Theorem (Cycle lemma) For every $x \in S_n$, set $I(x) = \{ i \in \mathbb{Z}/n\mathbb{Z} : x^{(i)} \in \bar{S}_n \}$. Then $\mathrm{Card}(I(x)) = 1$ and the element of $I(x)$ is the first time when the walk with jumps given by x reaches its infimum for the first time

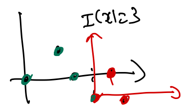
In the previous example $I(x) = \{3\}$.

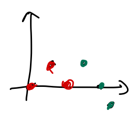
The proof is not complicated but a bit tedious to write: it is left to the reader.

We now see some a probabilistic consequence

Definition For $x \in S_n$, we define the Vervaat transform of x by $\mathcal{V}(x) = x^{(I(x))} \in \bar{S}_n$.

Example: $x = (1, -1, -1, 1, -1)$





Theorem The law of $\mathcal{V}(X_1, \ldots, X_n)$ under $\mathbb{P}(\cdot \mid W_n = -1)$ is equal to the law of $(X_1, \ldots, X_n)$ under $\mathbb{P}(\cdot \mid \zeta = n)$

Proof Set $\vec{X}_n = (X_1, \ldots, X_n)$. Take $x \in \bar{S}_n$, $0 \le k \le n-1$. Then

$$\mathbb{P}(\mathcal{V}(\vec{X}_n) = x, I(\vec{X}_n) = k, W_n = -1)$$
$$= \mathbb{P}(\vec{X}_n^{(k)} = x, I(\vec{X}_n) = k, W_n = -1)$$
$$= \mathbb{P}(\vec{X}_n^{(k)} = x, W_n = -1) \qquad \text{(Cycle lemma)}$$
$$= \mathbb{P}(\vec{X}_n = x, W_n = -1) \quad \text{because } \vec{X}_n^{(k)} \overset{(d)}{=} \vec{X}_n$$
$$= \mathbb{P}(\vec{X}_n = x, \zeta = n) \quad \text{because } x \in \bar{S}_n$$

Summing over k gives $\mathbb{P}(\mathcal{V}(\vec{X}_n) = x, W_n = -1) = n\, \mathbb{P}(\vec{X}_n = x, \zeta = n)$ $(*)$

Summing over $x \in \bar{S}_n$ gives $\mathbb{P}(W_n = -1) = n\, \mathbb{P}(\zeta = n)$. Dividing $(*)$ by this gives the result

To sum up, to understand $\mathcal{LW}(\mathcal{T}_n)$, one can study $(W_0, \dots, W_n)$ under $\mathbb{P}(\cdot \mid W_n=-1)$ and then apply the Vervaat transform.

Example The largest and second largest number of children of $\mathcal{T}_n$ is equal in law to the largest and second largest jump -1 of $(W_0, \dots, W_n)$ under $\mathbb{P}(\cdot \mid W_n=-1)$.

When μ has finite variance, W_n is of order $\mathbb{E}[W_1]n$ with fluctuations of order $\sqrt{n}$:

- When μ is not critical, the event $\{W_n=-1\}$ is a large deviation event.

In probability theory, many theorems concern "typical events", which have probability 1 or tending to one. Large deviations concern "atypical events" whose probability tends to 0. Typical questions then are:

- How fast is the convergence (rate of decay)?
- Given this atypical event, what are typical events of the system under the conditional law (known as the Gibbs conditioning principle in physics)?

II) A one-big-jump principle

Outline: 1) A local estimate

2) One big jump principle.

3) An application

We present a framework tailored to our application to random trees, but what follows can be extended to a more general context.

1) A local estimate

Let $(X_i)_{i\geq 1}$ be iid real-valued random variables. Set $S_0=0$ and $S_n = X_1 + \cdots + X_n$ for $n \geq 1$.

Assumption (H) $\mathbb{E}[X_1^2] < \infty$ and there exist $c>0$ and $\beta>2$ such that $\mathbb{P}(X_1 \in [u, u+1)) \underset{u\to\infty}{\sim} \frac{c}{u^{1+\beta}}$

It is not difficult to check that under (H), $\mathbb{P}(X_1 \geq u) \underset{u\to\infty}{\sim} \frac{c/\beta}{u^\beta}$ and that $\mathbb{E}[X_1^2]<\infty$.

