

Analytic urns of triangular form

Vincent Puyhaubert

September, 22 2003

General problem

- One urn containing balls of one of d colors
- Composition at $t = 0$ fixed
- At time n , a ball is randomly chosen in the urn, its color is inspected and the ball is placed back into the urn with some new balls. The number of added balls of each color only depends on the color of the drawn one.

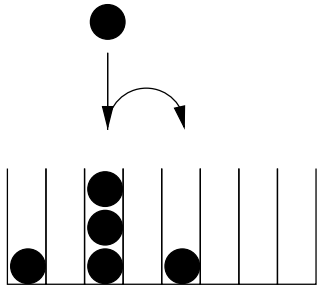
Drawn	Added			
\downarrow	c_1	c_2	\dots	c_d
c_1	$\delta_{1,1}$	\dots	\dots	$\delta_{1,d}$
\vdots	\vdots			\vdots
\vdots	\vdots			\vdots
c_d	$\delta_{d,1}$	\dots	\dots	$\delta_{d,d}$

Examples

Classical examples with two balls :

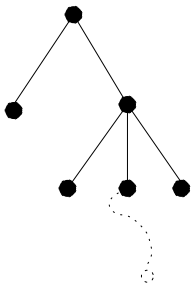
$$\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \quad \begin{bmatrix} 0 & s \\ s & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Hashing problems :



$$\begin{bmatrix} -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & & 0 \end{bmatrix}$$

Random generation of n -trees :



$$\begin{bmatrix} 0 & 1 & & & \\ 1 & -1 & 1 & & \\ \vdots & & \ddots & \ddots & \\ 1 & & & -1 & 1 \\ 1 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Probabilistic approach

$$\text{Matrix} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}$$

Introduce X_n , the r.v. giving the number of balls of type I . At time n , there are exactly $t_0 + 2n$ balls in the urn.

$$\underbrace{\Pr(X_{n+1} = k)}_{p_{n+1,k}} = \underbrace{\Pr(X_n = k)}_{p_{n,k}} \cdot \frac{k}{t_0 + 2n} + \underbrace{\Pr(X_n = k - 1)}_{p_{n,k-1}} \cdot \left(1 - \frac{k - 1}{t_0 + 2n}\right) \quad (1)$$

Set $p_n(u) = \sum_k p_{n,k} u^k$ and $F(z, u) = \sum_n p_n(u) z^n$.

Eq (1) $\xrightarrow{\text{Translate}}$ PDE satisfied by F

Enumerative approach (balanced urns only!)

Idea : Enumerate all possible “histories” starting from the initial urn

$$M = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \quad U_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Time 0	Draw	Time 1	Draw	Time 2
B : 1 W : 2	1	B : 1 3 4 W : 2	1	B : 1 3 4 5 6 W : 2
			2	B : 1 3 4 6 W : 2 5
			3	B : 1 3 4 5 6 W : 2
			4	B : 1 3 4 5 6 W : 2
	2	B : 1 4 W : 2 3	1	B : 1 4 5 6 W : 2 3
			2	B : 1 4 6 W : 2 3 5
			3	B : 1 4 6 W : 2 3 5
			4	B : 1 4 5 6 W : 2 3

Represent urn by monomials. We have

$$f_0(u, v) = uv \quad f_1(u, v) = u^3v + u^2v^2 \quad f_2(u, v) = 3u^5v + 3u^4v^2 + 2u^3v^3$$

and each history of length n has the same probability $2 \times 4 \times \cdots \times 2n$.

