Pólya processes

Nicolas Pouyanne

(1) Université de Versailles St-Quentin, LMV
(2) CNRS, UMR 8100
UFR de sciences
45, avenue des Etats-Unis
78035 Versailles cedex

pouyanne@math.uvsq.fr

I- Pólya-Eggenberger urns and Pólya processes

II- Asymptotics, general statements

III- Method

La Londe, journées de probabilités, 14sep7

Pólya-Eggenberger urns with 2 colours

One urn, red and black balls, initial composition $U_1 = {}^t(\#red, \#black)$.

Replacement matrix $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, a, b, c, d integers.

Composition at time n: vector $U_n = {}^t(\#red, \#black)$.

Assumptions
Balance: $a + b = c + d := B \geq 1$
Tenability: $\begin{cases} b \geq 0, \ c \geq 0 \\ a \geq 0 \text{ or } \big(a \mid c \text{ and } a \mid (\#red)_1\big) \\ d \geq 0 \text{ or } \big(d \mid b \text{ and } d \mid (\#black)_1\big) \end{cases}$

Pólya process: $(X_n)_n$, where $X_n = \frac{1}{B} U_n$.

Generalisation in dimension $d \geq 2$, replacement matrix with real entries (random walk in $\mathbb{R}^d$), possible linear change of coordinates: general definition of

Pólya process

Pólya process: definition

V: real vector space of finite dimension $d \geq 1$.
X_1, $w_1, \ldots, w_d$ vectors of V, $(l_k)_{1 \leq k \leq d}$ basis of V^* (linear forms).
$(X_n)_n$: random walk in V with increments in $\{w_1, \ldots, w_d\}$, defined by

$$\mathrm{Prob}\ (X_{n+1} = X_n + w_k | X_n) = \frac{l_k(X_n)}{n + \tau_1 - 1}$$

where τ_1 is the positive real number defined by $\tau_1 = \sum_{k=1}^{d} l_k(X_1)$.

Assumptions

i- Initialisation: $X_1 \neq 0$ and $\forall k \in \{1, \ldots, d\},\ l_k(X_1) \geq 0$;
ii- balance: for all $k \in \{1, \ldots, d\}$, $\sum_{j=1}^{d} l_j(w_k) = 1$;
iii- tenability: for all $k, k' \in \{1, \ldots, d\}$,

$$\begin{cases} k \neq k' \Longrightarrow l_k(w_{k'}) \geq 0, \\ l_k(w_k) \geq 0 \quad \text{or} \quad l_k(X_1)\mathbb{Z} + \displaystyle\sum_{j=1}^{d} l_k(w_j)\mathbb{Z} = l_k(w_k)\mathbb{Z}. \end{cases}$$

Pólya process

N.B.: urn process when $V = \mathbb{R}^d$ and $(l_k)_{1 \leq k \leq d}$ =canonical basis of V^*.
Vectors w_k: lines of the replacement matrix R (more exactly $\frac{1}{B}R$).

$\boxed{\text{Question 1: asymptotics of the random vector } X_n \text{ as } n \to \infty?}$

$\boxed{\text{Question 2: à quoi ça sert ?}}$

Transition endomorphism

Define $A \in \mathcal{L}(V)$ by $\forall v \in V,\ A(v) = \sum_{k=1}^{d} l_k(v) w_k$.

[Urn process: the matrix of A in the canonical basis is ${}^t R$.]

Define $u_1 = \sum_{k=1}^{d} l_k$, linear form, fixed by A.

Choose a Jordan basis of the process, *i.e.* a basis $(u_1, \dots, u_d)$ of linear forms in which ${}^t A$ has a (complex) block-diagonal form $\mathrm{Diag}\,(1, J_{p_1}(\lambda_{k_1}), \dots, J_{p_t}(\lambda_{k_t}))$ where $J_p(z)$ denotes the p-dimensional square matrix

$$J_p(z) = \begin{pmatrix} z & 1 & & \\ & z & \ddots & \\ & & \ddots & 1 \\ & & & z \end{pmatrix}.$$

Denote by $(v_1, \dots, v_d)$ its dual basis (of vectors).