Theorem (Doney '89, Nagaev '57)

Assume that X_1 satisfies (H) and that $\mathbb{E}[X_1]=0$. Fix $\varepsilon>0$. Then, uniformly in $m \geq \varepsilon n$,

$$\mathbb{P}(S_n \in [m, m+1)) \underset{n\to\infty}{\sim} n\, \mathbb{P}(X_1 \in [m, m+1))$$

That is, $$\sup_{m\geq \varepsilon n} \left| \frac{\mathbb{P}(S_n \in [m,m+1))}{n\, \mathbb{P}(X_1 \in [m,m+1))} - 1 \right| \xrightarrow[n\to\infty]{} 0$$

Intuition: $S_n \in [m, m+1)$ typically happens when one of the n jumps is in $[m,m+1)$ (this will be made precise later)

See the supplementary material for a proof.

2) A one-big jump principle

Notation • Set $V_n = \min\{1 \leq j \leq n : X_j = \max(X_1, \ldots, X_n)\}$

• Set $(\widehat{X}_1, \ldots, \widehat{X}_{n-1}) = (X_1, \ldots, X_{V_n - 1}, X_{V_n+1}, \ldots, X_n)$

Theorem (one big jump principle, Armendariz & Loulakis '11) Assume (H) and $\mathbb{E}[X_1]=0$.

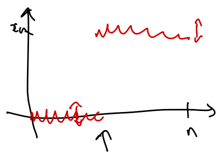
Fix $\varepsilon>0$ and a sequence (x_n) such that $x_n \geq \varepsilon n$ for all n sufficiently large.

We have $d_{TV}\big((\hat{X}_1,\dots,\hat{X}_{n-1})$ under $\mathbb{P}(\cdot \mid S_n\in[x_n,x_n+1)), (X_1,\dots,X_{n-1})\big) \xrightarrow[n\to\infty]{} 0$, that is

$$\sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} \left| \mathbb{P}((\hat{X}_1,\dots,\hat{X}_{n-1})\in A \mid S_n\in[x_n,x_n+1)) - \mathbb{P}((X_1,\dots,X_{n-1})\in A)\right| \xrightarrow[n\to\infty]{} 0$$

This means that under $\mathbb{P}(\cdot \mid S_n\in[x_n,x_n+1))$, once the biggest jump is removed, the remaining r.v are asymptotically iid with same law as X_1!

Thus $(S_0,S_1,\dots,S_n)$ under $\mathbb{P}(\cdot \mid S_n\in[x_n,x_n+1))$ looks like:



uniform time between 1 and n

fluctuations of order $\sqrt{n}$; second, third, etc. largest jumps of order $n^{1/\beta}$.
(this will be proved later)

In practice, to show that a property holds with probability tending to 0 or 1 for $(\hat{X}_1,\dots,\hat{X}_{n-1})$ under $\mathbb{P}(\cdot \mid S_n\in[x_n,x_n+1))$ one can show that it holds for $(X_1,\dots,X_{n-1})$ (which are iid!)

Proof

Let $\hat{\mu}_n$ be the law of $(\hat{X}_1,\dots,\hat{X}_{n-1})$ under $\mathbb{P}(\cdot \mid S_n\in[x_n,x_n+1))$

Let μ_n be the law of $(X_1,\dots,X_{n-1})$

To show that $\sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} |\mu_n(A)-\hat{\mu}_n(A)| \xrightarrow[n\to\infty]{} 0$ the idea is to find a "good" event E_n with

1. $\mu_n(E_n) \xrightarrow[n\to\infty]{} 1$
2. $\sup_{\substack{A\subset E_n\\ A\in\mathcal{B}(\mathbb{R}^{n-1})}} |\mu_n(A)-\hat{\mu}_n(A)| \xrightarrow[n\to\infty]{} 0$

Indeed, by 2 we then have $\hat{\mu}_n(E_n) \xrightarrow[n\to\infty]{} 1$, so

$$\sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} |\mu_n(A)-\hat{\mu}_n(A)| \leq \sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} |\mu_n(A\cap E_n)-\hat{\mu}_n(A\cap E_n)| + \mu_n(E_n^c)+\hat{\mu}_n(E_n^c) \xrightarrow[n\to\infty]{} 0$$

To do this, set $E_n = \{ a = (a_1,\dots,a_{n-1}) \in \mathbb{R}^{n-1} : |a_1+\dots+a_{n-1}| \le n^{3/4} \text{ and } \max_{1\le i\le n-1} a_i \le n^{3/4} \}$

Recall that $\mathbb{P}(X_1 \in [u,u+1)) \underset{u\to\infty}{\overset{(*)}{\sim}} \frac{c}{u^{1+\beta}}$ and $\mathbb{P}(X_1 \ge u) \sim \frac{c_0}{u^\beta}$ with $c_0 = c/\beta$.