Starting from one urn with (r, s) balls, the operator

$$u^r v^s \longrightarrow r u^{r+2} v^s + s u^{r+1} v^{s+1} = \left(u^2 \cdot u \frac{\partial}{\partial u} + uv \cdot v \frac{\partial}{\partial v} \right) u^r v^s$$

describes all possible events that can occur from that urn. By linearity

$$f_n = \Gamma f_{n-1} = \Gamma^n uv \quad \text{where} \quad \Gamma = u^2 \cdot u \frac{\partial}{\partial u} + uv \cdot v \frac{\partial}{\partial v}$$

Introduce $H = \sum f_n(u, v) \frac{z^n}{n!}$. H satisfies linear PDE

$$\frac{\partial H}{\partial z} = \Gamma H \quad H(0, u, v) = uv$$

General case :

$$\frac{\partial H}{\partial z} = x_1^{\delta_{1,1}} \cdots x_d^{\delta_{1,d}} x_1 \frac{\partial H}{\partial x_1} + \cdots + x_1^{\delta_{d,1}} \cdots x_d^{\delta_{d,d}} x_d \frac{\partial H}{\partial x_d}$$

$$H(0, x_1, \dots, x_d) = x_1^{a_1} \cdots x_d^{a_d}$$

Relations between the two methods

Let ℓ be the constant sum on the lines and t_0 the initial number of balls. With the previous notations, one has

$$p_n = \frac{1}{t_0(t_0+\ell)\cdots(t_0+\ell(n-1))} f_n$$

Recall that $F = \sum p_n z^n$ and $H = \sum f_n \frac{z^n}{n!}$. Let \mathcal{L}_α defined for any $\alpha > 0$ as

$$\mathcal{L}_\alpha : g \mapsto \frac{1}{\Gamma(\alpha)} \int_0^{+\infty} t^{\alpha-1} g(st) e^{-t} dt$$

Then $\mathcal{L}_\alpha(\sum a_n z^n) = \sum a_n \alpha(\alpha+1)\cdots(\alpha+n-1) s^n$. (Note : \mathcal{L}_1 is the classic Laplace transform). Then set $\alpha = t_0/\ell$. F and H satisfy

$$\mathcal{L}_1(H)(s) = \mathcal{L}_\alpha(F)(s\ell)$$

Solving linear PDE

Pb : find all functions that satisfy

$$(2) \quad P_1(x_1, \dots, x_d) \frac{\partial H}{\partial x_1} + \dots + P_d(x_1, \dots, x_d) \frac{\partial H}{\partial x_d} = 0$$

Consider the system :

$$(3) \quad x'_i(t) = P_i(x_1, \dots, x_d) \quad (1 \leq i \leq d)$$

Then H is a solution iff it is constant on each integral curve of (3).

Pb : If H_1, \dots, H_l are constants, then $\phi(H_1, \dots, H_l)$ is also constant for any ϕ . How do we get them all?

Theorem 1 *There exists $d - 1$ independent functions $\psi_1, \dots, \psi_{d-1}$ such that the solutions of (2) are given by $\phi(\psi_1, \dots, \psi_{d-1})$ for an arbitrary function ϕ .*

Independence means here that the dimension of $(\overrightarrow{\text{grad}}(\psi_i))_{1 \leq i \leq d-1}$ is $d - 1$ in at least one point. This implies that there is no relation $\psi_i = \varphi(\psi_1, \dots, \psi_{i-1}, \psi_{i+1}, \dots, \psi_{d-1})$.

Finding the first integrals of a system of differential equations

Usual notations

$$\frac{dx_1}{P_1} = \dots = \frac{dx_d}{P_d} = dt$$

No general method!!

“Solvable” cases : Suppose P_1 is a function of the only variable x_1 , P_2 is a function of x_1 and x_2 , etc ...

- Solve system $\frac{dx_1}{P_1(x_1)} = \frac{dx_2}{P_2(x_1, x_2)}$. This gives a relation of the form $G_1(x_1, x_2) = C_1$ where G_1 is a first integral.
- Invert this relation to get $x_1 = F_1(x_2, C_1)$. Use this relation to solve $\frac{dx_2}{P_2(x_1, x_2)} = \frac{dx_3}{P_3(x_1, x_2, x_3)}$.
- Repeat the process till the last equation ...

For triangular urns, our PDE is of this type. What a coincidence!!!!

