

Urnes de Pólya

Approche par la combinatoire analytique

Basile Morcrette
(avec Philippe Flajolet et Philippe Dumas)

Projet Algorithms, INRIA Rocquencourt.
LIP6, UPMC

ANR Boole
27 Juin 2011

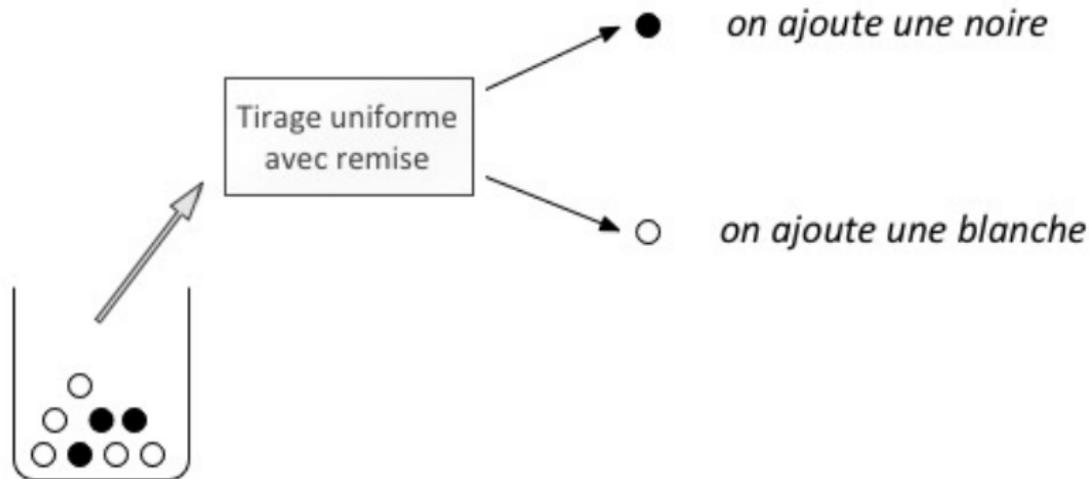
INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE



Plan

1. Modèles d'urnes
2. Approche exacte
formules booléennes
3. Analyse de singularité
urnes des k -arbres ordonnés croissants
4. Méthode de col
urnes de croissance préférentielle
5. Vers d'autres modèles d'urnes

1. Modèles d'urnes



- ▶ Une urne contenant des boules de deux couleurs différentes
- ▶ Des règles fixées pour l'évolution de l'urne

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

Urne de Pólya équilibrée

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \alpha, \delta \in \mathbb{Z}, \quad \beta, \gamma \in \mathbb{N}$$

Urne équilibrée : $\boxed{\alpha + \beta = \gamma + \delta}$ (nombre total de boules déterministe)

Configuration initiale fixée (a_0, b_0) : a_0 boules \bullet (comptées par x)

b_0 boules \circ (comptées par y)

Définition

Histoire de longueur n : une suite de n évolutions (n tirages)

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

$H_{n, a, b}$: nombre d'histoires de longueur n , débutant en (a_0, b_0) , et terminant en (a, b)

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.

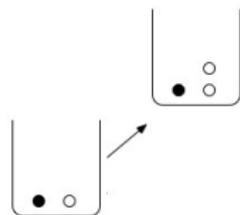


$$H(x, y, z) =$$

xyz

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.

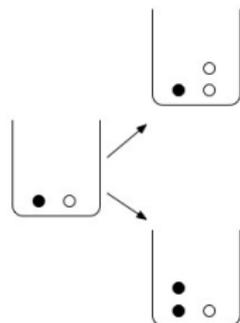


$$H(x, y, z) =$$

xyz

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



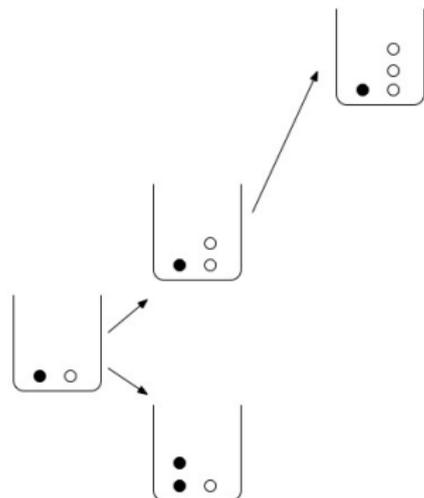
$$H(x, y, z) =$$

$$xyz$$

$$+ (xy^2 + x^2y) \frac{z^2}{2}$$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



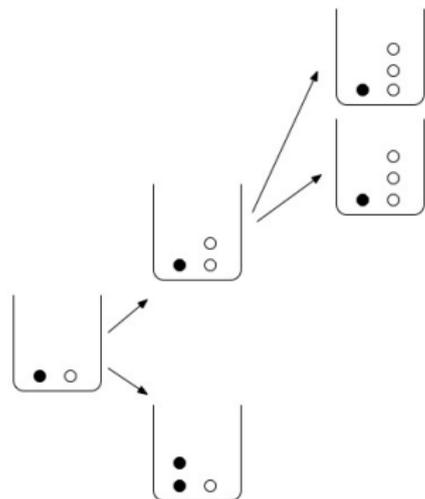
$$H(x, y, z) =$$

$$xyz$$

$$+ (xy^2 + x^2y) \frac{z^2}{2}$$

Compter les histoires - Exemple

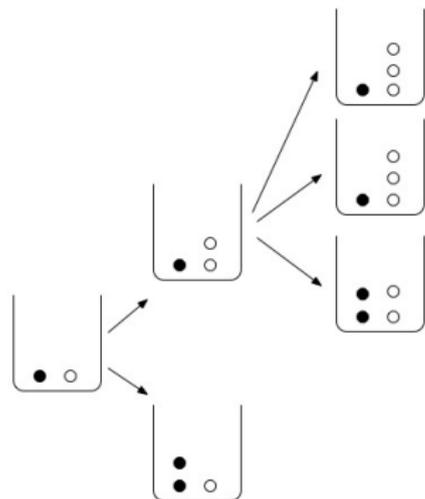
Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



$$H(x, y, z) =$$
$$xyz$$
$$+ (xy^2 + x^2y) \frac{z^2}{2}$$

Compter les histoires - Exemple

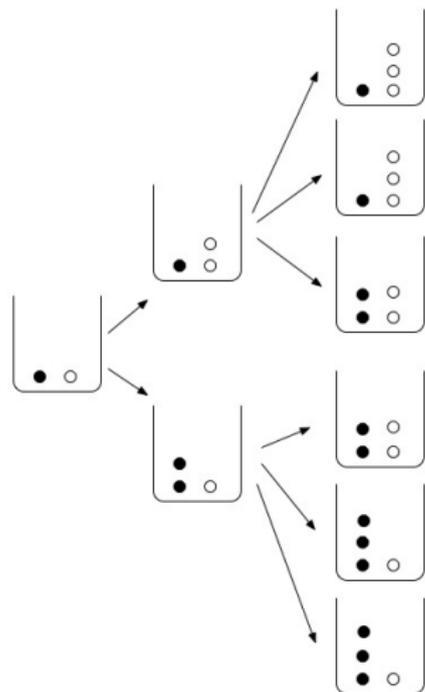
Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



