Pólya Urns
An analytic combinatorics approach

Basile Morcrette

Algorithms project, INRIA Rocquencourt.
LIP6, UPMC

CALIN Seminar
07/02/2012

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE | INRIA

UPMC SORBONNE UNIVERSITÉS

Outline

1. Urn model
2. An exact approach

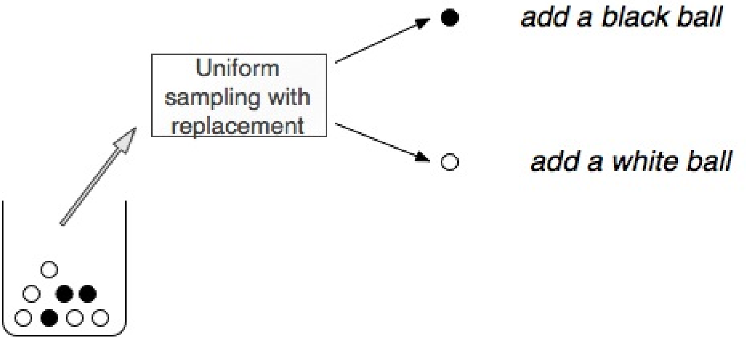
   boolean formulas
3. Singularity analysis

   family of k-trees
4. Saddle-point method

   preferential growth models
5. Towards other urn models

   unbalanced, with random entries

1. Urns models



- an urn containing balls of two colours
- rules for urn evolution

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

Balanced Pólya urns

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

Balanced urn : $\boxed{\alpha + \beta = \gamma + \delta}$ (deterministic total number of balls)

A given initial configuration (a_0, b_0) :

a_0 balls $\bullet$ (counted by x)

b_0 balls $\circ$ (counted by y)

Definition

History of length n : a sequence of n evolutions (n rules, n drawings)

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

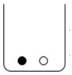
$H_{n,a,b}$: number of histories of length n, beginning in the configuration (a_0, b_0), and ending in (a, b)

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.

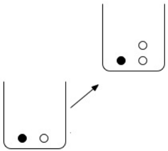
$$H(x, y, z) =$$

$$xy$$



Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.

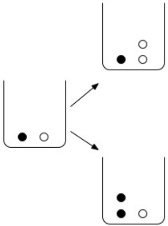


$H(x, y, z) =$

xy

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



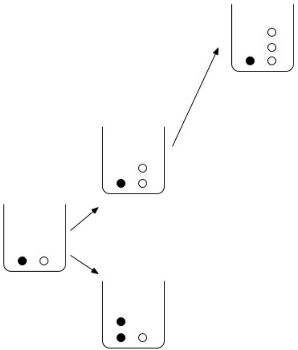
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



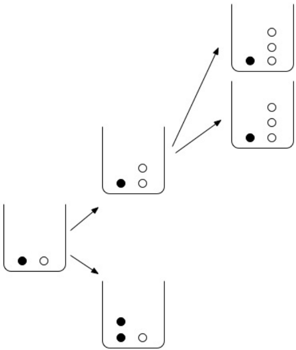
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



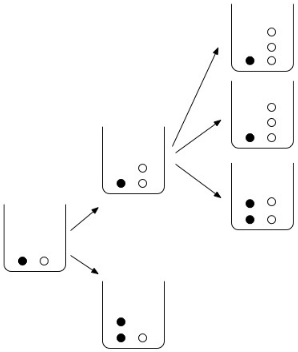
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



