UNIVERSITÁ COMMERCIALE LUIGI BOCCONI - MILANO

Istituto di Metodi Quantitativi

Dottorato di Ricerca in Statistica- XX Ciclo

Reinforced Urn Processes and Binary Tree

Coordinatore:

Ch.mo Prof. Pietro Muliere

Tesi di: Djilali Ait Aoudia

Matr. 1003862

UNIVERSITÁ COMMERCIALE LUIGI BOCCONI

ISTITUTO DI METODI QUANTITATIVI

The thesis **"Reinforced Urn Processes and Binary Tree"** by **Djilali Ait Aoudia** is recommended for acceptance by the members of the delegated committee, as stated by the enclosed reports, in partial fulfillment of the requirements for the degree of Doctor of Philisophy.

Dated: **January 2008**

Research Supervisor:

External Examiners:

UNIVERSITÁ COMMERCIALE ”LUIGI BOCCONI”

Dated: **January 2008**

Autor: **Djilali Ait Aoudia**
Title: **Reinforced Urn Processes and Binary Tree**
Departement: **Istituto di Metodi Quantitativi**

To my parents

≪ *Non vogliate negar l'esperienza di retro al sol, del mondo sanza gente. Considerate la vostra semenza: fatti non foste a viver come bruti, ma per seguir virtute e canoscenza.* ≫

Inferno XXVI, 116-120

ACKNOWLEDGEMENTS

Special thanks are due to Professor Pietro Muliere: not only has he been very generous with his time, but also his presence somehow cause me to math twice as well.
Thanks are also due to Professor Eric Marchand who helped me very much, and I am very thankful for this. Thanks for département de mathematiques de l'université de Sherbrooke for hospitality during my visit in particular Madame Marie-France Roy .
I am also grateful to Dr. Paolo Bulla for his help and I wish to thank friends, like Antonella, Laura, Lorenzo, Marta, Marilyna and my colleagues of PhD.
Finally many thanks to my parents for their continuous support and encouragement throughout this study.

List of Abbrevation

r.v	random variable
r.v.s	randon varibles
d.f.	distribution function
c.f.	characteristic function
a.s.	almost surely
i.i.d.	independ, identically distribution
iff	if and only if
CLT	Central Limit Theorem
WLLN	Weak Law of Large Numbers
SLLN	Strong Law of Large Numbers
m.g.f.	moment generating function

Notation

$(\Omega, \Im, \mathbb{P})$	Underliying probability space
$Ł^p(\Omega, \Im, \mathbb{P})$	The space of random variables, X such that $\|X\|^p$ is integrable
$\mathbb{E}(X)$	Expectation of X
$\mathbb{C}ov(X, Y)$	Covariance of X and Y
$\mathbb{V}ar(X)$	Variance of X
$\Im_n$	increassing σ fields $\subset \Im$
$\Im_0$	$\{\emptyset, \Omega\}$
$\mathbb{E}(X/\Im)$	conditional expectation
$\\|X\\|_p$	for $p-$ norme of X that is $(\mathbb{E}\|X\|^p)^{1/p}$, $p > 0$
$\mathbb{I}_A$	the indicator of A
A^c	the complement of event A

ABSTRACT

We consider a class of random processes that have a kind of reiforcement. The first chapter is to review Pólya urn schemes. We present basic results as they came by chronologically and we consider some generalizations.

In the second chapter, according to Muliere, Secchi and Walker (2005) a reinforced random process indexed by k-ary tree can be described as a stochastic process representing the outcomes of drawings in a system of urns whose composition determnide by the interaction of the geometrical structure (the tree) with a Pólya-like reinforcing rule. This process describes the evolution of the urn in draws to associate it with a tree structure.We are interested in a particular case of infinity binary tree. We then consider some generalizations .

In the last chapter, we develop new discrete distributions that describe the behavior of a sum of independent bernoulli random variables and we are interested a multidimensional case.

Contents

Chapter 1

Introduction

The intent of this chapter is to review Pólya urn schemes.We present basic results as they came by chronologically, and we sketch their original proof to provide hints on the broad array of methods employed, and the evolutionnary process leading to the current state of the art:

- Pólya -Eggenberger's urn .
  Eggenberger and Pólya (1923).
- Generalized Pólya urns
  -Bernard Friedman's urn
  Bernard Friedman (1943).
- Modern developments
  -Hill, Lane and Sudderth (1980)
  -Bagchi (1985)
  -Pemantle (1990)

The study of urn models has a long history. In 1657, Huygens proposed probelms about urns in his treatise, but he did not use the term "urn". James Bernoulli (1713) may have been the first person to montion problems in the language of urns who, in the third book of his *Ars Conjectandi*, discusses the problem of drawing "calculi" out of urns . The results on urn models up to 1977 were summarized in the book *Urn Models and Their Application.* Kotz and Balakrishnan (1991) .

1.1 Pólya's urn

The urn models described earlier in this section are all particular cases of the Pólya urn model. This was put forward by Eggenberger and Pólya (1923) as a model for contagious distribution, that is, for situation where occurence of an event, that is, has an afterffect; see also Jordan(1927).
In 1923, Eggenbeger and Pólya proposed the following urn scheme to model process such as the spread of infectious disease. In this scheme a single urn contains $a > 0$ white balls and $b > 0$ black balls. A ball is drawn at random and then replaced, together with $\alpha > 0$ balls of the same color. Here α is any fixed constant ($\alpha \in \mathbb{N}^*$). The procedure is repeated n times and we denote p_n the proportion of black balls after the n-th draw.
We define a sequence or random variables $\{X_n\}$, each equal to 0 or 1 according to the color black or white respectively of the ball sampled at stage n. Moreaver, for all $n \geqslant 1$, we set B_n and W_n equal to the number of black balls and white balls respectively in the urn before the $(n+1)$-th stage.
The dynamics of the processes $\{X_n\}, \{B_n\}$ and $\{W_n\}$ are governed by the following : X_1 has distribution

$$Bernoulli\left(\frac{b}{b+a}\right)$$

For all $n \geqslant 1$, conditionally on $X_1, \cdots, X_n$

$$X_{n+1} = \begin{cases} 1 & \text{with probability } \frac{B_n}{B_n+W_n} = p_n \\ 0 & \text{with probability } \frac{W_n}{B_n+W_n} = 1 - p_n \end{cases} \tag{1.1.1}$$

whereas

$$(B_{n+1}, W_{n+1}) = \begin{cases} (B_n + \alpha, W_n) & \text{with probability } \frac{B_n}{B_n+W_n} \\ (B_n, W_n + \alpha) & \text{with probability } \frac{W_n}{B_n+W_n} \end{cases} \tag{1.1.2}$$

The following is the basic theorem about down the first time it appeared in this form .

Theorem 1.1.1 *(Eggenberger and Pólya, 1923)*
As n grows to infinity, the propotion of black balls p_n converges almost surely to some random limit. Mereover, the distribution of the limit is a beta with parameters a/α and b/α.

Proof (**sketch**) : The successive fractions of black balls form a bounded martingale hence converge almost surely.
Let $\tilde{B_n}$ be the number of black balls splits in the Pólya-Eggenberger urn after n draws.
We have that

$$\mathbb{P}\left(\tilde{B_n} = k\right) = \binom{n}{k} \frac{\Gamma(k + b/\alpha)\Gamma(n - k + a/\alpha)\Gamma((a + b)/\alpha)}{\Gamma(b/\alpha)\Gamma(b/\alpha)\Gamma(n + (a + b)/\alpha)}$$

So for $x \in [0, 1]$, the distribution function of the black splits is

$$\mathbb{P}\left(\tilde{B_n} \leq nx\right) = \sum_{k=0}^{\lfloor nx \rfloor} \binom{n}{k} \frac{\Gamma(k + b/\alpha)\Gamma(n - k + a/\alpha)\Gamma((a + b)/\alpha)}{\Gamma(b/\alpha)\Gamma(a/\alpha)\Gamma(n + (a + b)/\alpha)}$$

Using Stirling's approximation to the gamma function and factorials in the binomial coefficient , proceed to the limit as $n \longmapsto \infty$ with

$$\mathbb{P}\left(\frac{\tilde{B_n}}{n} \leq x\right) \longmapsto \frac{\Gamma((a + b)/\alpha)}{\Gamma(b/\alpha)\Gamma(a/\alpha)} \int_0^x t^{b/\alpha - 1}(1 - t)^{a/\alpha - 1} dt$$

this show that the distribution of the limit is a beta with parameters a/α and b/α.

1.2 Generalization of Pólya's urn

A generalization of Pólya urn is the following urn scheme proposed by Friedman (1949); an urn initially contains $n_0(b)$ black balls and $n_0(w)$ white balls. At stage n, with $n \geqslant 1$, a ball is sampled from the urn and $1 + \alpha$ balls of the same color together with β balls of the other color are put beck in the urn. If $\alpha = 0$ and $\beta = 0$ we have usual extraction with replacement scheme which generate a sequence of i.i.d colors. If $\alpha = m - 1 > 0$ and $\beta = 0$ we recover the Pólya urn scheme considered above.

Theorem 1.2.1 *Let B_n be the number of white balls in nondegenerate Bernard Freidman's urn after n draws. The moment generating function $\phi_n(t) = \mathbb{E}\left[e^{tB_n}\right]$ satisfies the differential equation*

$$\phi_{n+1}(t) = e^{\beta t}\left[\phi_n(t) + \frac{e^{(\alpha - \beta)t} - 1}{\tau_n}\phi_n^{'}(t)\right] \tag{1.2.1}$$

with τ_n is the total number of balls after n drawns.

Proof :
Let I_n^B be the indicators of the events of drawing a black ball at the nth step.

Since the number of black balls after $n+1$ draws is what is was after n steps, plus the addition (possible negative) inccured by the ball sampled at step $n+1$.

$$B_{n+1} = B_n + (\alpha-\beta)I_n^B + \beta$$

Therefore

$$\begin{aligned}
\phi_{n+1}(t) &= \mathbb{E}\left(e^{tB_{n+1}}\right) \\
&= \mathbb{E}\left[\mathbb{E}\left(e^{tB_{n+1}} \mid \Im_n\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(e^{tB_{n+1}}\ \mathbb{I}_{\{B_{n+1}=B_n+\alpha\}} \mid \Im_n\right)\right] + \mathbb{E}\left[\mathbb{E}\left(e^{tB_{n+1}}\ \mathbb{I}_{\{B_{n+1}=B_n+\beta\}} \mid \Im_n\right)\right] \\
&= \mathbb{E}\left[e^{t(B_n+\alpha)}\mathbb{E}\left(\mathbb{I}_{\{B_{n+1}=B_n+\alpha\}} \mid \Im_n\right)\right] + \mathbb{E}\left[e^{t(B_n+\beta)}\mathbb{E}\left(\mathbb{I}_{\{B_{n+1}=B_n+\beta\}} \mid \Im_n\right)\right] \\
&= \mathbb{E}\left[e^{t(B_n+\alpha)}P\left(B_{n+1}=B_n+\alpha \mid \Im_n\right)\right] + \mathbb{E}\left[e^{t(B_n+\beta)}P\left(B_{n+1}=B_n+\beta \mid \Im_n\right)\right] \\
&= e^{\alpha t}\mathbb{E}\left[e^{tB_n}\left(\frac{B_n}{\tau_n}\right)\right] + e^{\beta t}\mathbb{E}\left[e^{tB_n}\left(1-\frac{B_n}{\tau_n}\right)\right] \\
&= e^{\beta t}\phi_n(t) + \frac{e^{\alpha t}}{\tau_n}\phi_n^{'}(t) - \frac{e^{\beta t}}{\tau_n}\phi_n^{'}(t) \\
&= e^{\beta t}\phi_n(t) + e^{\beta t}\left[\frac{e^{(\alpha-\beta)t}-1}{\tau_n}\phi_n^{'}(t)\right] \\
&= e^{\beta t}\left[\phi_n(t) + \frac{e^{(\alpha-\beta)t}-1}{\tau_n}\phi_n^{'}(t)\right]
\end{aligned}$$

This show that

$$\phi_{n+1}(t) = e^{\beta t}\left[\phi_n(t) + \frac{e^{(\alpha-\beta)t}-1}{\tau_n}\phi_n^{'}(t)\right] \tag{1.2.2}$$

with τ_n is the total number of balls after n drawns.

Twentieth centry research paid attention to simplfing results by focusing on the essential elements or asymptotics. David freedman (1965) develops an asymptotic theory by Bernard Friedman's. and was able to prove the following limit laws via martingale analysis.

Theorem 1.2.2 *(Freedman)*
Let $\rho = (\alpha-\beta)/(\alpha+\beta)$ and $\sigma^2 = (\alpha-\beta)^2/(1-2\rho)$. Then

$$(p_n - 1/2)\sqrt{n} \longrightarrow N(0,\sigma^2) \ \textit{if}\ \rho < 1/2$$

$$(p_n - 1/2)\sqrt{n\log(n)} \longrightarrow N(0,(\alpha-\beta)^2 \ \textit{if}\ \rho = 1/2$$

$$(p_n - 1/2)n^{1-\rho} \longrightarrow Z \ \ \textit{if}\ \rho > 1/2$$

where Z is a non-degenerate random variable and the convergence is in distribution.

David Freedman uses a moment calculation to obtain the precise rate of convergence of the succussive fractions of black balls, $p_0, p_1 \cdots$.

Theorem 1.2.3 *If* $\alpha \geqslant 0$ *and* $\beta > 0$, *the proportion of black balls* p_n *converge to* $1/2$ *with probability one whatever the initial composition* $(n_0(b), n_0(w))$ *of the urn.*

1.3 Modern developments

By classical development we refer to all the relevant materiel that can be found in textbook. Johnson and Kotz (1977) is a classical in this field. Kotz and Balakrishnan (1997)is a companion survey of a Pólya urn model that goes into many more offshoots and derived urns. Athreya and Ney (1972) puts a generalized model in the perspective of the branching process.

In 1980, Hill, Lane and Sudderth proposed a generalized Pólya-Eggenberger urn model. An urn containing black and white balls has a given initial composition. At each time, a ball is drawn from the urn and replaced along with another ball of the same color . The draws are not exactly representative of the contents of the urn but are determined by the contents in the following manner. Let the number of black and white balls at time n be B_n and W_n, respectively, instead of drawing a black ball with probability $\frac{B_n}{W_n+B_n}$, draw a black ball with probability $f\left(\frac{B_n}{W_n+B_n}\right)$, where f is any function mapping $[0,1]$ into itself.

Hill , Lane and Sudderth (1980) have then show that, under a condition on f at p_0, where $p_0 \in \{x : f(x) = x\}$, the fraction of black balls converge to p_0 with positive probability. Moreover, if p_0 is an unstable fixed point of f (i.e $f(p_0) = p_0$ and, in some neighborhood of $p_0, f(x) < x$ for $x < p_0$ and $f(x) > x$ for $x > p_0$) then $\mathbb{P}\left(B_n/(W_n + B_n) \longrightarrow p_0\right) = 0$.

Pemantle (1990) has generalized the original Pólya-Eggenberger urn process by replacing α (the number of extra balls added of the color drawn) with a function of time. He has suggested this model for the american presidential primary election procedure. To this end, let us assume an initial amount of popular support for each candidate that dictates that condidate's chance of winning the first primary and then assume that the support increases proportionally to the size of the states won by the candidate in each primary.

Formaly, $F : \mathbb{Z} \longrightarrow (0,1)$ be any function. Let $\{p_1, p_2, \cdots\}$ be the successive proportion of red balls in an urn that begins with B black balls and W white balls and evolves at follows : at discrete times $n = 1, 2, \cdots$ a ball is drawn and replaced in the urn along with $F(n)$ balls of the some color. The usual Polya urn scheme is the case where $F(n) = \alpha$ for all n. Pemantle (1990) has then shown that the proportions of balck balls in the urn forms a martingale, so that it must converge for any F and that the limit has no atoms except possibly at 0 and 1. Here, p_n is the proportion of black balls at time n and $\delta_n = \frac{F(n)}{B+W+\sum_{i=1}^{n}}$ are fractional additions. Rigorously stated, the limit p is such that $\mathbb{P}(p=0) = 1 - \mathbb{P}(p=1) = \frac{B}{W+B}$ if and only if $\sum_{i=1}^{\infty} \delta_n^2 = \infty$ and that the distribution of p has no atoms in $(0,1)$. As a counterexample, let us consider $B = W = 1$ and $F(n) = n$. In this case case the probability that all draws of the some color is $\frac{2}{3} \times \frac{6}{7} \times \frac{15}{16} \times \cdots > 0$, but in view of the above stated result is not entirely concentrated on $\{0,1\}$.