$$H(x, y, z) =$$
$$xyz$$
$$+ (xy^2 + x^2y)\frac{z^2}{2}$$

Compter les histoires - Exemple

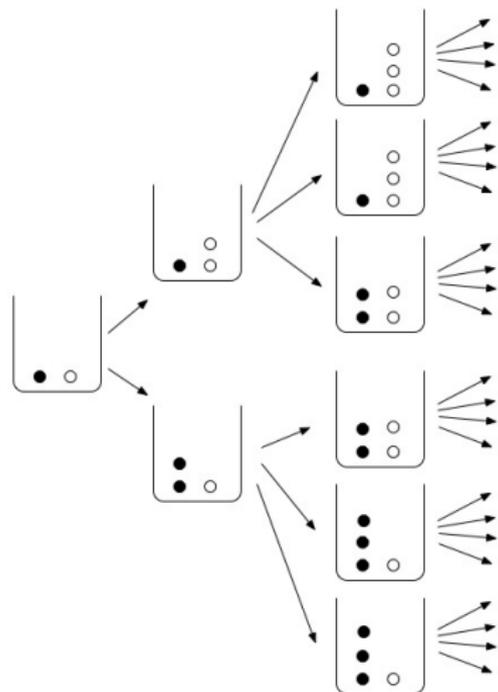
Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



$$\begin{aligned} H(x, y, z) = & \\ & xyz \\ & + (xy^2 + x^2y) \frac{z^2}{2} \\ & + (2xy^3 + 2x^2y^2 + 2x^3y) \frac{z^3}{6} \end{aligned}$$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



$$\begin{aligned} H(x, y, z) = & \\ & xyz \\ & + (xy^2 + x^2y) \frac{z^2}{2} \\ & + (2xy^3 + 2x^2y^2 + 2x^3y) \frac{z^3}{6} \\ & + \dots \end{aligned}$$

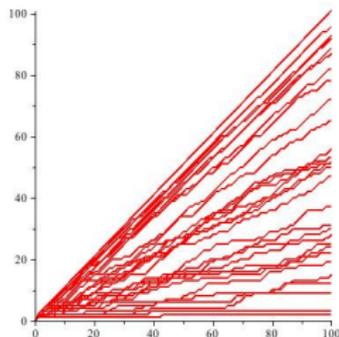
Des comportements très divers

Problème : Comprendre la composition de l'urne après n étapes, lorsque n tend vers ∞ .

Des comportements très divers

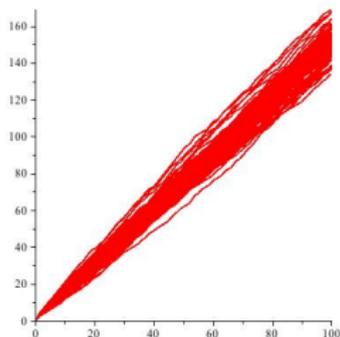
Problème : Comprendre la composition de l'urne après n étapes, lorsque n tend vers ∞ .

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



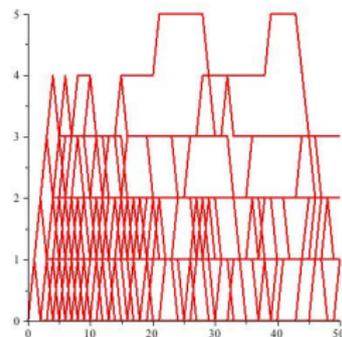
Urne de Pólya

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



Urne à croissance préférentielle

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



Urne triangulaire 3×3

Résultats probabilistes

$$\text{Urne } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

$$\text{Rapport } \rho = \frac{\alpha - \gamma}{\alpha + \beta}$$

- ▶ **Petites urnes:** $\rho \leq \frac{1}{2}$

Loi limite **gaussienne** [Smythe96] [Janson04]

- ▶ **Grandes urnes :** $\rho > \frac{1}{2}$

Lois **non gaussiennes** [Mahmoud] [Janson04]
[Chauvin–Pouyanne–Sahnoun11]

Outils :

- Plongements en temps continu [Jan04] [ChPoSa11]
- Martingales, théorème de la limite centrale

Urnes équilibrées et analyse

- ▶ Premiers pas [Flajolet–Gabarro–Pekari05] : *Analytic urns*
- ▶ [Flajolet–Dumas–Puyhaubert06], sur les urnes avec coefficients négatifs, et cas triangulaire
- ▶ [Kuba–Panholzer–Hwang07], urnes non équilibrées

Approche analytique Théorème [FIDuPu06]

$$\text{Urne } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \text{ et } \begin{cases} (a_0, b_0) \\ \alpha + \beta = \gamma + \delta \end{cases} \implies \text{où } \begin{cases} \dot{X} = X^{\alpha+1} Y^\beta \\ \dot{Y} = X^\gamma Y^{\delta+1} \end{cases} \quad H = X^{a_0} Y^{b_0}$$

Preuve de l'isomorphisme

Différencier = Piocher

$$\partial_x[xx \dots x] = (\cancel{x}x \dots x) + (x\cancel{x} \dots x) + \dots + (xx \dots \cancel{x})$$

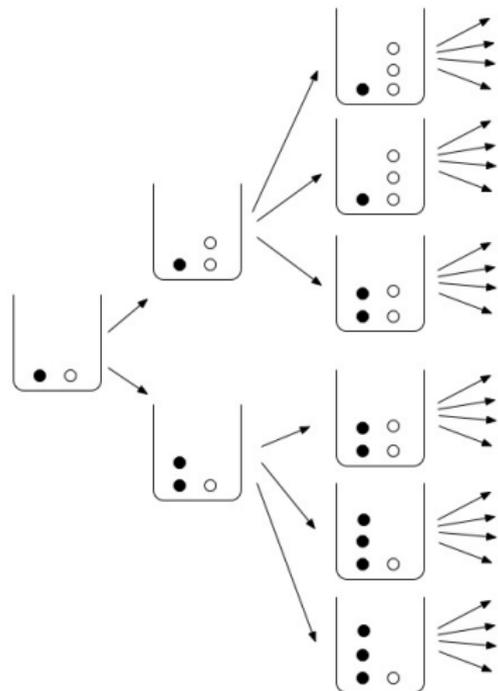
$$x\partial_x[xx \dots x] = (\underline{x}x \dots x) + (x\underline{x} \dots x) + \dots + (xx \dots \underline{x})$$

Posons $\mathfrak{D} = x^{\alpha+1}y^\beta\partial_x + x^\gamma y^{\delta+1}\partial_y$

Alors $\boxed{\mathfrak{D}[x^a y^b] = ax^{a+\alpha}y^{b+\beta} + bx^{a+\gamma}y^{b+\delta}}$

Compter les histoires - Exemple

Prenons l'urne $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ et $(a_1, b_1) = (1, 1)$.



$$\begin{aligned}
 H(x, y, z) = & \\
 & xyz \\
 & + (xy^2 + x^2y) \frac{z^2}{2} \\
 & + (2xy^3 + 2x^2y^2 + 2x^3y) \frac{z^3}{6} \\
 & + \dots
 \end{aligned}$$

Preuve de l'isomorphisme

Différencier = Piocher

$$\partial_x[xx \dots x] = (\cancel{x}x \dots x) + (x\cancel{x} \dots x) + \dots + (xx \dots \cancel{x})$$

$$x\partial_x[xx \dots x] = (\underline{x}x \dots x) + (x\underline{x} \dots x) + \dots + (xx \dots \underline{x})$$

Posons $\mathfrak{D} = x^{\alpha+1}y^{\beta}\partial_x + x^{\gamma}y^{\delta+1}\partial_y$

Alors $\boxed{\mathfrak{D}[x^a y^b] = ax^{a+\alpha}y^{b+\beta} + bx^{a+\gamma}y^{b+\delta}}$

$$\mathfrak{D}^n[x^{a_0}y^{b_0}] = \sum_{a,b} H_{n,a,b}x^a y^b$$

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

Preuve de l'isomorphisme

Différencier = Piocher

$$x \partial_x [xx \dots x] = (\underline{x}x \dots x) + (x\underline{x} \dots x) + \dots + (xx \dots \underline{x})$$

$$\mathcal{D} = x^{\alpha+1} y^\beta \partial_x + x^\gamma y^{\delta+1} \partial_y$$

$$\mathcal{D}[x^a y^b] = a x^{a+\alpha} y^{b+\beta} + b x^{a+\gamma} y^{b+\delta}$$

$$\mathcal{D}^n [x^{a_0} y^{b_0}] = \sum_{a,b} H_{n,a,b} x^a y^b$$

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

$$H(X(t), Y(t), z) = \sum_{n \geq 0} \partial_t^n [X(t)^{a_0} Y(t)^{b_0}] \frac{z^n}{n!} = X(t+z)^{a_0} Y(t+z)^{b_0}$$

Enfin $t = 0$, et c'est gagné !!

