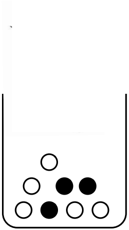
Aspects symboliques en combinatoire des urnes de Pólya

Basile Morcrette

Journées Boole

♪ 21 juin 2013 ♪
Paris

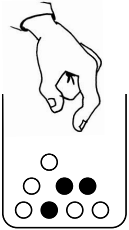
Urnes de Pólya



- Une urne contenant des boules de deux couleurs différentes
- Tirage uniforme avec remise
- Des règles fixées pour l'évolution de l'urne

$$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$$

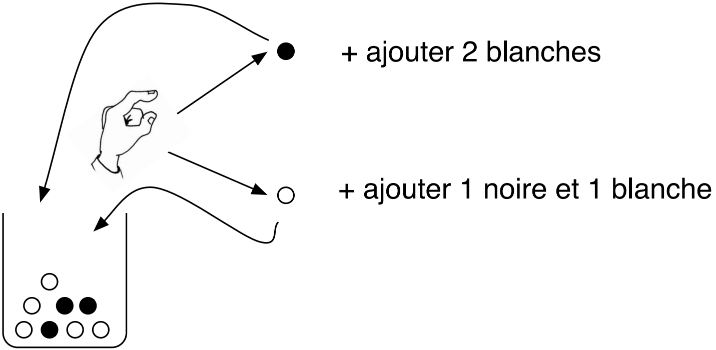
Urnes de Pólya



- Une urne contenant des boules de deux couleurs différentes
- Tirage uniforme avec remise
- Des règles fixées pour l'évolution de l'urne

$$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$$

Urnes de Pólya



- Une urne contenant des boules de deux couleurs différentes
- Tirage uniforme avec remise
- Des règles fixées pour l'évolution de l'urne

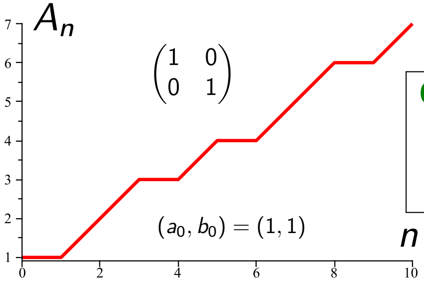
$$\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$$

Objectif général

Processus d'urne de Pólya

- configuration initiale (a_0, b_0)
- matrice de règles $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ $\quad a, d \in \mathbb{Z}, \quad b, c \in \mathbb{N}\,.$

(A_n, B_n) vecteur aléatoire, composition de l'urne après n tirages.
A_n nombre de boules noires dans l'urne au temps n.

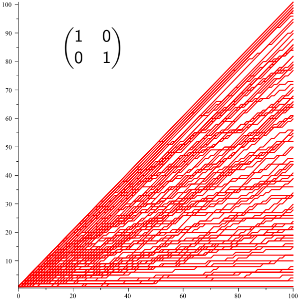


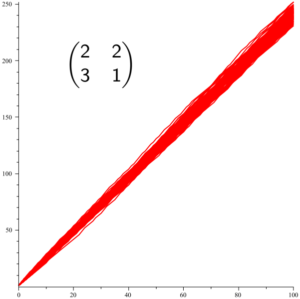
Question : Loi de (A_n, B_n) :

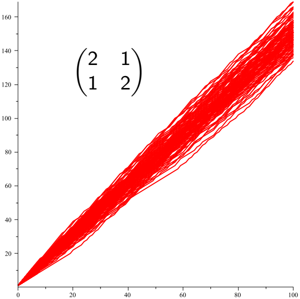
- pour n fixé ?
- pour n très grand, $n \to \infty$?

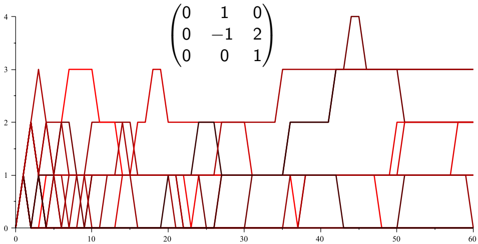
Comportements variés

Problème : Comprendre précisément les comportements limites de l'urne d'une urne de Pólya.









Urnes équilibrées - Réponse probabiliste

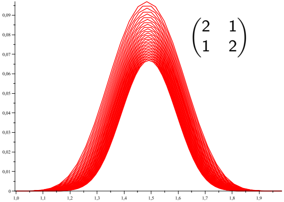
[Smythe 1996, Janson 2004, Pouyanne, Chauvin, Mailler, …2011, 2013]
Méthodes : plongement en temps continu, martingales, ...

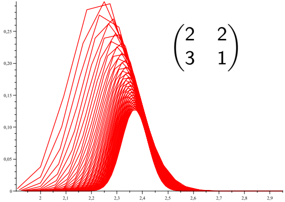
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \boxed{\begin{array}{llc} \sigma & = a+b = c+d & \text{balance} \\ p & = c-a = b-d & \text{dissimilarité} \end{array}}$$

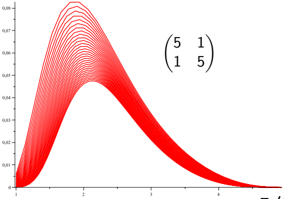
Quantité remarquable : $\lambda = \dfrac{-p}{\sigma}$: rapport des valeurs propres.