1.4 Exchangeable urn process :

1.4.1 Exchangeable random variables

The notion of exchangeability was investigated by Bruno de Finetti (1931,1937), who recognized its fundamental role for subjective probability and Bayesian statistics. Since then, many authors contributed to the understanding of the properties of exchangeable families of random variables. The result on exchangeable of families of random variables up to 1983 were summarized in the book Exchangeability and related topics (Aldous-D, Ecole d'été de probabilités de Saint-Flour,XIII-1983. This book stimulated many probabilists and statisticians to investigate different kinds of exchangeability .

Definition 1.4.1 *A finite sequence* $(Z_1, \cdots, Z_N)$ *of random variables is called exchangeable (or* N*-exchangeable, to indicate the number of random variables) if*

$$(Z_1, \cdots, Z_N) \stackrel{d}{=} \left(Z_{\pi(1)}, \cdots, Z_{\pi(N)}\right) \tag{1.4.1}$$

for each permutation π *of* $\{1, \cdots, N\}$*. An infinite sequence* $(Z_1, Z_2, \cdots)$ *is called exchangeable if*

$$(Z_1, Z_2, \cdots) \stackrel{d}{=} \left(Z_{\pi(1)}, Z_{\pi(2)}, \cdots\right) \tag{1.4.2}$$

for each finite permutation π *of* $\{1, 2, \cdots\}$*.*

The following Representation Theorem, due to de Finetti (1937), is our first fundamental result for an infinite exchangeable sequence of random variables.

Theorem 1.4.2 *(Bruno de Finetti. 1937)*
Let $Z = (Z_1, Z_2, \cdots)$ *be a sequence of* $\{0, 1\}$*-valued random variables.*
If the sequence Z *is exchangeable, then there is a random variable* $\Theta \in [0, 1]$ *such that given* Θ*, the variables* $Z_1, Z_2, \cdots$ *are independent and identically distributed* $Bernoulli(\Theta)$*.*

Lemma 1.4.3

Let $Z = (Z_1, Z_2, \cdots)$ be a exchangeable sequence of $\{0,1\}$-valued random variables. Then for all i, j

$$\mathbb{C}ov\,(Z_i, Z_j) \geqslant 0$$

Proof

We suppose that, the sequence $\{Z_i\}$ is exchangeable then by de Finetti's Representation Theorem there exists a random variable $\Theta \in [0,1]$ such that, the random variables $X_1, X_2, \cdots$ are conditionally i.i.d. given the random variable Θ.

Therefore

$$\begin{aligned} \mathbb{E}\,[Z_i Z_j] &= \mathbb{E}\,[Z_1 Z_2] \qquad \text{(by exchangeability)} \\ &= \mathbb{E}\,[\mathbb{E}\,(Z_1 Z_2|\Theta)] \\ &= \mathbb{E}\,[\mathbb{E}\,(Z_1|\Theta)]\,\mathbb{E}\,[\mathbb{E}\,(Z_2|\Theta)] \qquad \text{(conditional independence)} \\ &= \mathbb{E}\,[\mathbb{E}\,(Z_1|\Theta)]^2 \qquad \text{(conditionally identically dist'd)} \end{aligned}$$

And of course $\mathbb{E}\,[Z_i] = \mathbb{E}\,[\mathbb{E}\,(Z_1|\Theta)] = \mathbb{E}\,[Z_j]$, so that

$$\begin{aligned} \mathbb{C}ov\,(Z_i, Z_j) &= \mathbb{E}\,[\mathbb{E}\,(Z_1|\Theta)]^2 - [\mathbb{E}\,[\mathbb{E}\,(Z_1|\Theta)]]^2 \\ &= \mathbb{V}ar\,[\mathbb{E}\,(Z_1|\Theta)] \\ &\geqslant 0. \quad \blacksquare \end{aligned}$$

For a Pólya urn, we define a sequence or random variables $\{X_n\}$, each equal to 0 or 1 according to the color black or white respectively of the ball sampled at stage n. Moreaver, for all $n \geqslant 1$, we denote p_n the proportion of black balls after the nth draw. We know that the dynamics of the processes $\{X_n\}$ and $\{p_n\}$ are governed by the following : X_1 has distribution

$$Bernoulli\left(\frac{b}{b+w}\right)$$

For all $n \geqslant 1$, conditionally on $X_1, \cdots, X_n$

$$X_{n+1} = \begin{cases} 1 & \text{with probability} \quad p_n \\ 0 & \text{with probability} \quad 1 - p_n \end{cases} \tag{1.4.3}$$

There are several ways to see that the sequence X_n in the original Pólya's urn converges almost surely. The prettiest analysis of Pólya's urn is based on the following theorem.

Theorem 1.4.4 *The sequence* $\{X_n\}$ *generated by a Pólya urn is exchangeable and its de Finetti measure is a beta with parameters* $(\frac{a}{\alpha}, \frac{b}{\alpha})$

Proof :

Let $1 \leq k \leq n$ and $(i_1, \cdots, i_n)$ such that $i_j \in \{0,1\}$ and $\sum_{j=1}^{n} i_j = k$. Then

$$\begin{aligned}
\mathbb{P}(X_1 = i_1, \cdots, X_n = i_n) &= \frac{\prod_{j=0}^{k-1}(a+\alpha j) \cdot \prod_{j=0}^{n-k-1}(b+\alpha j)}{\prod_{j=0}^{n-1}(a+b+\alpha n)} \\
&= \frac{\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}\right)\Gamma\left(\frac{a+k}{\alpha}\right)\Gamma\left(\frac{b+n-k}{\alpha}\right)}{\Gamma\left(\frac{b}{\alpha}\right)\Gamma\left(\frac{a}{\alpha}\right)\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}+n\right)} \\
&= \int_0^1 x^k(1-x)^{n-k}\left[\frac{\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}\right)}{\Gamma\left(\frac{a}{\alpha}\right)\Gamma\left(\frac{b}{\alpha}\right)} x^{\frac{a}{\alpha}-1}(1-x)^{\frac{b}{\alpha}-1}\right] dx
\end{aligned}$$

Because of de Finetti's Representation Theoreme, this shows that the sequence is exchangeable. Moreover, the unicity of the representation implies that de Finetti measure of the sequence $\{X_n\}$ is a beta with parametrers $(\frac{a}{\alpha}, \frac{b}{\alpha})$.

1.4.2 Orthogonal sequence and The Strong Law of Large Numbers .

Definition 1.4.5 *(Orthogonal sequence)*

The basic sequence $\{U_n, n \geqslant 1\}$ *is said to be orthogonal if*

1) $U_n \in L^2(\Omega, \Im, P)$, *for all* $n \geqslant 1$.

2) $\mathbb{C}ov(U_i, U_j) = 0$, *for all* $i \neq j$

We have the following result.

Theorem 1.4.6 *(P.Hall -1980)*
Let $\{U_n, n \geqslant 1\}$ *be a sequence of random variables such that :*

1) *The sequence* $\{U_n, n \geqslant 1\}$ *is orthogonal .*
2) *For all* $n \geqslant 1,\ \mathbb{E}(U_n) = 0$.
3) *There exists a sequence* $\{c_n, n \geqslant 1\} \subset \mathbb{R}$, *such that*

$$\sum_{n=1}^{\infty} c_n^2 (\log n)^2 \mathbb{E}\left(U_n^2\right) < \infty$$

Then the series

$$\sum_{n=1}^{\infty} c_n U_n \qquad \textit{converges a.s} \tag{1.4.4}$$

Theorem 1.4.6 can be applied to obtain the following Strong Law of Large Numbers convergence theorem .

Theorem 1.4.7 *(S.L.L.N)*
Let $X = (X_1, X_2, \cdots)$ *be a sequence of* $\{0,1\}$*-valued random variables.*
If the sequence $X = (X_1, X_2, \cdots)$ *is exchangeable then there is random variable* $\Theta \in [0,1]$ *such that*

$$\frac{\sum_{i=1}^{n} X_i}{n} \underset{n\to\infty}{\longrightarrow} \Theta \qquad \textit{almost surely} \tag{1.4.5}$$

Proof :
We suppose that, the sequence $\{X_i\}$ is exchangeable then by de Finetti's Representation Theorem there exists a random variable $\Theta \in [0,1]$ such that given $\Theta = \theta$, the random variable $X_1, X_2, \cdots$ are independent Bernoulli(θ).
Let $\{U_n,\ n \geqslant 0\}$, the sequence of random variables such that $U_n = X_n - \Theta$.
We have that
1) $\mathbb{E}(U_n) = \mathbb{E}(X_n) - \mathbb{E}(\Theta) = \mathbb{E}(\mathbb{E}(X_n/\Theta)) - \mathbb{E}(\Theta) = \mathbb{E}(\Theta) - \mathbb{E}(\Theta) = 0$

2) For all i, j , such that $i \neq j$

$$\begin{aligned}
\mathbb{C}ov\left(U_i, U_j\right) &= \mathbb{E}\left(U_i U_j\right) \\
&= \mathbb{E}\left(\left(X_i - \Theta\right)\left(X_j - \Theta\right)\right) \\
&= \mathbb{E}\left(X_i X_j\right) - E\left(X_i \Theta\right) - \mathbb{E}\left(X_j \Theta\right) + \mathbb{E}\left(\Theta^2\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(X_i X_j \mid \Theta\right)\right) - \mathbb{E}\left(\mathbb{E}\left(X_i \Theta \mid \Theta\right)\right) - \left(\mathbb{E}\left(X_j \Theta \mid \Theta\right)\right) + \mathbb{E}\left(\Theta^2\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(X_i \mid \Theta\right)\mathbb{E}\left(X_i \mid \Theta\right)\right) - \mathbb{E}\left(\mathbb{E}\left(X_i \Theta \mid \Theta\right)\right) - \left(\mathbb{E}\left(X_j \Theta \mid \Theta\right)\right) + \mathbb{E}\left(\Theta^2\right) \\
&= \mathbb{E}\left(\Theta^2\right) - \mathbb{E}\left(\Theta \mathbb{E}\left(X_i \mid \Theta\right)\right) - \mathbb{E}\left(\Theta \mathbb{E}\left(X_j \mid \Theta\right)\right) + \mathbb{E}\left(\Theta^2\right) \\
&= 2\mathbb{E}\left(\Theta^2\right) - 2\mathbb{E}\left(\Theta^2\right) = 0
\end{aligned}$$

This shows that the sequence $\{U_n\}$ is orthogonal, with zero mean.

3) For $c_n = \frac{1}{n}$, we have

$$\sum_{n=1}^{\infty} c_n^2 (\log n)^2 \mathbb{E}\left(U_n^2\right) = \sum_{n=1}^{\infty} \frac{(\log n)^2 \mathbb{E}\left(U_n^2\right)}{n^2} < \infty \tag{1.4.6}$$

Because

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha (\log n)^\beta} < \infty \Longleftrightarrow \begin{cases} \alpha > 1 \quad \text{and} \quad \beta \in \mathbb{R} \\ \qquad\quad \text{or} \\ \alpha = 1 \quad \text{and} \quad \beta > 1 \end{cases}$$

Therefore, by P.Hall's theorem

$$\sum_{i=1}^{\infty} \frac{U_i}{i} \qquad \text{converge a.s} \tag{1.4.7}$$

We conclude by **Kronecker′s Lemma**[1] that

$$\frac{\sum_{i=1}^{n} U_i}{n} = \frac{\sum_{i=1}^{n} \left(X_i - \Theta\right)}{n} \underset{n \to \infty}{\longrightarrow} 0 \qquad a.s \tag{1.4.8}$$

This show that

$$\frac{\sum_{i=1}^{n} X_i}{n} \underset{n \to \infty}{\longrightarrow} \Theta \qquad \text{almost surely} \tag{1.4.9}$$

[1]**Kronecker Lemma** : Let $\{x_n, n \geqslant 1\}$ be a sequence of real numbers such that $\sum_{i=1}^{n} x_n$ converges, and let $\{b_n\}$ be a monotone sequence of positive constants with $b_n \uparrow \infty$. Then $b_n^{-1} \sum_{i=1}^{n} b_i x_i \longrightarrow 0$

1.4.3 Interpolation and Approximation by Bernstein Polynomial

One of the greatest pleausures in mathematics is the surprising connection that appear between apparently disconnected ideas and theories. One of the simplest is the elegant proof of the Weierstrass approximation theorem by S-Bernstein :
Given a function f on the closed $[0,1]$ the Bernstein polynomial of order n of f is defined by

$$\mathcal{B}(x,n,f)=\sum_{k=0}^{n} f(\frac{k}{n})\binom{n}{k}x^k(1-x)^{n-k}$$

converge uniformly to f on $[0,1]$.
However, the structure and origine of the polynomials become much clearer when given a probabilistic interpretation. We can write

$$\mathcal{B}(x,n,f)=\mathbb{E}\left(f(\frac{S_n}{n})\right)$$

where $S_n=X_1+\cdots+X_n$ for any $n-$uple (X_n) of independent random variables with Bernoulli law such that

$$\mathbb{P}(X_n=1)=x, \qquad \mathbb{P}(X_n=0)=1-x$$

Since, the sequence $\{p_n, n\geqslant 1\}$ of random variable describing the proportion of black balls after the nth draw converges almost surely to a random variable $p_\infty\in[0,1]$, and the sequence $\{X_n\}$ generated by a Polya urn is exchangeable and its de Finetti measure is a beta with parameters $(\frac{a}{\alpha},\frac{b}{\alpha})$. Therefore

$$\begin{aligned}
\varphi_n(t) &= \mathbb{E}\left(e^{itp_n}\right)\\
&= \mathbb{E}\left(e^{it(\frac{a+\alpha\sum_{i=0}^{n}X_i}{a+b+\alpha n})}\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(e^{it(\frac{a+\alpha\sum_{i=0}^{n}X_i}{a+b+\alpha n})}\mid\Theta\right)\right)\\
&= \int_\Omega \mathbb{E}\left(e^{it(\frac{a+\alpha\sum_{i=0}^{n}X_i}{a+b+\alpha n})}\mid\Theta\right)dP_\Theta\\
&= e^{\frac{iat}{a+b+\alpha n}}\int_0^1 \mathbb{E}\left(e^{it(\frac{\alpha\sum_{i=0}^{n}X_i}{a+b+\alpha n})}\mid\Theta=x\right)f_\Theta(x)dx\\
&= e^{\frac{iat}{a+b+\alpha n}}\int_0^1 \sum_{k=0}^{n} e^{it(\frac{\alpha k}{a+b+\alpha n})}\mathbb{P}\left(\sum_{i=0}^{n}X_{\sigma(i)}=k\mid\Theta=x\right)f_\Theta(x)dx\\
&= e^{\frac{iat}{a+b+\alpha n}}\int_0^1 \sum_{k=0}^{n} e^{it(\frac{\alpha k}{a+b+\alpha n})}\binom{n}{k}x^k(1-x)^{n-k}f_\Theta(x)dx
\end{aligned}$$

by the monotone convergence's theorem

$$
\begin{aligned}
\lim_{n\longrightarrow\infty}\varphi_n(t) &= \lim_{n\longrightarrow\infty}\mathbb{E}\left(e^{itp_n}\right)\\
&= \lim_{n\longrightarrow\infty} e^{\frac{iat}{a+b+\alpha n}}\int_0^1\sum_{k=0}^{n} e^{it(\frac{\alpha k}{a+b+\alpha n})}\binom{n}{k}x^k(1-x)^{n-k}f_\Theta(x)dx\\
&= \int_0^1 \underbrace{\lim_{n\longrightarrow\infty}\sum_{k=0}^{n} e^{it(\frac{\alpha k}{a+b+\alpha n})}\binom{n}{k}x^k(1-x)^{n-k}}_{\text{Bernstein Polynome approximation}} f_\Theta(x)dx\\
&= \int_0^1 e^{itx}f_\Theta(x)dx\\
&= \mathbb{E}\left(e^{it\Theta}\right)
\end{aligned}
$$

This shows that the distribution of the limit of p_n is beta with parameters $(\frac{a}{\alpha},\frac{b}{\alpha})$

1.4.4 Exchangeable urn process

Let $Y = (Y_1, Y_2, \cdots)$ be the sequence of $\{0,1\}$ -valued random variables. We know that if the sequence Y is generated by a Pólya urn, then it is exchangeable, with beta de Finetti measures.

We consider an urn with initial composition (a, b) of a white and b black balls and let f be a mapping from the unit interval to itself . Let the number of black and white balls at time n be B_n and W_n, respectively, instead of drawing a black ball with probability $\frac{B_n}{W_n+B_n}$, draw a black ball with probability $f\left(\frac{B_n}{W_n+B_n}\right)$. As before, let Y_n be the indicator of the event that the n-th ball added is black.

The process $Y = (Y_1, Y_2, \cdots)$ is a urn process with initial composition (a, b) and urn function f (Hill, Lane and Sudderth 1980).

Notice that the process Y for the Polya urn is an urn process with urn function $f(x) = x$ and is exchangeable.