Do it yourself

$$M = \begin{bmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,d} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \delta_{d-1,d} \\ 0 & \cdots & 0 & \delta_{d,d} \end{bmatrix} \quad U_0 = \begin{bmatrix} a_1 \\ \vdots \\ \vdots \\ a_d \end{bmatrix}$$

Step (1) : Find the PDE :

$$\frac{\partial H}{\partial z} = x_1^{\delta_{1,1}} \cdots x_d^{\delta_{1,d}} \cdot x_1 \frac{\partial H}{\partial x_1} + \cdots + x_d^{\delta_{d,d}} \cdot x_d \frac{\partial H}{\partial x_d}$$

Step (2) : Solve the system of d differential equations

$$-dz = \frac{dx_d}{x_d^*} = \frac{dx_{d-1}}{x_{d-1}^* x_d^*} = \frac{dx_{d-2}}{x_{d-2}^* x_{d-1}^* x_d^*} \cdots$$

The general solution of the PDE is given by

$$H = \phi(c_1, c_2, \dots, c_d)$$

for any arbitrary function ϕ .

Step (3) : Set $z = 0$. One has $H[z \leftarrow 0] = x_1^{a_1} \cdots x_d^{a_d}$. Solve the system

- $x_d = f_d(c_d[z \leftarrow 0])$
- $x_{d-1} = f_{d-1}(c_{d-1}[z \leftarrow 0], c_d[z \leftarrow 0])$
- ...

The generating function is then given by

$$H(x_1, \dots, x_d, z) = [f_1(c_1, \dots, c_d)]^{a_1} \cdots [f_d(c_d)]^{a_d}$$

Triangular urns of size 2

We consider the matrix $\begin{pmatrix} a & 0 \\ a-b & b \end{pmatrix}$, with initial conditions (a_0, b_0) (also introduce $t_0 = a_0 + b_0$). Following the previous methodology, one has

$$H(z, u, v) = u^{a_0} v^{b_0} (1 - azu^a)^{-a_0/a} \left(1 - v^{b_0} (1 - (1 - azu^a)^{b/a}) \right)^{-b_0/b}$$

Let $\Delta_n = \frac{n!}{t_0 \cdots (t_0 + (n-1)a)}$. Let X_n be the random variable giving the number of balls of the second type. As a classical application of bivariate generating functions, one has

$$E[X_n(X_n - 1) \cdots (X_n - \ell + 1)] = \Delta_n [z^n] \frac{\partial^\ell H}{\partial v^\ell} \Big|_{u,v=1}$$

By a small recurrence, we find that

$$E[(X_n)^\ell] = c_{\ell,0} (1 - az)^{-(t_0 + a\ell) \cdot b/a} + c_{\ell,1} (1 - az)^{-(t_0 + a(\ell-1)) \cdot b/a} + \dots$$

with $c_{\ell,0} = b_0 \cdots (b_0 + (\ell - 1)b)$.

Since for $\alpha < \beta$ one has $[z^n](1-z)^{-\alpha} = o([z^n](1-z)^{-\beta})$, only the first term gives the asymptotics and

$$E[(X_n)^\ell] = \frac{1}{b^\ell} \frac{\Gamma\left(\frac{b_0 + \ell b}{b}\right) \Gamma\left(\frac{t_0}{a}\right)}{\Gamma\left(\frac{b_0}{b}\right) \Gamma\left(\frac{t_0 + \ell b}{a}\right)} n^{\ell b/a} + O(n^{(\ell-1)b/a}).$$

\implies Standard deviation and expectation of the same order of growth.

Non gaussian law!

Limit laws for the composition of the urn

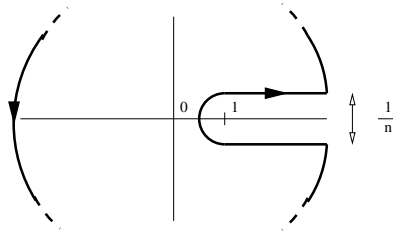
Let Z_n be the r.v. giving the number of times a ball of type II has been drawn ($X_n = b_0 + bZ_n$). Then

$$\Pr(Z_n = k) = a^n \Delta_n[v^k z^n](1-z)^{-a_0/a} \left(1 - v(1 - (1-z)^{b/a})\right)^{-\frac{b_0}{b}}$$

