markov random walk*. As the X_n are positive, $(\xi_n, S_n)_{n\geq 0}$ is called *Markov renewal process*.
$\rightsquigarrow$ Markov renewal theory

Markov random walk

With $X_0 := 0$, $(\xi_n, X_n)_{n\geq 0}$ is a *Markov modulated sequence*, i.e. a time-homogeneous MC on $\{0,1\} \times \mathbb{R}$ with the specific property

$$\mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot \,|\xi_n, X_n) = \mathbb{P}((\xi_{n+1}, X_{n+1}) \in \cdot \,|\xi_n).$$

The corresponding $(\xi_n, S_n)_{n\geq 0}$ is called *Markov random walk*. As the X_n are positive, $(\xi_n, S_n)_{n\geq 0}$ is called *Markov renewal process*.
$\rightsquigarrow$ Markov renewal theory

Important quantity: *Stationary mean/drift* $\mu = \sum_{i\in\{0,1\}} \pi_i \mathbb{E}_i X_1 > 0$

Approach

Distributional equality $D_n \stackrel{\mathrm{d}}{=} v(X_0^{(n)}, \log n)$ still holds w.r.t. every $\mathbb{P}_i$.

Find results for $v(0,t)$ and proceed as in the iid case.

Approach

Distributional equality $D_n \overset{\mathrm{d}}{=} v(X_0^{(n)}, \log n)$ still holds w.r.t. every $\mathbb{P}_i$.
Find results for $v(0,t)$ and proceed as in the iid case.

Idea: Combine theory of discrete MCs with ordinary renewal theory via *cyclic decomposition*:

- Let $\sigma_0(i) = 0$ and $\sigma_n(i) = \inf\{n > \sigma_{n-1}(i) : \xi_n = i\}$ be sequence of successive recurrence times of state i.
- Then cycles $Z_n = (\sigma_{n+1}(i) - \sigma_n(i), (\xi_k, X_{k+1})_{\sigma_n(i) \le k < \sigma_{n+1}(i)})$, $n \ge 0$, are independent and even identically distributed (w.r.t. every $\mathbb{P}_j$) for $n \ge 1$.
- $(S_{\sigma_n(i)})_{n \ge 0}$ is SRW/SRP with respect to $\mathbb{P}_i$.

Lattice-type

Observation: If $S_{\sigma_n(i)}$ has lattice span $d(i)$ w.r.t. $\mathbb{P}_i$, then $d = d(i)$ does not depend on i.

But: In general S_n is not concentrated on $d\mathbb{Z}$ w.r.t. $\mathbb{P}_i$!

Lattice-type

Observation: If $S_{\sigma_n(i)}$ has lattice span $d(i)$ w.r.t. $\mathbb{P}_i$, then $d = d(i)$ does not depend on i.
But: In general S_n is not concentrated on $d\mathbb{Z}$ w.r.t. $\mathbb{P}_i$!

Example

Ξ MC on $\{0,1\}$, β function on $\{0,1\}$. Define

$$X_n := 1 + \beta(\xi_n) - \beta(\xi_{n-1}), \quad X_0 := 0.$$

$(\xi_n, S_n)_{n\geq 0}$ is a Markov random walk, $S_{\sigma_n(i)} \in \mathbb{Z}$ $\mathbb{P}_i$-a.s., and

$$S_n = n + \beta(\xi_n) - \beta(i) \quad \mathbb{P}_i\text{-a.s},$$

generally not concentrated on $\mathbb{Z}$ ($\beta(x) = x$ and ξ_n iid, uniform on $[-1, 1]$).

Lattice-type

Lemma (Shurenkov)

If $d > 0$, then there exists a shift function $\beta: \{0, 1\} \to [0, d)$ such that $\mathbb{P}_i(S_n \in \beta(\xi_n) - \beta(i) + d\mathbb{Z}) = 1$ for all i.

We call $(\xi_n, S_n)_{n\geq 0}$ *d-arithmetic with shift function* β, if such β exists, otherwise *non-arithmetic*.

Lattice-type

Lemma (Shurenkov)

If $d > 0$, then there exists a shift function $\beta: \{0,1\} \to [0,d)$ such that $\mathbb{P}_i(S_n \in \beta(\xi_n) - \beta(i) + d\mathbb{Z}) = 1$ for all i.

We call $(\xi_n, S_n)_{n\ge 0}$ *d-arithmetic with shift function* β, if such β exists, otherwise *non-arithmetic*. Here $(\xi_n, S_n)_{n\ge 0}$ is arithmetic, if

$$\frac{\log(1-p_{0,0})}{\log p_{1,1}} + \frac{\log(1-p_{1,1})}{\log p_{1,1}} \text{ and } \frac{\log(1-p_{0,0})}{\log p_{0,0}} + \frac{\log(1-p_{1,1})}{\log p_{0,0}}$$

are both rational.