Transition endomorphism, continued

Let σ_2 be defined as

$$\sigma_2 = \begin{cases} 1 \text{ if } 1 \text{ is multiple eigenvalue of } A; \\ \\ \max\{\Re\lambda,\ \lambda \in \mathrm{Sp}(A),\ \lambda \neq 1\} \text{ otherwise.} \end{cases}$$

[Hypotheses imply $\sigma_2 \leq 1$.]

Last two definitions before statement : the principal blocks of A are the $J_p(z)$ where

- $\Re(z) = \sigma_2$;
- p is maximal for this property.

The process is called principally semisimple (pss) when the principal blocks are of size 1.

Asymptotics of *small* Pólya processes

Theorem 1 (Athreya, Karlin, Janson)
With some irreducibility assumption...
1- If $\sigma_2 < 1/2$,

$$\frac{X_n - nv_1}{\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \text{Gaussian vector.}$$

2- If $\sigma_2 = 1/2$ and if $\nu + 1$ is the size of the principal blocks,

$$\frac{X_n - nv_1}{\sqrt{n \log^{2\nu+1} n}} \xrightarrow[n\to\infty]{\mathcal{D}} \text{Gaussian vector.}$$

Proof: continuous time embedding as multitype branching process, martingale considerations, stopping time, back to discrete process.

Asymptotics of *large* Pólya processes

Theorem 2
Suppose $\sigma_2 > 1/2$ *+pss (for simplicity, no irreducibility assumption).*
Let $\lambda_2, \ldots, \lambda_r$ *be the eigenvalues of* A *such that* $\Re(\lambda_k) = \sigma_2$*.*
1- There exist unique (complex-valued) random variables $W_2, \ldots, W_r$ *s.t.*

$$X_n = nv_1 + \sum_{k=2}^{r} n^{\lambda_k} W_k v_k + o(n^{\sigma_2})$$

as n *tends to infinity,* o *is a.s. and* $\mathrm{L}^{\geq 1}$*.*
2- $\forall \alpha \in \mathbb{N}^{r-1}$*,*

$$E\left(W_2^{\alpha_2} \ldots W_r^{\alpha_r}\right) = \frac{\Gamma(\tau_1)}{\Gamma(\tau_1 + \langle \alpha, \lambda \rangle)} Q_\alpha(X_1)$$

where $\tau_1 = u_1(X_1)$*,*
$\langle \alpha, \lambda \rangle = \sum_{k=2}^{r} \alpha_k \lambda_k$*,*
$Q_\alpha =$ *the* α*-th reduced polynomial.*

Non pss case: principal nilpotents of index ν multiply the n^{λ_k}'s by $\log^\nu n$.

Large processes, prototypes

• When the process has only one real principal block, *i.e.* when $A \sim \begin{pmatrix} 1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix}$,

then $\boxed{\dfrac{X_n - nv_1}{n^{\sigma_2}} \xrightarrow[n\to\infty]{\mathrm{L}^{\geq 1}+\mathrm{ps}} W_2 v_2}$ + moments of W_2.

• When the process has two non real one-dim. principal blocks, *i.e.* when $A \sim \begin{pmatrix} 1 & & & \\ & \lambda_2 & & \\ & & \overline{\lambda_2} & \\ & & & \ddots \end{pmatrix}$, then $\boxed{\dfrac{X_n - nv_1}{n^{\sigma_2}} = 2\Re\left(n^{i\sigma_2'} W_2 v_2\right) + o_{\mathrm{ps}+\mathrm{L}^{\geq 1}}(1)}$

+ joint moments of W_2 and $\overline{W_2}$,

i.e.

$$X_n = nv_1 + n^{\sigma_2}\rho\Big[v_2' \cos(\sigma_2' \log n + \varphi) + v_2'' \sin(\sigma_2' \log n + \varphi)\Big] + o(n^{\sigma_2}).$$