Extending in v , then using Cauchy's formula yields

$$\Pr(Z_n = k) = C_{n,k} \oint (1-z)^{-a_0/a} \left(1 - (1-z)^{b/a}\right)^k \frac{dz}{z^{n+1}}$$

The contour is extended as follows



From now, set $k = \lfloor xn^{b/a} \rfloor$ and $z = 1 - t/n$:

$$\Pr(Z_n = \lfloor xn^{b/a} \rfloor) \sim C_{n,x} n^* \int_{\mathcal{H}} t^{-a_0/a} e^{-x t^{b/a}} e^t dt$$

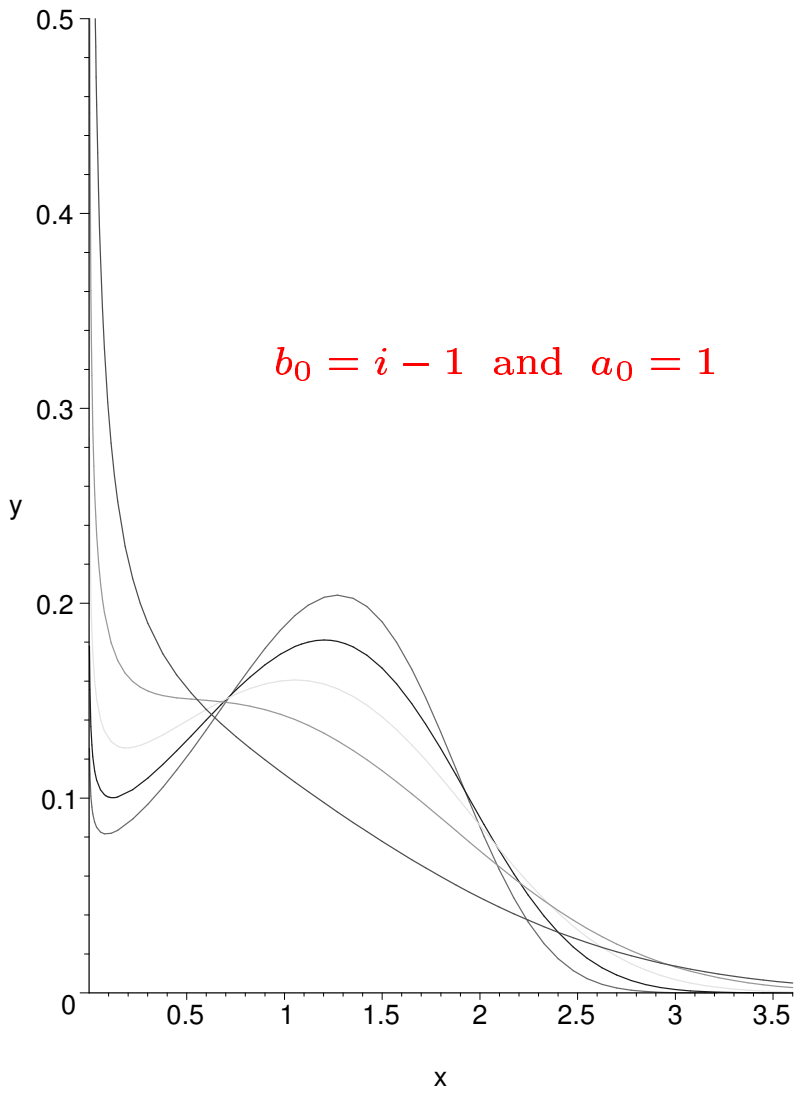
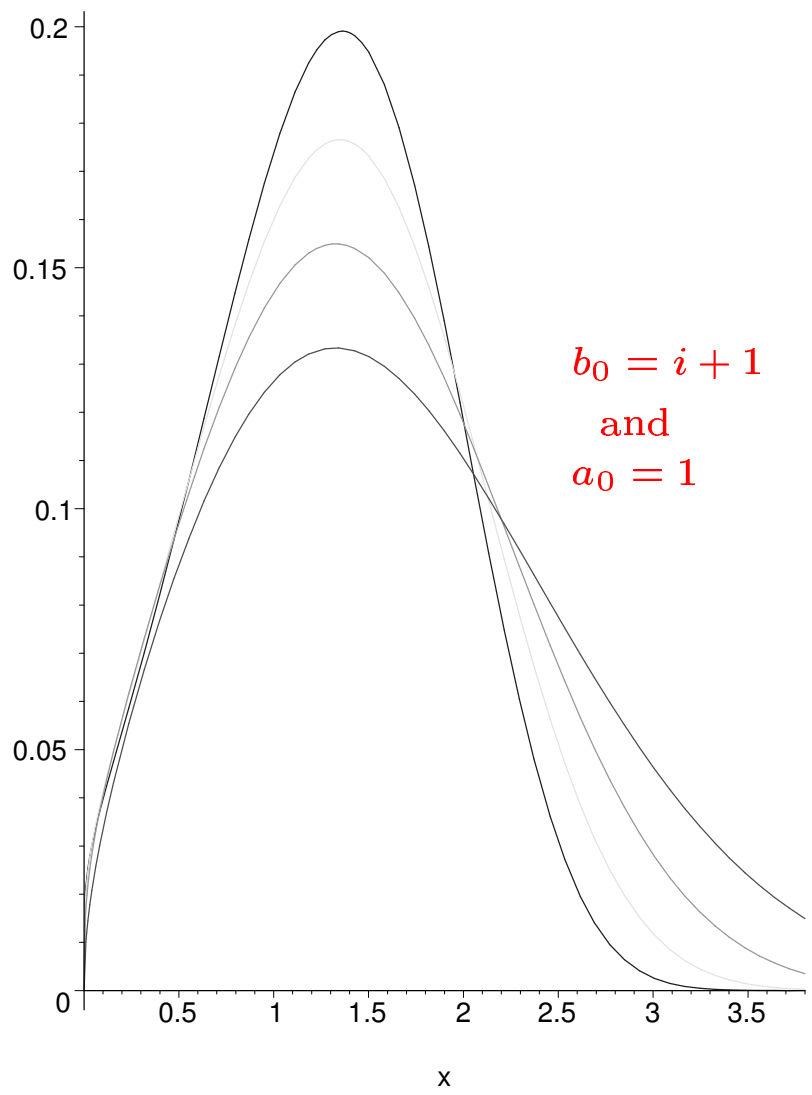
where \mathcal{H} is a clockwise loop around the negative real axis.