Soit $(X(t), Y(t))$ solution de

$$\begin{cases} \dot{X} = X^{\alpha+1} Y^\beta & X(t=0) = x \\ \dot{Y} = X^\gamma Y^{\delta+1} & Y(t=0) = y \end{cases}$$

$$\partial_t (X^a Y^b)$$

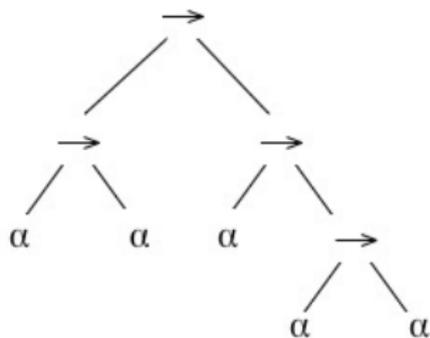
$$= a X^{a-1} \dot{X} Y^b + b X^a Y^{b-1} \dot{Y}$$

$$= a X^{a+\alpha} Y^{b+\beta} + b X^{a+\gamma} Y^{b+\delta}$$

$$\partial_t^n (X^a Y^b) = \mathcal{D}^n [x^a y^b] \quad \begin{array}{l} x \rightarrow X \\ y \rightarrow Y \end{array}$$

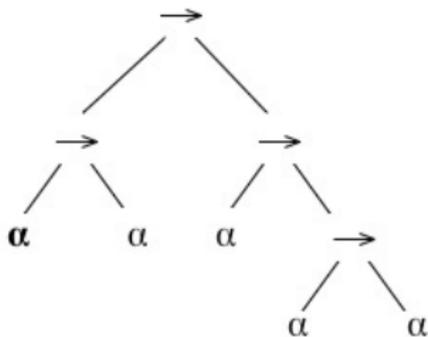
2. Urnes et branches d'arbres aléatoires

- ▶ **Motivation** : quantifier la proportion de tautologies parmi les formules logiques ayant comme seul connecteur l'implication.
- ▶ **Modèle probabiliste** : croissance uniforme aux feuilles (modèle ABR)
 - ▶ tirer une feuille au hasard uniformément
 - ▶ la remplacer par un noeud binaire et deux feuilles



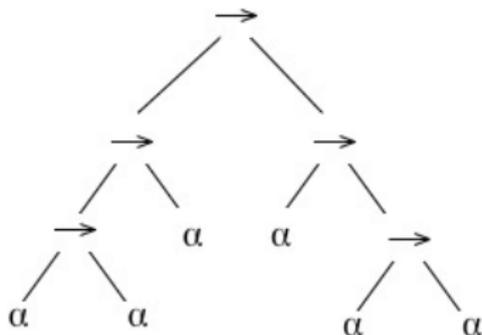
2. Urnes et branches d'arbres aléatoires

- ▶ **Motivation** : quantifier la proportion de tautologies parmi les formules logiques ayant comme seul connecteur l'implication.
- ▶ **Modèle probabiliste** : croissance uniforme aux feuilles (modèle ABR)
 - ▶ tirer une feuille au hasard uniformément
 - ▶ la remplacer par un noeud binaire et deux feuilles



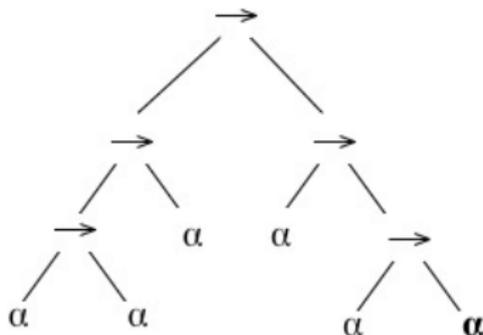
2. Urnes et branches d'arbres aléatoires

- ▶ **Motivation** : quantifier la proportion de tautologies parmi les formules logiques ayant comme seul connecteur l'implication.
- ▶ **Modèle probabiliste** : croissance uniforme aux feuilles (modèle ABR)
 - ▶ tirer une feuille au hasard uniformément
 - ▶ la remplacer par un noeud binaire et deux feuilles



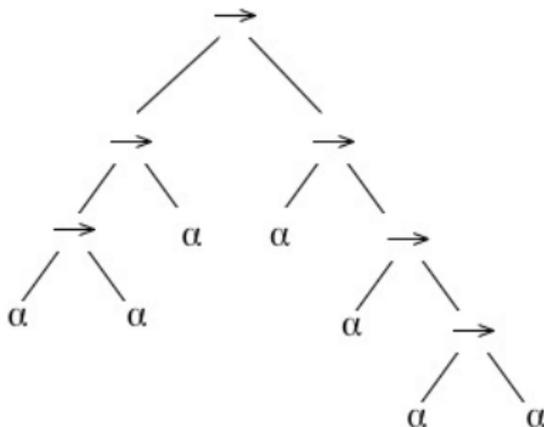
2. Urnes et branches d'arbres aléatoires

- ▶ **Motivation** : quantifier la proportion de tautologies parmi les formules logiques ayant comme seul connecteur l'implication.
- ▶ **Modèle probabiliste** : croissance uniforme aux feuilles (modèle ABR)
 - ▶ tirer une feuille au hasard uniformément
 - ▶ la remplacer par un noeud binaire et deux feuilles

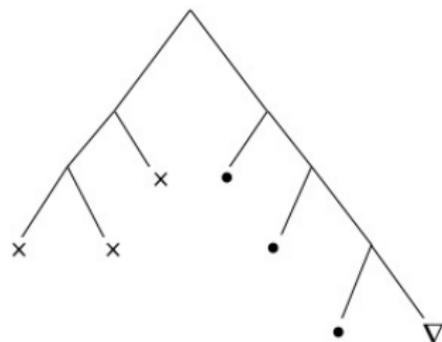


2. Urnes et branches d'arbres aléatoires

- ▶ **Motivation** : quantifier la proportion de tautologies parmi les formules logiques ayant comme seul connecteur l'implication.
- ▶ **Modèle probabiliste** : croissance uniforme aux feuilles (modèle ABR)
 - ▶ tirer une feuille au hasard uniformément
 - ▶ la remplacer par un noeud binaire et deux feuilles



Modélisation par une urne 3×3



3 couleurs,
avec les règles :

$\nabla \rightarrow \bullet \nabla$
 $\bullet \rightarrow \times \times$
 $\times \rightarrow \times \times$

Urne
correspondante :

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

Fonction génératrice des histoires

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

z compte la longueur de l'histoire,
 y compte le nombre de boules \bullet .