A constant function $f(x) = p$ generates a Bernoulli process $Y_1, Y_2, \cdots$ of independent , $Bernoulli(p)$ variables. Such a process is clearly exchangeable and has a de Finetti measure concentrated at thr single point p. Another trivial collection of exchangeable urn process are the determinitic ones. Is an urn process such that initially a black ball is added with probability p and a white with probability $1-p$, and that $a = b$. This scheme corresponds to an urn function f such that

$$f(x) = \begin{cases} 0 & \text{if} \quad 0 \leq x < \frac{b}{a+b} \\ p & \text{if} \quad x = \frac{b}{a+b} \\ 1 & \text{if} \quad \frac{b}{a+b} < x \leq 1 \end{cases} \tag{1.4.10}$$

In 1987, Hill, Lane and Sudderth proposed the following problem:

Which urn processes are exchangeable ?

Theorem 1.4.8 *(Hill, Lane and Sudderth, 1987)*

The only exchangeable processes $Y = (Y_1, Y_2, \cdots)$ produced by their urn scheme are the Pólya's process, the i.i.d. Bernoulli process and the deterministic one.

1.5 A Pólya urn for the Dirichlet process

In this section we study a very useful family of prior distributions introduced by Ferguson (1973). Ferguson introduced the Dirichlet processes,univered many of their basic properties, and applied them to a variety of nonparametric estimation problems, thus providing for the first time a Bayesian interpretation for some of the commonly used nonparametric procedures.

1.5.1 Dirichlet Distribution

For $\alpha > 0$ let Gamma$(\alpha, 1)$ denote the gamma distribution on $(0, \infty)$ with mean α, whose density at x is $\Gamma(\alpha)^{-1}x^{\alpha-1}e^{-x}$, and define Gamma $(0,1)$ to be the distribution degenerate at 0.
We begin with the definition of the Dirichlet distribution. For $n \geqslant 2$ and $i = 1, \cdots, n$, let $\alpha_i \geqslant 0$ and such that $\sum_{i=1}^{n} \alpha_i > 0$, and set $Z_1, \cdots, Z_n$ to be independent real random variables with distribution Gamma $(\alpha_1, 1), \cdots$,Gamma$(\alpha_n, 1)$, respectively.

Definition 1.5.1 *The probability distribution on* $([0,1]^n, \mathcal{B}^n[0,1])$ *of the random vector*

$$\left(\frac{Z_1}{\sum_{i=1}^{n} Z_i}, \cdots, \frac{Z_n}{\sum_{i=1}^{n} Z_i}\right)$$

is said to be Dirichlet with parameters $(\alpha_1, \cdots, \alpha_n)$ *and denote by* $\mathcal{D}[.;\alpha_1, \cdots, \alpha_n]$.

Following are some properties of the Dirichlet distribution.
Properties :
1)

$$\left(\frac{Z_1}{\sum_{i=1}^{n} Z_i}, \cdots, \frac{Z_n}{\sum_{i=1}^{n} Z_i}\right)$$

is independent of $\sum_{i=1}^{n} Z_i$.
2) $\mathcal{D}[.;\alpha_1, \cdots, \alpha_n]$ is singular with respect to the Lebesgue measure on $([0,1]^n, \mathcal{B}^n[0,1])$; however, if $\alpha_i > 0$ for $i = 1, 2, \cdots, n$, then the probability distribution induced by $\mathcal{D}[.|\alpha_1, \cdots, \alpha_n]$ on $([0,1]^{n-1}, \mathcal{B}^{n-1}[0,1])$ via the orthogonal projection which drops the k-th coordinate , with $k \in \{1, \cdots, n\}$, is absolutely continuous with respect to the Lebesgue measure and has density with $\alpha_i > 0$ for $i = 1, 2, \cdots, n$ is said have *Dirichlet distribution* with parameter $(\alpha_1, \cdots, \alpha_n)$ if has density

$$d_{\alpha_1, \cdots, \alpha_n}(y_1, \cdots, y_{k-1}, y_{k+1}, \cdots, y_n)$$

$$= \frac{\Gamma(\sum_{i=1}^{n} \alpha_i)}{\prod_{i=1}^{n} \Gamma(\alpha_i)} [\prod_{\{i \in \{1, \cdots, n\}: i \neq k\}} y_i^{\alpha_i - 1}][1 - \sum_{\{i \in \{1, \cdots, n\}: i \neq k\}} y_i]^{\alpha_k - 1} I_{A_{n-1}}(y_1, \cdots, y_{k-1}, y_{k+1}, \cdots, y_n)$$

where

$$A_{n-1} = \left\{ (y_1, \cdots, y_{n-1}) : y_i \geqslant 0 \text{ for } i = 1, 2, \cdots, n-1, \sum_{i=1}^{n-1} y_i < 1 \right\}$$

Note that for $n = 2$ and $k = 2$ this is the density of a Beta distribution with parameters (α_1, α_2).

1.5.2 Dirichlet Distribution via Pólya Urn Scheme

We will now extend the Pólya urn scheme of the previous section by considering an urn which possibly initially a "continuum" of colors a given distribution. This generalization is due to Blackwell and MacQueen (1973) who introduced it in order to generate an exchangeable sequence of random variables whose de Fenitti measure is a Dirichlet process. For a recent use this approch, see Cifarelli et al (1999) or Mauldin et al (1992) .

Let $\mathbf{X}$ be a metric space, separable and complete, endowed with its Borel σ-field $\mathcal{X}$ and let α be a finite measure on $(\mathbf{X}, \mathcal{X})$.

Consider a Polya urn with $\alpha(\mathcal{X})$ balls of which $\alpha(i)$ are of color i; $i = 1, 2, \cdots, r$. Drawn balls at random from the urn, replacing each drawn by two balls of the same color. Let $X_n = j$ if j is the color sampled at the n-th stage. Therefore

$$\mathbb{P}(X_1 = j) = \frac{\alpha(j)}{\alpha(\mathcal{X})} \tag{1.5.1}$$

whereas for all $n \geqslant 1$, the conditional distribution of X_{n+1} given $X_1, \cdots, X_n$, is such that

$$\mathbb{P}(X_{n+1} = j \mid X_1, X_2, \cdots, X_n) = \frac{\alpha(j) + \sum_{i=1}^{n} \delta_{X_i(j)}}{\alpha(\mathcal{X} + n)} := m_n(j) \tag{1.5.2}$$

Hence, m_n describes the random probability distribution of colors in the urn after the n-th ball has been sampled.

Definition 1.5.2 *The sequence $\{X_n\}$ is called a Polya sequence on $\mathbf{X}$ with parameter α.*

We know that the Pólya urn generates an infinity exchangeable sequence of Bernoulli random variables which are conditionally independent and identically distributed given the random probability of succes which is Beta distribution. This has been generalized by the elegant result of Blackwell and MacQueen (1973).

Theorem 1.5.3 *(Blackwell and MacQueen)*
1) The sequence $\{X_n\}$ is exchangeable and its de Fenitti measure is a Dirichlet distribution with parameters $(\alpha(0), \cdots, \alpha(r))$.
2) As n grows to infinity, the random vector $(m_n(0), \cdots, m_n(r))$ converge with probability one to a random vector $\alpha^* = (\alpha^*(0), \cdots, \alpha^*(r))$.
3) The random vector α^* has Dirichlet distribution with parmeters $(\alpha(0), \cdots, \alpha(r))$.

Proof (sketch) :
1) For all $n \geqslant 1$ and $x_1, \cdots, x_n \in (\mathcal{X})$

$$\begin{aligned}
\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) \\
&= \frac{\alpha(x_1)}{\alpha(\mathcal{X})} \prod_{i=1}^{n-1} \frac{\alpha + \delta_{\sum_{j=1}^{i-1} x_j}}{\prod_{j=1}^{i-1} (\alpha(\mathcal{X}) + j)} \\
&= \frac{\{\alpha(1)(\alpha(1)+1) \cdots (\alpha(1)+n_1-1)\}\{\alpha(2)(\alpha(2)+1) \cdots (\alpha(2)+n_2-1)\} \cdots}{\alpha(\mathcal{X})(\alpha(\mathcal{X})+1) \cdots (\alpha(\mathcal{X})+n-1)} \\
&= \prod_{j=0}^{r} \frac{\alpha(j)^{[n_j]}}{\alpha(\mathcal{X})^{[n]}}
\end{aligned}$$

where $n_i = \#\{X_j = i\}$ for $i = 1, \cdots, n$ and $a^{[k]}$ is the ascending factorial given by $a^{[k]} = a(a+1) \cdots (a+k-1)$ if $k \geqslant 1$ and $a^{[0]} = 1$. This shows that the sequence $\{X_n\}$ is exchangeable and de Fenitti's Representation Theorem implies that there is a unique distribution G on

$$A_r = \left\{ (y_1, \cdots, y_r) : y_i \geqslant 0 \text{ for } i = 1, 2, \cdots, r, \sum_{i=1}^{r} y_i < 1 \right\}$$

such that, for all $n \geqslant 1$ and $x_1, \cdots, x_n \in \mathbf{X}$,

$$\mathbb{P}(X_1 = x_1, \cdots, X_n = x_n) = \int_{A_r} \left(1 - \sum_{i=1}^{r} y_i\right)^{n_0} \prod_{i=1}^{r} y^{n_i} dG(y_1, \cdots, y_r).$$

2) If G is the distribution induced on A_r by Dirichlet distribution on $[0,1]^{r+1}$ with parameters $(\alpha(0), \cdots, \alpha(r))$, then

$$\int_{A_r} \left(1 - \sum_{i=1}^{r} y_i\right)^{n_0} \prod_{i=1}^{r} y^{n_i} dG(y_1, \cdots, y_r) = \prod_{j=0}^{r} \frac{\alpha(j)^{[n_j]}}{\alpha(\mathcal{X})^{[n]}}$$

Moreover, the unicity of the representation implies that the de Fenitti measure of the sequence $\{X_n\}$ is Dirichlet with parameter $(\alpha(0), \cdots, \alpha(r))$. Therefore the same de Fenitti's Representation Theorem implies that the random vector

$$\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}(0),\cdots,\frac{1}{n}\sum_{i=1}^{n}\delta_{X_r}(0)\right)$$

converges with probability one to a random vector $\alpha^* = (\alpha^*(0 =,\cdots,\alpha^*(r))$ whose distribution is Dirichlet with parameter $(\alpha(0),\cdots,\alpha(r))$ Finally we note that the limiting behavior of the random vector

$$(m_n(0),\cdots,m_n(r))$$

is the same as that of

$$\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{X_i}(0),\cdots,\frac{1}{n}\sum_{i=1}^{n}\delta_{X_r}(0)\right)$$

Chapter 2

Reinforced urn processes indexed by binary tree

Abstract :
Based on reinforced urn process indexed by tree introduced by Muliere, Secchi and Walker (2005) we propose a study of a particular case of infinity binary tree.
Keywords: Reinforced process, urn schemes, Martingale.
.

2.1 Introduction

According to Muliere, Secchi and Walker (2005) a reinforced random process indexed by $k-$ary tree can be described as a stochastic process representing the outcomes of drawings in a system of urns whose composition is determinede by the interaction of the geometrical structure (the tree) with a Pólya-like reinforcing rule.
For this reason an effective comprehension of this model can not do without recalling these two concepts.

Indeed, a traditional two-color Pólya urn is characterized by an initial composition of balls of colors 0 and 1 and most important by a reinforcement rule such that when a ball of given color is sampled, the composition of the urn is updated returning that ball in the urn with another one of the same color. By this way, the sequence of the random variables keeping track of the successive drawings outcomes is exchangeable.

On the other hand, a tree T is a connected graph that contains no cycles. Often a distinguished vertex θ, the root, is identified and the tree is considered as a directed graph where the vertices go in the direction away from θ. Given a vertex $\sigma \in T$, there is a unique path $\pi(\theta, \sigma)$ from θ to σ .

The number of edges in $\pi(\theta, \sigma)$ is the level number of σ and indicated with $|\sigma|$. Notice that $|\theta| = 0$. For all the vertices in $\sigma \in T$ but the root, there exists a vertex $\sigma = \overleftarrow{\tau}$ called the parent of level $|\tau|$-1 and with an edge to τ . τ is a child of σ and to vertices with the same parents are said siblings.

When the number of vertices is infinite and all the vertices have the same number of children, say k, T is an infinite k-ary tree.

In a reinforced random process indexed by a tree, the urns are allocated in the vertices of a tree and the following a sampling scheme is run: all the urns have the same initial composition of the two colors balls and, starting with an extraction from the urn in the root, we keep on sampling from the urns in the successive levels in such a way as the compositions of the children urns are reinforced by the result of the draw in parent, conformly with the tree s genealogy.

This framework allows to model, moving along the different branches, dependent sequences of random variables. Actually dependence stems from geometry: the closer the branches, the higher the dependence between the sequences. By the definition of de Finetti (1937) the collection of the sequences is said partially exchangeable.

2.2 The geometry of a binary tree

Let T be an infinity binary tree, that is a connected graph with a distinguished node called the roote indicated with θ, no cycles and a countable number of node such that every $\sigma \in T$ has two children and one parent except for θ that has no parent.

Given a node $\sigma \in T$ there is a unique path $\pi(\theta, \sigma)$ in the tree T connecting the node σ with the root θ. The number of nodes of $\pi(\theta, \sigma)$ is called the level number $|\sigma|$. The parent of σ is the unique node $\tau \in T$ with level number $|\sigma| - 1$ and with the an edge to σ. σ is child of τ and we note the parent of σ by $\overleftarrow{\sigma}$.

We will consider tree T with an infinitely number of nodes and such that every node in the tree has two children, these are called infinite tree.

2.2.1 Chauvin-Neuveu representation

Let $U = \bigcup_{n \geq 0} \{0,1\}^n$, where by convention $\{0,1\}^0 = \theta$.

An element u of $\{0,1\}^n$ is written $u = u_1 \cdots u_n$, with $u_i \in \{0,1\}$, for $i = 1, \cdots, n$.

If $u = u_1 \cdots u_n$ and $v = v_1 \cdots v_n$ belong to U, we write $uv = u_1 \cdots u_n v_1 \cdots v_n$ for contenation of u and v, in particular $u\theta = \theta u = u$.

If $u = u_1 \cdots u_n$, $(u_k \in \{0,1\}, n \leq \infty)$ we write $|u| = n$ and $u/k = u_1 \cdots u_k$.

(u/k is parent of k-th generation of u).

If $u \neq \theta$ say that uv a descendant of u (ie :$u = \overleftarrow{uv}$) and u is a ancestor of uv. Moreover $u0$ (resp $u1$) is called **left** (resp **right**) child of u .

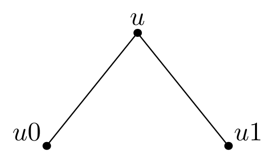

Figure 2.1: the childs of u

A complete binary tree T is subset of U such that:

$$\begin{cases} \theta \in T \\ uv \in T \Rightarrow u \in T \\ \forall u \in T : u0 \in T \Leftrightarrow u1 \in T \end{cases} \tag{2.2.1}$$

Let $E = \{u_1 u_2 \cdots : \forall n \in \mathbb{N} : u_1 \cdots u_n \in T\}$. The set E is called the space of ends of T .

We define $\tilde{T} = T \cup E$, endowed with the distance d such that for every $u, v \in \tilde{T}$

$$d(u,v) = \begin{cases} e^{-(u \wedge v)} & \text{if } u \neq v \\ 0 & \text{otherwise} \end{cases} \tag{2.2.2}$$

Where $u \wedge v = \max\{k \in \mathbb{N} : u/k = v/k\}$.

We note that, for all u, v and $\tau \in T$, $u \wedge v \geq \min(u \wedge \tau, v \wedge \tau)$, this prove that d is a metric on $\tilde{T}$.

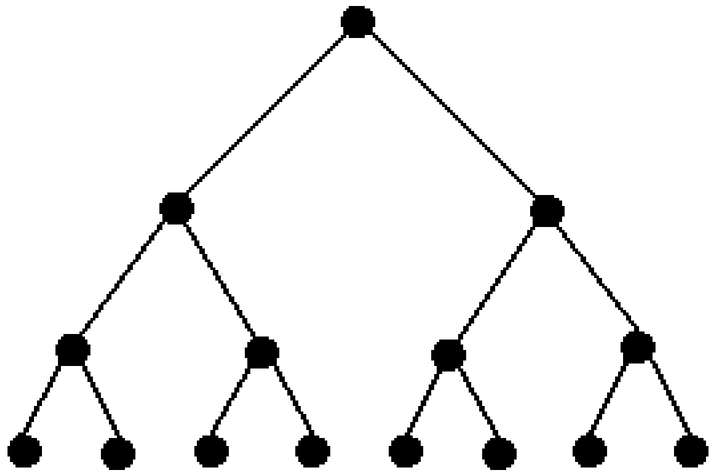

Figure 2.2: the binary tree T, n=3

The following result can be derived from Section 1.6 in [5], $(pp219-221)$. A direct proof appears in Lemma 7.3 in Woess [46].

