Phénomènes de condensation dans les arbres aléatoires

Question : une population se reproduit de manière asexuée et homogène dans le temps, partant d'un individu, et s'éteignant après avoir produit n individus.
Quel est l'ordre de grandeur du nombre max d'enfants d'un individu lorsque $n \to \infty$?

Historique rapide (étoffé plus tard)

- 19è siècle : Bienaymé, Galton & Watson : étude de l'extinction de noms nobles (proba d'extinction d'un processus de branchement)
- Aldous années 90 : étude de la forme de l'arbre généalogique, lien avec le mouvement brownien.
- Jonsson–Stefansson '11 : identification d'un régime avec phénomène de condensation
- Nagaev, Dovey, Denisov–Dieker–Schneer '07 : estimées sur des marches aléatoires dans le régime "one big jump"
- Armendariz–Loulakis '11 : théorème limite pour la trajectoire d'une marche aléatoire dans le régime "one big jump"

Plan : I) Arbres de Bienaymé-Galton-Watson et leur codage

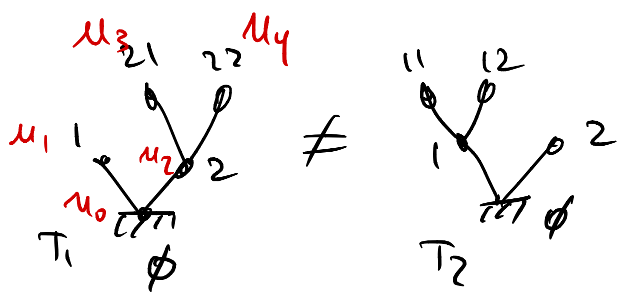
II) Le "one-big jump principle" pour des marches aléatoires

III) (=I+II) Applications aux arbres aléatoires

I) Arbres de BGW, codage

a) Arbres

Les arbres seront planaires enracinés finis



$T_1 = \{ \emptyset, 1, 2, 21, 22 \}$

$T_2 = \{ \emptyset, 1, 2, 11, 12 \}$

On les voit comme un ensemble d'étiquettes, qui sont des suites d'entiers strictement positifs

Notations : • $k_u(T)$ = nb d'enfants du sommet u dans T

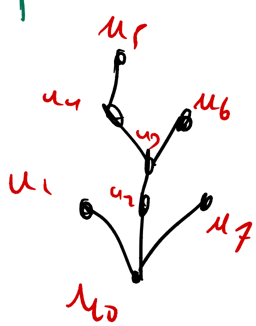
• $|T|$: taille de l'arbre (nombre de sommets)

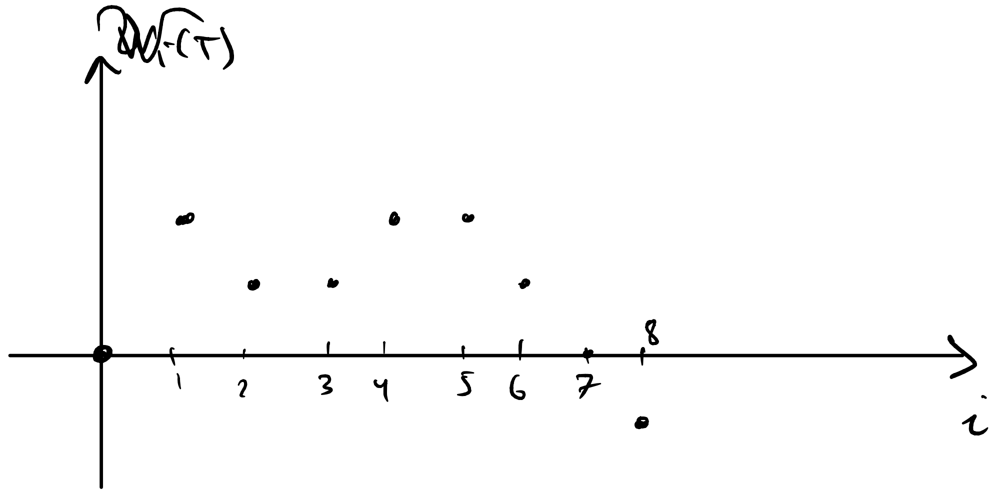
• $\Delta(T) = \max_{u \in T} k_u(T)$

Définition Soit T un arbre avec n sommets. On les ordonne en utilisant l'ordre lexicographique : $u_0 < u_1 < \cdots < u_{n-1}$. La marche de Lukasiewicz de T, $\mathcal{W}(T) = (\mathcal{W}_0(T), \dots, \mathcal{W}_n(T))$ est définie par :

- $\mathcal{W}_0(T) = 0$
- $\mathcal{W}_{i+1}(T) = \mathcal{W}_i(T) + k_{u_i}(T) - 1$ pour $0 \leq i \leq n-1$

Exemple :





Proposition L'application

$$\{\text{arbres avec } n \text{ sommets}\} \longrightarrow \overline{S}_n$$

$$T \longmapsto (k_{u_i}(T) - 1 \; ; \; 0 \leq i \leq n-1)$$

est une bijection, où

$$\overline{S}_n = \{ (x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \cdots + x_n = -1, \; x_1 + \cdots + x_i > -1 \text{ pour } 1 \leq i \leq n \}$$

On peut le démontrer par récurrence.

b) Arbres de Bienaymé

Soit $\mu = (\mu(i))_{i\geq 0}$ une mesure de proba sur $\mathbb{N}$.