<u>We check ①</u> $\mu_n(E_n^c) \le \mathbb{P}(|S_{n-1}| > n^{3/4}) + \mathbb{P}(\max_{1\le j\le n-1} X_j > n^{3/4})$

- the first term is $o(1)$ since $\frac{S_{n-1}}{\sqrt{n}}$ converges in distribution
- The second term is $1 - (1 - \underbrace{\mathbb{P}(X_1 > n^{3/4})}_{\sim \frac{c_0}{n^{\frac{3}{4}\beta}}})^{n-1} \xrightarrow[n\to\infty]{} 0$ since $\frac{n}{n^{\frac{3}{4}\beta}} \to 0$ $(\beta > 2)$

<u>We check ②</u>: to simplify, first assume that X_1 is a discrete random variable. Set $E_n' = E_n \cap \text{Support}(X_1)^{n-1}$

for $a = (a_1,\dots,a_{n-1}) \in E_n'$ observe that

$$\{ (\hat{X}_1,\dots,\hat{X}_{n-1}) = a,\ S_n \in [x_n, x_n+1) \} = \bigcup_{i=1}^{n} \{ (X_1,\dots,X_{i-1},X_{i+1},\dots,X_n) = a,\ X_i \in [x_n - a_1 - \dots - a_{n-1}, x_n - a_1 - \dots - a_{n-1}+1) \}$$

and that the union is disjoint. Thus

$$\hat{\mu}_n(a) = \mu_n(a) \times \underbrace{\frac{n\,\mathbb{P}(X_1 \in [x_n - a_1 - \dots - a_{n-1}, x_n - a_1 - \dots - a_{n-1}+1))}{\mathbb{P}(S_n \in [x_n, x_n+1))}}$$

$$= 1 + \varepsilon_n(a) \quad \text{with} \quad \sup_{a\in E_n} |\varepsilon_n(a)| \xrightarrow[n\to\infty]{} 0$$ by the theorem (local estimate) and $(*)$

Then for $A \subset E_n$, $|\hat{\mu}_n(A) - \mu_n(A)| \le \sum_{a\in E_n'} |\hat{\mu}_n(a) - \mu_n(a)| \le \sum_{a\in E_n'} \mu_n(a)\varepsilon_n(a) \le \sup_{a\in E_n'} |\varepsilon_n(a)|$.

We conclude that $\sup_{A\subset E_n'} |\hat{\mu}_n(A) - \mu_n(A)| = 1 + o(1)$.

- Let us now treat the general case.

Take $A \subset E_n$. As above,

$$\{ (\hat{X}_1,\dots,\hat{X}_{n-1}) \in A,\ S_n \in [x_n, x_n+1) \} = \bigcup_{i=1}^{n} \{ (X_1,\dots,X_{i-1},X_{i+1},\dots,X_n) \in A,\ X_i \in [x_n - X_1 - \dots - X_{n-1}, x_n - X_1 - \dots - X_{n-1}+1) \}$$

and the union is disjoint. Thus

$$\hat{\mu}_n(A) = \frac{n \int_A \mathbb{P}_{X_1}(da_1)\cdots\mathbb{P}_{X_1}(da_{n-1})\, \mathbb{P}(X_1 \in [x_n - a_1 - \dots - a_{n-1}, x_n - a_1 - \dots - a_{n-1}+1))}{\mathbb{P}(S_n \in [x_n, x_n+1))}$$

As above, uniformly in $(a_1,\dots,a_{n-1}) \in E_n$, $\mathbb{P}(X_1 \in [x_n - a_1 - \dots - a_{n-1}, x_n - a_1 - \dots - a_{n-1}+1)) \underset{n\to\infty}{\sim} \mathbb{P}(X_1 \in [x_n, x_n+1))$,

so $\hat{\mu}_n(A) = \mu_n(A) \dfrac{n\,\mathbb{P}(X_1 \in [x_n, x_n+1))}{\mathbb{P}(S_n \in [x_n, x_n+1))} (1 + \varepsilon_n(A))$ with $\sup_{A\subset E_n} |\varepsilon_n(A)| \xrightarrow[n\to\infty]{} 0$

The conclusion follows