Loi de Poisson dans les sous-arbres

Soit $U_{k,n}$ le nombre de sous-arbres gauches de taille k directement accrochés à la branche droite d'un arbre aléatoire de taille n .

Théorème

- ▶ $U_{1,n}$ converge en loi, $U_{1,n} \xrightarrow[n \rightarrow \infty]{} U_1$,
- ▶ $U_1 \sim \mathcal{Poisson}(1)$, avec vitesse de convergence en $O\left(\frac{2^n}{n!}\right)$.

Loi de Poisson dans les sous-arbres

Soit $U_{k,n}$ le nombre de sous-arbres gauches de taille k directement accrochés à la branche droite d'un arbre aléatoire de taille n .

Théorème

- ▶ $U_{1,n}$ converge en loi, $U_{1,n} \xrightarrow[n \rightarrow \infty]{} U_1$,
- ▶ $U_1 \sim \mathcal{Poisson}(1)$, avec vitesse de convergence en $O\left(\frac{2^n}{n!}\right)$.

Généralisation Utilisation d'une urne $(k+2) \times (k+2)$

Théorème

- ▶ $U_{k,n}$ converge en loi, $U_{k,n} \xrightarrow[n \rightarrow \infty]{} U_k$,
- ▶ $U_k \sim \mathcal{Poisson}\left(\frac{1}{k}\right)$, avec vitesse de convergence en $O\left(\frac{(2k)^n}{n!}\right)$.

3. Urnes des k -arbres

Motivation : modélisation de réseaux sociaux [Panholzer–Seitz 2010]

Définition

Un k -arbre T est

- ▶ soit une k -clique
- ▶ soit il existe un sommet f ayant une k -clique comme voisin et $T \setminus f$ est un k -arbre

Ordonné : fils distinguables.

Croissant : sommets numérotés dans l'ordre d'apparition.

3. Urnes des k -arbres

Motivation : modélisation de réseaux sociaux [Panholzer–Seitz 2010]

Définition

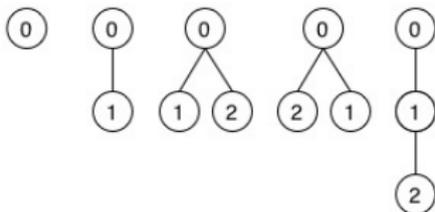
Un k -arbre T est

- ▶ soit une k -clique
- ▶ soit il existe un sommet f ayant une k -clique comme voisin et $T \setminus f$ est un k -arbre

Ordonné : fils distinguables.

Croissant : sommets numérotés dans l'ordre d'apparition.

1-arbres ordonnés croissants (dits PORT)



3. Urnes des k -arbres

Motivation : modélisation de réseaux sociaux [Panholzer–Seitz 2010]

Définition

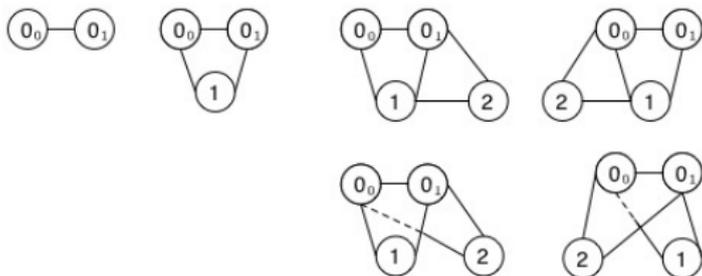
Un k -arbre T est

- ▶ soit une k -clique
- ▶ soit il existe un sommet f ayant une k -clique comme voisin et $T \setminus f$ est un k -arbre

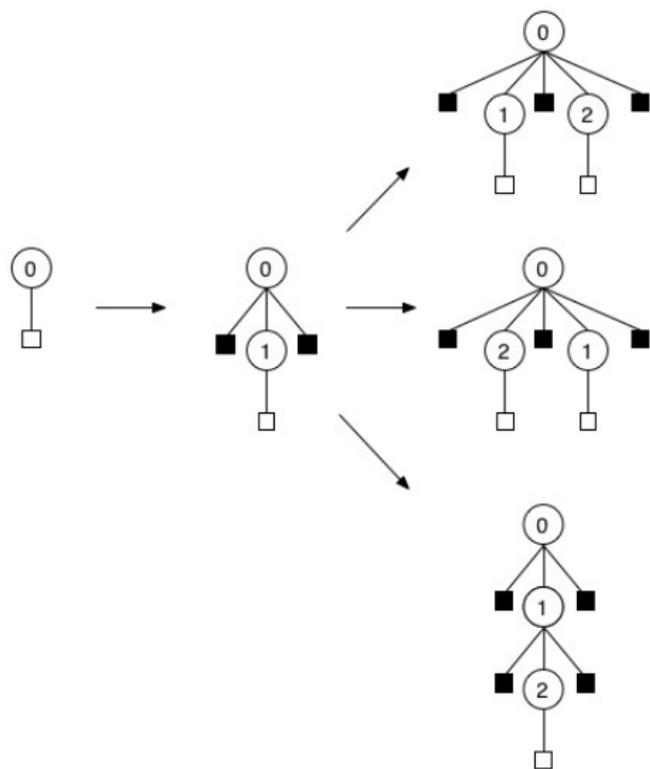
Ordonné : fils distinguables.

Croissant : sommets numérotés dans l'ordre d'apparition.