Expansion of expectation

Starting point is

$$\mathbb{E}_i v(0,t) = \sum_{n=0}^{\infty} \mathbb{P}_i(v(0,t) > n) = \sum_{n=0}^{\infty} \mathbb{P}_i(S_n \le t) = \sum_{j\in\{0,1\}} \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = j, S_n \le t)$$

Expansion of expectation

Starting point is

$$\mathbb{E}_i v(0,t) = \sum_{n=0}^{\infty} \mathbb{P}_i(v(0,t) > n) = \sum_{n=0}^{\infty} \mathbb{P}_i(S_n \le t) = \sum_{j \in \{0,1\}} \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = j, S_n \le t)$$

and asymptotics for the right side are known if non-arithmetic or d-arithmetic with shift 0:

$$\mathbb{U}^{ij}(B) := \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = j, S_n \in B) = \begin{cases} \sum_{n=1}^{\infty} \mathbb{P}_i(S_{\sigma_n(j)} \in B) & \text{if } j \neq i, \\ \sum_{n=0}^{\infty} \mathbb{P}_i(S_{\sigma_n(i)} \in B) & \text{otherwise.} \end{cases}$$

$\mathbb{U}^{ij}$ is renewal measure of $(S_{\sigma_n(j)})_n$ w.r.t. $\mathbb{P}_i$.

Expansion of expectation

Starting point is

$$\mathbb{E}_i v(0,t) = \sum_{n=0}^{\infty} \mathbb{P}_i(v(0,t) > n) = \sum_{n=0}^{\infty} \mathbb{P}_i(S_n \le t) = \sum_{j\in\{0,1\}} \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = j, S_n \le t)$$

and asymptotics for the right side are known if non-arithmetic or d-arithmetic with shift 0:

$$\mathbb{U}^{ij}(B) := \sum_{n=0}^{\infty} \mathbb{P}_i(\xi_n = j, S_n \in B) = \begin{cases} \sum_{n=1}^{\infty} \mathbb{P}_i(S_{\sigma_n(j)} \in B) & \text{if } j \neq i, \\ \sum_{n=0}^{\infty} \mathbb{P}_i(S_{\sigma_n(i)} \in B) & \text{otherwise.} \end{cases}$$

$\mathbb{U}^{ij}$ is renewal measure of $(S_{\sigma_n(j)})_n$ w.r.t. $\mathbb{P}_i$.

Observation: If $(\xi_n, S_n)_n$ is d-arithmetic with shift function β, then $(\xi_n, S_n - \beta(\xi_n) + \beta(\xi_0))_n$ is a d-arithmetic MRW with shift 0 (and same stationary drift as $(\xi_n, S_n)_n$).

Expansion of expectation

Theorem

(a) If $(\xi_n, S_n)_{n\geq 0}$ is non-arithmetic, then

$$\mathbb{E}_i v(0,t) = \frac{t}{\mu} + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S_{\sigma_1(j)}^2 - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

(b) If $(\xi_n, S_n)_{n\geq 0}$ is d-arithmetic with shift 0, then

$$\mathbb{E}_i v(0,t) = \frac{t}{\mu} + \frac{d}{\mu}\left(\frac{1}{2} - \left\{\frac{t}{d}\right\}\right) + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S_{\sigma_1(j)}^2 - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

(c) If $(\xi_n, S_n)_{n\geq 0}$ is d-arithmetic with shift β, then (with $\beta_{ij} = \beta(j) - \beta(i) \in (-d,d)$)

$$\mathbb{E}_i v(0,t) = \frac{t}{\mu} + \frac{d}{\mu}\left(\frac{1}{2} - \sum_{j\in\{0,1\}} \pi_j \left\{\frac{t-\beta_{ij}}{d}\right\}\right) + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S_{\sigma_1(j)}^2 - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

Expansion of expectation

Theorem

(a) If $(\xi_n, S_n)_{n\geq 0}$ is non-arithmetic, then

$$\mathbb{E}_i D_n = \frac{\log n}{\mu} + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S^2_{\sigma_1(j)} + \frac{\gamma}{\mu} - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

(b) If $(\xi_n, S_n)_{n\geq 0}$ is d-arithmetic with shift function 0, then (with ψ continuous and d-periodic)

$$\mathbb{E}_i D_n = \frac{\log n}{\mu} + \psi(\log n) + \frac{\gamma}{\mu} + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S^2_{\sigma_1(j)} - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

(c) If $(\xi_n, S_n)_{n\geq 0}$ is d-arithmetic with shift β, then

$$\mathbb{E}_i D_n = \frac{\log n}{\mu} + \sum_{j\in\{0,1\}} \pi_j \psi(\log n - \beta_{ij}) + \frac{\gamma}{\mu} + \frac{1}{2\mu^2} \sum_{j\in\{0,1\}} \pi_j^2 \mathbb{E}_j S^2_{\sigma_1(j)} - \frac{1}{\mu} \sum_{j\neq i} \pi_j \mathbb{E}_i S_{\sigma_1(j)} + o(1).$$

Further renewal theoretic arguments can be used in the analysis of a variety of other trie-related functionals such as

- imbalance factor
- size
- size of b-trie
- size of PATRICIA-trie
- depth in PATRICIA-trie
- fringe tries
- protected nodes
- external path length
- external path length of PATRICIA-trie
- ...

Further renewal theoretic arguments can be used in the analysis of a variety of other trie-related functionals such as

- imbalance factor
- size
- size of *b*-trie
- size of PATRICIA-trie
- depth in PATRICIA-trie
- fringe tries
- protected nodes
- external path length
- external path length of PATRICIA-trie
- ...

Thank you for your attention.