Pour un arbre fini T on pose

$$\mathbb{P}_\mu(T) = \prod_{u\in T} \mu(k_u(T)) \quad (\ast)$$

Théorème (Bienaymé 1845, Galton-Watson 1875, Steffensen 1930)
On suppose $\mu(0)+\mu(1)<1$ (toujours supposé implicitement dans la suite). Alors

$$\sum_{i\geq 0} i\mu(i) \leq 1 \iff \mathbb{P}_\mu \text{ est une mesure de probabilité sur les arbres finis}$$

(le μ-processus de branchement s'éteint p.s $\iff \sum i\mu(i)\leq 1$)

Preuve en étudiant des points fixes de fonctions génératrices

Un B_μ arbre est un arbre aléatoire de loi $\mathbb{P}_\mu$.

Pour faire le lien avec la marche de Lukasiewicz on introduit la marche aléatoire $(W_n)_{n\geq 0}$:

Soit $(X_i)_{i\geq 1}$ iid de loi: $\mathbb{P}(X_1 = k) = \mu(k+1)$ pour $k\geq -1$

On pose $W_n = X_1 + \dots + X_n$

On pose $\zeta = \inf\{k \geq 1 : W_k = -1\} \in \mathbb{N} \cup \{+\infty\}$.

Proposition Soit $\mathcal{T}$ un B_μ arbre. Alors

$$(\mathcal{W}_0(\mathcal{T}), \dots, \mathcal{W}_{|\mathcal{T}|}(\mathcal{T})) \overset{loi}{=} (W_0, W_1, \dots, W_\zeta)$$

La preuve est assez directe par (∗)

Dans la suite on note $\mathcal{T}_n$ est un B_μ arbre conditionné à avoir n sommets (on se restreint toujours implicitement aux n tels que $\mathbb{P}(|\mathcal{T}|=n)>0$).

Corollaire • $|\mathcal{T}| \overset{loi}{=} \zeta$

• $(\mathcal{W}_0(\mathcal{T}_n), \dots, \mathcal{W}_n(\mathcal{T}_n)) \overset{loi}{=} (W_0, \dots, W_n)$ sous $\mathbb{P}(\cdot \mid \zeta = n)$

difficulté : conditionnement "non local". Pour le rendre "local" on va utiliser le lemme cyclique.

c) Lemme cyclique

Soit $S_n = \{(x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \dots + x_n = -1\}$

On rappelle que

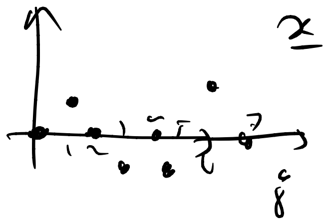
$$\overline{S}_n = \{(x_1, \dots, x_n) \in \{-1, 0, 1, \dots\}^n : x_1 + \dots + x_n = -1, \; x_1 + \dots + x_i > -1 \text{ pour } 1 \leq i \leq n\}$$

On identifie $\mathbb{Z}/n\mathbb{Z}$ avec $\{0, 1, \dots, n-1\}$.

Pour $\underline{x} = (x_1, \dots, x_n) \in S_n$, $i \in \mathbb{Z}/n\mathbb{Z}$, on définit

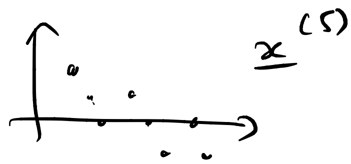
$\underline{x}^{(i)} = (x_{i+1}, x_{i+2}, \dots, x_{i+n})$ avec addition modulo n.

Exemple :



$\underline{x} = (1, -1, -1, 1, -1, 2, -1, -1)$

$\underline{x}^{(5)} = (2, -1, -1, 1, -1, -1, 1, -1)$



$\underline{x}^{(3)} = \dots$

Théorème (lemme cyclique) Pour tout $\underline{x} \in S_n$, on pose $I(\underline{x}) = \{ i \in \mathbb{Z}/n\mathbb{Z} : \underline{x}^{(i)} \in \bar{S}_n \}$. Alors $\mathrm{Card}(I(\underline{x})) = 1$ et $I(\underline{x})$ est constitué du 1er instant où la marche de sauts $\underline{x}$ touche son infimum pour la première fois.

ex : $I(\underline{x}) = \{3\}$

Preuve : pas trop compliqué.