Proposition 2.2.1 (**Woess**)
The space $\tilde{T}$ with the distance d is compact, T is a discrete subspace of $\tilde{T}$ and E is a compact subspace of $\tilde{T}$.

2.3 A reinforced dichotomous process indexed by a binary tree

We are now ready for the introduction of a stochastic process $X = \{X_\tau : \tau \in T\}$ of Bernoulli random variables defined on a rich probability space $(\Omega, \Im, \mathbb{P})$ and indexed by the nodes of a binary tree T. The law of the process X is defined as follows.

We fixe two positive real numbers a and b and set X_θ to be a Bernoulli random variable with parameter $p_\theta = \frac{a}{a+b}$.
For $n \geq 0$, let $\Im_n \subseteq \Im$ be the sigma-field generated by the random variable X_τ with $|\tau| < n$, $\Im_n = \sigma(X_\tau \ : \ \tau < n)$
Given $\Im_n$, assume that the random variables X_τ with $|\tau| = n$ are conditionally independent such that

$$X_\tau \mid \Im_n \sim Ber\left(p_\tau = \frac{a + \alpha \sum\limits_{i=0}^{n-1} X_{\sigma(i)}}{a+b+\alpha n} \right)$$

If $\pi(\theta, \tau) = (\theta = \sigma_{(0)}, \sigma_{(1)}, \cdots, \sigma_{(n-1)}, \tau)$.

In fact, every node of the tree T labels an urn containing balls of colours 1 and 0 and each node in the tree corresponds to a state of the urn. The root θ corresponds to the initial state and we suppose containing a ball of color 1 and b balls of color 0. We sample a ball from this urn and we denote X_θ the color of the ball extracted and replace it along with α additional balls of the same color to each urn corresponding to the two children of the root and repeat this procedure a level $1, 2, \ldots$
The compositon of the urn corresponding to a node $\sigma \in T$ is the same as that of its parent plus α extra balls of color $X_{\overleftarrow{\sigma}}$ sampled from the urn of its parent $\overleftarrow{\sigma}$.
We define the process $p = \{p_\sigma : \sigma \in T\}$ (resp: $Y = \{Y_\sigma : \sigma \in T\}$) which describes the proportion (resp: the total number) of balls of color 1 contained in the urns indexed by the nodes of the tree T.
Let $T_n = \{\sigma \in T : |\sigma| = n\}$, we write $\sigma(n)$ if $\sigma \in T_n$ and $\pi(\theta, \sigma) = (\theta = \sigma_{(0)}, \sigma_{(1)}, \cdots, \sigma_{(n)})$.
Then $p_{\sigma(n)}$ (resp $Y_{\sigma(n)}$) describes the proportion (resp: the total number) of balls of color 1 contained in the urn indexed by the node σ at generation n.
Now, we have the following result.

Theorem 2.3.1
For every fixed $k \geq 0$

$$Z_n = \prod_{i=0}^{k-1} \left(\frac{Y_{\sigma(n)} + i\alpha}{a+b+\alpha(n+i)} \right) \tag{2.3.1}$$

is a bounded martingale with respect to the filtration $\Im_n$*, hence it converge almost surely to the random* Z_∞*, and* $P(Z_\infty < \infty) = 1$ *a.s.*

Proof :

Trivially we have $Z_n \in [0,1]$ and Z_n is $\Im_n$ measurable for all $n \geqslant 0$.

$$\begin{aligned}
\mathbb{E}\left(Z_{n+1}/\Im_n\right) &= \mathbb{E}\left(Z_{n+1}\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}\right\}} \mid \Im_n\right) + \mathbb{E}\left(Z_{n+1}\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}+\alpha\right\}} \mid \Im_n\right) \\
&= \mathbb{E}\left(\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+i\alpha}{a+b+\alpha(n+1+i)}\right)\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}\right\}} \mid \Im_n\right) + \\
&\qquad +\mathbb{E}\left(\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+\alpha+i\alpha}{a+b+\alpha(n+1+i)}\right)\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}+\alpha\right\}} \mid \Im_n\right) \\
&= \mathbb{P}(Y_{\sigma(n+1)} = Y_{\sigma(n)} \mid \Im_n)\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+i\alpha}{a+b+\alpha(n+1+i)}\right) + \\
&\qquad +\mathbb{P}(Y_{\sigma(n+1)} = Y_{\sigma(n)}+\alpha \mid \Im_n)\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+\alpha(i+1)}{a+b+\alpha(n+1+i)}\right) \\
&= \left(\frac{a+b+\alpha n-Y_{\sigma(n)}}{a+b+\alpha n}\right)\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+i\alpha}{a+b+\alpha(n+1+i)}\right) + \\
&\qquad +\left(\frac{Y_{\sigma(n)}}{a+b+\alpha n}\right)\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+\alpha+i\alpha}{a+b+\alpha(n+1+i)}\right) \\
&= \left(\frac{a+b+\alpha n-Y_{\sigma(n)}+Y_{\sigma(n)}+\alpha k}{a+b+\alpha(n+k)}\right)\prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)}+i\alpha}{a+b+\alpha(n+i)}\right) = Z_n.
\end{aligned}$$

This shows that $\{Z_n\}$ is bounded martingale with respect to filtration $\Im_n$, hence it converges almost surely to the random $Z_\infty \in [0,1]$.

Now, let $M_n := p_{\sigma(n)}$, the sequence of random variable describing the proportion of balls of color 1 contained in the urns indexed by the nodes σ at generation n.

The theorem(2.3.1) can be applied to obtain the following

Theorem 2.3.2

i) *The sequence* $\{M_n\}$ *is martingale with respect to the filtration* $\{\Im_n\}$ *with value in* $[0,1]$, *therefore it converges almost surely to a random variable* $M_\infty \in [0,1]$.

ii) *For every fixed* k,

$$\lim_{n\to\infty} \mathbb{E}\left[M_n^k\right] = \mathbb{E}\left[M_\infty^k\right].$$

iii) *The law of* M_∞ *is beta* $\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)$.

proof :

i)Let $n \geq 0$, then $M_n \in [0,1]$ and

$$\begin{aligned}
\mathbb{E}\left(M_{n+1}|\Im_n\right) &= \mathbb{E}\left[p_{\sigma(n+1)} \mid \Im_n\right] = \mathbb{E}\left[\frac{(a+b+\alpha n)p_{\sigma(n)} + \alpha X_{\sigma(n)}}{a+b+\alpha n+\alpha} \mid \Im_n\right] \\
&= \frac{[a+b+\alpha n]p_{\sigma(n)} + \alpha\mathbb{E}[X_{\sigma(n)} \mid \Im_n]}{a+b+\alpha n+\alpha} = \frac{[a+b+\alpha n]p_{\sigma(n)} + \alpha p_{\sigma(n)}}{a+b+\alpha(n+1)} \\
&= \frac{(a+b+\alpha(n+1))p_{\sigma(n)}}{a+b+\alpha(n+1)} = M_n \ \text{ almost surely.}
\end{aligned}$$

This shows that $\{M_n\}$ is a $[0,1]$ -bounded martingale with respect to the filtration $\Im_n$, hence there exists a random variable $M_\infty \in [0,1]$ such that

$$\lim_{n\to\infty} M_n = M_\infty.$$

ii) For $k \geq 1$: part **i)** and Dominated Convergences theorem implies that:

$$\lim_{n\to\infty} E\left[M_n^k\right] = E\left[M_\infty^k\right].$$

iii) We have

$$p_{\sigma(n)} = \frac{Y_{\sigma(n)}}{a+b+\alpha n} \underset{n\to\infty}{\longrightarrow} M_\infty \ a.s$$

Then for every fixed $j \geq 0$

$$\frac{Y_{\sigma(n)} + j\alpha}{a+b+\alpha(n+j)} \underset{n\to\infty}{\longrightarrow} M_\infty \quad a.s$$

Therefore

$$Z_n = \prod_{i=0}^{k-1}\left(\frac{Y_{\sigma(n)} + i\alpha}{a+b+\alpha(n+i)}\right) \underset{n\to\infty}{\longrightarrow} Z_\infty = M_\infty^k \quad a.s$$

Note that by the previous theorem $\{Z_n\}$ is bounded martingale with respect to the filtration $\{\Im_n\}$.

Then

$$\mathbb{E}\left(M_\infty^k\right) = \mathbb{E}\left(Z_\infty\right) = E\left(Z_0\right) = \prod_{i=0}^{k-1}\left(\frac{a+j\alpha}{a+b+\alpha j}\right) = \prod_{i=0}^{k-1}\left(\frac{\frac{a}{\alpha}+j}{\frac{a}{\alpha}+\frac{b}{\alpha}+j}\right) \quad a.s$$

Finally

$$\begin{aligned}
\varphi(t) = \mathbb{E}\left(e^{itM_\infty}\right) &= \sum_{k=0}^{\infty}\frac{t^k\varphi^{(k)}(0)}{k!} \\
&= \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\mathbb{E}\left(M_\infty^k\right) \quad , \quad \left(\varphi^{(k)}(0) = i^k\mathbb{E}\left(M_\infty^k\right)\right) \\
&= \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\prod_{i=0}^{k-1}\left(\frac{\frac{a}{\alpha}+j}{\frac{a}{\alpha}+\frac{b}{\alpha}+j}\right) = \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\frac{B(\frac{a}{\alpha}+k,\frac{b}{\alpha})}{B(\frac{a}{\alpha},\frac{b}{\alpha})} \\
&= \frac{1}{B(\frac{a}{\alpha},\frac{b}{\alpha})}\sum_{k=0}^{\infty}\frac{(it)^k}{k!}\int_0^1 x^{\frac{a}{\alpha}+k-1}\left(1-x\right)^{\frac{b}{\alpha}-1}dx \\
&= \frac{1}{B(\frac{a}{\alpha},\frac{b}{\alpha})}\int_0^1 x^{\frac{a}{\alpha}-1}\left(1-x\right)^{\frac{b}{\alpha}-1}\sum_{k=0}^{\infty}\frac{(itx)^k}{k!}dx \\
&= \frac{1}{B(\frac{a}{\alpha},\frac{b}{\alpha})}\int_0^1 e^{itx}x^{\frac{a}{\alpha}-1}\left(1-x\right)^{\frac{b}{\alpha}-1}dx \\
&= \mathbb{E}\left(e^{itC}\right).
\end{aligned}$$

With $B(a,b)$ is the beta function evaluated in (a,b), that is $B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, with Γ the usual gamma function and C is a random variable having a beta distribution with parameters $\frac{a}{\alpha}$ and $\frac{b}{\alpha}$.
We conclude by Lévy's theorem.
Then next corollary is a simple consequence of the theorem(2.3.2).

Corollary 2.3.3
For every $\sigma \in T$

$$X_\sigma \sim Ber\left(\frac{a}{a+b}\right) \tag{2.3.2}$$

proof :

Let $\sigma \in T$

$$\begin{aligned}
\mathbb{E}\left[e^{itX_\sigma}\right] &= \mathbb{E}\left[\mathbb{E}\left(e^{itX_\sigma} \mid \Im_n\right]\right) \\
&= \mathbb{E}\left[p_\sigma e^{it} + 1 - p_\sigma\right] \qquad\qquad ((X_\sigma \mid \Im_n) \sim Ber(p_\sigma)) \\
&= e^{it}\mathbb{E}\left[p_\sigma\right] + 1 - \mathbb{E}\left[p_\sigma\right] \\
&= e^{it}\left[\frac{a}{a+b}\right] + 1 - \left[\frac{a}{a+b}\right] = \mathbb{E}\left[e^{itU}\right]
\end{aligned}$$

with U is $Ber\left(\frac{a}{a+b}\right)$

2.4 Application to exchangeable random variables:

This section reviews some classic two-color urn schemes. There are a basic theorems about the process $\left\{X_{\sigma(n)}\right\}$ and $\left\{p_{\sigma(n)}\right\}$ wich are both consequence of de Finetti's Representation Theorem.

Theorem 2.4.1

The sequence $\{X_{\sigma(n)}\}$ generated by a Pólya urn is exchangeable and its de Finetti measure is a beta with parameters $(\frac{a}{\alpha}, \frac{b}{\alpha})$

Proof :

Let $1 \leq k \leq n$ and $(i_1, \cdots, i_n)$ such that $i_j \in \{0,1\}$ and $\sum_{j=1}^{n} i_j = k$. Then

$$\begin{aligned}
\mathbb{P}\left(X_{\sigma(1)} = i_1, \cdots, X_{\sigma(n)} = i_n\right) &= \frac{\prod_{j=0}^{k-1}(a+\alpha j)\cdot\prod_{j=0}^{n-k-1}(b+\alpha j)}{\prod_{j=0}^{n-1}(a+b+\alpha n)} \\
&= \frac{\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}\right)\Gamma\left(\frac{a+k}{\alpha}\right)\Gamma\left(\frac{b+n-k}{\alpha}\right)}{\Gamma\left(\frac{b}{\alpha}\right)\Gamma\left(\frac{a}{\alpha}\right)\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}+n\right)} \\
&= \int_0^1 x^k(1-x)^{n-k}\left[\frac{\Gamma\left(\frac{a}{\alpha}+\frac{b}{\alpha}\right)}{\Gamma\left(\frac{a}{\alpha}\right)\Gamma\left(\frac{b}{\alpha}\right)}x^{\frac{a}{\alpha}-1}(1-x)^{\frac{b}{\alpha}-1}\right]dx
\end{aligned}$$

Because of de Finetti's Representation Theoreme, this shows that the sequence is exchangeable. Moreover, the unicity of the representation implies that de Finetti measure of the sequence $\{X_{\sigma(n)}\}$ is a beta with parametrers $(\frac{a}{\alpha}, \frac{b}{\alpha})$.

Corollary 2.4.2

For all $\sigma(n) \in T(n)$,

$$\lim_{n\to\infty} \frac{Y_{\sigma(n)}}{\sum_{i=0}^{n} p_{\sigma(i)}} = \alpha \qquad \textit{almost surely.} \tag{2.4.1}$$

Proof

We know that, the sequence $\{X_{\sigma(n)}\}$ is generated by a Pólya urn, hence it's exchangeable and its de Finetti measure is *beta* $\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)$.

Then the law of large numbers (Theorem 1.4.7) imply that,

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n} X_{\sigma(i)}}{n} = \Theta \qquad \text{almost surely.} \qquad \Theta \sim beta\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right).$$

Since

$$\left\{\omega \in \Omega : \lim_{n\to\infty} p_{\sigma(n)} = \Theta \right\} \subseteq \left\{\omega \in \Omega : \lim_{n\to\infty} \frac{\sum_{i=0}^{n} p_{\sigma(i)}}{n} = \Theta \right\}$$

However Theorem 2.3.2 implies that:

$$\lim_{n\to\infty} \frac{\sum_{i=0}^{n} p_{\sigma(i)}}{n} = \Theta \qquad \text{almost surely.} \qquad \Theta \sim beta\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)$$

We conclude that:

$$\lim_{n\to\infty} \frac{Y_{\sigma(n)}}{\sum_{i=0}^{n} p_{\sigma(i)}} = \lim_{n\to\infty} \frac{a + \alpha \sum_{i=0}^{n} X_{\sigma(i)}}{\sum_{i=0}^{n} p_{\sigma(i)}} = \alpha \qquad \text{almost surely.}$$

Lemma 2.4.3

$\forall n \in \mathbb{N}, \forall \sigma, \nu \in T(n)$

$$Y_{\sigma(n)} \stackrel{\mathcal{L}}{=} Y_{\nu(n)} \qquad \textit{and} \qquad p_{\sigma_n} \stackrel{\mathcal{L}}{=} p_{\nu_n}$$

where the symbol $\stackrel{\mathcal{L}}{=}$ *means equality in law.*

Proof :

Let $\sigma, \nu \in T(n)$

We introduce the generating function of $Y_{\sigma(n)}$ and $Y_{\nu(n)}$, which we denote by f_n and g_n

Then we have

$$\begin{aligned}
f_{n+1}(t) &= \mathbb{E}\left[e^{itY_{\sigma(n+1)}}\right] = \mathbb{E}\left[E\left(e^{it(Y_{\sigma(n)}+\alpha X_{\sigma(n)})}\right) \mid \Im_n\right] \\
&= \mathbb{E}\left[e^{itY_{\sigma(n)}}\mathbb{E}\left(e^{it\alpha X_{\sigma(n)}} \mid \Im_n\right)\right] = \mathbb{E}\left[e^{itY_{\sigma(n)}}\left(p_{\sigma(n)}e^{i\alpha t} + 1 - p_{\sigma(n)}\right)\right] \\
&= \left(e^{i\alpha t} - 1\right)\mathbb{E}\left[p_{\sigma(n)}e^{itY_{\sigma(n)}}\right] + \mathbb{E}\left[e^{itY_{\sigma(n)}}\right] \\
&= \frac{e^{i\alpha t} - 1}{i(a+b+\alpha n)} f_n^{'}(t) + f_n(t)
\end{aligned}$$

Finally f_n and g_n solve the system :

$$\begin{cases} h_{n+1}(t) = \frac{e^{i\alpha t}-1}{i(a+b+\alpha n)} h_n^{'}(t) + h_n(t) \\ h_0 = e^{ita} \end{cases}$$

We conclude by induction and by continuity that $f \equiv g$ which proves the theorem.