Transition operator Φ

Conditional expectation: if $f : V \to W$ is any measurable function,

$$\begin{aligned} E^{\mathcal{F}_n} f(X_{n+1}) &= \sum_{k=1}^{d} \frac{l_k(X_n)}{n+\tau_1-1} \times f(X_n + w_k) \\ &= \left(\mathrm{Id} + \frac{\Phi}{n+\tau_1-1} \right)(f)(X_n) \end{aligned}$$

where Φ is the difference operator defined for any $v \in V$ by

$$\boxed{\Phi(f)(v) = \sum_{k=1}^{d} l_k(v) \Big[f(v+w_k) - f(v) \Big].}$$

Induction leads to $\boxed{Ef(X_n) = \gamma_{\tau_1,n}(\Phi)(f)(X_1)}$

where $\gamma_{\tau_1,n}$ is the polynomial defined by $\gamma_{\tau_1,1} = 1$ and, for any $n \geq 2$, $\gamma_{\tau_1,n}(t) = \prod_{k=1}^{n-1} \left(1 + \frac{t}{k+\tau_1-1}\right)$.

Reduction of Φ

Recall: $\Phi(f)(v) = \sum_{k=1}^{d} l_k(v)\Big[f(v+w_k) - f(v)\Big].$

Φ stabilizes the space of polynomials (d variables) of degree $\leq e$, any $e \geq 0$.

ENCORE MIEUX :

for any $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d$, denote $\mathbf{u}^\alpha = u_1^{\alpha_1} \ldots u_d^{\alpha_d}$.

On the α's: degree-antialphabetical order, *i.e.* for $d = 3$

$$(1,0,0) < (0,1,0) < (0,0,1) < (2,0,0) < (1,1,0) < (1,0,1) < (0,2,0) < \ldots$$

Define $S_\alpha = \mathrm{Vect}\{\mathbf{u}^\beta,\ \beta \leq \alpha\}$; then $\boxed{\Phi(S_\alpha) \subseteq S_\alpha.}$

$\longrightarrow$ Φ-stable filtration of the space of polynomials.

Reduced polynomials

For any $z \in \mathbb{C}$, denote

$$\ker(\Phi - z)^{\infty} = \bigcup_{n \geq 0} \ker(\Phi - z\,\mathrm{Id})^{n}$$

the characteristic space of Φ relative to z (on polynomials).

The eigenvalues of Φ on polynomials are the complex numbers

$$\langle \alpha, \lambda \rangle = \alpha_1 \lambda_1 + \cdots + \alpha_d \lambda_d.$$

Definition : for any $\alpha \in \mathbb{N}^d$, the α-th reduced polynomial Q_α is the projection of $\mathbf{u}^\alpha$ on $\ker(\Phi - \langle \alpha, \lambda \rangle)^{\infty}$ parallel to $\bigoplus_{z \neq \langle \alpha, \lambda \rangle} \ker(\Phi - z)^{\infty}$.

Basis $(Q_\alpha)_{\alpha \in \mathbb{N}^d}$ of polynomials.
Sometimes, closed formula. Always, recursive computation.

Asymptotics of reduced moments $EQ_\alpha(X_n)$

• If Q_α is eigenfunction of Φ (eigenvalue $\langle \alpha, \lambda \rangle$), then

$$\begin{aligned} EQ_\alpha(X_n) &= \gamma_{\tau_1,n}(\langle \alpha, \lambda \rangle) \times Q_\alpha(X_1) \\ &\underset{n\to\infty}{\sim} \frac{\Gamma(\tau_1)}{\Gamma(\tau_1 + \langle \alpha, \lambda \rangle)} n^{\langle \alpha, \lambda \rangle} Q_\alpha(X_1) \quad \text{(Stirling)}. \end{aligned}$$

• If not, let ν_α be the index of nilpotence of Q_α for Φ. Taylor + logarithmic derivative of $\gamma_{\tau_1,n}$ lead to $\log n$-term:

$$EQ_\alpha(X_n) \underset{n\to\infty}{\sim} \frac{n^{\langle \alpha, \lambda \rangle} \log^{\nu_\alpha} n}{\nu_\alpha!} \frac{\Gamma(\tau_1)}{\Gamma(\tau_1 + \langle \alpha, \lambda \rangle)} (\Phi - \langle \alpha, \lambda \rangle)^{\nu_\alpha}(Q_\alpha)(X_1).$$