⚠ Change $I_{\underline{x}} \to I(\underline{x})$

d) Applications

Soit $F : \mathbb{Z}^n \to \mathbb{R}$ une fonction invariante sous permutations cycliques (i.e. $\forall x \in \mathbb{Z}^n$, $\forall i \in \mathbb{Z}/n\mathbb{Z}$, $F(\underline{x}) = F(\underline{x}^{(i)})$)

Exemples : $F(x_1, \dots, x_n) = x_1 + \dots + x_n$
$= x_1 \cdots x_n$
$= \max(x_1, \dots, x_n)$

Proposition Soit $F : \mathbb{Z}^n \to \mathbb{R}$ une fonction invariante sous permutations cycliques. On utilise la marche aléatoire $W_n = X_1 + \dots + X_n$ précédemment introduite. Alors

1) $\mathbb{E}[F(X_1, \dots, X_n) \mathbb{1}_{\zeta = n}] = \frac{1}{n} \mathbb{E}[F(X_1, \dots, X_n) \mathbb{1}_{W_n = -1}]$

2) $\mathbb{P}(\zeta = n) = \frac{1}{n} \mathbb{P}(W_n = -1)$

3) La loi de $F(X_1, \dots, X_n)$ sachant que $\zeta = n$ est = loi de de $F(X_1, \dots, X_n)$ sach $W_n = -1$

Preuve 2) provient de 1) (prendre $F \equiv 1$)
3) provient de 1 et 2)

1) Quelques notations : $\mathbb{X}_n = (X_1, \dots, X_n)$. Alors

- $\{\zeta = n\} = \{\mathbb{X}_n \in \overline{S}_n\}$

- $\{W_n = -1\} = \{\mathbb{X}_n \in S_n\}$

- $\forall i \in \mathbb{Z}/n\mathbb{Z}$, $\mathbb{X}_n^{(i)} \overset{\text{loi}}{=} \mathbb{X}_n$

$$\mathbb{E}\left[F(\mathbb{X}_n)\mathbb{1}_{\mathbb{X}_n\in\bar{S}_n}\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[F(\mathbb{X}_n^{(i)})\mathbb{1}_{\mathbb{X}_n^{(i)}\in\bar{S}_n}\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[F(\mathbb{X}_n)\mathbb{1}_{\mathbb{X}_n^{(i)}\in\bar{S}_n}\right]$$

$$= \frac{1}{n}\mathbb{E}\Big[F(\mathbb{X}_n)\underbrace{\sum_{i=1}^{n}\mathbb{1}_{\mathbb{X}_n^{(i)}\in\bar{S}_n}}_{=\mathbb{1}_{\mathbb{X}_n\in S_n}\text{ (lemme cyclique)}}\Big]$$

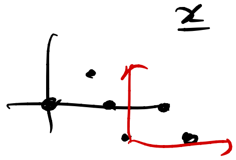
d'où le résultat.

Ainsi, $\Delta(\mathcal{T}_n) \overset{loi}{=} \max(X_1,\dots,X_n)+1$ sous $\mathbb{P}(\cdot\,|W_n=-1)$

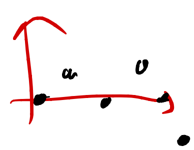
Définition (transformée de Vervaat) Soit $\underline{x}\in S_n$ On pose

$$\nu(\underline{x}) = \underline{x}^{(I(\underline{x}))}$$

Exemple: $\underline{x} = (1,-1,-1,1,-1$



$\nu(\underline{x})$



Théorème

1) Sous $\mathbb{P}(\cdot\,|W_n=-1)$, $\nu(X_1,\dots,X_n)$ a la loi de $(X_1,\dots,X_n)$ sous $\mathbb{P}(\cdot\,|\zeta=n)$

2) Sous $\mathbb{P}(\cdot\,|W_n=-1)$, $\nu(X_1,\dots,X_n)$ et $I((X_1,\dots,X_n))$

sont ⫫ et la loi de $I(x_1,\dots,x_n)$ sous $\mathbb{P}(\cdot \mid W_n=-1)$

est uniforme sur $\{0,1,\dots,n-1\}$

<u>Preuve</u> On pose $\mathbb{X}_n=(X_1,\dots,X_n)$. Soit

$\underline{x}\in\overline{S}_n$ et $0\leq k\leq n-1$.

$$\mathbb{P}(\gamma(\mathbb{X}_n)=\underline{x},\ I(\mathbb{X}_n)=k,\ W_n=-1)$$

$$=\mathbb{P}(\mathbb{X}_n^{(I(\mathbb{X}_n))}=\underline{x},\ I(\mathbb{X}_n)=k,\ W_n=-1)$$

$$=\mathbb{P}(\mathbb{X}_n^{(k)}=\underline{x},\ I(\mathbb{X}_n)=k,\ W_n=-1)$$

Mais $\{\mathbb{X}_n^{(k)}=\underline{x},\ I(\mathbb{X}_n)=k,\ W_n=-1\}$

$$=\{\mathbb{X}_n^{(k)}=\underline{x},\ W_n=-1\} \quad \text{car } \underline{x}\in\overline{S}_n$$

$$=\mathbb{P}(\mathbb{X}_n^{(k)}=\underline{x},\ W_n=-1)$$

$$=\mathbb{P}(\mathbb{X}_n=\underline{x},\ W_n=-1)$$

$$=\mathbb{P}(\mathbb{X}_n=\underline{x},\ \zeta=n) \quad \text{car } \underline{x}\in\overline{S}_n$$