Théorème

- si $\lambda \leq \frac{1}{2}$ comportement gaussien
- si $\lambda > \frac{1}{2}$ comportement non gaussien



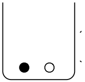




Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ et $(a_0, b_0) = (1, 1)$.

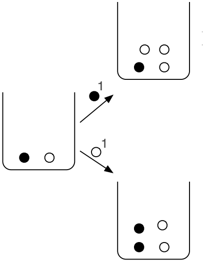
$$H(x, y, z) =$$



$$xy$$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ et $(a_0, b_0) = (1, 1)$.



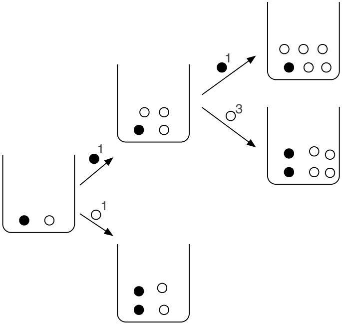
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^3 + x^2y^2)\,\frac{z}{1!}$$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ et $(a_0, b_0) = (1, 1)$.



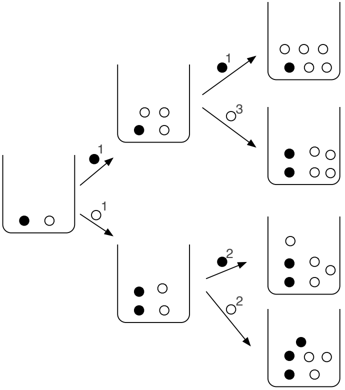
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^3 + x^2y^2)\,\frac{z}{1!}$$

Compter les histoires - Exemple

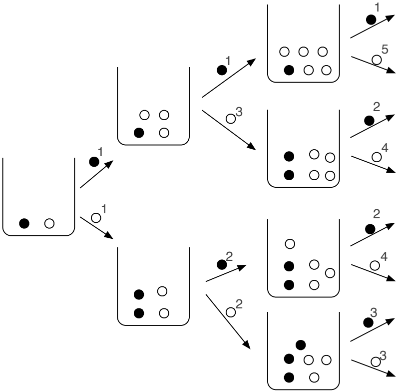
Prenons l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ et $(a_0, b_0) = (1, 1)$.



$$
\begin{aligned}
H(x,y,z) = &\\
& xy \\
+ \quad & (xy^3 + x^2y^2)\,\frac{z}{1!} \\
+ \quad & (xy^5 + 5x^2y^4 + 2x^3y^3)\,\frac{z^2}{2!}
\end{aligned}
$$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ et $(a_0, b_0) = (1, 1)$.



$$H(x,y,z) =$$

$$xy$$

$$+ \quad (xy^3 + x^2y^2)\,\frac{z}{1!}$$

$$+ \quad (xy^5 + 5x^2y^4 + 2x^3y^3)\,\frac{z^2}{2!}$$

$$+ \quad \dots$$

Fonction génératrice des histoires d'une urne équilibrée

Histoire de longueur n = une suite de n évolutions (n tirages)
$H_{n,i,j}$: nombre d'histoires de longueur n, débutant en (a_0, b_0), et terminant en (i, j).

$$H(x,y,z) = \sum_{n,i,j} H_{n,i,j}\, x^i y^j \frac{z^n}{n!}$$

Équation aux Dérivées Partielles [Flajolet - Gabarró - Pekari, 2005]

$$\partial_z H = x^{a+1} y^b \partial_x H + x^c y^{d+1} \partial_y H$$

ou encore $\quad \partial_z H = \mathcal{D} H \quad$ avec $\quad \mathcal{D} = x^a y^b\, x\partial_x + x^c y^d\, y\partial_y$

Système différentiel [Flajolet - Dumas - Puyhaubert, 2006]

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ et } \left\{ \begin{array}{c} (a_0, b_0) \\ a+b=c+d \end{array} \right. \Longrightarrow \quad \begin{array}{c} H = X^{a_0}\, Y^{b_0} \\ \text{avec } \left\{ \begin{array}{rcl} \dot{X} & = & X^{a+1}\, Y^b \\ \dot{Y} & = & X^c\, Y^{d+1} \\ X(0)=x, & & Y(0)=y \end{array} \right. \end{array}$$