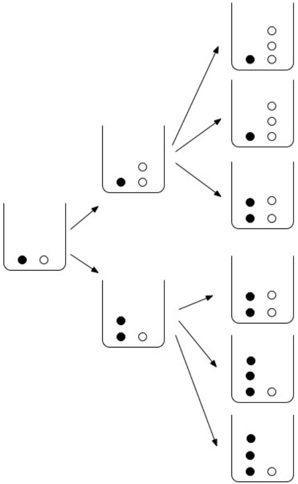
$$H(x, y, z) =$$

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



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

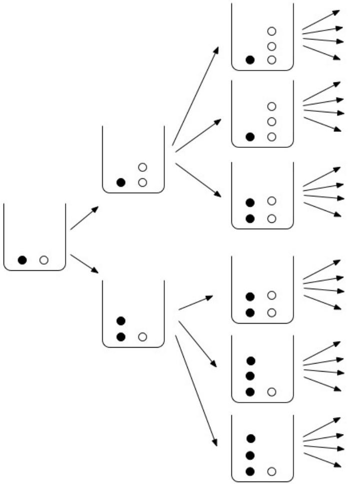
$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

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

Combinatorics of histories - Example

We consider this urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with $(a_0, b_0) = (1, 1)$.



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

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

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

$$+ \quad \dots$$

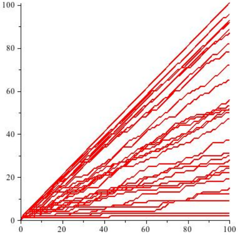
Various behaviours

Problem : Understand the urn composition after n steps, and asymptotically when n tends to ∞.

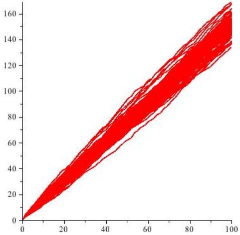
Various behaviours

Problem : Understand the urn composition after n steps, and asymptotically when n tends to ∞.

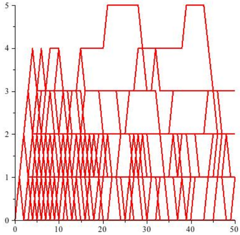
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$



Pólya urn



Preferential growth urn



Triangular 3×3 urn

Probabilistic results

Urn $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ Ratio $\rho = \dfrac{\alpha - \gamma}{\alpha + \beta}$

- Small urns : $\rho \leqslant \frac{1}{2}$

  Gaussian limit law [Smythe96] [Janson04]

- Large urns : $\rho > \frac{1}{2}$

  Non gaussian laws [Mahmoud] [Janson04] [Chauvin–Pouyanne–Sahnoun11]

Tools :

- embedding in continuous time [Jan04] [ChPoSa11]
- martingales, central limit theorem

Balanced urns and analysis

- First steps : [Flajolet–Gabarro–Pekari05], *Analytic urns*
- [Flajolet–Dumas–Puyhaubert06], on urns with negative coefficients, and triangular cases
- [Kuba–Panholzer–Hwang07], unbalanced urns

Analytic approach : theorem [FlDuPu06]

$$\text{Urn } \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \text{ and } \left\{ \begin{array}{c} (a_0, b_0) \\ \alpha+\beta=\gamma+\delta \end{array} \right. \Longrightarrow \quad \begin{array}{c} H = X^{a_0}\, Y^{b_0} \\ \text{with } \left\{ \begin{array}{rcl} \dot{X} & = & X^{\alpha+1}\, Y^{\beta} \\ \dot{Y} & = & X^{\gamma}\, Y^{\delta+1} \end{array} \right. \end{array}$$

Isomorphism proof

Differenciate = Pick

$$\partial_x[xx\dots x] = (\not{x}x\dots x) + (x\not{x}\dots x) + \dots + (xx\dots\not{x})$$

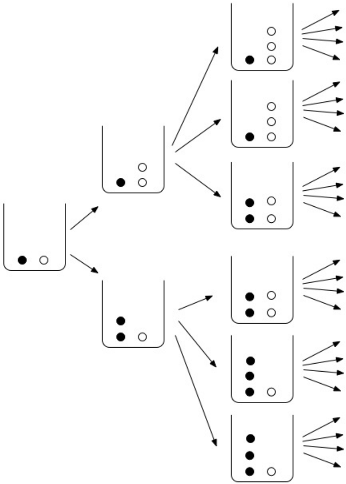
$$x\partial_x[xx\dots x] = (\underline{x}x\dots x) + (x\underline{x}\dots x) + \dots + (xx\dots\underline{x})$$

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

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

Counting histories - Example

Take the urn $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $(a_0, b_0) = (1, 1)$.