On divise par $\mathbb{P}(W_n=-1)=n\,\mathbb{P}(\zeta=n)$:

$$\mathbb{P}(\gamma(\mathbb{X}_n)=\underline{x},\ I(\mathbb{X}_n)=k \mid W_n=-1)=\frac{1}{n}\,\mathbb{P}(\mathbb{X}_n=\underline{x}\mid \zeta=n)$$

ce qui montre 2) et 1) (en sommant sur k)

∾

Interlude

Objectif d'aujourd'hui : étudier T_n, un B_μ arbre conditionné à avoir n sommets lorsque $n \to \infty$ dans le régime :

- $m = \sum_{i=1}^{\infty} i\mu(i) < 1$
- $\mu(i) \underset{i \to \infty}{\sim} \frac{c}{i^{1+\beta}}$ avec $\beta > 2$ $\quad (\mathrm{Var}(\mu) < \infty)$

Remarque : $\mathbb{E}[X_1] = \sum_{i \geq -1} i\mu(i+1) = \sum_{i=1}^{\infty} i\mu(i) - 1 = m - 1$

Ainsi $\mathbb{E}[X_1] = 0 \iff \mu$ critique $(\sum i\mu(i) = 1)$

Si μ n'est pas critique, le conditionnement par $W_n = -1$ est donc un conditionnement dans un régime de grandes déviations $(W_n = -1 \iff \underbrace{W_n + (1-m)n}_{\text{marche centrée}} = (1-m)n)$

Proposition Notons $G_\mu(x) = \sum_{i=0}^{\infty} \mu(i) x^i$. On suppose que $\tilde{\mu}(i) = a \cdot b^i \mu(i)$ pour $a, b > 0$ (on dit que $\tilde{\mu}$ est un tiltage exponentiel de μ)

Soit T_n un B_μ arbre conditionné à avoir n sommets
— $\tilde{T}_n$ — $B_{\tilde{\mu}}$ —

Alors $T_n \overset{\text{loi}}{=} \tilde{T}_n$

<u>Preuve</u> Soit Π_n = {arbres à n sommets} et $T \in \Pi_n$

$$P_{\tilde{\mu}}(T) = \prod_{u \in T} \tilde{\mu}(k_u(T)) = \prod_{u \in T} a\, b^{k_u(T)} \mu(k_u(T))$$
$$= a^n b^{n-1} P_\mu(T)$$

En sommant on a donc

$$P_{\tilde{\mu}}(\Pi_n) = a^n b^{n-1} P_\mu(\Pi_n)$$

Donc $$\frac{P_{\tilde{\mu}}(T)}{P_{\tilde{\mu}}(\Pi_n)} = \frac{P_\mu(T)}{P_\mu(\Pi_n)}$$

∞

<u>Remarque</u> : μ peut être exponentiellement tiltée en une mesure critique ssi, en notant R le rayon de convergence de G_μ on a $\lim\limits_{x \nearrow R} \frac{x\, G_\mu'(x)}{G_\mu(x)} \geq 1$.

Avant de poursuivre, quelques éléments historiques et de contexte concernant l'étude des limites de grands arbres aléatoires.

II) Estimées sur les marches aléatoires

On suppose ici que X est une v.a à valeurs dans $\{-1,0,1,2,\dots\}$ tq :

- $\mathbb{P}(X>0)>0$
- $\mathbb{E}[X]=0$
- $\mathrm{Var}(X)<\infty$

$(X_i)_{i\geq 1}$ iid de loi X

$S_n = X_1+\dots+X_n$

a) Une inégalité maximale

Proposition (Fuk-Nagaev '71, Denisov-Dieker-Shneer '08)

$\exists C>0$ tq $\forall n\geq 1$, $x>0$ et $c\geq 1$:

$$\mathbb{P}(S_n > x\sqrt{n},\, X_1\leq c\sqrt{n},\dots, X_n\leq c\sqrt{n}) \leq C\exp\left(-\frac{x}{c}\right)$$

Preuve l'idée est d'introduire la marche aléatoire tronquée

$$\tilde{S}_n = \sum_{i=1}^{n} X_i \mathbb{1}_{X_i\leq c\sqrt{n}} \quad \text{car on a}$$

$$\mathbb{P}(S_n > x\sqrt{n}, X_1\leq c\sqrt{n}, \dots, X_n\leq c\sqrt{n}) \leq \mathbb{P}(\tilde{S}_n \geq x\sqrt{n}),$$ qu'on

majore "à la Chernoff" : en posant $\lambda_n = \frac{1}{c\sqrt{n}}$:

$$\mathbb{P}(\tilde{S}_n \geq x\sqrt{n}) = \mathbb{P}\left(e^{\frac{\tilde{S}_n}{c\sqrt{n}}} \geq e^{\frac{x}{c}}\right) \leq e^{-\frac{x}{c}}\ \mathbb{E}\left[e^{\frac{\tilde{X}}{c\sqrt{n}}}\right]^n$$

avec $\tilde{X} = X\,\mathbb{1}_{X \leq c\sqrt{n}}$

On écrit ensuite

$$\mathbb{E}\left[e^{\frac{\tilde{X}}{c\sqrt{n}}}\right] = 1 + \underbrace{\mathbb{E}\left[\frac{\tilde{X}}{c\sqrt{n}}\right]}_{m_n} + \underbrace{\mathbb{E}\left[\left(\frac{\tilde{X}}{c\sqrt{n}}\right)^2 \cdot \frac{e^{\frac{\tilde{X}}{c\sqrt{n}}} - 1 - \frac{\tilde{X}}{c\sqrt{n}}}{(\tilde{X}/c\sqrt{n})^2}\right]}_{s_n}$$