Méthode symbolique pour la PDE - Cas équilibré

Pick and replace est *l'opérateur de pointage* $\Theta_x = x\partial_x$

$$x^i y^j \xrightarrow{?} i\, x^{i+a}\, y^{j+b} + j\, x^{i+c}\, y^{j+d}$$

Méthode symbolique pour la PDE - Cas équilibré

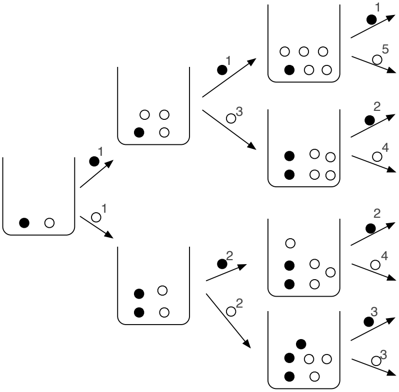
Pick and replace est *l'opérateur de pointage* $\Theta_x = x\partial_x$

$$x^i y^j \xrightarrow{\mathfrak{D}} i\, x^{i+a}\, y^{j+b} + j\, x^{i+c}\, y^{j+d}$$

$$\mathfrak{D} = x^a y^b\, \Theta_x + x^c y^d\, \Theta_y$$

Compter les histoires - Exemple

Soit l'urne $\begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}$ avec $(a_0, b_0) = (1, 1)$.



$$H(x,y,z) = xy + (xy^3 + x^2y^2)\,\frac{z}{1!} + (xy^5 + 5x^2y^4 + 2x^3y^3)\,\frac{z^2}{2!} + \cdots$$

Méthode symbolique pour la PDE - Cas équilibré

Pick and replace est *l'opérateur de pointage* $\Theta_x = x\partial_x$

$$x^i y^j \xrightarrow{\mathfrak{D}} i\, x^{i+a}\, y^{j+b} + j\, x^{i+c}\, y^{j+d}$$

$$\mathfrak{D} = x^a y^b\, \Theta_x + x^c y^d\, \Theta_y$$

Itération à partir de la configuration initiale,

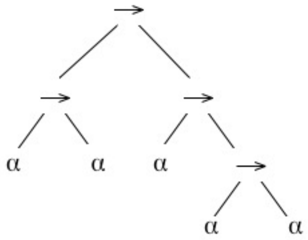
$$\mathfrak{D}^n \left[x^{a_0} y^{b_0}\right] = \sum_{i,j} H_{n,i,j}\, x^i y^j$$

$$H(x,y,z) = \sum_{n\geq 0} \mathfrak{D}^n [x^{a_0} y^{b_0}] \frac{z^n}{n!} = \left(e^{\mathfrak{D} z}\right) \circ [x^{a_0} y^{b_0}]$$

Preuve PDE Différentier par rapport à z : $\partial_z H = \mathfrak{D} H$

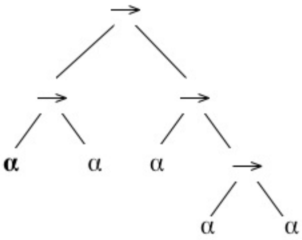
Boole dans les Urnes

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler, 2011]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



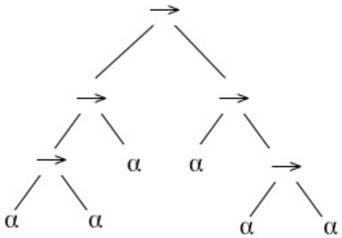
Boole dans les Urnes

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler, 2011]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



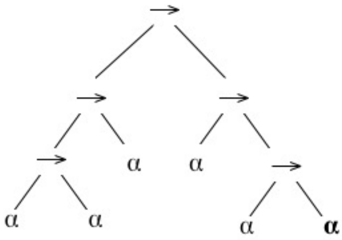
Boole dans les Urnes

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler, 2011]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



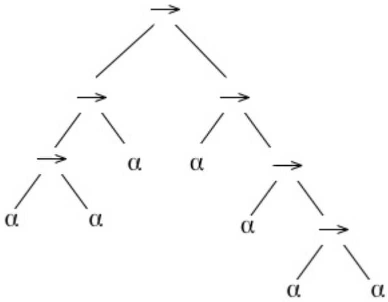
Boole dans les Urnes

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler, 2011]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves

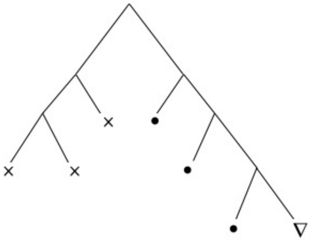


Boole dans les Urnes

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler, 2011]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



A 3×3 urn model



3 colors,
with rules :

$$\begin{array}{ccc} \nabla & \to & \bullet\,\nabla \\ \bullet & \to & \times\,\times \\ \times & \to & \times\,\times \end{array}$$

Corresponding
urn :

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

Generating function of histories

$$\boxed{H(y,z) = \exp\left(\ln\left(\frac{1}{1-z}\right) + (y-1)z\right)}$$

z counts the length of history,
y counts the number of $\bullet$ balls.

Poisson Law in subs-trees

Let $U_{k,n}$ be the number of left sub-trees of of size k directly hanging on the right branch of a random tree of size n.

Theorem

- $U_{1,n}$ converges in law, $U_{1,n} \underset{n\to\infty}{\longrightarrow} U_1$,
- $U_1 \sim \mathcal{P}\text{oisson}\,(1)$, with rate of convergence $O\left(\frac{2^n}{n!}\right)$.