Finally, expanding $e^{-xt^{c/a}}$ and integrating termwise

Theorem 2

$$Pr\left(Z_n = \lfloor xn^{b/a} \rfloor\right) \sim n^{-b/a} \frac{\Gamma\left(\frac{t_0}{a}\right)}{\Gamma\left(\frac{b_0}{b}\right)} x^{b_0/b-1} \sum_{k \geq 0} \frac{(-1)^k}{\Gamma\left(\frac{a_0 - kb}{a}\right)} \frac{x^k}{k!}$$

Limit laws for Urn $\binom{i+2 \ 0}{2 \ i}$ for i from 2 to 6



Mellin transform of the limit law

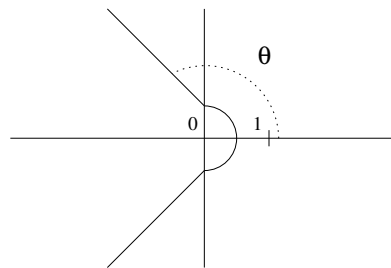
Why? Strong relations with moments of the random variable.

$$f^*(s) = \int_0^{+\infty} x^{s-1} f(x) dx$$

We start with the expression of the limit law as an integral

$$g(x) = \frac{\Gamma\left(\frac{t_0}{a}\right)}{\Gamma\left(\frac{b_0}{b}\right)} \int_{\mathcal{H}} G(x, t) dt$$

with $G(x, t) = x^{b_0/b-1} e^{-x t^{b/a}} t^{-a_0/a} e^t$. We choose \mathcal{H} as follows



where $\frac{\pi}{2} < \theta < \frac{\pi}{2} \cdot \frac{a}{b}$.