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

$$xy$$

$$+ \quad (xy^2 + x^2y)z$$

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

$$+ \quad \dots$$

Isomorphism proof

Differenciate = Pick

$$\partial_x[xx\dots x] = (\not{x}x\dots x) + (x\not{x}\dots x) + \dots + (xx\dots\not{x})$$

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

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

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

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

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

Isomorphism proof

Differenciate = Pick

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

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

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

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

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

Let $(X(t), Y(t))$ be solution of

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

$$\begin{aligned} &\partial_t(X^a Y^b) \\ = \; & aX^{a-1}\dot{X}Y^b + bX^aY^{b-1}\dot{Y} \\ = \; & \boxed{aX^{a+\alpha}Y^{b+\beta} + bX^{a+\gamma}Y^{b+\delta}} \end{aligned}$$

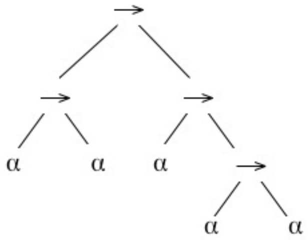
$$\partial_t^n(X^aY^b) = \mathfrak{D}^n\left[x^a y^b\right] \quad \begin{matrix} x \to X \\ y \to Y \end{matrix}$$

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

Then $t=0$, and it's over !!

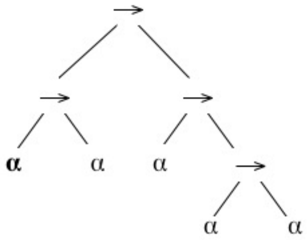
2. Urns and random trees

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler11]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



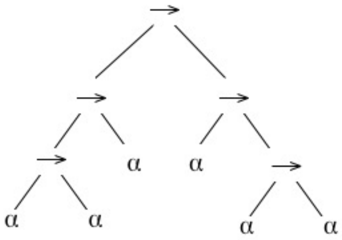
2. Urns and random trees

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler11]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



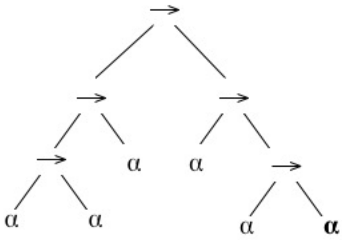
2. Urns and random trees

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler11]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves



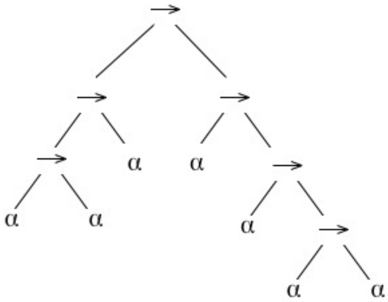
2. Urns and random trees

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler11]
- Probabilistic model : uniform growth in leaves (BST model)
  - choose randomly a leave
  - replace it by a binary node and two leaves

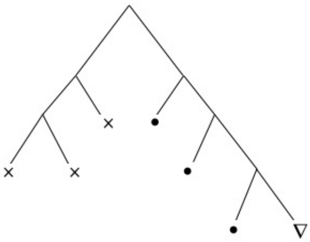


2. Urns and random trees

- Motivation : quantify the fraction of tautologies among all logic formulas having only one logic operator : implication. [Mailler11]
- 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 :

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

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 $O\left(\frac{(2k)^n}{n!}\right)$.