Asymptotics of joint principal moments $\mathbf{u}^{\alpha}(X_n)$

Develop any $\mathbf{u}^{\alpha}$ in the $(Q_\alpha)_\alpha$ basis (complex coordinates):

$$\mathbf{u}^{\alpha} = Q_\alpha + \sum_{\beta<\alpha,\ \langle\beta,\lambda\rangle\neq\langle\alpha,\lambda\rangle} q_{\alpha,\beta} Q_\beta. \tag{1}$$

Recall: $EQ_\beta(X_n) \sim C n^{\langle\beta,\lambda\rangle} \log^{\nu_\beta} n$.

Questions :

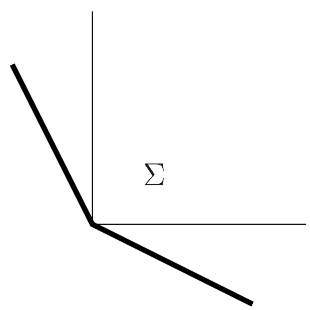
1- which $q_{\alpha,\beta}$ are zero?
2- For a given α, which $\Re\langle\beta,\lambda\rangle$ is maximal among indices $\beta<\alpha$ such that $q_{\alpha,\beta}\neq 0$?

$\longrightarrow$ Refine indices in formula (1)?

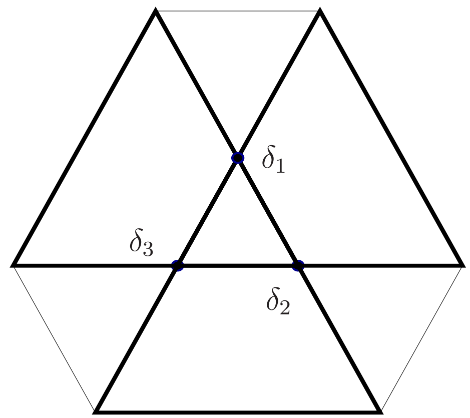
Polytopes in the space of exponents

Let Σ be the cone in $\mathbb{R}^d$ defined by ($(\delta_k)_{1\leq k\leq d}$ is the canonical basis of $\mathbb{R}^d$)

$$\Sigma = \sum_{1\leq i\neq j\leq d} \mathbb{R}_{\geq 0}(2\delta_i - \delta_j),$$



Dimension 2



Dimension 3

...and A_α some suitable compact rational polyhedral polytope ...

Fin de l'histoire des moments

Some geometrical work, and formula (1) is refined:

$$\mathbf{u}^\alpha = Q_\alpha + \sum_{\beta \in A_\alpha - \Sigma} q_{\alpha,\beta} Q_\beta.$$

Property of A_α's and Σ: is α is such that

$$\forall k, \quad \Big(\Re(\lambda_k) \leq 1/2 \Big) \Longrightarrow \Big(\alpha_k = 0 \Big),$$

then $EQ_\alpha(X_n)$ is the winner in the asymptotics of $E\,\mathbf{u}^\alpha(X_n)$ (and other ones. . .).

$\longrightarrow$ Asymptotics of joint principal moments $E\,\mathbf{u}^\alpha(X_n)$ (phase transition $1/2$ appears there).

Fin de l'histoire

Decompose X_n as sum of its projections on the characteristic spaces for A:

$$X_n = (n + \tau_1 - 1)v_1 + \pi_{]1/2,1[}(X_n) + \pi_{\leq 1/2}(X_n).$$

• $\pi_{]1/2,1[}$-term
Renormalize to obtain a martingale $\left(\gamma_{\tau_1,n}(\pi_{]1/2,1[}A)^{-1}\pi_{]1/2,1[}(X_n)\right)_n$.
By evaluation of joint principal moments, convergence in $\mathrm{L}^{\geq 1}$ (Burkholder inequality) and computation of all moments of its limit.

• $\pi_{\leq 1/2}$-term
By evaluation of joint principal moments, $o(n^{\sigma_2})$ almost surely and in $\mathrm{L}^{\geq 1}$.

Et voilà.