Poisson Law in subs-trees

Let $U_{k,n}$ be the number of left sub-trees of of size k directly hanging on the right branch of a random tree of size n.

Theorem

- $U_{1,n}$ converges in law, $U_{1,n} \underset{n\to\infty}{\longrightarrow} U_1$,
- $U_1 \sim \mathcal{P}\text{oisson}\,(1)$, with rate of convergence $O\left(\frac{2^n}{n!}\right)$.

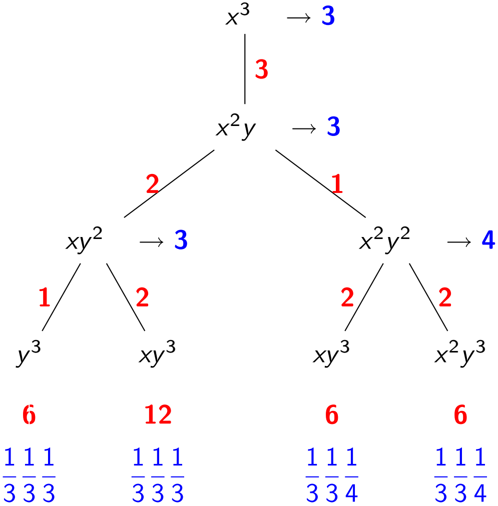
Generalisation With a $(k+2) \times (k+2)$ urn

Theorem

- $U_{k,n}$ converges in law, $U_{k,n} \underset{n\to\infty}{\longrightarrow} U_k$,
- $U_k \sim \mathcal{P}\text{oisson}\left(\frac{1}{k}\right)$, with rate of convergence $O\left(\frac{(2k)^n}{n!}\right)$.

$\boxed{\text{Urnes non équilibrées}}$

Fin des histoires...



$\begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$ débutant en $(3, 0)$

Plus de :

- balance
- nombre total de boules déterministes
- équiprobabilité des histoires

Combinatoire des histoires $\neq$ Probabilités.

Cas général non équilibré : probabilités et symbolique I

$$p_n(x,y) = \sum_{i,j} \mathbb{P}\left\{A_n = i, B_n = j\right\} x^i y^j$$

$$x^i y^j \xrightarrow{?} \frac{i}{i+j} x^{i+a} y^{j+b} + \frac{j}{i+j} x^{i+c} y^{j+d}$$

Cas général non équilibré : probabilités et symbolique I

$$p_n(x, y, \mathbf{t}) = \sum_{i,j} \mathbb{P}\left\{A_n = i, B_n = j\right\} \, x^i y^j \mathbf{t}^{i+j}$$

$$x^i y^j t^{i+j} \xrightarrow{\mathbf{?}} \frac{i}{i+j} x^{i+a} y^{j+b} t^{i+j+a+b} + \frac{j}{i+j} x^{i+c} y^{j+d} t^{i+j+c+d}$$

Cas général non équilibré : probabilités et symbolique I

$$p_n(x, y, \mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\; x^i y^j \mathbf{t}^{i+j}$$

$$x^i y^j t^{i+j} \xrightarrow{\;?\;} \frac{i}{i+j} x^{i+a} y^{j+b} t^{i+j+a+b} + \frac{j}{i+j} x^{i+c} y^{j+d} t^{i+j+c+d}$$

Introduisons deux operateurs, intégration et différentiation,

$$\mathfrak{I}[x^i y^j t^{i+j}] = \int_0^t x^i y^j w^{i+j} \frac{dw}{w} = x^i y^j \frac{t^{i+j}}{i+j}$$

$$\mathfrak{D} = x^a y^b t^{a+b} \Theta_x + x^c y^d t^{c+d} \Theta_y$$

Cas général non équilibré : probabilités et symbolique I

$$p_n(x, y, \mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\ x^i y^j \mathbf{t}^{i+j}$$

$$x^i y^j t^{i+j} \xrightarrow{\mathfrak{D} \circ \mathfrak{I}} \frac{i}{i+j} x^{i+a} y^{j+b} t^{i+j+a+b} + \frac{j}{i+j} x^{i+c} y^{j+d} t^{i+j+c+d}$$

Introduisons deux operateurs, intégration et différentiation,

$$\mathfrak{I}[x^i y^j t^{i+j}] = \int_0^t x^i y^j w^{i+j} \frac{dw}{w} = x^i y^j \frac{t^{i+j}}{i+j}$$

$$\mathfrak{D} = x^a y^b t^{a+b} \Theta_x + x^c y^d t^{c+d} \Theta_y$$

Cas général non équilibré : probabilités et symbolique I

$$p_n(x, y, \mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\ x^i y^j \mathbf{t}^{i+j}$$

$$x^i y^j t^{i+j} \xrightarrow{\mathfrak{D} \circ \mathfrak{I}} \frac{i}{i+j} x^{i+a} y^{j+b} t^{i+j+a+b} + \frac{j}{i+j} x^{i+c} y^{j+d} t^{i+j+c+d}$$

Introduisons deux operateurs, intégration et différentiation,

$$\mathfrak{I}[x^i y^j t^{i+j}] = \int_0^t x^i y^j w^{i+j} \frac{dw}{w} = x^i y^j \frac{t^{i+j}}{i+j}$$

$$\mathfrak{D} = x^a y^b t^{a+b} \Theta_x + x^c y^d t^{c+d} \Theta_y$$

$$p_{n+1} = \mathfrak{D} \circ \mathfrak{I}[p_n] = \ldots = (\mathfrak{D} \circ \mathfrak{I})^{n+1} [x^{a_0} y^{b_0} t^{a_0+b_0}]$$ –> ugly !