On montre que $m_n = \mathcal{O}\left(\frac{1}{n}\right)$ et $s_n = \mathcal{O}\left(\frac{1}{n}\right)$.

- $m_n = \frac{1}{c\sqrt{n}}\mathbb{E}[X\,\mathbb{1}_{X \leq c\sqrt{n}}] = -\frac{1}{c\sqrt{n}}\mathbb{E}[X\,\mathbb{1}_{X > c\sqrt{n}}]$

Alors $n m_n = -\frac{\sqrt{n}}{c}\mathbb{E}[X\,\mathbb{1}_{X > c\sqrt{n}}]$ car $\mathbb{E}[X] = 0$

Or $t\,\mathbb{E}[X\,\mathbb{1}_{X>t}] \leq \mathbb{E}[X^2\,\mathbb{1}_{X>t}] \xrightarrow[t \to \infty]{} 0$ par convergence dominée

- $x \mapsto \frac{e^x - 1 - x}{x^2}$ est bornée sur $[-1,1]$, donc

$$s_n = \mathcal{O}\left(\mathbb{E}\left[\frac{\tilde{X}^2}{cn}\right]\right)$$

Mais $\mathbb{E}\left[\frac{\tilde{X}^2}{c^2 n}\right] = \frac{1}{c^2 n}\mathbb{E}[X^2\,\mathbb{1}_{X \leq c\sqrt{n}}] = \mathcal{O}\left(\frac{1}{n}\right)$ car $\mathbb{E}[X^2] < \infty$

∞

b) Une estimée locale

On suppose de plus que

- $\mathbb{P}(X=n) \sim \frac{c}{n^{1+\beta}}$ avec $\beta > 2$

Théorème (Doney '89, Nagaev '57)

Soit $\varepsilon > 0$. Uniformément en $m \geq \varepsilon n$: $\mathbb{P}(S_n = m) \sim n\, \mathbb{P}(X_1 = m)$ $\left(= \frac{c\, n}{m^{1+\beta}}\right)$

Preuve : on pose $l_m = \frac{m}{(\ln m)^3}$. On écrit

$$n P_1^{m,n} \leq \mathbb{P}(S_n = m) \leq n P_1^{m,n} + P_2^{m,n} + P_3^{m,n} \qquad \text{avec}$$

$$P_1^{m,n} = \mathbb{P}(S_n = m,\ X_n \geq l_m,\ \max_{1 \leq k \leq n-1} X_k < l_m)$$

$$P_2^{m,n} = \mathbb{P}(S_n = m,\ \max_{1 \leq k \leq n} X_k < l_m)$$

$$P_3^{m,n} = \mathbb{P}\Big(S_n = m,\ \bigcup_{1 \leq i < k \leq n} \{X_i \geq l_m,\ X_k \geq l_m\}\Big)$$

On montre que $P_1^{m,n} \sim \mathbb{P}(X_1 = m)$ et $P_2^{m,n} = o(n\, \mathbb{P}(X_1 = m))$, $P_3^{m,n} = o(n\, \mathbb{P}(X_1 = m))$ (toujours uniformément en $m \geq \varepsilon n$)

- Pour borner $P_3^{m,n}$:

$$P_3^{m,n} \leq \binom{n}{2} \mathbb{P}(S_n = m,\ X_1 \geq l_m,\ X_2 \geq l_m)$$

$$\leq n^2\, \mathbb{P}(X_1 \geq l_m)\, \mathbb{P}(X_2 \geq l_m)$$

$$\leq C' \frac{n^2}{\ell_m^{2\beta}} = C' \frac{n^2}{m^{2\beta}} \ln(m)^3 = o\left(\frac{n}{m^{1+\beta}}\right)$$

car $\frac{n^2}{m^{2\beta}} \ln(m)^3 \cdot \frac{m^{1+\beta}}{n} = \frac{n}{m^{\beta-1}} \ln(m)^3 \longrightarrow 0$ car $\beta > 2$

Pour borner $P_2^{m,n}$ on utilise 1) :

$$P_2^{m,n} \leq \mathbb{P}(S_n \geq m, \max_{1 \leq k \leq n} X_k < \ell_m) \leq C \exp\left(-\frac{m}{\ell_m}\right)$$

$$\leq C \exp(-\ln(m)^3)$$

$$= o\left(\frac{n}{m^{1+\beta}}\right)$$

- Pour borner $P_1^{m,n}$ c'est plus délicat, il faut re-tronquer en fonction de S_{n-1}