Then for any $t \in \mathcal{H}$ and any $x > 1 - b_0/b$, $\int_0^{+\infty} |x^{s-1}G(x, t)| dx$ is convergent and equal to

$$\int_0^{+\infty} x^{s-1}G(x, t)dx = \Gamma\left(\frac{b_0}{b} + s - 1\right) t^{-\frac{t_0}{a} - (s-1)\frac{b}{a}} e^t$$

Integration with respect to variable t then gives theorem

Theorem 3 *The limit law admits a Mellin transform defined for any $s > 1 - b_0/b$ by*

$$g^*(s) = \frac{\Gamma\left(\frac{b_0+(s-1)b}{b}\right) \Gamma\left(\frac{t_0}{a}\right)}{\Gamma\left(\frac{b_0}{b}\right) \Gamma\left(\frac{t_0+(s-1)b}{a}\right)}$$

Triangular urn of size 3

$$M = \begin{bmatrix} s_1 & \delta & s_3 - s_1 - \delta \\ 0 & s_2 & s_3 - s_2 \\ 0 & 0 & s_3 \end{bmatrix}$$

Diff. eq. $\frac{\partial H}{\partial z} = u^{s_1+1} v^\delta w^{s_3-s_1-\delta} \frac{\partial H}{\partial u} + v^{s_2+1} w^{s_3-s_2} \frac{\partial H}{\partial z} + w^{s_3+1} \frac{\partial H}{\partial w}$

Prime Integrals

$$-dz = \frac{dw}{w^{s_3+1}} \Rightarrow c_1 = z - \frac{1}{s_3 w^{s_3}}$$

$$\frac{dv}{v^{s_2+1} w^{s_3-s_2}} = \frac{dw}{w^{s_3+1}} \Rightarrow c_2 = \frac{1}{v^{s_2}} - \frac{1}{w^{s_2}}$$

$$\frac{du}{u^{s_1+1} v^\delta w^{s_3-s_1-\delta}} = \frac{dv}{v^{s_2+1} w^{s_3-s_2}} \Rightarrow \frac{du}{u^{s_1+1}} = \frac{1}{v^{s_1+1}} (c_2 + v^{s_2})^{\frac{s_1+\delta-s_2}{s_2}} dv$$

$$c_3 = \frac{1}{u^{s_1}} + c_2^{\frac{s_1}{s_2}} g \left(v c_2^{\frac{1}{s_2}} \right) \quad \text{with} \quad g = s_1 \int \frac{1}{t^{s_1+1}} (1+t^{s_2})^{\frac{s_1+\delta-s_2}{s_2}} dt$$

Generating function

$$f_3 = w (1 - s_3 z w^{s_3})^{-1/s_3}$$

$$f_2 = v \left(1 - v^{s_2} \left(\frac{1}{w^{s_2}} - \frac{1}{f_3^{s_2}} \right) \right)^{-1/s_2}$$

$$f_1 = u \left[1 - u^{s_1} \left[\left(\frac{1}{w^{s_2}} - \frac{1}{v^{s_2}} \right)^{s_1/s_2} g \left(v \left(\frac{1}{w^{s_2}} - \frac{1}{v^{s_2}} \right)^{1/s_2} \right) \right. \right. \\ \left. \left. - \left(\frac{1}{f_3^{s_2}} - \frac{1}{f_2^{s_2}} \right)^{s_1/s_2} g \left(f_2 \left(\frac{1}{f_3^{s_2}} - \frac{1}{f_2^{s_2}} \right)^{1/s_2} \right) \right] \right]$$

$$H = f_1^{a_1} f_2^{a_2} f_3^{a_3}$$

Examples with closed form generating functions

Sum = 2

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

$$T_1 \sim \alpha_0 n^{1/2}$$

$$T_2 \sim \beta_0 n^{1/2}$$

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

$$T_1 \sim \alpha_0 n^{1/2}$$

$$T_2 \sim \alpha_0 n^{1/2} \ln n$$

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

$$T_1 \sim \alpha_0 n^{1/2}$$

$$T_2 \sim \alpha_0 n^{1/2}$$

Sum = 3

$s_1 < s_2$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{1/3}$$

$$T_2 \sim (\alpha_0 + \beta_0) n^{2/3}$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{1/3}$$