Remark 2.4.4 *We note that:*
$\{\overleftarrow{\sigma} = \overleftarrow{\nu}\} \Rightarrow \left\{Y_\sigma \overset{a.s}{=} Y_\nu \text{ and } p_\sigma \overset{a.s}{=} p_\nu\right\}$

The next result is and its proof are analogues of those Theorem 2.3.1.

Proposition 2.4.5

We suppose $\alpha = 1$.

1)We fixed $0 < \theta < 1$ and for $\sigma \in T(n)$ we define:

$$N_n^\theta := \frac{(a+b+n-1)!}{(Y_{\sigma(n)}-1)!(a+b+(n-1)-Y_{\sigma(n)})!}\theta^{Y_{\sigma(n)}}(1-\theta)^{a+b+n-Y_{\sigma(n)}}$$

N_n^θ is a non-negative martingale, then $N_\infty^\theta := lim N_n^\theta$ exists almost surely.

Proof :

For all $n \in \mathbb{N}$

$\mathbb{E}\left[N_n^\theta\right] \le (a+b+n-1)! < \infty \ \ a.s$

And

$$
\begin{aligned}
\mathbb{E}\left[N_{n+1}^{\theta}|\Im_n\right] &= \mathbb{E}\left[N_{n+1}^{\theta}\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}\right\}} \mid \Im_n\right] + \mathbb{E}\left[N_{n+1}^{\theta}\mathbb{I}_{\left\{Y_{\sigma(n+1)}=Y_{\sigma(n)}+1\right\}} \mid \Im_n\right] \\
&= N_n^{\theta}\left[\frac{(a+b+n)(1-\theta)}{a+b+n-Y_{\sigma(n)}}\mathbb{P}\left(Y_{\sigma(n+1)} = Y_{\sigma(n)} \mid \Im_n\right)\right] \\
&\qquad + N_n^{\theta}\left[\frac{(a+b+n)\theta}{Y_{\sigma(n)}}\mathbb{P}\left(Y_{\sigma(n+1)} = Y_{\sigma(n)}+1 \mid \Im_n\right)\right] \\
&= N_n^{\theta}\left[\frac{(a+b+n)(1-\theta)}{(a+b+n-Y_{\sigma(n)})}\frac{(a+b+n-Y_{\sigma(n)})}{(a+b+n)} + \frac{\theta(a+b+n)}{Y_{\sigma(n)}}\frac{Y_{\sigma(n)}}{(a+b+n)}\right] \\
&= N_n^{\theta}
\end{aligned}
$$

But this result states that:

N_n^{θ} is a non-negative martingale, then $X_\infty := lim X_n$ exists almost surely.

Example ($Wright - Fisher\ model$)

Consider a population comprising a fixed number $m = a+b+N-1$ of genes in any generation. Each genes is of one out ot two genetic types, A_1 and A_2; each generation will consist of i genes of type A_1 and $(m-i)$ genes of type A_2.

Suppose the genetic composition of a daugthher generation is derived by binomial sampling from the genes of the parent generation. Then, if in the parent generation there are i genes of type A_1, the probability that in the daugther generation there will be j such genes is

$$
\mathbb{P}(i,j) = \binom{a+b+N-1}{j}\left(\frac{i}{a+b+N-1}\right)^j\left(1-\frac{i}{a+b+N-1}\right)^{a+b+N-1-j}
$$

ie if we fixed $i, \mathbb{P}(i,\bullet)$ is binomial$\left(a+b+N-1, \frac{i}{a+b+N-1}\right)$

We consider the Markov chain X such that,

let P the transition matrix in $\mathbb{F} = \{a+b, a+b+1, \cdots, a+b+N-1\}$ defined such that

$$
\mathbb{P}(i,j) = \binom{a+b+N-1}{j}\left(\frac{i}{a+b+N-1}\right)^j\left(1-\frac{i}{a+b+N-1}\right)^{a+b+N-1-j}
$$

Then the Markov chain $X = (\Omega, \Im, (\Im_n)_{n\geq 0}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in\mathbb{F}})$ is a P_x-martingale and $X_\infty := lim X_n$ exist almost surely.

In terms of urn models, we imagine an urn containing m balls, each ball being wither, say, 1 or 0. A second urn is then filled with m balls as follows, balls are drawn one by one, with replacement, from the first urn. As each ball is drawn, a new ball of same color as that drawn is placed in the second urn; this process continues until there are m balls in the second urn. $\mathbb{P}(i,j)$ yields the probability that the second urn contain j balls of color 1, given that the first contained i balls of color 0.

2.4.1 Eggenberger-Pólya distribution

We shall establish the following theorem which is also from P.Hall and C.C.Heyde (1980) theorem 7.2.

Theorem 2.4.6
For $\sigma(n) \in T(n)$:

$$i) \quad Y_{\sigma(n)} \overset{\mathcal{L}}{=} a + \alpha B_n$$

$$ii) \quad p_{\sigma(n)} \overset{\mathcal{L}}{=} \tfrac{a+\alpha B_n}{a+b+\alpha n}$$

With B_n be distributed as binomial(n,ξ) and ξ having a beta distribution on $[0,1]$ with parameters $\frac{a}{\alpha}$ and $\frac{b}{\alpha}$. And the symbol $\overset{\mathcal{L}}{=}$ means equality in law.

Proof :
Let $\sigma(n) \in T(n)$:
i)The sequence $\left\{X_{\sigma(n)}\right\}_{n\geq 0}$ is generated by a Pólya urn, hence is exchangeable and by de Finetti's Representation Theorem there exists a random variable ξ, such that :

$$X_{\sigma(n)} \mid \xi \overset{i.i.d}{\sim} Ber\left(\xi\right) \qquad \text{with} \qquad \xi \sim beta\left(\tfrac{a}{\alpha}, \tfrac{b}{\alpha}\right).$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left(e^{itY_{\sigma(n)}}\right) &= \mathbb{E}\left(E\left(e^{it(a+\alpha\sum_{k=0}^{n-1}X_{\sigma(k)})} \mid \xi\right)\right)\\
&= e^{ita}\mathbb{E}\left(E\left(e^{i\alpha t(\sum_{k=0}^{n-1}X_{\sigma(k)})} \mid \xi\right)\right)\\
&= e^{ita}\mathbb{E}\left(\xi e^{i\alpha t}+1-\xi\right)^{n}\ , \qquad \left(X_{\sigma(n)} \mid \xi \overset{i.i.d}{\sim} Ber\left(\xi\right)\right)\\
&= e^{ita}\int_0^1 \left(xe^{it\alpha}+1-x\right)^{n} f_{\xi}(x)dx\\
&= e^{ita}\int_0^1 \mathbb{E}\left(e^{it\alpha B_n} \mid \xi = x\right) f_{\xi}(x)dx\\
&= \mathbb{E}\left(e^{it(a+\alpha B_n)}\right)\ , \qquad \left(B_n \sim \text{binomial}\left(n,\xi\right)\right)
\end{aligned}$$

We conclude by Lèvy theorem's that,

$$Y_{\sigma(n)} \sim a+\alpha B_n, \qquad \text{with} \quad B_n \sim \text{binomial}\left(n,\xi\right) \quad \textit{and} \quad \xi \sim \textit{beta}\left(\frac{a}{\alpha},\frac{b}{\alpha}\right).$$

The following result concerning the distribution of $Y_{\sigma(n)}$ is due to Eggenberger and Pólya.

Corollary 2.4.7 (*Eggenberger – Pólya distribution*)
For all $\sigma \in T(n)$:

$$\mathbb{P}(Y_{\sigma(n)} = a+\alpha k) = \binom{n}{k}\frac{B(k+\frac{a}{\alpha}, n-k+\frac{b}{\alpha})}{B(\frac{a}{\alpha},\frac{b}{\alpha})}, \qquad 0 \le k \le n.$$

Proof : We have that

$$\begin{aligned}
\mathbb{P}(Y_{\sigma(n)} = a+\alpha k) &= \mathbb{P}(B_n = k) = \int_0^1 P(B_n = k \mid \xi = x) f_{\xi}(x)dx\\
&= \int_0^1 \binom{n}{k}x^k(1-x)^{n-k}\frac{x^{\frac{a}{\alpha}-1}(1-x)^{\frac{b}{\alpha}-1}}{B(\frac{a}{\alpha},\frac{b}{\alpha})}dx = \int_0^1 \binom{n}{k}\frac{x^{\frac{a}{\alpha}+k-1}(1-x)^{n-k+\frac{b}{\alpha}-1}}{B(\frac{a}{\alpha},\frac{b}{\alpha})}dx\\
&= \binom{n}{k}\frac{B(k+\frac{a}{\alpha}, n-k+\frac{b}{\alpha})}{B(\frac{a}{\alpha},\frac{b}{\alpha})}, \qquad 0 \le k \le n.
\end{aligned}$$

Other ways of expressing the probabilities are discussed in Johnson. N.L, Kotz. S, and Kemp A.W (1992). The pgf is

$$G(z) = \frac{{}_2F_1\left(-n, a/\alpha; -n+1-b/\alpha, z\right)}{{}_2F_1\left(-n, a/\alpha; -n+1-b/\alpha, 1\right)}$$

where

$$ {}_2F_1(\lambda,\gamma;\eta,z) = 1 + \frac{\lambda\gamma}{\eta}\cdot\frac{z}{1!} + \frac{\lambda(\lambda+1)\gamma(\gamma+1)}{\eta(\eta+1)}\cdot\frac{z^2}{2!} + \cdots \tag{2.4.2} $$

is a Gaussian hypergeometric series (see N.L. Johnson e al (1992) , chapter 6).

Theorem 2.4.6 can be applied to obtain the following lemma.

Lemma 2.4.8

$$ i)\quad \mathbb{E}\left(Y_{\sigma(n)}\right) = a + \frac{\alpha an}{a+b} \quad \text{and} \quad \mathbb{V}ar\left(Y_{\sigma(n)}\right) = \frac{\alpha^2 nab(\frac{a+b}{\alpha}+n)}{(a+b)^2(\frac{a+b}{\alpha}+1)} $$

$$ ii)\quad \mathbb{E}\left(p_{\sigma(n)}\right) = \frac{a}{a+b} \quad \text{and} \quad \mathbb{V}ar\left(p_{\sigma(n)}\right) = \frac{\alpha^2 nab(\frac{a+b}{\alpha}+n)}{(a+b+\alpha n)^2(a+b)^2(\frac{a+b}{\alpha}+1)} $$

Proof :

First method :

Let $\sigma \in T(n)$, we have

$$ \begin{aligned} m_n &= \mathbb{E}\left(Y_{\sigma(n)}\right) = \mathbb{E}\left(\mathbb{E}\left[Y_{\sigma(n)} \mid \Im_{n-1}\right]\right) \\ &= \mathbb{E}\left(\mathbb{E}\left[Y_{\sigma(n)} I_{\left\{Y_{\sigma(n)}=Y_{\sigma(n-1)}+\alpha\right\}} \mid \Im_{n-1}\right]\right) + \mathbb{E}\left(\mathbb{E}\left[Y_{\sigma(n)} I_{\left\{Y_{\sigma(n)}=Y_{\sigma(n-1)}\right\}} \mid \Im_{n-1}\right]\right) \\ &= \mathbb{E}\left(\left[Y_{\sigma(n-1)}+\alpha\right]\mathbb{P}\left(Y_{\sigma(n)}=Y_{\sigma(n-1)}+\alpha \mid \Im_{n-1}\right)\right) + \mathbb{E}\left(Y_{\sigma(n-1)}\mathbb{P}\left(Y_{\sigma(n)}=Y_{\sigma(n-1)} \mid \Im_{n-1}\right)\right) \\ &= \mathbb{E}\left((Y_{\sigma(n-1)}+\alpha)\left[\frac{Y_{\sigma(n-1)}}{a+b+\alpha(n-1)}\right]\right) + \mathbb{E}\left(Y_{\sigma(n-1)}\left[\frac{a+b+\alpha(n-1)-Y_{\sigma(n-1)}}{a+b+\alpha(n-1)}\right]\right) \\ &= \mathbb{E}\left(\frac{Y_{\sigma(n-1)}(a+b+\alpha n)}{a+b+\alpha(n-1)}\right) = \left(\frac{a+b+\alpha n}{a+b+\alpha(n-1)}\right)\mathbb{E}\left(Y_{\sigma(n-1)}\right) \\ &= \left(\frac{a+b+\alpha n}{a+b+\alpha(n-1)}\right) m_{n-1} \end{aligned} $$

Then

$$ \begin{aligned} m_n &= \left(\frac{a+b+\alpha n}{a+b+\alpha(n-1)}\right) m_{n-1} = \left(\frac{a+b+\alpha n}{a+b+\alpha(n-1)}\right)\left(\frac{a+b+\alpha n-1}{a+b+\alpha(n-2)}\right) m_{n-2} \\ &= \left(\frac{a+b+\alpha n}{a+b+\alpha(n-1)}\right)\left(\frac{a+b+\alpha(n-1)}{a+b+\alpha(n-2)}\right)\cdots\left(\frac{a+b+\alpha}{a+b}\right) m_0 \\ &= \left(\frac{a+b+\alpha n}{a+b}\right) a = a + \frac{\alpha an}{a+b} \qquad \left(\ P(Y_{\sigma(0)}=a)=1\ a.s\ \right) \end{aligned} $$

And by a similar computation of $\mathbb{E}\left(Y^2_{\sigma(n)}\right)$

$$ \mathbb{V}ar\left(Y_{\sigma(n)}\right) = \frac{\alpha^2 nab(a+b+n)}{(a+b)^2(a+b+1)} $$

Second method :

We have,

$$\begin{aligned}
\mathbb{E}(B_n) &= \int_0^1 \mathbb{E}(B_n \mid \xi = x) f_\xi(x) = \\
&= \frac{1}{B(\frac{a}{\alpha}, \frac{a}{\alpha})} \int_0^1 n x^{\frac{a}{\alpha}} (1-x)^{\frac{b}{\alpha}-1} dx \\
&= \frac{na}{a+b}
\end{aligned}$$

by similar computation of $\mathbb{E}\left(B_n^2\right)$.

$$\mathbb{V}ar\,(B_n) = \frac{nab(\frac{a+b}{\alpha}+n)}{(a+b)^2(\frac{a+b}{\alpha}+1)}$$

Therefore

$$\mathbb{E}\left(Y_{\sigma(n)}\right) = a + \frac{\alpha a}{a+b}.$$

and

$$\mathbb{V}ar\left(Y_{\sigma(n)}\right) = \alpha^2 \frac{nab(\frac{a+b}{\alpha}+n)}{(a+b)^2(\frac{a+b}{\alpha}+1)}.$$

ii) Since

$p_{\sigma(n)} = \frac{Y_{\sigma(n)}}{a+b+\alpha n}$ then $(i) \Leftrightarrow (ii)$.

2.5 On the number of consecutive successes in the n trials

In this section we are interested in the distributions of the number of two consecutive successes in the n: trials , in other words, the distributions of the number of runs of length 2.

For $n \geqslant 1$, let use define

$$S_n = \#\{k \geqslant 0 : Y_{\sigma(k+2)} = Y_{\sigma(k)} + 2,\ k \leq n-1\} = \sum_{i=0}^{n-1} X_{\sigma(i)} X_{\sigma(i+1)}$$

to be the number of two consecutive successes until level n, with $Y_{\sigma(n)})$ describing the the total number of balls of color 1 contained in the urn indexed by the node σ at generation n. Hence we have the following result.

Theorem 2.5.1

$$\mathbb{P}\left(S_n = i\right) = \sum_{r=i}^{\infty}(-1)^{r-i}\binom{r}{i}\sum_{k=1}^{r}\binom{r-1}{r-k}\binom{n-r}{k}\frac{B\left(\frac{a}{\alpha}+k+r,\frac{b}{\alpha}\right)}{B\left(\frac{a}{\alpha},\frac{b}{\alpha}\right)} \tag{2.5.1}$$

Proof : (see chapter 4, theorem(4.2.2)).