Tout d'abord, $\frac{S_n}{\sqrt{n}}$ converge en loi (TCL) donc

$\frac{S_n}{n^{3/4}} \xrightarrow{\mathbb{P}} 0$. On écrit

$P_1^{m,n} \leq Q_1^{m,n} + Q_2^{m,n} + Q_3^{m,n}$ avec

- $Q_1^{m,n} = \mathbb{P}(S_n = m, X_n \geq \ell_m, \max_{1 \leq k \leq n-1} X_k < \ell_m, S_{n-1} > \frac{m}{\ln(m)})$
- $Q_2^{m,n} = \mathbb{P}(S_n = m, X_n \geq \ell_m, \max_{1 \leq k \leq n-1} X_k < \ell_m, S_{n-1} < -n^{3/4})$
- $Q_3^{m,n} = \mathbb{P}(S_n = m, X_n \geq \ell_m, \max_{1 \leq k \leq n-1} X_k < \ell_m, -n^{3/4} < S_{n-1} < \frac{m}{\ln m})$

Pour $Q_1^{m,n}$ par inégalité maximale :

$$Q_1^{m,n} \leq \mathbb{P}\left(S_{n-1} > \frac{m}{\ln m}, \max_{1 \leq k \leq n-1} X_k < \ell_m\right) \leq C \exp(-\ln m^2)$$

$$= o\left(\frac{n}{m^{1+\beta}}\right)$$

Pour $Q_2^{m,n}$:

$$Q_2^{m,n} \leq \sum_{j<-n^{3/4}} \mathbb{P}(S_{n-1}=j,\ X_n = m-j)$$
$$\leq \underbrace{\mathbb{P}(S_{n-1}<-n^{3/4})}_{o(1)} \underbrace{\sup_{j>n^{3/4}} \mathbb{P}(X_1 = m+j)}_{O(\mathbb{P}(X_1=m))}$$

Pour $Q_3^{m,n}$ on écrit

$$Q_3^{m,n} = \mathbb{P}\left(X_n = m-S_{n-1},\ X_n \geq \ell_m,\ \max_{1\leq k\leq n-1} X_k < \ell_m,\ -n^{3/4} < S_{n-1} < \frac{m}{\ln m}\right)$$

Mais $m - \frac{m}{\ln m} \gg \ell_m$ et $\mathbb{P}(X_n = m-j) \sim \mathbb{P}(X_1=m)$ uniformément en $-n^{3/4} \leq j < \frac{m}{\ln m}$. Donc

$$Q_3^{m,n} \sim \mathbb{P}\left(\max_{1\leq k\leq n-1} X_k < \ell_m,\ -n^{3/4} \leq S_{n-1} \leq \frac{m}{\ln m}\right)\mathbb{P}(X_1=m)$$

Mais $\mathbb{P}\left(-n^{3/4} \leq S_{n-1} \leq \frac{m}{\ln m}\right) \longrightarrow 1$

et $\mathbb{P}\left(\max_{1\leq k\leq n-1} X_k < \ell_m\right) \xrightarrow[n\to\infty]{} 1$ car $n\,\mathbb{P}(X_1 \geq \ell_m) \sim \frac{Cn}{\ell_m^{\beta}} \longrightarrow 0$.

$\sim$

Remarque : le résultat reste vrai pour $\beta > 1$ (preuve à adapter légèrement)

c) Un one-big jump principle

On suppose ici que X est une v.a à valeurs dans $\{-1,0,1,2,\dots\}$ tq :

- $\mathbb{P}(X>0)>0$
- $\mathbb{E}[X]=0$
- $\mathbb{P}(X=n) \sim \frac{c}{n^{1+\beta}}$ avec $\beta>2$

$(X_i)_{i\geq 1}$ iid de loi X

$S_n = X_1+\dots+X_n$

On fixe une suite (x_n) tq $\liminf_{n\to\infty} \frac{x_n}{n}>0$

Notations : $\cdot\, V_n = \inf\{1\leq j\leq n : X_j = \max\{X_i : 1\leq i\leq n\}\}$

$\cdot$ Soit $(X_1^{(n)},\dots,X_{n-1}^{(n)})$ une v.a dont la loi est celle de $(X_1,\dots,X_{V_n-1},X_{V_n+1},\dots,X_n)$ sous $\mathbb{P}(\cdot\,|S_n=x_n)$

Théorème (Armendáriz-Loulakis '11)

On a $d_{VT}((X_1^{(n)},\dots,X_{n-1}^{(n)}),(X_1,\dots,X_{n-1})) \xrightarrow[n\to\infty]{} 0$

i.e
$$\sup_{A\in\mathcal{B}(\mathbb{R}^{n-1})} \left|\mathbb{P}((X_1^{(n)},\dots,X_{n-1}^{(n)})\in A)-\mathbb{P}((X_1,\dots,X_{n-1})\in A)\right| \xrightarrow[n\to\infty]{} 0$$

Donnons tout de suite un corollaire qui explique le nom "one-big jump principle" :

Corollaire Soit Δ_n le plus grand élément de $(X_1,\dots,X_n)$ sachant $S_n=x_n$ et $\Delta_n^{(2)}$ le 2è plus grand.