3. An urn for k-trees

Motivation : model of graphs [Panholzer–Seitz 2010]

Definition

A k-tree T is

- either a k-clique
- or there exists a vertex f with a k-clique as neighbor and $T \backslash f$ is a k-tree

Ordered : distinguishable children.
Increasing : vertices labelled in apparition order

3. An urn for k-trees

Motivation : model of graphs [Panholzer–Seitz 2010]

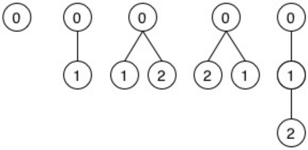
Definition

A k-tree T is

- either a k-clique
- or there exists a vertex f with a k-clique as neighbor and $T \backslash f$ is a k-tree

Ordered : distinguishable children.
Increasing : vertices labelled in apparition order

increasing ordered 1-tree (or PORT)



3. An urn for k-trees

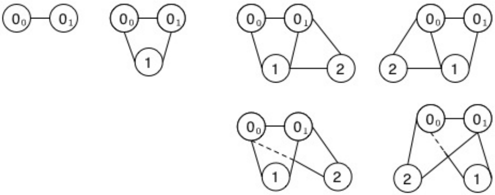
Motivation : model of graphs [Panholzer–Seitz 2010]

Definition

A k-tree T is

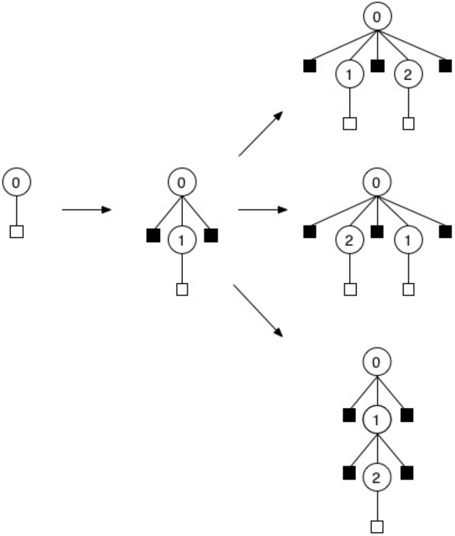
- either a k-clique
- or there exists a vertex f with a k-clique as neighbor and $T \backslash f$ is a k-tree

Ordered : distinguishable children.
Increasing : vertices labelled in apparition order