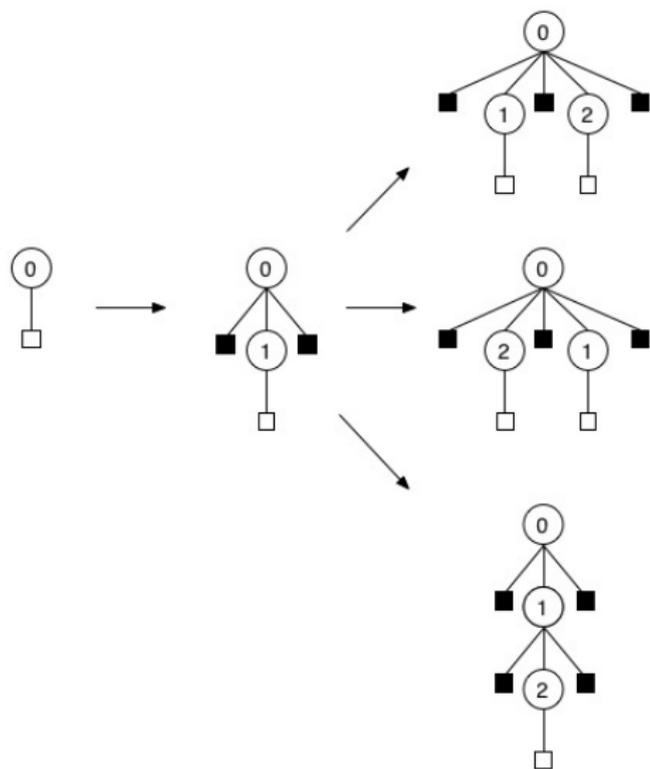
2-arbres ordonnés croissants



Modélisation



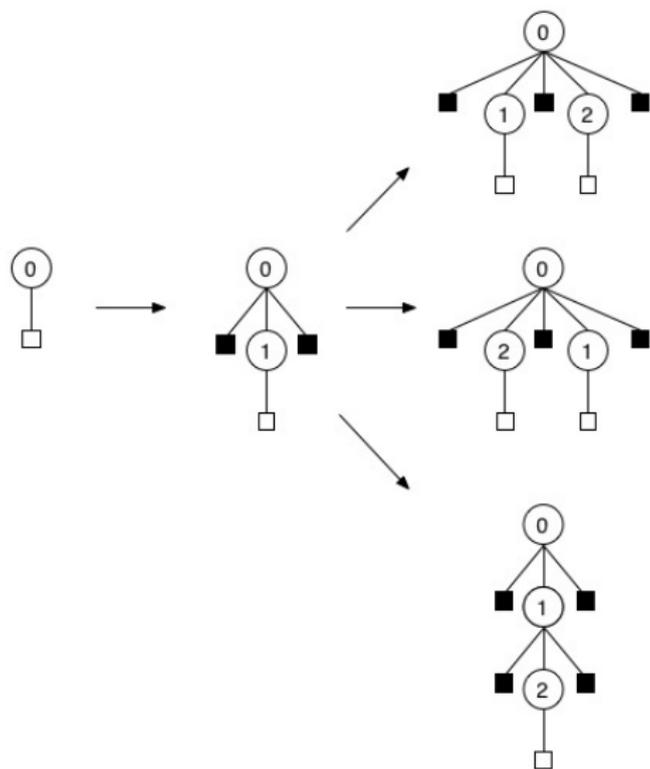
Modélisation



Pour les 1-arbres

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

Modélisation



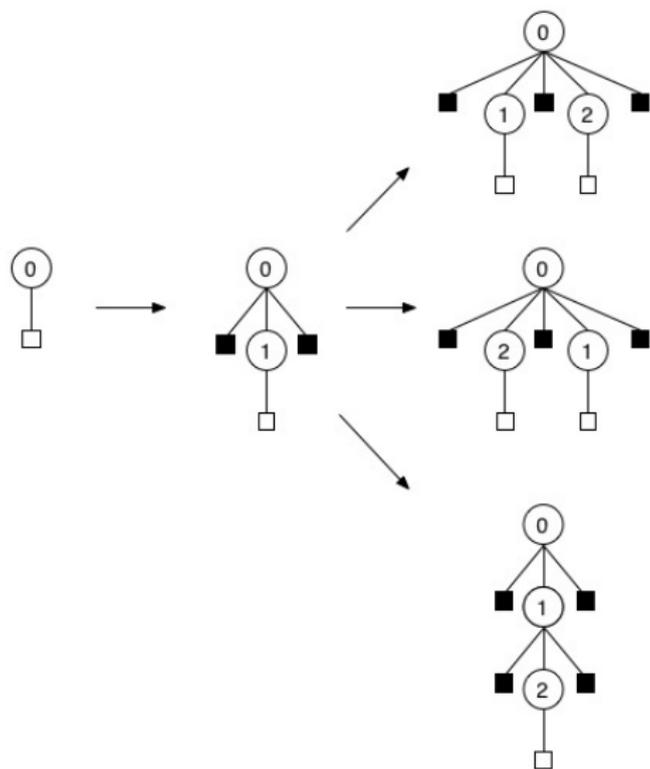
Pour les 1-arbres

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

Pour les k -arbres

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

Modélisation



Pour les 1-arbres

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

Pour les k -arbres

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

$$\frac{\frac{d}{dz} H(x, z)}{H(x, z)^k (H(x, z) - x + 1)^2} = 1$$

Analyse de singularité

$$\frac{\frac{d}{dz}H(x, z)}{H(x, z)^k (H(x, z) - x + 1)^2} = 1$$

Ingrédients :

- ▶ décomposition en éléments simples
- ▶ intégration
- ▶ changements de variable

$$x^k e^{-x} \prod_{i=1}^{k-1} \exp\left(\left(1 - \frac{k}{i}\right) (1 - x^{-1})^i\right) = \exp(-1 - b^{k+1} (K_k(b) - z))$$

un peu d'analyse...

$$p_n(b) = \left(\frac{1}{(k+1)K_k(b)}\right)^n \left(1 + O\left(\frac{1}{n}\right)\right)$$

Loi normale limite

Soit X_n la variable aléatoire comptant le nombre de boules \square dans l'urne U_1 après n étapes.

Cas $k = 1$: nombre de feuilles dans un PORT de taille n .

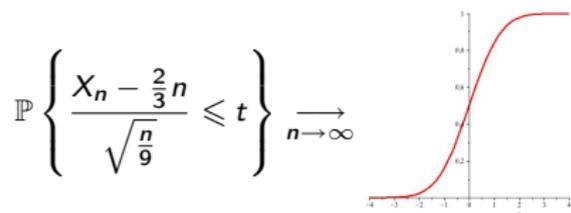
Théorème

$$\mathbb{P} \left\{ \frac{X_n - \frac{2}{3}n}{\sqrt{\frac{n}{9}}} \leq t \right\} = \Phi(t) + O\left(\frac{1}{\sqrt{n}}\right).$$

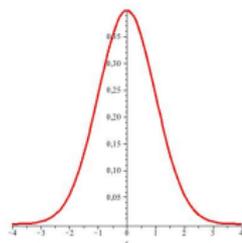
Théorème

Notons $p_{n,k} = \mathbb{P}\{X_n = k\}$. La distribution des X_n satisfait une **loi locale limite** de type **gaussienne** avec vitesse de convergence $O\left(\frac{1}{\sqrt{n}}\right)$, i.e.

$$\sup_{t \in \mathbb{R}} \left| \frac{\sqrt{n}}{3} p_{n, \lfloor 2n/3 + t\sqrt{n}/3 \rfloor} - \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \right| \leq \frac{1}{\sqrt{n}}.$$



$$\frac{\sqrt{3n}}{2} \mathbb{P} \left\{ X_n = \left\lfloor \frac{2n}{3} + t \frac{\sqrt{n}}{3} \right\rfloor \right\} \xrightarrow{n \rightarrow \infty}$$

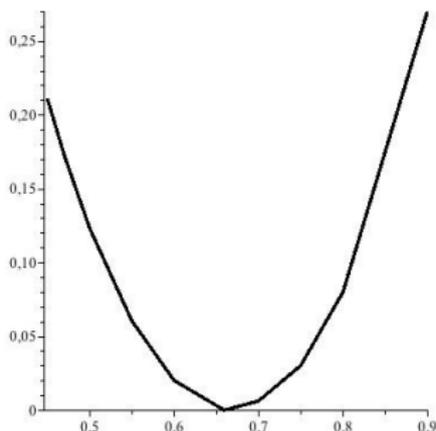


Grandes déviations

- ▶ borne exponentiellement petite sur la grande déviation par rapport à la valeur moyenne : quantification sur les **événements rares**

Théorème

- ▶ si $0.42 < t < 2/3$, $\mathbb{P}(X_n \leq tn) \approx e^{-nW(t)}$ (queue gauche)
- ▶ si $2/3 < t < 0.73$, $\mathbb{P}(X_n \geq tn) \approx e^{-nW(t)}$ (queue droite)



4. Urnes à croissance préférentielle

Motivation : caractériser les urnes 2×2 **additives** (coefficients positifs).

Approche : avoir une classe d'urnes avec des fonctions génératrices "exploitables".

Théorème [FIM.11]

La classe d'urnes équilibrées $\begin{pmatrix} 2\alpha & \beta \\ \alpha & \alpha + \beta \end{pmatrix}$, avec $\alpha > 0$, $\beta \geq 0$, a une fonction génératrice bvariée **algébrique**.