Alors

(1) $\frac{\Delta_n}{x_n} \xrightarrow{\mathbb{P}} 1$

(2) $\frac{\Delta_n - x_n}{\sigma\sqrt{n}} \xrightarrow{loi} N(0,1)$ où $\sigma = Var(X_1)$.

(3) $\mathbb{P}\left(\frac{\Delta_n^{(1)}}{n^{1/\beta}} \leq u\right) \xrightarrow[n\to\infty]{} \exp\left(-\frac{c}{\beta}\frac{1}{u^\beta}\right)$

Preuve : (3) Par le thm il suffit de mq

$\mathbb{P}\left(\frac{\max(X_1, \dots, X_{n-1})}{n^{1/\beta}} \leq u\right) \xrightarrow[n\to\infty]{} \exp\left(-\frac{c}{\beta}\frac{1}{u^\beta}\right)$

Cette proba vaut $\left(1 - \mathbb{P}(X_1 \geq u\, n^{\frac{1}{\beta}})\right)^{n-1}$

Comme $\mathbb{P}(X \geq n) \sim \frac{c}{\beta}\frac{1}{n^\beta}$, ceci exp $\sim \exp\left(-(n-1)\frac{c}{\beta}\cdot\frac{1}{(u\, n^{1/\beta})^\beta}\right)$

(1) et (2) Par (3) avec grande proba

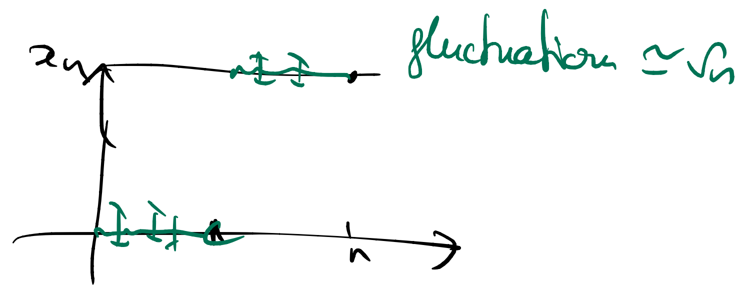
$$\Delta_n = x_n - \left(X_1^{(n-1)} + \dots + X_{n-1}^{(n-1)}\right)$$

On en déduit le résultat car

$$\frac{X_1^{(n-1)} + \dots + X_{n-1}^{(n-1)}}{\sigma\sqrt{n}} \xrightarrow{loi} N(0,1)$$

$\sim$

Image : sous $\mathbb{P}(\cdot \mid S_n = x_n)$:



Preuve du thm

Soit μ_n la loi de $(X_1^{(n)}, \dots, X_{n-1}^{(n)})$

Soit $\tilde{\mu}_n$ la loi de $(X_1, \dots, X_{n-1})$

Pour montrer que $\sup_A |\mu_n(A) - \tilde{\mu}_n(A)| \to 0$ l'idée est essentiellement de trouver un évènement E_n tq

① $\tilde{\mu}_n(E_n) \to 1$

② $\sup_{A \subset E_n} |\mu_n(A) - \tilde{\mu}_n(A)| \xrightarrow[n\to\infty]{} 0$

(en effet, on a alors $\mu_n(\bar{E}_n) \to 0$ et $\tilde{\mu}_n(\bar{E}_n) \to 0$)

Posons $E_n = \{a = (a_1, \dots, a_{n-1}) : |a_1 + \dots + a_{n-1}| \leq n^{3/4}$ et $\max_{1 \leq i \leq n-1} a_i < x_n - n^{3/4}\}$

Si $a = (a_1, \dots, a_{n-1}) \in E_n$, on a

$$\{(X_1^{(n)}, \dots, X_{n-1}^{(n)}) = a, S_n = x_n\}$$

$$= \bigcup_{i=1}^{n} \{(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n) = a, \ X_i = x_n - a_1 - \dots - a_{n-1}\}$$

Donc $\mu_n(a) = \tilde{\mu}_n(a) \times \underbrace{\dfrac{n\mathbb{P}(X_1 = x_n - a_1 - \dots - a_{n-1})}{\mathbb{P}(S_n = x_n)}}_{1 + \varepsilon_n(a)}$ avec $\varepsilon_n(a) \to 0$ unif en $a \in E_n$

Ainsi $\sup_{A \subset E_n} |\mu_n(A) - \tilde{\mu}_n(A)| = 1 + o(1)$.

Or $\tilde{\mu}_n(\overline{E}_n) \leq \mathbb{P}(|S_{n-1}| > n^{3/4}) + \mathbb{P}(\max_{1 \leq j \leq n-1} X_j > x_n - n^{3/4})$

$\underline{\text{1er terme}} = o(1)$

$\underline{\text{2ème terme}} = 1 - (1 - \underbrace{\mathbb{P}(X_1 > x_n - n^{3/4})}_{\sim \frac{c}{\beta} \times \frac{1}{x_n^\beta}})^{n-1} \longrightarrow 0$ car $\frac{n}{x_n^\beta} \to 0$.