Cas général non équilibré : probabilités et symbolique II

$$p_n(x, y, \mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\ x^i y^j \mathbf{t}^{i+j}$$

$$p_{n+1} = \mathfrak{D} \circ \mathfrak{I}[p_n]$$

$$\boxed{\begin{gathered} \mathfrak{D} = x^a y^b t^{a+b} \Theta_x + x^c y^d t^{c+d} \Theta_y \\ \mathfrak{I}[f(t)] = \int_0^t f(w) \frac{dw}{w} \end{gathered}}$$

Cas général non équilibré : probabilités et symbolique II

$$p_n(x,y,\mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\; x^i y^j \mathbf{t}^{i+j}$$

$$p_{n+1} = \mathfrak{D} \circ \mathfrak{I}[p_n]$$

$$\boxed{\begin{gathered}\mathfrak{D} = x^a y^b t^{a+b}\Theta_x + x^c y^d t^{c+d}\Theta_y \\ \mathfrak{I}[f(t)] = \int_0^t f(w)\frac{dw}{w}\end{gathered}}$$

Introduisons une nouvelle Fonction Génératrice, $\psi_n = \mathfrak{I}[p_n]$

$$\psi_n(x,y,t) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\; x^i y^j \frac{t^{i+j}}{i+j}$$

$$p_n = t\partial_t\psi_n \quad \text{et} \quad \boxed{t\partial_t\psi_{n+1} = \mathfrak{D}(\psi_n)}$$

Cas général non équilibré : probabilités et symbolique II

$$p_n(x,y,\mathbf{t}) = \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\; x^i y^j \mathbf{t}^{i+j}$$

$$p_{n+1} = \mathfrak{D} \circ \mathfrak{I}[p_n]$$

$$\boxed{\begin{aligned} \mathfrak{D} &= x^a y^b t^{a+b}\Theta_x + x^c y^d t^{c+d}\Theta_y \\ \mathfrak{I}[f(t)] &= \int_0^t f(w)\frac{dw}{w} \end{aligned}}$$

Introduisons une nouvelle Fonction Génératrice, $\quad \psi_n = \mathfrak{I}[p_n]$

$$\psi_n(x,y,t) \;=\; \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\}\; x^i y^j \frac{t^{i+j}}{i+j}$$

$$p_n = t\partial_t \psi_n \quad \text{et} \quad \boxed{t\partial_t \psi_{n+1} = \mathfrak{D}(\psi_n)}$$

Enfin $t\partial_t = x\partial_x + y\partial_y$, ainsi $\Psi = \sum_n \psi_n z^n$ vérifie

$$\boxed{\left[\left(1 - zx^a y^b\right)\Theta_x + \left(1 - zx^c y^d\right)\Theta_y\right] \cdot \Psi(x,y,z) = x^{a_0} y^{b_0}}$$

pour toute urne additive !

...mais pondérer, c'est gagné !

$$\text{(PGF } \textit{pondérée}\text{)} \qquad \Psi(x,y,z) = \sum_{n\geq 0} \sum_{i,j} \mathbb{P}\{A_n = i, B_n = j\} \, \frac{x^i \, y^j}{i+j} \, z^n$$

Théorème

$$\text{(PDE)} \qquad \left[(1 - zx^a y^b)\,\Theta_x + (1 - zx^c y^d)\,\Theta_y\right] \cdot \Psi(x,y,z) = x^{a_0} y^{b_0}$$

Solution méthode des caractéristiques, et équation différentielle ordinaire

$$\text{(ODE)} \qquad \frac{dx}{dy} = \frac{x(1 - zx^a y^b)}{y(1 - zx^c y^d)}$$

Théorème Soit $x = g(y,z,u)$ la solution générale de (ODE), avec u constante d'intégration, et $U(x,y,z)$ (intégrale première) telle que $g(y,z,U(x,y,z)) = x$. Alors,

$$\Psi(x,y,z) = \int_0^y \frac{g(t,z,U(x,y,z))^{a_0} \, t^{b_0 - 1}}{1 - z\, g(t,z,U(x,y,z))^c \, t^d} \, dt$$

[Morcrette 2013]

Application : Knuth *cutting loops*

2.1. **Cutting loops.** The following problem by Daniel Shine was published on pages 144–145 of the *Journal of Recreational Mathematics*, Volume 33:

> **2680**. One thousand loops of string, each of unit length, are placed in a box. One piece of string is removed at random. It is cut at a random location and put back into the box. This select-and-cut process is repeated 1000 times. All pieces, whether loops or not, are equally likely to be selected. After the 1000 repetitions, what is the average length of the pieces of string in the box?

$$\mathcal{K} = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$(a_0, b_0) = (m, 0)$$

Débutons avec m boucles dans une boîte.
À chaque étape, piocher, découper et replacer dans la boîte.
Q : Longueur moyenne d'un brin après m coupes ?