La FG des histoires $H(x, 1, z)$ est racine du polynôme en Y

$$(z - a - b(x)) Y^{2\alpha+\beta} + b(x) Y^\alpha + a$$

$$\text{avec } b(x) = \frac{x^{-\alpha} - 1}{\alpha + \beta} \text{ et } a = (2\alpha + \beta)^{-1}$$

Preuve

$$\text{Système différentiel : } \begin{cases} \dot{X} = X^{2\alpha+1} Y^\beta \\ \dot{Y} = X^\alpha Y^{\alpha+\beta+1} \end{cases} \quad \dot{X} = \frac{\partial}{\partial z} X$$

$$\frac{\dot{X}}{X^{\alpha+1}} = \frac{\dot{Y}}{Y^{\alpha+1}} = X^\alpha Y^\beta$$

$$X^{-\alpha} - Y^{-\alpha} = x^{-\alpha} - y^{-\alpha}$$

$$\frac{\dot{Y}}{Y^{\alpha+\beta+1}} (Y^{-\alpha} + x^{-\alpha} - y^{-\alpha}) = 1$$

$$\frac{1}{2\alpha + \beta} Y^{-(2\alpha+\beta)} + \frac{x^{-\alpha} - y^{-\alpha}}{\alpha + \beta} Y^{-(\alpha+\beta)} = - \left(z - \frac{x^{-\alpha} - y^{-\alpha}}{\alpha + \beta} - \frac{1}{2\alpha + \beta} \right)$$

Urne équilibrée $a + b = a_0 + b_0 + n\sigma$. On pose $y = 1$.

$$\left(z - \frac{x^{-\alpha} - 1}{\alpha + \beta} - \frac{1}{2\alpha + \beta} \right) Y^{2\alpha+\beta} + \frac{x^{-\alpha} - 1}{\alpha + \beta} Y^\alpha + \frac{1}{2\alpha + \beta} = 0$$

Premières observations

$$\text{Équilibre } \sigma = 2\alpha + \beta \quad \rho = \frac{\alpha}{2\alpha + \beta} \leq \frac{1}{2}$$

Pour $x = 1$, l'équation devient : $(z - \sigma^{-1})Y^\sigma + \sigma^{-1} = 0$

Ainsi pour $(a_0, b_0) = (0, 1)$

$$H(1, 1, z) = (1 - \sigma z)^{-1/\sigma} \quad h_n \sim \frac{\sigma^n n^{1/\sigma-1}}{\Gamma(1/\sigma)}$$

Proposition

Soit X_n la variable aléatoire comptant le nombre de boules de couleur x dans l'urne au temps n . Alors

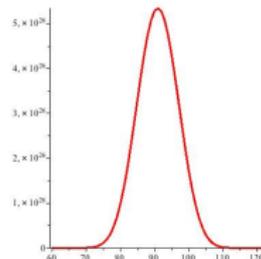
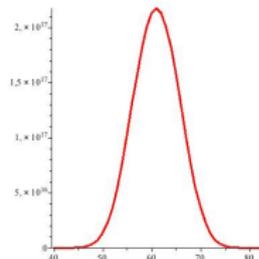
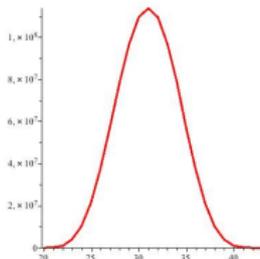
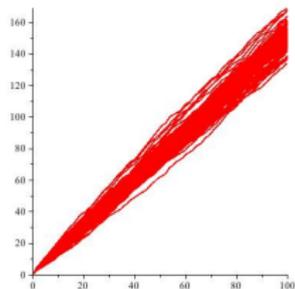
$$\mathbb{E}(X_n) = \frac{\alpha(2\alpha + \beta)}{\alpha + \beta} n + \frac{\alpha}{\alpha + \beta} \frac{\Gamma(\frac{1}{2\alpha + \beta})}{\Gamma(\frac{\alpha + 1}{2\alpha + \beta})} n^{\frac{\alpha}{2\alpha + \beta}} + \frac{\alpha}{\alpha + \beta} + O\left(n^{\frac{\alpha}{2\alpha + \beta} - 1}\right),$$

$$\mathbb{V}(X_n) = \frac{\alpha^3(2\alpha + \beta)}{(\alpha + \beta)^2} n + O\left(n^{\frac{\alpha + \beta}{2\alpha + \beta}}\right).$$

Exemple $\alpha = 1, \beta = 1$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \begin{array}{l} x \rightarrow x x y \\ y \rightarrow x y y \end{array}$$

croissance préférentielle



$$\left(z - \frac{x^{-1} - 1}{2} - \frac{1}{3} \right) Y^3 + \frac{x^{-1} - 1}{2} Y + \frac{1}{3} = 0$$

Méthode de col pour $x = 1$

$$\left(z - \frac{1}{3}\right) Y^3 + \frac{1}{3} = 0$$

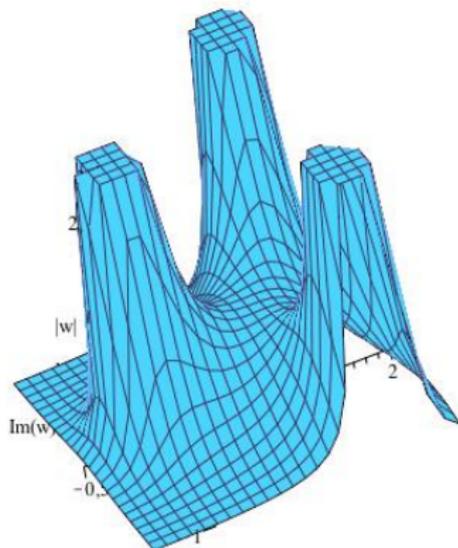
$$y_n = \frac{1}{2i\pi} \oint \frac{Y(z)}{z^{n+1}} dz$$

$$y_n = \frac{3^{n+1}}{2i\pi} \oint a(w) h(w)^{n+1} dw$$

$$\begin{cases} a(w) = 1 - w \\ h(w) = \frac{1}{w(w^2 - 3w + 3)} \end{cases}$$

$$h'(w) = \frac{-3(w-1)^2}{w^2(w^2 - 3w + 3)^2}$$

évaluer l'intégrale selon le bon contour...



$$w \mapsto |h(w)|$$

3 pôles

1 col double en $w = 1$

Méthode de col $x=1$ (suite)

$$t \in [0..L]$$

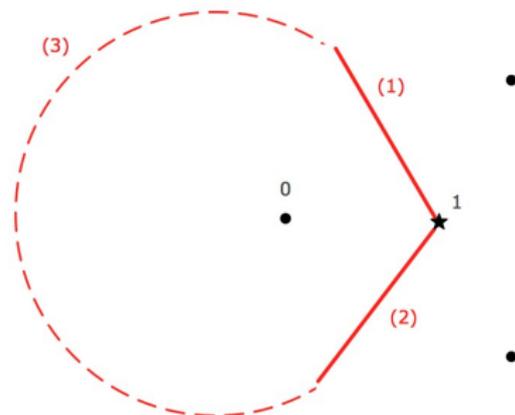
$$(1) w(t) = 1 + te^{i2\pi/3}$$

$$(2) w(t) = 1 + te^{-i2\pi/3}$$

$$h(w(t))^n = \exp(-n(t^3 + O(t^6)))$$

Choisir $L...$ $nL^3 \rightarrow \infty$ et $nL^6 \rightarrow 0$

Prenons $L \sim n^{-1/4}$



$$\int_{(1)} + \int_{(2)} : \int_0^\infty ue^{-u^3} du \text{ et } \int_{(3)} \text{ exponentiellement négligeable}$$