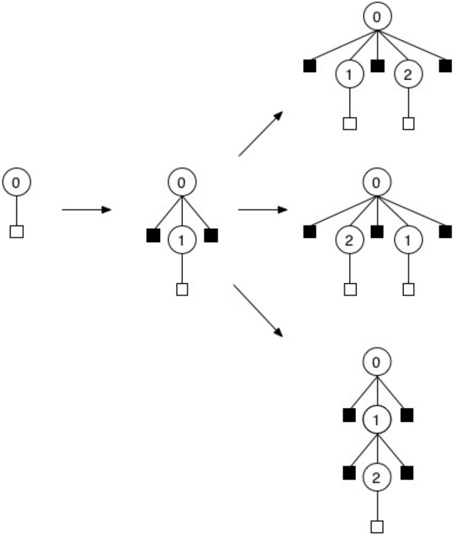


increasing ordered 2-tree

Urn Model



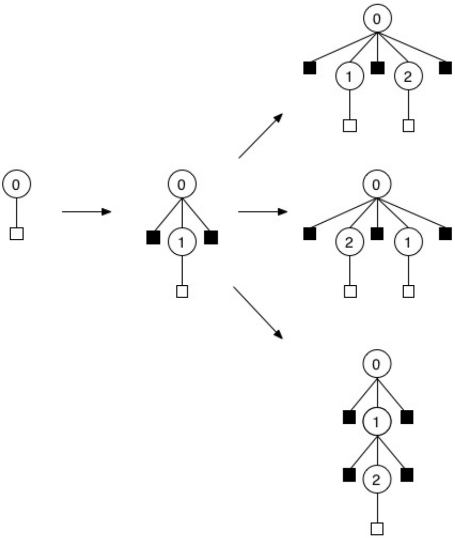
Urn Model



For 1-trees

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

Urn Model



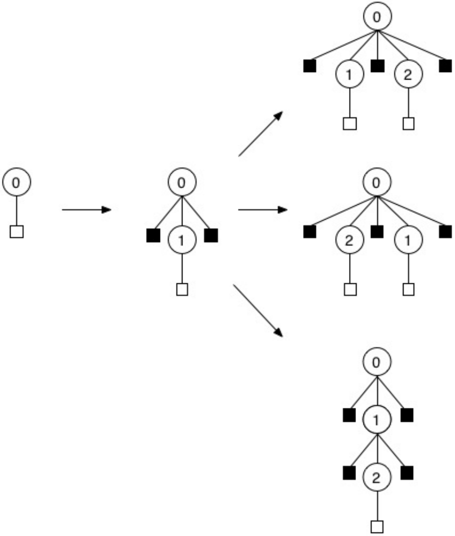
For 1-trees

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

For k-trees

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

Urn Model



For 1-trees

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

For k-trees

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

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

Singularity analysis

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

Ingredients :

- partial fraction expansion
- integration
- variable substitution

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

some analysis...

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

Normal limit law

If X_n counts the number of $\square$ balls in the urn U_1 after n draws.
Case $k = 1$: number of leaves in a PORT of size n.

Theorem

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

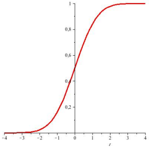
Local limit law

Theorem

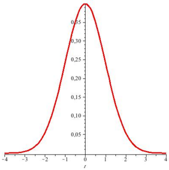
Let $p_{n,k} = \mathbb{P}\{X_n = k\}$. The X_n distribution satisfies a gaussienne local limit law with rate of convergence $O\left(\frac{1}{\sqrt{n}}\right)$, i.e.

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

$$\mathbb{P}\left\{ \frac{X_n - \frac{2}{3}n}{\sqrt{\frac{n}{9}}} \leqslant t \right\} \underset{n\to\infty}{\longrightarrow}$$



$$\frac{\sqrt{3n}}{2}\, \mathbb{P}\left\{ X_n = \left\lfloor \frac{2n}{3} + t\frac{\sqrt{n}}{3} \right\rfloor \right\} \underset{n\to\infty}{\longrightarrow}$$

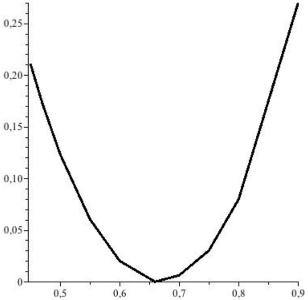


Large deviations

- exponentially small bound on the devation from the mean : quantify rare events.

Theorem

- if $0.42 < t < 2/3$, $\mathbb{P}(X_n \leqslant tn) \approx e^{-nW(t)}$ (left tail)
- if $2/3 < t < 0.73$, $\mathbb{P}(X_n \geqslant tn) \approx e^{-nW(t)}$ (right tail)



4. Preferential growth urns

Motivation : characterization of additive 2×2 urns (positive coefficients).
Approach : finding a class of urns with "nice" generating functions.

Theorem [M12]

The balanced urns class $\begin{pmatrix} 2\alpha & \beta \\ \alpha & \alpha+\beta \end{pmatrix}$, with $\alpha > 0$, $\beta \geqslant 0$, has an algebraic bivariate generating function.

The histories GF $H(x, 1, z)$ cancels the following polynomial in Y

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

$$\text{with } b(x) = \frac{x^{-\alpha} - 1}{\alpha + \beta} \text{ and } a = (2\alpha + \beta)^{-1}.$$

Proof

Differential system : $\left\{ \begin{array}{rcl} \dot{X} & = & X^{2\alpha+1}\, Y^{\beta} \\ \dot{Y} & = & X^{\alpha}\, Y^{\alpha+\beta+1} \end{array} \right.$ $\qquad \dot{X} = \frac{\partial}{\partial z} X$

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

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

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

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

Balanced urn $a + b = a_0 + b_0 + n\sigma$. We set $y = 1$.