Système caractéristique : $$\frac{dx}{x(1 - zx^{-1}y)} = \frac{dy}{y(1 - zy)} = \frac{dw}{x^m}$$

EDO 1er ordre : $$\frac{dx}{dy} = \frac{x - zy}{y(1 - zy)}$$

Solution générale : $$g(y, z, u) = \frac{y}{1 - zy}\left(u - z\ln(y)\right)$$

Intégrale première : $$U(x, y, z) = \frac{x(1 - zy)}{y} + z\ln(y)$$

$$\Psi_{\mathcal{K}}(x, 1, z) = \int_0^1 \frac{t^{m-1}}{(1 - zt)^{m+1}} \left(x(1 - z) - z\ln(t)\right)^m dt$$

Urnes à coefficients aléatoires

Mélanger Pólya et Friedman

$$\begin{pmatrix} \mathcal{B}_p & 1-\mathcal{B}_p \\ 1-\mathcal{B}_p & \mathcal{B}_p \end{pmatrix}, \quad \text{où } \mathcal{B}_p \sim \mathrm{Bernoulli}(p)$$

Configuration initiale : $(a_0, b_0) = (1, 1)$.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ avec proba p

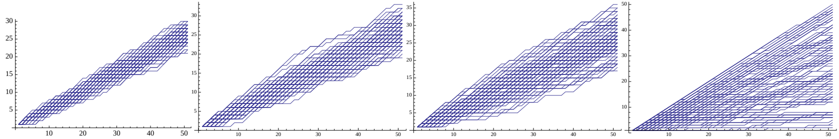
urne de Pólya–Eggenberger,
distribution limite *uniforme*

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ avec proba $1-p$

urne de Friedman,
distribution limite *gaussienne*

Mélanger Pólya et Friedman

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ avec proba } p \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ avec proba } 1-p$$

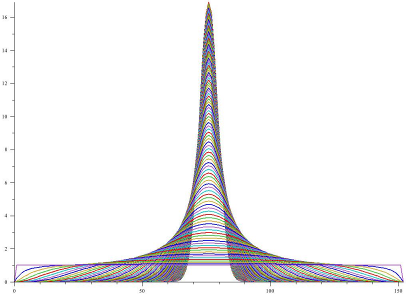


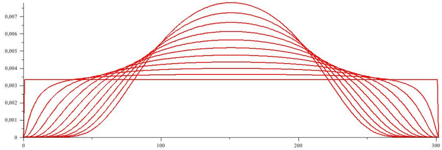
$p = 0$ Friedman

$p = 0.4$

$p = 0.7$

$p = 1$ Pólya





Système différentiel pour les urnes aléatoires équilibrées

$$\begin{pmatrix} 1-\mathcal{B}_p & \mathcal{B}_p \\ \mathcal{B}_p & 1-\mathcal{B}_p \end{pmatrix}, \quad \text{avec } \mathcal{B}_p \sim \mathrm{Bernoulli}(p)$$

$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ avec proba p $\qquad$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ avec proba $1-p$

Ainsi, $H(x,y,z) = X(x,y,z)^{a_0}\ Y(x,y,z)^{b_0}$, avec

$$\begin{cases} \dot{X} &= p\,X\,Y + (1-p)\,X^2 \\ \dot{Y} &= p\,X\,Y + (1-p)\,Y^2 \end{cases}$$

Probabilité d'avoir a boules noires et b boules blanches après n tirages :

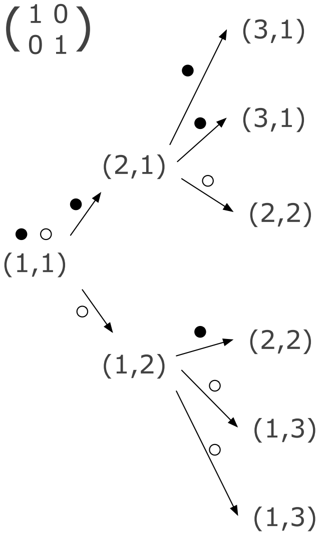
$$p_{n,a,b} = \frac{[x^a y^b z^n]H(x,y,z)}{[z^n]H(1,1,z)} \qquad \text{(histoires pondérées)}$$

Théorème La série génératrice $H(x,y,z)$ s'exprime par un système différentiel pour toute urne équilibrée $\left(\begin{smallmatrix} \mathcal{A} & \sigma-\mathcal{A} \\ \sigma-\mathcal{B} & \mathcal{B} \end{smallmatrix}\right)$, avec σ constante, et $\mathcal{A}, \mathcal{B}$ deux variables aléatoires à valeurs dans $\{-1, 0, 1, \ldots, \sigma\}$.