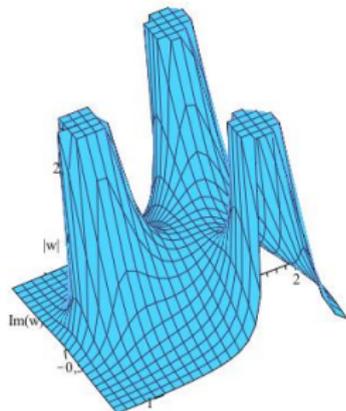
$$y_n = \frac{3^n}{\Gamma(1/3)} \left(n^{-2/3} + O\left(n^{-11/12}\right) \right)$$

Méthode de col $x \neq 1$

$$\left(z - \frac{x^{-1} - 1}{2} - \frac{1}{3}\right) Y^3 + \frac{x^{-1} - 1}{2} Y + \frac{1}{3} = 0$$

$$y_n = \frac{3^{n+1}}{2i\pi} \oint a_x(w) h_x(w)^{n+1} dw$$

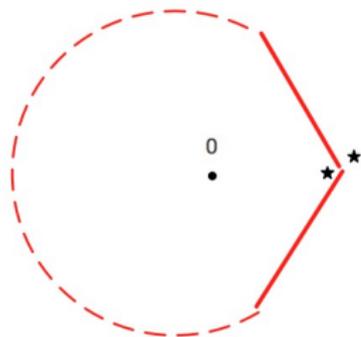
$$h'_x(1) = h'_x(x^{-1}) = 0$$



$w \mapsto |h_x(w)|$

3 pôles

2 cols en $w = 1$ et $w = x^{-1}$



$$x = 1 + \frac{\tilde{x}}{\sqrt{n}}, \quad |\tilde{x}| < 1$$

$$y_n(x) \sim \frac{3^n n^{-2/3}}{\Gamma(1/3)} \exp\left(\frac{3}{2} \sqrt{n} \tilde{x} - \frac{3}{8} \tilde{x}^2\right)$$

$$\boxed{p_n(x) = \frac{y_n(x)}{y_n(1)} \sim \exp\left(\frac{3}{2} \sqrt{n} \tilde{x} - \frac{3}{8} \tilde{x}^2\right)}$$

Loi normale limite

Soit X_n la variable aléatoire comptant le nombre de boules \bullet dans l'urne après n étapes.

Théorème

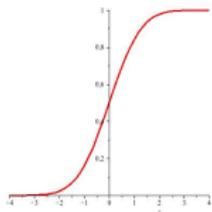
$$\mathbb{P} \left\{ \frac{X_n - \frac{3}{2}n}{\sqrt{\frac{3n}{4}}} \leq t \right\} = \Phi(t) + O\left(\frac{1}{\sqrt{n}}\right).$$

Théorème

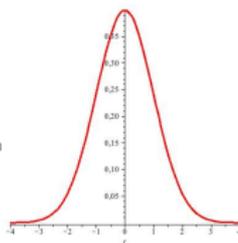
Notons $p_{n,k} = \mathbb{P}\{X_n = k\}$. La distribution des X_n satisfait une **loi locale limite** de type **gaussienne** avec vitesse de convergence $O\left(\frac{1}{\sqrt{n}}\right)$, i.e.

$$\sup_{t \in \mathbb{R}} \left| \frac{\sqrt{3n}}{2} p_{n, \lfloor 3n/2 + t\sqrt{3n}/2 \rfloor} - \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \right| \leq \frac{1}{\sqrt{n}}.$$

$$\mathbb{P} \left\{ \frac{X_n - \frac{3}{2}n}{\sqrt{\frac{3n}{4}}} \leq t \right\} \xrightarrow{n \rightarrow \infty}$$



$$\frac{\sqrt{3n}}{2} \mathbb{P} \left\{ X_n = \left\lfloor \frac{3n}{2} + t \frac{\sqrt{3n}}{2} \right\rfloor \right\} \xrightarrow{n \rightarrow \infty}$$

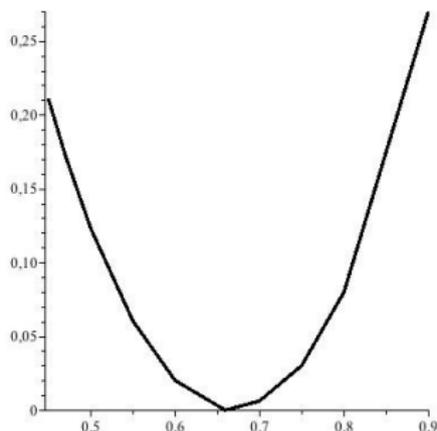


Grandes déviations

- ▶ Borne exponentiellement petite sur la grande déviation par rapport à la valeur moyenne : quantification sur les **événements rares**

Théorème

- ▶ si $0.42 < t < 2/3$, $\mathbb{P}(X_n \leq tn) \approx e^{-nW(t)}$ (queue gauche)
- ▶ si $2/3 < t < 0.73$, $\mathbb{P}(X_n \geq tn) \approx e^{-nW(t)}$ (queue droite)



Cas général

$$\left(z - \frac{x^{-\alpha} - 1}{\alpha + \beta} - \frac{1}{2\alpha + \beta}\right) Y^{2\alpha + \beta} + \frac{x^{-\alpha} - 1}{\alpha + \beta} Y^\alpha + \frac{1}{2\alpha + \beta} = 0$$

$$y_n(x) = \frac{\sigma^{n+1}}{2i\pi} \oint a_x(w) h_x(w)^{n+1} dw$$

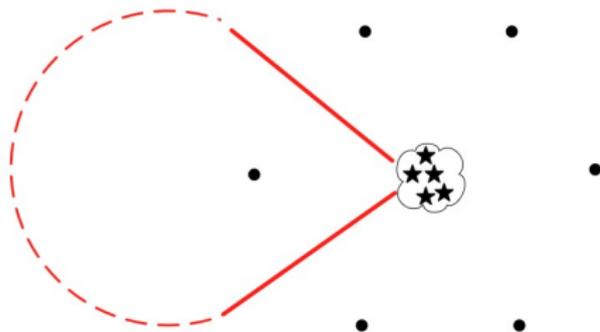
$h_x(w) : \sigma = 2\alpha + \beta$ pôles

Col en **1** avec multiplicité $\alpha + \beta - 1$

On a α autres cols en $1 - (1 - x^{-\alpha})^{1/\alpha}$

$x \sim 1 + O(n^{-1/2})$

$L \sim n^{-\frac{1}{\sigma+1}}$



$$y_n(x) \sim \frac{\sigma^n n^{\frac{1-\sigma}{\sigma}}}{\Gamma(1/\sigma)} \exp\left(\frac{\alpha\sigma}{\alpha + \beta} \sqrt{n\tilde{x}} - \frac{\alpha^3\sigma}{2(\alpha + \beta)^2} \tilde{x}^2\right)$$

$$p_n(x) \sim \exp\left(\frac{\alpha\sigma}{\alpha + \beta} \sqrt{n\tilde{x}} - \frac{\alpha^3\sigma}{2(\alpha + \beta)^2} \tilde{x}^2\right)$$

Jusqu'ici... sur les urnes équilibrées

Modèle d'urne	Objets	Fonction Géné.	Outils	Lois
$\begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ triangulaire	formules booléennes	explicite	formules exactes	Poisson avec vitesse de convergence
$\begin{pmatrix} k-1 & 2 \\ k & 1 \end{pmatrix}$ additive 1 paramètre	k -arbres ord. croissants	implicite	analyse de singularités	lois limite et locale (gauss.), et grandes déviations
$\begin{pmatrix} 2\alpha & \beta \\ \alpha & \alpha+\beta \end{pmatrix}$ additive 2 paramètres	croissance préférentielle	implicite, algébrique	méthode de col, cols coalescents	lois limite et locale (gauss.), et grandes déviations