2.6 First Passage Time and Branching process:

2.6.1 First Passage Time

We define the first passage time :

$$T(k) = \inf\left\{r > 0 : Y_{\sigma(r)} = a + \alpha k\right\} \qquad k = 0, 1, \cdots, r.$$

with $Y_{\sigma(n)} = a + \alpha\sum_{i=0}^{n-1} X_{\sigma(i)}$.
We note :

$$D_{\sigma(n)} := n - \sum_{i=0}^{n-1} X_{\sigma(i)} = n - B_{\sigma(n)} :$$

the total number of balls of color 0 added in the urn indexed by the node $\sigma(n)$ all the way to generation $n-1$.
We set

$$R_{\sigma(n)} = B_{\sigma(n)} - D_{\sigma(n)} :$$

the difference of the total number of balls of color 1 and color 0 added in the urn indexed by the node $\sigma(n)$ all the way to generation $n-1$.
As a prelude to the results of this section we give the next lemmas.

Lemma 2.6.1 *Let* j, k, r, s *be integers with* $0 \le r \le s,\ 0 \le r \le j \le k \le n.$ *Then*

$$\mathbb{P}\left(Y_{\sigma(j)} = a + \alpha r \mid Y_{\sigma(k)} = a + \alpha s\right) = \frac{\binom{j}{r}\binom{k-j}{s-r}}{\binom{k}{s}} \tag{2.6.1}$$

Proof :
Since

$$\mathbb{P}(Y_{\sigma(n)} = a + \alpha k) = \binom{n}{k}\frac{B(k+\frac{a}{\alpha}, n-k+\frac{a}{\alpha})}{B(\frac{a}{\alpha},\frac{a}{\alpha})}, \qquad 0 \le k \le n.$$

and

$$\begin{aligned}
\mathbb{P}\left(Y_{\sigma(j)} = a + \alpha r, \quad Y_{\sigma(k)} = a + \alpha s\right) &= \mathbb{P}\left(\sum_{i=1}^{j} X_{\sigma(i)} = r, \sum_{i=1}^{k} X_{\sigma(i)} = s\right) \\
&= \mathbb{P}\left(\sum_{i=1}^{j} X_{\sigma(i)} = r, \sum_{i=j+1}^{k} X_{\sigma(i)} = s - r\right) \\
&= \binom{j}{r}\binom{k-j}{s-r}\frac{B(k+\frac{a}{\alpha}, n-k+\frac{b}{\alpha})}{B(\frac{a}{\alpha}, \frac{b}{\alpha})}
\end{aligned}$$

The conclusion in the lemma now follows from the definition of conditional probability.

Lemma 2.6.2 *Let $\{X_i, 1 \le i \le n\}$ be nonnegative integer-valued $\mathcal{L}_1$ exchangeables random variables and set*

$$B_n = \sum_{i=1}^{n} X_i$$

Then

$$\mathbb{P}\left(B_j < j, 1 \le j \le n \mid B_n\right) = \left(1 - \frac{B_n}{n}\right)^{+} \tag{2.6.2}$$

Proof(see Y.S.Chow and H.Teicher "Probability theory" page 242).

Theorem 1.4.2 can be applied to obtain the following result

Theorem 2.6.3

For $k = 0, 1, 2, \cdots$,

$$\mathbb{P}\left(T(k) = n\right) = \begin{cases} \binom{n-1}{k-1}\frac{B\left(n-k+\frac{b}{\alpha}, k+\frac{a}{\alpha}\right)}{B\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)} & \textit{if} \;\; k \le n \\ 0 & \textit{if} \;\; k > n \end{cases} \qquad n \geqslant 1 \tag{2.6.3}$$

First proof

We have,

$$\begin{aligned}
\{\omega : T(k)=n\} &= \{\omega : Y_{\sigma(0)} < a+\alpha k, \cdots, Y_{\sigma(n-1)} < a+\alpha k, Y_{\sigma(n)} = a+\alpha k\} \\
&= \{\omega : Y_{\sigma(0)} < a+\alpha k, \cdots, Y_{\sigma(n-1)} = a+\alpha(k-1), Y_{\sigma(n)} = a+\alpha k\} \\
&= \{\omega : Y_{\sigma(n-1)} = a+\alpha(k-1), Y_{\sigma(n)} = a+\alpha k\} \\
&= \{\omega : \sum_{i=0}^{n-2} X_{\sigma(i)} = k-1, \sum_{i=0}^{n-1} X_{\sigma(i)} = k\} \\
&= \{\omega : \sum_{i=0}^{n-2} X_{\sigma(i)} = k-1, X_{\sigma(n-1)} = 1\}
\end{aligned}$$

Therefore, if $k > n, \{\omega : T(k)=n\} = \emptyset$, and $\mathbb{P}(T(k)=n) = 0$.

If $k \leq n$,

$$\begin{aligned}
\mathbb{P}(T(k)=n) &= \mathbb{P}\left(\sum_{i=0}^{n-2} X_{\sigma(i)} = k-1, X_{\sigma(n-1)} = 1\right) \\
&= \int_\Omega \mathbb{P}\left(\sum_{i=0}^{n-2} X_{\sigma(i)} = k-1, X_{\sigma(n-1)} = 1 \mid \Theta\right) dP_\Theta \\
&= \int_\Omega \mathbb{P}\left(\sum_{i=0}^{n-2} X_{\sigma(i)} = k-1 \mid \Theta\right) \mathbb{P}\left(X_{\sigma(n-1)} = 1 \mid \Theta\right) dP_\Theta \\
&= \int_\Omega \binom{n-1}{k-1} \Theta^{k-1} (1-\Theta)^{n-k} \Theta dP_\Theta \\
&= \int_\Omega \binom{n-1}{k-1} \Theta^{k} (1-\Theta)^{n-k} dP_\Theta \\
&= \int_0^1 \binom{n-1}{k-1} x^{k} (1-x)^{n-k} f_\Theta(x) dx \\
&= \frac{\binom{n-1}{k-1}}{B\left(\frac{a}{\alpha}, \frac{a}{\alpha}\right)} \int_0^1 x^k (1-x)^{n-k} x^{\frac{a}{\alpha}-1} (1-x)^{\frac{b}{\alpha}-1} dx \\
&= \frac{\binom{n-1}{k-1}}{B\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)} \int_0^1 x^{\frac{a}{\alpha}+k-1} (1-x)^{n-k+\frac{b}{\alpha}-1} dx \\
&= \binom{n-1}{k-1} \frac{B\left(n-k+\frac{b}{\alpha}, k+\frac{a}{\alpha}\right)}{B\left(\frac{b}{\alpha}, \frac{a}{\alpha}\right)}
\end{aligned}$$

Second proof

We have,

$$\begin{aligned}
\{\omega : T(k)=n\} &= \{\omega : Y_{\sigma(0)} < a+\alpha k, \cdots, Y_{\sigma(n-1)} < a+\alpha k, Y_{\sigma(n)} = a+\alpha k\} \\
&= \{\omega : Y_{\sigma(0)} < a+\alpha k, \cdots, Y_{\sigma(n-1)} = a+\alpha(k-1), Y_{\sigma(n)} = a+\alpha k\} \\
&= \{\omega : Y_{\sigma(n-1)} = a+\alpha(k-1), Y_{\sigma(n)} = a+\alpha k\}
\end{aligned}$$

Therefore, if $k > n, \{\omega : T(k) = n\} = \emptyset$, and $\mathbb{P}\left(T(k) = n\right) = 0$ If $k \leq n$,

$$\begin{aligned}
\mathbb{P}\left(T(k) = n\right) &= \mathbb{P}\left(Y_{\sigma(n-1)} = a + \alpha(k-1), Y_{\sigma(n)} = a + \alpha k\right) \\
&= \mathbb{P}\left(Y_{\sigma(n-1)} = a + \alpha(k-1) / Y_{\sigma(n)} = a + \alpha k\right) \mathbb{P}\left(Y_{\sigma(n)} = a + \alpha k\right) \\
&= \frac{\binom{n-1}{k-1}\binom{n-(n-1)}{k-(k-1)}}{\binom{n}{k}} \mathbb{P}\left(Y_{\sigma(n)} = a + \alpha k\right) \qquad \text{(lemma 2.6.1)} \\
&= \frac{\binom{n-1}{k-1}}{\binom{n}{k}} \binom{n}{k} \frac{B(k + \frac{a}{\alpha}, n - k + \frac{a}{\alpha})}{B(\frac{a}{\alpha}, \frac{a}{\alpha})} \\
&= \binom{n-1}{k-1} \frac{B\left(n - k + \frac{b}{\alpha}, k + \frac{a}{\alpha}\right)}{B\left(\frac{a}{\alpha}, \frac{a}{\alpha}\right)}
\end{aligned}$$

Third proof

If $k > n, \{\omega : T(k) = n\} = \emptyset$ then $\mathbb{P}\left(T(k) = n\right) = 0$. If $k \leq n$, we have that

$$\begin{aligned}
\mathbb{P}\left(T(k) = n\right) &= \mathbb{P}\left(r - D_{\sigma(r)} < r \text{ for } 0 \leq r < n \text{ and } D_{\sigma(n)} = n - k\right) \\
&= \mathbb{P}\left(D_{\sigma(n)} - D_{\sigma(r)} < n - r \text{ for } 0 \leq r < n \text{ and } D_{\sigma(n)} = n - k\right) \\
&= \mathbb{P}\left(D_{\sigma(i)} < i \text{ for } 0 \leq i \leq n \text{ and } D_{\sigma(n)} = n - k\right) \\
&= \frac{k}{n} \mathbb{P}\left(D_{\sigma(n)} = n - k\right) \qquad \text{(lemma 2.6.2)} \\
&= \frac{k}{n} \binom{n}{n-k} \frac{B\left(n - k + b, n - (n-k) + a\right)}{B\left(b, a\right)} \\
&= \binom{n-1}{k-1} \frac{B\left(n - k + b, k + a\right)}{B\left(b, a\right).}
\end{aligned}$$

2.6.2 Branching process

Around 1873, Galton and Watson came up with a model for explaining the desappearance of certain family names in England. Their model , now known as the Galton-Watson process, is extremely simple (see the historical survey by Kandall, 1966). They describe the evolution in discrete time of a population of individuals who reproduce themselves according to an offpring distribution μ on $\mathbb{N}$, the associated Galton-Watson process is the Markov chain $(N_k, k \geqslant 0)$ with values in $\mathbb{N}$ such that, conditionally on N_n,

$$N_{n+1} \stackrel{(d)}{=} \sum_{i=1}^{N_n} \xi_i$$

where the variables ξ_i are i.i.d, with distribution μ and the symbol $\stackrel{(d)}{=}$ means equality in distribution. We consider a two-type branching process in which there are no deaths and each individuals gives birth at rate 1. If at some time there are a individuals of type 1 and b individuls of type 2, then the probability that the next individuals born will have type 1 is $a/(a+b)$. It follows that the distribution of the number type 1 individuals when the population size reaches n is the same as the distribution the number of white balls in the urn σ at generation n, equal to $Y_{\sigma(n)}$. We then have the following result.

Proposition 2.6.4
For $n \in \mathbb{N}$ and $\sigma \in T$, we define:

$$H_n := \# \left\{ r \geq 0 : R_{\sigma(r)} = n+1, R_{\sigma(r-1)} = n \right\}.$$

Then conditionally to ξ , $H_n, n \in \mathbb{N}$ is a Galton-Watson process with a geometric reproduction distribution with parameter ξ. With ξ having a beta distribution on $[0,1]$ with parameters $\frac{a}{\alpha}$ and $\frac{a}{\alpha}$.

Proof :
Since the sequence $\left\{X_{\sigma(n)}\right\}$ is generated by a Polya urn, hence it's exchangeable and its de Finetti measure is *beta* $\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)$.
Fix $n \in \mathbb{N}$. We define the following stopping times:

$$\begin{aligned} \sigma_0 &= 0 \\ \sigma_i &= \inf\left\{s > \sigma_{i-1} : R_s = n+1 \ \text{ and } \ R_{s-1} = n\right\} \\ \tau_i &= \inf\left\{s > \sigma_i : R_s = n\right\} \end{aligned}$$

Observe that $\tau < \sigma_{i+1}$ and $((\sigma_i < \infty) \Leftrightarrow (H_n \geq i))$.

We note:

$$H_{[s,t]}^{n+1,n+2} = \{r : R_r = n+1, R_{r+1} = n+2, s \leq r \leq t\}$$

Obviously, $H_{[\sigma_i,\sigma_{i+1}]}^{n+1,n+2} = H_{[\sigma_i,\tau_i]}^{n+1,n+2}$. Moreover by Markov property,the random variables $H_{[\sigma_i,\tau_i]}^{n+1,n+2}$ are independend and identically distributed as N_1 and we have:

$$H_{n+1} = \sum_{i=1}^{H_n} H_{[\sigma_i,\tau_i]}^{n+1,n+2} \tag{2.6.4}$$

and this proves the branching property of N.

To identify the law of the Galton-Watson process, we only need to compute the distribution of H_1. Let us consider the following stopping times

$$S = \inf\{s > 0, R_s = 2 \text{ and } R_s = 1\}$$

$$T = \inf\{s > S,\ R_s = 1\}$$

One has $H_1 = 0$ if $S = \infty$.

By the Markov property

$$\mathbb{P}(H_1 = 0 \mid \xi) = \mathbb{P}(S = +\infty \mid \xi) = \xi$$

$$\begin{aligned} \forall n > 0,\ \mathbb{P}(H_1 = n \mid \xi) &= \mathbb{P}(S = \infty \mid \xi)\, \mathbb{P}(H_1 = n-1 \mid \xi) \\ &= \xi \mathbb{P}(H_1 = n-1 \mid \xi) \\ &= \xi^n \end{aligned}$$

2.7 Martingales Product

Our last result in this chapter is given in the following proposition.

Let use define

$$W_n = \prod_{i=0}^{n} \frac{X_{\sigma(i)}}{p_{\sigma(i)}}, \qquad n \in \mathbb{N}.$$

Proposition 2.7.1

The process $(W_n, n \geq 0)$ *:*

i)Is a non-negative martingale with respect to filtration $\Im_n$*, so that*

$$W_\infty := \lim_{n\to\infty} W_n \quad \textit{exists} \quad \textit{a.s} \quad \textit{and} \qquad P(W_\infty = 0) = 1.$$

ii)Is a Markov chain with value in

$$\tilde{E} = \left\{ \prod_{i=0}^{n} \frac{a+b+i\alpha}{a+i\alpha}, n \geq 0 \right\} \bigcup \{0\}$$

The transition kernel of the Markov chain is described as follows

$\forall a_n, b_n \in E$

$$Q\left(a_n, b_n\right) = \begin{cases} \prod_{i=0}^{n} \frac{a+i\alpha}{a+b+i\alpha} & \textit{if} \;\; b_n(a+\alpha n) = a_n(a+b+\alpha n) \\ 1 - \prod_{i=0}^{n} \frac{a+i\alpha}{a+b+i\alpha} & \textit{if} \;\; b_n = 0 \end{cases}$$

And $\quad Q\left(0, 0\right) = 1.$

Proof :

i)for $n \geq 0$, we have (a.s)

$$\begin{aligned}
\mathbb{E}[W_n] &= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=0}^{n}\frac{X_{\sigma(i)}}{p_{\sigma(i)}} \mid \Im_{n-1}\right)\right] \\
&= \mathbb{E}\left[\left(\prod_{i=0}^{n-1}\frac{X_{\sigma(i)}}{p_{\sigma(i)}}\right) E\left(\frac{X_n}{p_{\sigma(n)}} \mid \Im_{n-1}\right)\right] \\
&= \mathbb{E}\left[\left(\prod_{i=0}^{n-1}\frac{X_{\sigma(i)}}{p_{\sigma(i)}}\right) \frac{1}{p_{\sigma(n)}} E\left(X_n \mid \Im_{n-1}\right)\right] \\
&= \mathbb{E}[W_{n-1}] = \cdots = \mathbb{E}[W_0] \\
&= \mathbb{E}\left[\left(\frac{a+b}{a}\right) X_{\sigma(0)}\right] = 1.
\end{aligned}$$

And

$$\begin{aligned}
\mathbb{E}(W_{n+1}/\Im_n) &= \mathbb{E}\left(\prod_{i=0}^{n+1}\frac{X_{\sigma(i)}}{p_{\sigma(i)}} \mid \Im_n\right) \\
&= \left(\prod_{i=0}^{n}\frac{X_{\sigma(i)}}{p_{\sigma(i)}}\right) E\left(\frac{X_{n+1}}{p_{\sigma(n+1)}}\right) \\
&= \left(\prod_{i=0}^{n}\frac{X_{\sigma(i)}}{p_{\sigma(i)}}\right) \left(\frac{1}{p_{\sigma(n+1)}}\right) \mathbb{E}(X_{n+1}/\Im_n) \\
&= W_n
\end{aligned}$$

So that W is non-negative martingale, then $W_\infty := \underset{n\to\infty}{lim} W_n$ exists almost surely in $[0,\infty]$.

ii)We have for all $(a_0, a_2, \cdots, a_n) \in E^{n+1}$:

$$\mathbb{P}(W_n = a_n/W_{n-1} = a_{n-1}, \cdots, W_0 = a_0) = \mathbb{P}(W_0 = a_0)\prod_{i=1}^{n}\mathbb{P}(W_i = a_i/W_{i-1} = a_{i-1}).$$

and

$$\begin{aligned}
\{W_n \neq 0\} &\Leftrightarrow \left\{X_{\sigma(i)} = 1, i = 0, 1, \cdots, n\right\}. \\
&\Leftrightarrow p_{\sigma(i)} = \frac{a+\alpha i}{a+b+\alpha i}, \forall i = 0, 1, \cdots, n. \\
&\Leftrightarrow W_{n-1} \neq 0 \text{ and } W_n = W_{n-1}\frac{a+b+\alpha n}{a+\alpha n}
\end{aligned}$$

We conclude that $(W_n)_{n\geq 0}$, is a Markov chain with value in

$$E = \left\{\prod_{i=0}^{n}\frac{a+b+i\alpha}{a+i\alpha}, n \geq 0\right\} \bigcup \{0\}$$

with the transition kernel is describe as follows

$\forall a_n, b_n \in E$

$$Q\left(a_n, b_n\right)=\begin{cases}\prod_{i=0}^{n} \frac{a+i\alpha}{a+b+i\alpha} & \text{if } \ b_n(a+\alpha n)=a_n(a+b+\alpha n) \\ 1-\prod_{i=0}^{n} \frac{a+i\alpha}{a+b+i\alpha} & \text{if } \ b_n=0\end{cases}$$

Finally we have , 0 is the unique recurrent state, in fact a state 0 is absorbing, so $\mathbb{P}\left(W_{\infty} \in\{0, \infty\}\right)=1.\ a.s$

But by (i) $\{W_n\}$ is Martingale with respect to the filtration $\Im_n$.