$$T_2 \sim (2\alpha_0 + \beta_0) n^{2/3}$$

Sum = 3

$$s_1 = s_2$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{1/3}$$

$$T_2 \sim \alpha_0 n^{1/3} \ln n$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{1/3}$$

$$T_2 \sim 2\alpha_0 n^{1/3}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{2/3}$$

$$T_2 \sim \alpha_0 n^{2/3} \ln n$$

$$s_1 > s_2$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

$$T_1 \sim \alpha_0 n^{2/3}$$

$$T_2 \sim \alpha_0 n^{2/3} \ln n$$

Special cases of size d

$$M = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 2 & \ddots & \vdots \\ & & \ddots & 1 \\ & & & d \end{bmatrix}$$

Diff. eq. $\frac{\partial H}{\partial z} = x_1^2 x_2 \cdots x_d \frac{\partial H}{\partial x_1} + x_2^3 x_3 \cdots x_d \frac{\partial H}{\partial x_2} + \cdots + x_d^{d+1} \frac{\partial H}{\partial x_d}$

Prime integrals

$$\frac{dx_1}{x_1^2 x_2 \cdots x_d} = \frac{dx_2}{x_2^3 x_3 \cdots x_d} = \cdots = \frac{dx_d}{x_d^{d+1}} = -dz$$

$$c_1 = \frac{1}{x_1} - \frac{1}{x_2} \quad c_2 = \frac{1}{x_2^2} - \frac{1}{x_3^2} \quad \cdots \quad c_d = z - \frac{1}{dx_d^d}$$

General solution

$$\Phi \left(\frac{1}{x_1} - \frac{1}{x_2}; \frac{1}{x_2^2} - \frac{1}{x_3^2}; \cdots; z - \frac{1}{dx_d^d} \right)$$

Initial conditions :

$$H \left(\frac{1}{x_1} - \frac{1}{x_2}; \frac{1}{x_2^2} - \frac{1}{x_3^2}; \cdots; \frac{1}{dx_d^d} \right) = x_1^{a_1} \cdots x_d^{a_d}$$

⇒ Define $f_d = x_d (1 - dzx_d^d)^{-1/d}$

⋮

$$f_i = x_i \left(1 - x_i^i \left(\frac{1}{x_{i+1}^i} - \frac{1}{f_{i+1}^i} \right) \right)^{-1/i}$$

Then

$$H = f_1^{a_1} \cdots f_d^{a_d}$$

Finally, $\forall i \in \llbracket 1; n - 1 \rrbracket$

$$E(X_i) \sim (a_1 + \dots + a_i) \frac{\Gamma\left(\frac{t_0+i}{n}\right)}{\Gamma\left(\frac{t_0}{n}\right)} n^{i/n}$$

Urn of the hashing problem

Let us begin with size 3

$$M = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad U_0 = \begin{bmatrix} n_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Diff. eq.

$$\frac{\partial H}{\partial z} = x_2 \frac{\partial H}{\partial x_1} + x_3 \frac{\partial H}{\partial x_2} + x_3 \frac{\partial H}{\partial x_3}$$

Prime integrals

$$-dz = \frac{dx_3}{x_3} \Rightarrow c_3 = z - \ln x_3$$

$$\frac{dx_2}{x_3} = \frac{dx_3}{x_3} \Rightarrow c_2 = x_2 - x_3$$

$$\frac{dx_1}{x_2} = \frac{dx_2}{x_3} \Rightarrow dx_1 = \frac{x_2}{x_2 - c_2} dx_2$$

$$c_1 = x_1 - x_2 + (x_2 - x_3) \ln x_3$$

$$\Rightarrow H(z, x_1, x_2, x_3) = (x_1 + x_2 z + x_3(e^z - z - 1))^{n_0}$$

General case :

$$H(z, x_1, \dots, x_d) = \left(x_1 + \dots + x_{d-1} \frac{z^d}{d!} + x_d (e^z - 1 - z - \dots - \frac{z^d}{d!}) \right)^{n_0}$$