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

First observations

The balance $\sigma = 2\alpha + \beta$ The ratio $\rho = \frac{\alpha}{2\alpha+\beta} \leqslant \frac{1}{2}$

For $x = 1$, equation becomes : $(z - \sigma^{-1})Y^{\sigma} + \sigma^{-1} = 0$

Thus, for $(a_0, b_0) = (0, 1)$

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

Proposition

Let X_n be the random variable counting the number of x-colored balls in the urn after n steps. Then,

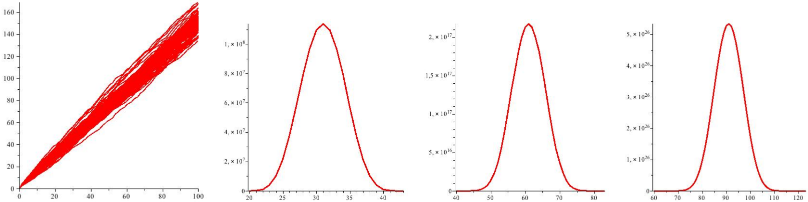
$$\mathbb{E}(X_n) = \frac{\alpha(2\alpha+\beta)}{\alpha+\beta} n + \frac{\alpha}{\alpha+\beta} \frac{\Gamma(\frac{1}{2\alpha+\beta})}{\Gamma(\frac{\alpha+1}{2\alpha+\beta})} n^{\frac{\alpha}{2\alpha+\beta}} + \frac{\alpha}{\alpha+\beta} + O\left(n^{\frac{\alpha}{2\alpha+\beta}-1}\right) ,$$

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

Example $\alpha = 1$, $\beta = 1$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad \begin{array}{lcl} \mathtt{x} & \rightarrow & \mathtt{x\,x\,y} \\ \mathtt{y} & \rightarrow & \mathtt{x\,y\,y} \end{array}$$

Preferential growth



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

Saddle-point method for x=1

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

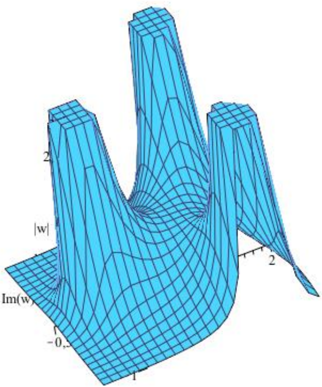
$$y_n = \frac{1}{2i\pi} \oint \frac{Y(z)}{z^{n+1}} \mathrm{d}z$$

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

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

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

integrate with a right contour...



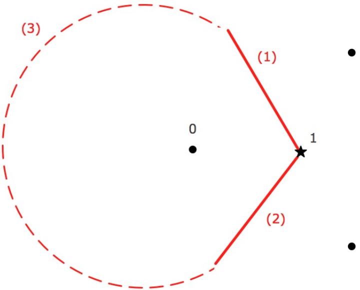
$w \mapsto |h(w)|$

3 poles

1 double saddle-point in $w = 1$

Saddle-point method for x=1 (next)

$t \in [0..L]$
(1) $w(t) = 1 + te^{i2\pi/3}$
(2) $w(t) = 1 + te^{-i2\pi/3}$



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

Choose L... $nL^3 \to \infty$ and $nL^6 \to 0$
We set $L \sim n^{-1/4}$

$$\int_{(1)} + \int_{(2)} : \int_0^\infty ue^{-u^3} du \text{ and } \int_{(3)} \text{exponentially small}$$

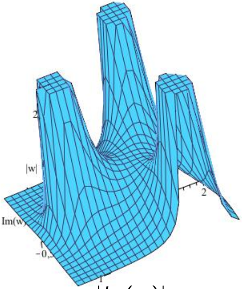
$$\boxed{y_n = \frac{3^n}{\Gamma(1/3)}\left(n^{-2/3} + O\left(n^{-11/12}\right)\right)}$$