So

$$\mathbb{E}\left(W_{\infty}\right)=\mathbb{E}\left(W_n\right)=\mathbb{E}\left(W_0\right)=1.$$

Therefore

$$\mathbb{P}\left(W_{\infty}=0\right)=1.\ a.s$$

2.8 A Dirichlet reinforced process indexed by binary tree

In this section we study the connection between the Dirichlet distribution and the reinforced stochastic process $X=\{X_\tau, \sigma \in T\}$ of random variable indexed by the vertices of tree T and with value in a Polish space S endowed with its Borel sigma-field $\mathcal{S}$.

As before, we assume that the random variable of the process X are defined on a rich enough probability space $(\Omega, \Im, P)$, and we specify the law of X recursively on the levels of the tree T.

Let G_0 be a probability distribution on S and $c>0$ a constant.

Set X_θ to be a random variable with value in S and probability distribution G_0 .

For $n \geqslant 0$ let $\Im_n \subseteq \Im$ be the sigma-field generated by the random variables X_σ with $|\sigma| \leq n$. Given $\Im_n$, assume that the $n+1$ random variables X_τ has values in S and probability distribution

$$G_\tau=\frac{cG_0+\sum_{i=0}^{n} \delta_{X_{\sigma_i}}}{c+n+1} \tag{2.8.1}$$

If $\pi(\theta, \tau)=(\theta=\sigma_{[0]}, \sigma_{[1]}, \cdots, \sigma_{(n-1]}, \tau)$; for $s \in S$, δ_s indicates the point mass at s.

We set $S=\mathcal{R}$, $G_0=\frac{b}{a+b}\delta_0+\frac{a}{a+b}\delta_1$ and $c=a+b$, for every $B \in \mathcal{S}$, the process $\{I_{X_\sigma \in B}: \sigma \in T\}$ is a dichotomous reinforced process indexed by the tree T.

Given an end $\varepsilon = (\theta = \varepsilon_0, \varepsilon_1, \cdots) \in E$, the sequence of random variables $X_\varepsilon = (X_{\varepsilon_0}, X_{\varepsilon_1}, \cdots)$ is a Polya sequence with parameter $\alpha = cG_0$. In fact the space S will be that for the colors of the balls in the urn which possibly contains a continuum of color , the intial color X_θ is drawn from G_0. At each level $n = 1, 2, \cdots$, as the ball is sampled from the urn and replaced in it along with another ball of the same color. For $n \geqslant 1$, let $X_{\varepsilon_n} = x$ if x is the color sampled in the urn indexed by the node ε_n at the n-th generation. Therefore, for all $B \in \mathcal{X}$

$$\mathbb{P}\left(X_{\varepsilon_1} \in B\right) = \frac{cG_0(B) + \delta_{X_0}}{c+1}$$

whereas for all $n \geqslant 1$, the conditional distribution of $X_{\varepsilon_{n+1}}$ given $X_{\varepsilon_0}, \cdots, X_{\varepsilon_n}$ is such that

$$\mathbb{P}\left(X_{\varepsilon_{n+1}} \in B \mid X_{\varepsilon_0}, \cdots, X_{\varepsilon_n}\right) = \frac{G_0(B) + \sum_{i=0}^{n} \delta_{X_{\varepsilon_i}}}{c+n}$$

for all $B \in \mathcal{X}$.

The sequence of the random variables $X_\varepsilon = (X_{\varepsilon_0}, X_{\varepsilon_1}, \cdots)$ is exchangeable and that the random distribution function of the sequence G_{ε_n} weakly converges to a Dirichlet process G_ε with parameter cG_0on a set of probability one. Moreover, given G_ε, the random variables of the Polya sequence $X_\varepsilon = (X_{\varepsilon_0}, X_{\varepsilon_1}, \cdots)$ are independent and identically distributed with distribution G_ε; see Muliere et al (2005) for details.

Chapter 3

Generalization of binary tree:

3.1 introduction

In this chapter we propose a new stochastic process indexed by the vertices of a different kind of the tree T. As seen in the Introduction T is obtained by starting from a binary tree so that each vertex splits into two children, but vertices at the side of each other have a child in common. Hence it turns out that each vertex has two parents and two children except the root (no parents) and the vertices long the left and right extreme branches (just one parent).In fact the tree T is similar to the tree defined in chapter 1 but for this tree, evry node $\sigma \in T$, the left (resp: the right) child of σ (resp:of neighbour of σ) go the same point.
We note that ,given a node σ in T the path $\pi(\theta, \sigma)$ in the tree T connecting the node σ with the root is not unique.

Can we utilised the Chauvin-Neveu representation we have:

The tree T is a sub set of $U = \bigcup_{n\geq 0} \{0,1\}^n$ such that:

$$\begin{cases} \theta \in T \\ uv \in T \Rightarrow u \in T \\ u0 \in T \Leftrightarrow u1 \in T \\ \forall u, \nu \in T: \text{ if } \; u = u_1 \cdots u_n, \nu = \nu_1 \cdots \nu_n, \text{then: } \left(\sum_{i=1}^n u_i = \sum_{i=1}^n \nu_i\right) \Leftrightarrow (u = \nu) \,. \end{cases}$$

.

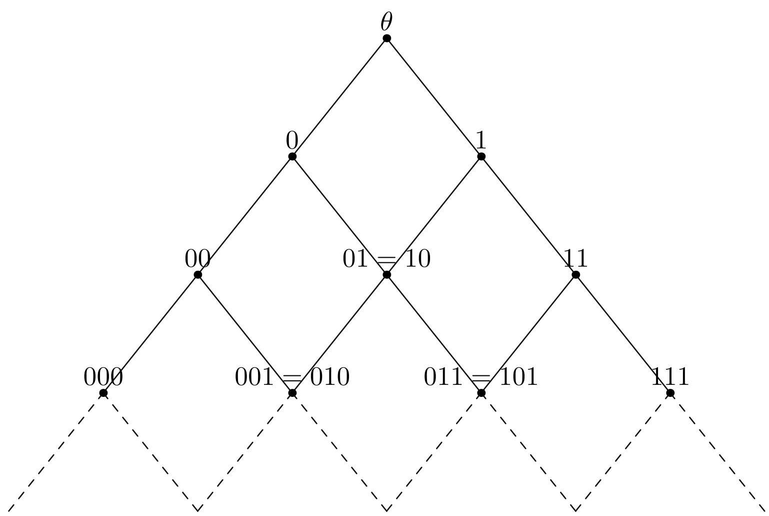

Figure 3.1: The binary tree T

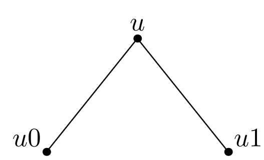

Figure 3.2: The children of u

Let $A(\sigma) := \{\tau \in T : \tau \text{ is ancestor of } \; \sigma\} = \{\tau \in T : \tau \in \pi(\theta, \sigma) \;\text{ for some }\; \pi(\theta, \sigma)\}$

We are now ready for introduction the Muliere-Secchi-Walker process $X = \{X_\sigma\}$ indexed by the node of binary tree T.

The definition of the law of the process X is recursive on the levels of T. Let a and b be positives reals a positive reals numbers and set

$$X_\theta \sim Ber\left(p_\theta = \frac{a}{a+b}\right)$$

For $n \geq 0$, let $\Im_n \subseteq \Im$ be the sigma-field generated by the random variable X_σ with $|\sigma| < n$, asume that the random variable X_σ with $|\sigma| = n$ are conditionally independent such that

$$X_\sigma \mid \Im_n \sim Ber\left(p_\sigma = \frac{a + \sum\limits_{\tau \in A(\sigma)} X_\tau}{a + b + \#A(\sigma)}\right)$$

where $A(\sigma) := \{\tau \in T : \tau \text{ is ancestor of } \sigma\} = \{\tau \in T : \tau \in \pi(\theta, \sigma) \text{ for some } \pi(\theta, \sigma)\}$.

We note $T(n) := \{\sigma \in T : |\sigma| = n\}$, and for $v = v_1 \cdots v_n \in T(n)$, we write $v(i,n)$ if $\sum_{j=1}^{n} v_j = i$ and by convention $v(0,0) = \theta$.

We fixe n and we define

$$\begin{aligned} v(\cdot, n) : \quad \{0, \cdots, n\} &\longrightarrow T(n) \\ i &\longrightarrow v(i,n) \end{aligned}$$

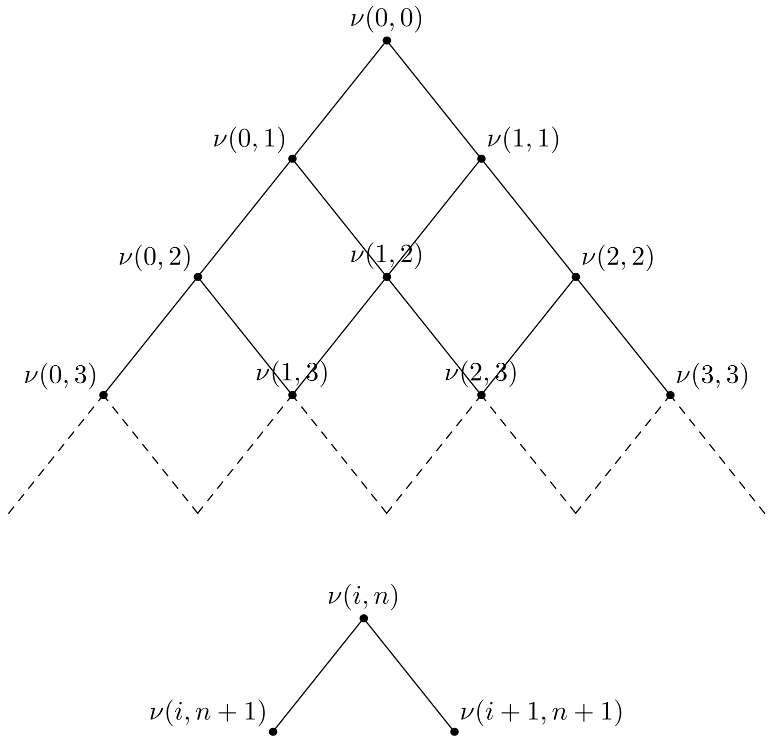

Figure 3.3: The binary tree T ordered

Lemma 3.1.1 *For any $n \geq 0$ and $0 \leq i \leq n$.*

i)

$$\begin{aligned} v(\cdot, n): \quad \{0, \cdots, n\} &\longrightarrow T(n) \\ i &\longrightarrow v(i,n). \end{aligned}$$

is bijiction.

ii) $$A(v(i,n)) = \{v(k, n-j) : k = (i-j)^+ \cdots i \wedge (n-j), \quad j = 1, \cdots, n\}$$

iii) $$Card\ A(v(i,n)) = i(n-i) + n.$$

Proof :

i)For all n, $v(\cdot,n)$ is injective and $Card\,\{0,1,\cdots,n\} = n+1 = Card\ T(n)$.
Therefore

$$\begin{array}{rll} v(\cdot,n): & \{0,\cdots,n\} & \longrightarrow T(n) \\ & i & \longrightarrow v(i,n). \end{array}$$

Is a bijiction.

ii)By construction (see Figure (3.3)).

iii) Trivialy we have that

$$Card\ A(v(0,n)) = n$$

And we have

$$A(v(i,n)) = A(v(i-1,n-1)) \bigcup \left\{ \bigcup_{k=0}^{n-i} \{v(i,i+k)\} \right\}$$

Therefore

$$\begin{aligned} Card\ A(v(i,n)) &= Card\ A(v(i-1,n-1)) + n - i + 1 \\ &= Card\ A(v(i-2,n-2)) + 2(n-i+1) \\ &= \cdots = Card\ A(v(0,n-i) + i(n-i+1) \\ &= n - i + i(n-i+1) \\ &= i(n-i) + n \end{aligned}$$

Using the notation $v(i,n)$ we can write:

$$X_{v(0,0)} \sim Ber\left(p_{v(0,0)}\right) \qquad \text{with} \qquad p_{v(0,0)} = \frac{a}{a+b}.$$

and for any $n > 0$ and $0 \leq i \leq n$:

$$X_{v(i,n)}/\Im_n \overset{\perp}{\sim} Ber\left(p_{v(i,n)}\right) \qquad \text{with} \qquad p_{v(i,n)} = \left(\frac{a + \sum\limits_{\substack{k=(i-j)^+ \cdots i\wedge(n-j) \\ j=1\cdots n}} X_{v(k,n-j)}}{a+b+i(n-i)+n} \right)$$

Lemma 3.1.2 $\forall n \in \mathbb{N}$, $\forall i \leq n$:

$$E\left(p_{v(i,n)}\right) = \tfrac{a}{a+b}.$$

Proof

Trivially we have $E\left(p_{v(0,0)}\right) = \frac{a}{a+b}$, because $p_{v(0,0)} = \frac{a}{a+b}$.

By induction :

Let

$$P(n) = \left(E\left[p_{v(i,m)}\right] = \frac{a}{a+b} \quad , \forall i \leq m \quad , \forall m \leq n \right)$$

We show that :

$$\left(P(n) \overset{?}{\Rightarrow} P(n+1) \right).$$

We suppose $P(n)$ is true. $(\star)$

Since $\{p_{v(n,n)}, \Im_n, n \geqslant 0\}$ is a martingale, therefore for all $n \geqslant 0$, $E\left(p_{v(n+1,n+1)}\right) = \frac{a}{a+b}$.

For $i \leq n$ fixed :

$$
\begin{aligned}
E\left(p_{v(i,n+1)}\right) &= E\left(\frac{a+\sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} X_{v(k,n+1-j)}}{a+b+i(n+1-i)+n+1}\right)\\
&= \left(\frac{a+\sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} E\left(X_{v(k,n+1-j)}\right)}{a+b+i(n+1-i)+n+1}\right)\\
&= \left(\frac{a+\sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} E\left(E\left(X_{v(k,n+1-j)}|\Im_n\right)\right)}{a+b+i(n+1-i)+n+1}\right)\\
&= \left(\frac{a+\sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} E\left(p_{v(k,n+1-j)}\right)}{a+b+i(n+1-i)+n+1}\right)\\
&\stackrel{(\star)}{=} \left(\frac{a+\sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} \frac{a}{a+b}}{a+b+i(n+1-i)+n+1}\right) = \left(\frac{a+\frac{a}{a+b}\left(i(n+1-i)+n+1\right)}{a+b+i(n+1-i)+n+1}\right)\\
&= \frac{a}{a+b}.
\end{aligned}
$$

We conclude that $P(n+1)$ is true.

As a direct consequence the following lemma is got.

Lemma 3.1.3

For all $n \in \mathbb{N}$, *and for all* $l \leq n$

$$X_{v(l,n)} \sim Ber\left(\frac{a}{a+b}\right).$$

Proof

Let $n \in \mathbb{N}$ and $0 \leq l \leq n$

$$\begin{aligned}
\mathbb{E}\left[e^{itX_{v(l,n)}}\right] &= \mathbb{E}\left[\mathbb{E}\left(e^{itX_{v(l,n)}}|\Im_n\right]\right) \\
&= \mathbb{E}\left[p_{v(l,n)}e^{it} + 1 - p_{v(l,n)}\right] \qquad \left(\left(X_{v(l,n)}|\Im_n\right) \sim Ber(p_{v(l,n)})\right) \\
&= e^{it}\mathbb{E}\left[p_{v(l,n)}\right] + 1 - \mathbb{E}\left[p_{v(l,n)}\right] \\
&= e^{it}\left[\frac{a}{a+b}\right] + 1 - \left[\frac{a}{a+b}\right] = \mathbb{E}\left[e^{itU}\right]
\end{aligned}$$

with U is $Ber\left(\frac{a}{a+b}\right)$.