5. Des perspectives riches

1. Urnes décroissantes (*diminishing*)
2. Urnes non équilibrées
3. Urnes “stochastiques” équilibrées

1. Urnes décroissantes (*diminishing*) équilibrées

$$\mathcal{K} = \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix}, \quad \alpha + \beta = \gamma + \delta \quad \alpha, \beta, \gamma, \delta \geq 0.$$

$$\begin{cases} \dot{X} &= X^{-\alpha+1} Y^{-\beta} \\ \dot{Y} &= X^{-\gamma} Y^{-\delta+1} \end{cases} \quad K = X^{a_0} Y^{b_0}$$

$$X = X(x, y, z) \quad Y = Y(x, y, z)$$

$$\tilde{X} = X(x, y, -z)^{-1} \quad \tilde{Y} = Y(x, y, -z)^{-1}$$

Alors

$$\begin{cases} \dot{\tilde{X}} &= \tilde{X}^{\alpha+1} \tilde{Y}^{\beta} \\ \dot{\tilde{Y}} &= \tilde{X}^{\gamma} \tilde{Y}^{\delta+1} \end{cases} \quad \mathcal{H} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad H = \tilde{X}^{a_0} \tilde{Y}^{b_0}$$

$$\boxed{K(x, y, z) = [x^{\geq 0} y^{\geq 0}] H(x^{-1}, y^{-1}, -z)^{-1}}$$

$$K(x, y, z) = \frac{1}{2i\pi} \oint \oint \frac{H(u^{-1}, v^{-1}, -z)^{-1}}{(x-u)(y-v)} du dv$$

1. Urnes décroissantes (*diminishing*) équilibrées

$$\mathcal{K} = \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix}, \quad \alpha + \beta = \gamma + \delta \quad \alpha, \beta, \gamma, \delta \geq 0.$$

$$\begin{cases} \dot{X} &= X^{-\alpha+1} Y^{-\beta} \\ \dot{Y} &= X^{-\gamma} Y^{-\delta+1} \end{cases} \quad K = X^{a_0} Y^{b_0}$$

$$X = X(x, y, z) \quad Y = Y(x, y, z)$$

$$\tilde{X} = X(x, y, -z)^{-1} \quad \tilde{Y} = Y(x, y, -z)^{-1}$$

Alors

$$\begin{cases} \dot{\tilde{X}} &= \tilde{X}^{\alpha+1} \tilde{Y}^{\beta} \\ \dot{\tilde{Y}} &= \tilde{X}^{\gamma} \tilde{Y}^{\delta+1} \end{cases} \quad \mathcal{H} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad H = \tilde{X}^{a_0} \tilde{Y}^{b_0}$$

$$\boxed{K(x, y, z) = [x^{\geq 0} y^{\geq 0}] H(x^{-1}, y^{-1}, -z)^{-1}}$$

$$K(x, y, z) = \frac{1}{2i\pi} \oint \oint \frac{H(u^{-1}, v^{-1}, -z)^{-1}}{(x-u)(y-v)} du dv$$

To be continued...

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y) = \sum_{a,b} p_{n,a,b} x^a y^b$$

$$x^a y^b \longrightarrow \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta}$$

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \xrightarrow{?} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \stackrel{?}{\Rightarrow} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

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

$$\mathfrak{D} = x^{\alpha+1} y^{\beta} t^{\alpha+\beta} \partial_x + x^{\gamma} y^{\delta+1} t^{\gamma+\delta} \partial_y$$

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \xrightarrow{\mathfrak{D} \circ \mathfrak{J}} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

$$\mathfrak{J}[x^a y^b t^{a+b}] = \int_0^t x^a y^b w^{a+b} \frac{dw}{w} = x^a y^b \frac{t^{a+b}}{a+b}$$

$$\mathfrak{D} = x^{\alpha+1} y^{\beta} t^{\alpha+\beta} \partial_x + x^{\gamma} y^{\delta+1} t^{\gamma+\delta} \partial_y$$

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \xrightarrow{\mathfrak{D} \circ \mathfrak{I}} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

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

$$\mathfrak{D} = x^{\alpha+1} y^{\beta} t^{\alpha+\beta} \partial_x + x^{\gamma} y^{\delta+1} t^{\gamma+\delta} \partial_y$$

$$\phi_{n+1} = \mathfrak{D} \circ \mathfrak{I}(\phi_n)$$

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \xrightarrow{\mathfrak{D} \circ \mathfrak{I}} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

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

$$\mathfrak{D} = x^{\alpha+1} y^{\beta} t^{\alpha+\beta} \partial_x + x^{\gamma} y^{\delta+1} t^{\gamma+\delta} \partial_y$$

$$\phi_{n+1} = \mathfrak{D} \circ \mathfrak{I}(\phi_n)$$

On pose $\psi_n = \mathfrak{I}(\phi_n)$. i.e. $\psi_n = \sum_{a,b} p_{n,a,b} x^a y^b \frac{t^{a+b}}{a+b}$

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

2. Urnes non équilibrées

Le système différentiel ne tient plus...

$$\phi_n(x, y, t) = \sum_{a,b} p_{n,a,b} x^a y^b t^{a+b}$$

$$x^a y^b t^{a+b} \xrightarrow{\mathfrak{D} \circ \mathfrak{I}} \frac{a}{a+b} x^{a+\alpha} y^{b+\beta} t^{a+b+\alpha+\beta} + \frac{b}{a+b} x^{a+\gamma} y^{b+\delta} t^{a+b+\gamma+\delta}$$

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

$$\mathfrak{D} = x^{\alpha+1} y^{\beta} t^{\alpha+\beta} \partial_x + x^{\gamma} y^{\delta+1} t^{\gamma+\delta} \partial_y$$

$$\phi_{n+1} = \mathfrak{D} \circ \mathfrak{I}(\phi_n)$$

On pose $\psi_n = \mathfrak{I}(\phi_n)$. i.e. $\psi_n = \sum_{a,b} p_{n,a,b} x^a y^b \frac{t^{a+b}}{a+b}$

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

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

$$\boxed{[(x - z x^{\alpha+1} y^{\beta}) \partial_x + (y - z x^{\gamma} y^{\delta+1}) \partial_y] (\Psi(x, y, z)) = x^{a_0} y^{b_0}}$$

3. Urnes "stochastiques" équilibrées

$$\begin{pmatrix} 1 - \mathcal{B} & \mathcal{B} \\ \mathcal{B} & 1 - \mathcal{B} \end{pmatrix}, \text{ avec } \mathcal{B} \sim \text{Bern}(p)$$

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

Alors $H(x, y, z) = X(x, y, z)^{a_0} Y(x, y, z)^{b_0}$ où (X, Y) est la solution de

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

La probabilité d'avoir a boules x et b boules y après n tirages est donnée par :

$$p_{n,a,b} = \frac{n! [x^a y^b z^n] H(x, y, z)}{n! [z^n] H(1, 1, z)}.$$

Ce principe reste valable pour une urne équilibrée $\begin{pmatrix} \mathcal{A} & \sigma - \mathcal{A} \\ \sigma - \mathcal{B} & \mathcal{B} \end{pmatrix}$, avec σ constant, et \mathcal{A}, \mathcal{B} variables aléatoires à valeurs entières sur un espace d'états finis.