Saddle-point method for $x \neq 1$

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

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

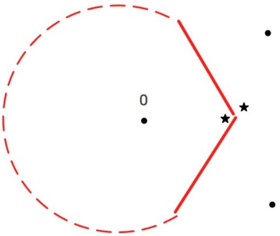
$$h_x'(1) = h'(x^{-1}) = 0$$



$w \mapsto |h_x(w)|$

3 poles

2 saddle-points in $w = 1$ and $w = x^{-1}$



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

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

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

Gaussian limit law

Let X_n be the random variable couting the number of • balls in the urn after n steps.

Theorem

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

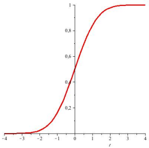
Local limit law

Theorem

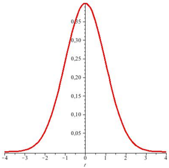
We set $p_{n,k} = \mathbb{P}\{X_n = k\}$. The X_n distribution satisfies a local limit law of gaussian type with speed of convergence $O\left(\frac{1}{\sqrt{n}}\right)$, i.e.

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

$$\mathbb{P}\left\{ \frac{X_n - \frac{3}{2}n}{\sqrt{\frac{3n}{4}}} \leqslant t \right\} \underset{n\to\infty}{\longrightarrow}$$



$$\frac{\sqrt{3n}}{2}\,\mathbb{P}\left\{ X_n = \left\lfloor \frac{3n}{2} + t\frac{\sqrt{3n}}{2} \right\rfloor \right\} \underset{n\to\infty}{\longrightarrow}$$

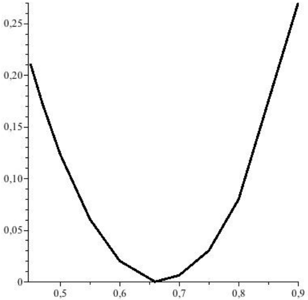


Large deviations

- Exponentially small bound on the large deviation with regards to the mean : quantification on rare events

Theorem

- si $0.42 < t < 2/3$, $\mathbb{P}(X_n \leqslant tn) \approx e^{-nW(t)}$ (left tail)
- si $2/3 < t < 0.73$, $\mathbb{P}(X_n \geqslant tn) \approx e^{-nW(t)}$ (right tail)



General case

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

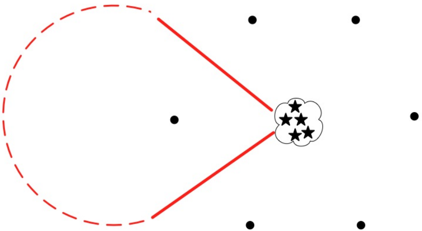
$$y_n(x) = \frac{\sigma^{n+1}}{2i\pi} \oint a_x(w) h_x(w)^{n+1} \mathrm{d}w$$

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

Saddle-point in 1
with multiplicity $\alpha + \beta - 1$

The other α saddle-points
in $1 - (1 - x^{-\alpha})^{1/\alpha}$

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



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

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

Until now... on balanced urns

Urn Model	Objects	Gen. Fun.	Tools	Laws
$\begin{pmatrix} 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ triangular	boolean formulas	explicit	exact formulas	Poisson Law with rate of convergence
$\begin{pmatrix} k-1 & 2 \\ k & 1 \end{pmatrix}$ additive 1 parameter	increasing ord. k-trees	implicit	singularity analysis	limit and local (gauss.) laws, and large deviations
$\begin{pmatrix} 2\alpha & \beta \\ \alpha & \alpha+\beta \end{pmatrix}$ additive 2 parameter	preferential growth	implicit, algebraic	(coalescing) saddle-point method	limit and local (gauss.) laws, and large deviations

a generic approach for all algebraic balanced additive urn models (Guess'N'Prove from A. Bostan)