[Morcrette-Mahmoud, 2012]

Compter les histoires de Pólya

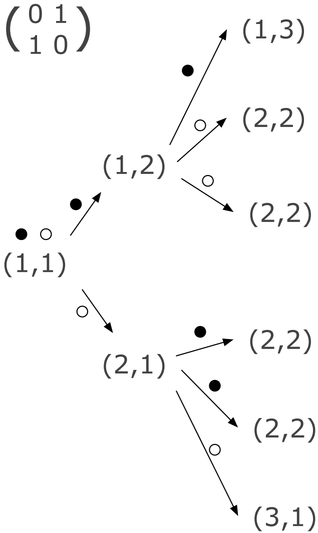
Soit l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(b_0, w_0) = (1, 1)$.



$$
\begin{array}{rc}
 & H(x,y,z) = \\
 & xy \\
+ & (xy^2 + x^2y)z \\
+ & (2x^3y + 2x^2y^2 + 2xy^3)\frac{z^2}{2} \\
+ & \dots
\end{array}
$$

Compter les histoires de Friedman

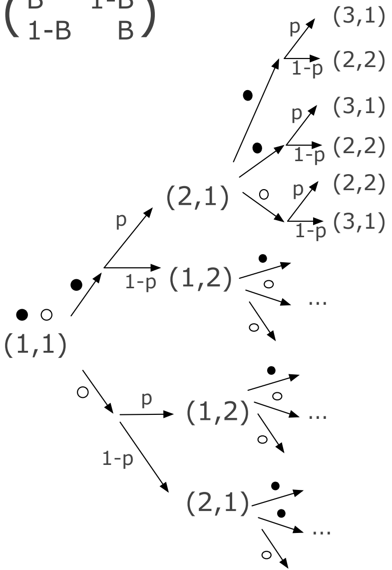
Soit l'urne $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $(b_0, w_0) = (1, 1)$.



$$
\begin{aligned}
H(x, y, z) = & \\
& xy \\
+ & (xy^2 + x^2y)z \\
+ & (xy^3 + 4x^2y^2 + x^3y)\frac{z^2}{2} \\
+ & \dots
\end{aligned}
$$

Compter les histoires de Pólya–Friedman

$$\begin{pmatrix} B & 1-B \\ 1-B & B \end{pmatrix}$$



p^2

$p(1-p)$

$$
\begin{aligned}
Q(x,y,z) = {} & xy \\
& + (xy^2 + x^2y)z \\
& + \Big((2p^2 + (1-p)^2 + 3p(1-p)) \quad xy^3 \\
& + (2p^2 + 4(1-p)^2 + 6p(1-p)) \quad x^2y^2 \\
& + (2p^2 + (1-p)^2 + 3p(1-p)) \quad x^3y \Big) \frac{z^2}{2} \\
& + \dots
\end{aligned}
$$

Application : More than 2 colors

$\mathcal{B}_p$ a Bernoulli(p), and (b_0, r_0, g_0) the initial condition.

$$\begin{pmatrix} -1 & \mathcal{B}_p & 1-\mathcal{B}_p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \left\{ \begin{array}{lcl} X' & = & p\,Y + (1-p)\,H\,; \\ Y' & = & Y\,; \\ H' & = & H\,. \end{array} \right.$$

$$Q\,(x,y,h,z) = X^{b_0}(z)\,Y^{r_0}(z)\,H^{g_0}(z)\,,$$

$$\begin{array}{lcl} X(z) & = & (p\,y + (1-p)\,h)\,(e^z - 1) + x\,; \\ Y(z) & = & y\,e^z\,; \\ H(z) & = & h\,e^z\,. \end{array}$$

The distribution of red balls R_n is

$$\mathbb{P}(R_n = r) = \left(\frac{p}{(1-p)}\right)^r \sum_{j=0}^{b_0} (-1)^j \binom{j}{r} \binom{b_0}{j} (1-p)^j \sum_{k=0}^{j} \binom{j}{k} (-1)^k \left(\frac{k}{s_0}\right)^n .$$

Mélanger règles aléatoires et absence d'équilibre

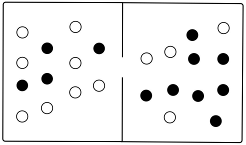
Exemple : graphes aléatoires sur n sommets

Commençons avec $\binom{n}{2}$ boules noires (arêtes potentielles)
À chaque étape : piocher une boule. Si elle est noire, choisir avec proba. p de la remplacer par une blanche (une véritable arête), et avec proba. $1-p$ la supprimer définitivement

$$\begin{pmatrix} -1 & \mathcal{B}_p \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \mathbb{1}_{\mathcal{B}_p=1} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \mathbb{1}_{\mathcal{B}_p=0}$$

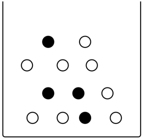
$\Psi(x,y,z)$ vérifie $\quad (x-(py+1-p)z)\,\partial_x\Psi + (y-zy)\,\partial_y\Psi = x^{\binom{n}{2}}$