$\sim$

III) Application : condensation dans les arbres de Bienaymé sous-critiques

Soit μ une loi de de production tq

(1) $m = \sum i\mu(i) < 1$

(2) $\mu(n) \sim \frac{c}{n^{1+\beta}}$ avec $\beta > 2$.

Soit T_n un B_μ arbre conditionné à avoir n sommets.
Notons $\Delta(T_n)$ le plus grand degré de T_n.

Jonsson & Stefansson '11 $\frac{\Delta(T_n)}{(1-m)n} \xrightarrow{\mathbb{P}} 1$

Théorème (K '15)

(1) $\frac{\Delta(T_n) - (1-m)n}{\sigma\sqrt{n}} \to N(0,1)$ où $\sigma = \mathrm{Var}(\mu)$

(2) Si $\Delta^{(2)}(T_n)$ désigne le 2ème plus grand degré,
$\mathbb{P}\left(\frac{\Delta^{(2)}(T_n)}{n^{1/\beta}} \leq u \right) \xrightarrow[n\to\infty]{} \exp\left(-\frac{c}{\beta}\frac{1}{u^\beta}\right)$

idée : one-big jump principle pour la marche de Lukasiewicz

Soit $(X_i)_{1\leq i\leq n}$ iid de loi $\mathbb{P}(X_1=i)=\mu(i+1)$ $i\geq -1$.
$W_n = X_1+\dots+X_n$
On a $\mathbb{E}[X_1]=m-1$ et $Var(X_1)=Var(\mu)$

<u>Preuve</u>: le + grand saut et 2è plus grand saut d'une suite sont invariants par permutation cyclique.
Donc $(|\Delta(T_n)|, \Delta^{(2)}(T_n)) \overset{loi}{=}$ (+ gd saut, 2è plus grand saut) de $(X_1,\dots,X_n)$ sous $\mathbb{P}(\cdot \mid W_n=-1)$
En posant $\overline{X}_i = X_i + 1-m$, $\overline{W}_n = W_n + n(1-m)$ on a une marche centrée: $W_n=-1 \Leftrightarrow \overline{W}_n = n(1-m)-1$:
on peut appliquer le one big-jump principle et son corollaire

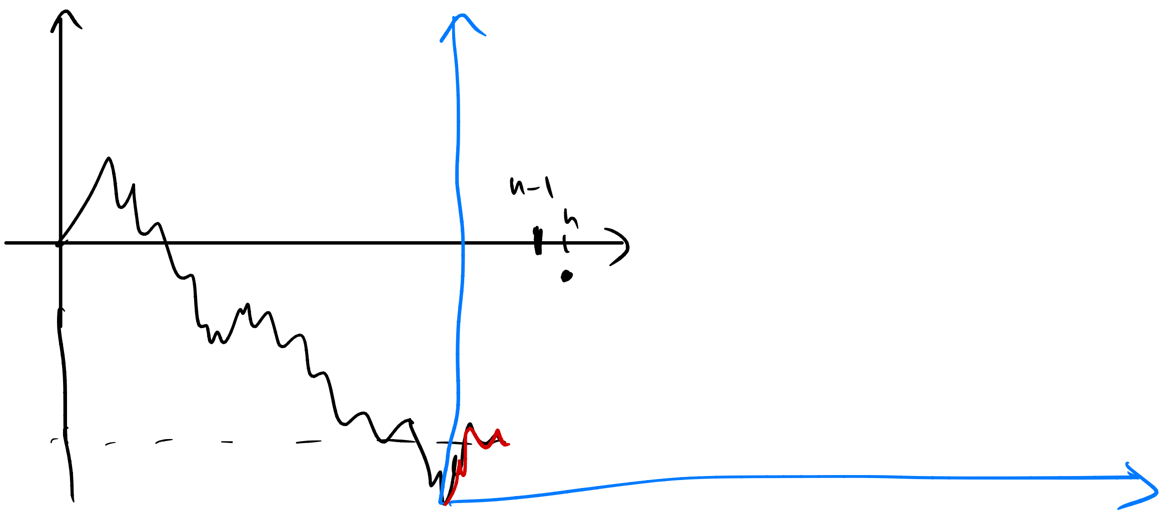
$\sim$

En fait le one-big jump permet de contrôler en fait toute la marche de Lukasiewicz !

On sait que $\mathcal{V}\mathcal{T}(\tau_n) \overset{\text{loi}}{=} \nu(X_1, \dots, X_n)$ sous $\mathbb{P}(\cdot \mid W_n = -1)$

Proposition

$$d_{VT}(\mathcal{V}\mathcal{T}(\tau_n), \nu(X_1, \dots, X_{n-1}, -1)) \xrightarrow[n\to\infty]{} 0$$



Ceci permet par exemple de montrer que

$$\frac{\text{Hauteur}(\tau_n)}{\ln(n)} \xrightarrow{\mathbb{P}} \frac{1}{\ln(1/m)}$$

et que τ_n "ressemble" à une épine finie sur laquelle se branchent des B_μ arbres $\mathcal{M}$, avec un nombre infini branché en haut de l'épine.