5. What's next?

1. *Diminishing* urns
2. Unbalanced urns
3. Balanced urns with random entries

1. *Diminishing* balanced urns

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

$$\left\{ \begin{array}{rcl} \dot{X} & = & X^{-\alpha+1}\, Y^{-\beta} \\ \dot{Y} & = & X^{-\gamma}\, Y^{-\delta+1} \end{array} \right. \qquad K = X^{a_0} Y^{b_0}$$

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

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

Then

$$\left\{ \begin{array}{rcl} \dot{\tilde{X}} & = & \tilde{X}^{\alpha+1}\, Y^{\beta} \\ \dot{\tilde{Y}} & = & \tilde{X}^{\gamma}\, Y^{\delta+1} \end{array} \right. \qquad \mathcal{H} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad H = \tilde{X}^{a_0} \tilde{Y}^{b_0}$$

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

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

1. *Diminishing* balanced urns

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

$$\left\{ \begin{array}{lcl} \dot{X} & = & X^{-\alpha+1}\, Y^{-\beta} \\ \dot{Y} & = & X^{-\gamma}\, Y^{-\delta+1} \end{array} \right. \qquad K = X^{a_0} Y^{b_0}$$

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

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

Then

$$\left\{ \begin{array}{lcl} \dot{\tilde{X}} & = & \tilde{X}^{\alpha+1}\, Y^{\beta} \\ \dot{\tilde{Y}} & = & \tilde{X}^{\gamma}\, Y^{\delta+1} \end{array} \right. \qquad \mathcal{H} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad H = \tilde{X}^{a_0} \tilde{Y}^{b_0}$$

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

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

To be continued...

2. Unbalanced urns

The differential system does not hold anymore...

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

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

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

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

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

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

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

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

2. Unbalanced urns

The differential system does not hold anymore...

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

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

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

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

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

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

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

Finally $t\partial_t = x\partial_x + y\partial_y$, thus $\Psi = \sum_n \psi_n z^n$ verifies

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

3. Balanced urns with random entries

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

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

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

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

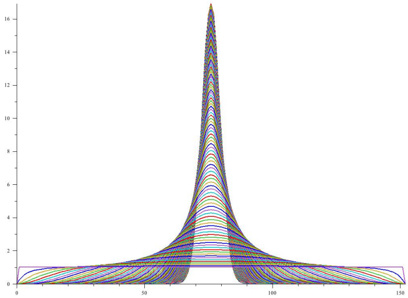
Probability to have a black balls and b white balls after n draws:

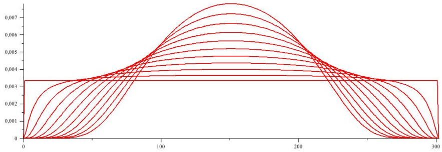
$$p_{n,a,b} = \frac{[x^a y^b z^n]H(x,y,z)}{[z^n]H(1,1,z)}.$$

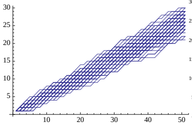
True for any balanced urn $\left(\begin{smallmatrix} \mathcal{A} & \sigma-\mathcal{A} \\ \sigma-\mathcal{B} & \mathcal{B} \end{smallmatrix}\right)$, with σ constant, and $\mathcal{A}, \mathcal{B}$ random variables on a finite state space $\{-1,\, 0,\, 1, \ldots, \sigma\}$.
[M., Mahmoud, 2012]

3. Balanced urns with random entries

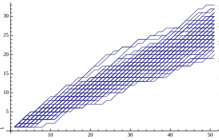
$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ with proba p $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ with proba $1-p$



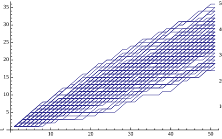




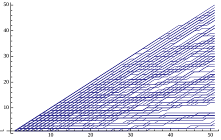
$p = 1$



$p = 0.6$



$p = 0.3$



$p = 0$