Urns with multiple draws : the Bernoulli-Laplace process



Chamber A
a white
b black

Chamber B
b white
a black



One urn
a white, b black

$$\begin{array}{c} \\ WW \\ BW \\ BB \end{array} \begin{array}{c} \begin{array}{cc} W & B \end{array} \\ \begin{pmatrix} -1 & 1 \\ 0 & 0 \\ 1 & -1 \end{pmatrix} \end{array}$$

$$a + b = N$$

$$x^a y^b \xrightarrow{\mathcal{G}} a^2\, x^{-1} y^1\, x^a y^b + b^2\, x^1 y^{-1}\, x^a y^b + 2ab\, x^0 y^0\, x^a y^b$$

$$\Theta_u := u\partial_u \text{ (\textit{pick \& replace} a ball)}$$

$$\mathcal{G} := x^{-1} y^1\, \Theta_x^2 + x^1 y^{-1}\, \Theta_y^2 + 2\, x^0 y^0 \Theta_x \Theta_y$$

$$H_n(x, y) := \mathcal{G}^n \circ x^{a_0} y^{b_0}$$

$$\Pr\{(a_0, b_0) \rightsquigarrow (a, b)\} = \frac{1}{N^{2n}} \left[x^a y^b\right] H_n(x, y)$$

[Flajolet, July 2010]

Perspectives

- Applications aux statistiques locales de processus de graphes aléatoires (Séries -Parallèles [Mahmoud 2013], Erdös-Rényi ?, autres ?)
- Analyse asymptotique des urnes à coefficients aléatoires et urnes non équilibrées
- Urnes non équilibrées : le principe symbolique fonctionne pour les règles aléatoires, pour les pioches multiples

Conclusion

a uniform urn $\begin{pmatrix} \mathcal{U}_\theta & \theta - \mathcal{U}_\theta \\ \theta - \mathcal{U}_\theta & \mathcal{U}_\theta \end{pmatrix}$

$$\mathbb{P}(B_n = b) = \frac{T_{\theta,n,b-b_0}}{(\theta+1)^n}, \qquad Q(x,1,z) = \frac{x^{b_0}}{\left(1 - \frac{\theta}{\theta+1}\left(\sum_{l=0}^{\theta} x^l\right) z\right)^{\frac{b_0+w_0}{\theta}}}$$

where $T_{\theta,n,k} := [x^k]\,(1 + x + \ldots + x^\theta)^n$.

Knuth *cutting loops*

$\mathcal{K} = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad (a_0, b_0) = (m, 0)$

$$\Psi_{\mathcal{K}}(x,1,z) = \int_0^1 \frac{t^{m-1}}{(1-zt)^{m+1}} \left(x(1-z) - z\ln(t)\right)^m \, dt$$

Coupon collector with delay $\begin{pmatrix} -\mathcal{B}_p & \mathcal{B}_p \\ 0 & 0 \end{pmatrix}$

$$Q(x,1,z) = \left(e^z + (x-1)e^{(1-p)z}\right)^{b_0} e^{w_0 z}.$$

$$\mathbb{P}(B_n = b) = \sum_{j=b}^{b_0} (-1)^{j-b} \binom{b_0}{j}\binom{j}{b}\left(\frac{s_0 - pj}{s_0}\right)^n$$

If you can specify it,
you can analyze it !

$\mathcal{BM} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ $\quad a, d > 0 \quad (a_0, b_0)$

$$\Psi_{\mathcal{BM}}(x,y,z) = \int_0^y \left[(x^{-a} - z)\left(\frac{t^{-d} - z}{y^{-d} - z}\right)^{a/d} + z\right]^{-a_0/a} \frac{t^{b_0 - 1}}{1 - zt^d}\, dt$$

$$\begin{pmatrix} -1 & \mathcal{B}_p & 1 - \mathcal{B}_p \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbb{P}(R_n = r) = \left(\frac{p}{(1-p)}\right)^r \sum_{j=0}^{b_0} (-1)^j \binom{j}{r}\binom{b_0}{j}(1-p)^j \sum_{k=0}^{j} \binom{j}{k}(-1)^k \left(\frac{k}{s_0}\right)^n.$$

For $(a_0, b_0) = (0, 1)$, $\quad \Psi_{\mathcal{KMR}}(1,y,z) =$ $\quad \mathcal{KMR} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

$$\int_0^1 \frac{z}{(1-zt)\left(1 - (1-zt)\left[\frac{1}{1-z} - \frac{z}{y(1-z)} + \ln(1-z) - \ln(1-zt)\right]\right)}\, dt$$

Example : random graph of n vertices

Start with $\binom{n}{2}$ black balls (for potential edges)

$$\begin{pmatrix} -1 & \mathcal{B}_p \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \mathbb{1}_{\mathcal{B}_p = 1} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \mathbb{1}_{\mathcal{B}_p = 0}$$

$\Psi(x,y,z)$ verifies $\quad (x - (py + 1 - p)z)\,\partial_x \Psi + (y - zy)\,\partial_y \Psi = x^{\binom{n}{2}}$