3.2 symmetry of T

The processes $p = \{p_{\upsilon(i,n)}, 0 \leq i \leq n, 0 \leq n\}$ enjoy the property of simmetry in that the random variables associated to vertices symmetric with respect to the vertical axis have the same law.

Lemma 3.2.1 *(Symmetry)*

$\forall n \in \mathbb{N}$, $\forall i \leq n$:

$$i)\ p_{v(i,n)} \stackrel{\mathcal{L}}{=} p_{v(n-i,n)}.$$

$$ii)\ Y_{v(i,n)} \stackrel{\mathcal{L}}{=} Y_{v(n-i,n)}.$$

with $Y_{v(i,n)}$ describe the total number of ball of color1 contained in the urn indexed by the node $v(i,n)$, and the symbol $\stackrel{\mathcal{L}}{=}$ means equality in law.

Proof :

Recall that $A(\tau)$ is the ancestors set of the node τ . Note that if $\tau' \in A(\tau)$, then $A(\tau') \subseteq A(\tau)$. For a node τ such that $|\tau| = n$, let define $P^1(\tau)$ the set of the parents of τ that is the ancestors of τ of level $n-1$ and, more generally, for $j = 1, ..., n$, $P^j(\tau)$ as the set of ancestors of τ at level $n-j$. We have, for such τ with $|\tau| = n$, $A(\tau) = \bigcup_{j=1}^{n} P^j(\tau)$ and also ,for $k = 1, \cdots, n-1$ the recursive relation $\bigcup_{\eta \in P^k(\tau)} A(\eta) = \bigcup_{j=k+1}^{n} P^j(\tau)$.

Now let $s : T(n) \longrightarrow T(n), s(\upsilon(i,n)) = \upsilon(n-i,n)$ be the function relating a node to its symmetric with respect to the vertical symmetry axis of T. Notice that $A(\upsilon(n-i,n)) = s(\tau), \tau \in A(\upsilon(i,n))$. By the definition of the process in section 1, we have, for fixed $\tau \in T$ and given $(x, x_\eta) \in \{0,1\}^2, \eta \in A(\tau))$,

$$\mathbb{P}\left(X_\tau = x / X_\eta = x_\eta, \eta \in A(\tau)\right) = \left(\frac{a + \sum_{\eta \in A(\tau)} x_\eta}{a+b+\#A(\tau)}\right)^x \left(1 - \frac{a + \sum_{\eta \in A(\tau)} x_\eta}{a+b+\#A(\tau)}\right)^x$$

$$\mathbb{P}\left(X_\tau = x | X_{s(\eta)} = x_\eta, \eta \in A(\tau)\right) = \left(\frac{a + \sum_{\eta \in A(\tau)} x_\eta}{a+b+\#A(\tau)}\right)^x \left(1 - \frac{a + \sum_{\eta \in A(\tau)} x_\eta}{a+b+\#A(\tau)}\right)^x$$

so that

$$\mathbb{P}\left(X_\tau = x | X_\eta = x_\eta, \eta \in A(\tau)\right) = \mathbb{P}\left(X_\tau = x | X_{s(\eta)} = x_\eta, \eta \in A(\tau)\right)$$

Let us consider a given node $\upsilon(i,n)$ and fixed $\{x_\tau, \tau \in A(\upsilon(i,n))\}$

$$\begin{aligned}
\mathbb{P}\left(X_\tau = x_\tau, \tau \in A(\upsilon(i,n))\right) &= \prod_{j=1}^{n-1} \prod_{\tau \in P^j(\upsilon(i,n))} \mathbb{P}\left(X_\tau = x_\tau | X_\eta, \eta \in A(\tau)\right) \mathbb{P}\left(X_\theta = x_\theta\right) \\
&= \prod_{j=1}^{n-1} \prod_{\tau \in P^j(u(i,n))} \mathbb{P}\left(X_{s(\tau)} = x_\tau | X_{s(\eta)}, \eta \in A(\tau)\right) \mathbb{P}\left(X_\theta = x_\theta\right) \\
&= \mathbb{P}\left(X_{s(\tau)} = x_\tau, \tau \in A(\upsilon(i,n))\right)
\end{aligned}$$

(the x_η's are the different $x_\tau^{'}$ s corresponding to the different subsets of the above set). So the random vectors $(X_\tau, \tau \in A(\upsilon(i,n)))$ and $(X_\tau, \tau \in A(\upsilon(n-i,n)))$ have the same probability law. As $Y_{\upsilon(i,n)} = a + \sum_{\tau \in A(\upsilon(i,n))} X_\tau$ and $Y_{\upsilon(n-i,n)} = a + \sum_{\tau \in A(\upsilon(n-i,n))} X_\tau$, it turns out that $Y_{\upsilon(i,n)}$ and $Y_{\upsilon(n-i,n)}$ have the same law too. Analogously for $p_{\upsilon(i,n)}$ and $p_{\upsilon(n-i,n)}$..
Now, we will define the sequence $\{S_n, n \geq 0\}$, such that

$$\begin{cases} S_0 = 0 \\ S_n := a_n \left(p_{(n-1,n)} - p_{\upsilon(n,n)}\right) \end{cases} \tag{3.2.1}$$

with

$$a_n = \prod_{i=0}^{n-1} \left(\frac{a+b+2i+1}{a+b+2i}\right).$$

We then have the following result.

Proposition 3.2.2

The sequence $\{S_n\}$ *is a martingale with respect to the filtration* $\Im_n$.

Proof

i) $\forall n \in \mathbb{N},\ \mathbb{E}\left(S_n\right) = 0$.

ii)

$$\begin{aligned}\mathbb{E}\left(S_{n+1}|\Im_n\right) &= a_{n+1}\mathbb{E}\left(p_{v(n,n+1)} - p_{v(n+1,n+1)}|\Im_n\right) \\ &= a_{n+1}\left(\mathbb{E}\left(p_{v(n,n+1)} \mid \Im_n\right) - p_{v(n,n)}\right).\end{aligned}$$

Therefore

$$\begin{aligned}&\mathbb{E}\left(p_{v(n,n+1)} \mid \Im_n\right) - p_{v(n,n)} \\ &= \mathbb{E}\left(\frac{a + \sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n-j)\\ j=1\cdots n}} X_{v(k,n+1-j)}}{a+b+n(n+1-n)+n+1} \mid \Im_n\right) - p_{v(n,n)} \\ &= \mathbb{E}\left(\frac{a + \sum\limits_{\substack{k=(i-j)^+\cdots i\wedge(n+1-j)\\ j=1\cdots n+1}} X_{v(k,n-j)} + X_{v(n-1,n)} + X_{v(n,n)}}{a+b+n(n+1-n)+n+1} \mid \Im_n\right) - p_{v(n,n)} \\ &= \frac{(a+b+2n-1)\,p_{v(n-1,n)} + p_{v(n-1,n)} + p_{v(n,n)}}{a+b+2n+1} - p_{v(n,n)} \\ &= \frac{a+b+2n}{a+b+2n+1}\left(p_{v(n-1,n)} - p_{v(n,n)}\right).\end{aligned}$$

So

$$\mathbb{E}\left(S_{n+1} \mid \Im_n\right) = \left(\prod_{i=0}^{n}\frac{a+b+2i+1}{a+b+2i}\right)\frac{a+b+2n}{a+b+2n+1}\left(p_{v(n-1,n)} - p_{v(n,n)}\right) = S_n.$$

As a prelude to the results of this section we give the next theorem from P.Hall and C.C.Heyde (1980) which will often be applied.

Theorem 3.2.3 *Let $\{X_n, n \geq 1\}$ be a sequence of r.v. and $\{\Im_n, n \geq 1\}$ an increasing sequence of σ-fields with X_n measurable with respect to $\Im_n$ for each n. Let X be an r.v. and c a constant such that :*

1) $\mathbb{E} \mid X \mid< \infty$.
2) $\mathbb{P}\left(\mid X_n \mid> x\right) \leqslant c\mathbb{P}\left(\mid X \mid\right)$ *for each* $x \geq 0$ *and* ≥ 1.
3) $\mathbb{E}\left(\mid X \mid \log^{+} \mid X \mid\right) < \infty$.

then

$$n^{-1}\sum_{i=1}^{n}\left(X_i - \mathbb{E}\left(X_i \mid \Im_{i-1}\right)\right) \underset{n\to\infty}{\longrightarrow} 0 \qquad a.s.$$

Now we have $\left\{p_{\upsilon(i,n)}\right\}_n$ describes the proportion of balls of color 1 contained in the urn indexed by the vertices $\upsilon(i,n)$ in the tree T.
For $i \geq 0$, we define the processes $\left\{\rho_{(i,n)}\right\}_n$ such that

$$\rho_{(i,n)} = \begin{cases} p_{\upsilon(n,n)} & \text{if} \quad i > n \\ p_{\upsilon(i,n)} & \text{if} \quad i \leq n \end{cases} \tag{3.2.2}$$

We note that for $i \in \{n, n+1, \cdots, \} \cup \{0\}$ the processes $\left\{\rho_{(i,n)}\right\}_n$ is the Polya process.
The theorem (3.2.3) can be applied to obtain the following result

Theorem 3.2.4
For fixed $i \geq 0$, there exists $\varphi_i : \mathbb{N} \longrightarrow \mathbb{N}$, such that $\varphi_i(n) \nearrow \infty$ and

the process $\left\{\rho_{(i,\varphi_i(n))}\right\}_{n\geq 0}$ *converge a.s to the random variable* Θ *with* $\Theta \sim Beta(a,b)$

Proof :
We know that for $i = 0$, the process $\left\{\rho_{(0,n)}\right\}_n$ is Pólya, therefore $\rho_{(0,n)}$ converge almost surely to the random variable Θ with $\Theta \sim Beta(a,b)$.
Since, for fixed $i \geq 0$ the sequence $\left\{\rho_{\varphi_i(n)}\right\}_{n\geq 0}$ its bounded, then exists $\varphi_i : \mathbb{N} \longrightarrow \mathbb{N}$, such that $\varphi_i(n) \nearrow \infty$, and the process $\left\{\rho_{\varphi_i(n)}\right\}_{n\geq 0}$ converge almost surely.
Now, we show that the limit is always Θ with $\Theta \sim Beta(a,b)$.

i = 1 :

The theorem (3.2.3) can be applied to obtain the following

$$\begin{aligned}
\left\{\omega : \lim_{n\to+\infty}\rho_{(1,\varphi_1(n))} = L_1 \quad , \qquad a.s\right\} &\subseteq \left\{\omega : \lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n} p_{v(1,\varphi_1(j))} = L_1 \quad , \qquad a.s.\right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n} X_{v(1,\varphi_1(j))} = L_1 \ , \qquad a.s.\right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty}\frac{1}{\varphi_1(n)}\sum_{j=0}^{\varphi_1(n)} X_{v(1,\varphi_1(j))} = L_1 \quad , \qquad a.s.\right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty}\frac{1}{\varphi_1(n)}\sum_{j=0}^{\varphi_1(n)} X_{v(1,j)} = L_1 \quad , \qquad a.s.\right\}.
\end{aligned}$$

And we have

$$\rho_{(1,\varphi_1(n))} = \frac{a + \sum_{j=1}^{\varphi_1(n)} X_{v(1,j)} + \sum_{j=0}^{\varphi_1(n)} X_{v(0,j)}}{a+b+2\varphi_1(n)-1}$$

Therefore

$$\begin{aligned}
L_1 &= \lim_{n\to+\infty}\rho_{(1,\varphi_1(n))} = \lim_{n\to+\infty}\frac{a + \sum_{j=1}^{\varphi_1(n)} X_{v(1,j)} + \sum_{j=0}^{\varphi_1(n)} X_{v(0,j)}}{a+b+2\varphi_1(n)-1} \\
&= \lim_{n\to+\infty}\frac{\varphi_1(n)}{a+b+2\varphi_1(n)-1}\left\{\frac{a + \sum_{j=1}^{\varphi_1(n)} X_{v(1,j)}}{\varphi_1(n)} + \frac{\sum_{j=0}^{\varphi_1(n)} X_{v(0,j)}}{\varphi_1(n)}\right\} \\
&= \frac{L_1 + \Theta}{2} \qquad a.s.
\end{aligned}$$

We conclude that $\Theta = L_1$ almost surely.

General case.

By similar computation we have that

$$\begin{aligned}
\left\{\omega : \lim_{n\to+\infty} \rho_{(i,\varphi_i(n))} = L_i \quad , \qquad a.s \right\} &\subseteq \left\{\omega : \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n} p_{\upsilon(i,\varphi_i(j)} = L_i \quad , \qquad a.s.\right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n} X_{\upsilon(i,\varphi_i(j))} = L_i \ , \qquad a.s.\right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty} \frac{1}{\varphi_i(n)}\sum_{j=0}^{\varphi_i(n)} X_{\upsilon(i,\varphi_i(j))} = L_i \quad , \qquad a.s. \right\} \\
&\subseteq \left\{\omega : \lim_{n\to+\infty} \frac{1}{\varphi_i(n)}\sum_{j=0}^{\varphi_i(n)} X_{\upsilon(i,j)} = L_i \quad , \qquad a.s.\right\}.
\end{aligned}$$

and

$$\begin{aligned}
p_{\upsilon(i,\varphi_i(n))} &= \frac{\sum_{k=0}^{\varphi_i(n)-i} X_{(0,k)} + \sum_{k=0}^{\varphi_i(n)-i+1} X_{(1,k)} + \cdots + \sum_{k=0}^{\varphi_i(n)-1} X_{(i,k)}}{a+b+i(\varphi_i(n)-i)+\varphi_i(n)} \\
&= \frac{\varphi_i(n)}{a+b+i(\varphi_i(n)-i)+\varphi_i(n)} \left\{\frac{\sum_{k=0}^{\varphi_i(n)-i} X_{(0,k)}}{\varphi_i(n)} + \frac{\sum_{k=0}^{\varphi_i(n)-i+1} X_{(1,k)}}{\varphi_i(n)} \right. \\
&\qquad \left. + \cdots + \frac{\sum_{k=0}^{\varphi_i(n)-1} X_{(i,k)}}{\varphi_i(n)}\right\}
\end{aligned}$$

We conclude by induction that

$$\begin{cases} L_i = \frac{L_i + iL_{i-1}}{i+1} & \text{a.s.} \\ L_0 = \Theta & \text{a.s.} \end{cases} \tag{3.2.3}$$

Therefore we get the thesis.

The next result is less obvious.

Corollary 3.2.5

For fixed $i \geq 0$ and $k \geq 0$,

$$\mathbb{E}\left[\rho_{(i,n)}^k\right] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\left[\Theta^k\right] \tag{3.2.4}$$

with $\Theta \sim Beta(a,b)$.

Proof :

Since , for fixed $i \geq 0$, $\rho_{(i,n)} \in [0,1]$, therefore for all $k \geq 0, \mathbb{E}[\rho^k_{(i,n)}] \in [0,1]$, and $\lim\limits_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}]$ exist and it's finite.

Now, we show that this limit it's unique.

We suppose that $\lim\limits_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}] = \ell_1$, and $\lim\limits_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}] = \ell_2$.

The theorem(3.2.4) can be applied to obtain the following

$$\begin{cases} \ell_1 = \lim\limits_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}] = \mathbb{E}\left[\Theta^k\right] \\ \ell_2 = \lim\limits_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}] = \mathbb{E}\left[\Theta^k\right] \end{cases} \tag{3.2.5}$$

We conclude that

$$\mathbb{E}\left[\rho^k_{(i,n)}\right] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\left[\Theta^k\right] \tag{3.2.6}$$

We will prove the following results.

Theorem 3.2.6

For fixed $j \geq 0$

$$\rho_{(j,n)} \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \Theta \qquad \textit{with } \Theta \sim Beta(a,b). \tag{3.2.7}$$

Proof :

Let $\psi_{(j,n)}$ the characteristic function of $\rho_{(j,n)}$, then we have

$$\begin{aligned} \psi_{(j,n)}(t) &= \mathbb{E}\left[e^{it\rho_{(j,n)}}\right] \\ &= \sum_{k=0}^{\infty} \frac{t^k \psi^{(k)}_{(j,n)}(0)}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(it)^k \mathbb{E}\left[\rho^k_{(j,n)}\right]}{k!} \ , \qquad \left(\psi^{(k)}_{(n,j)}(0) = i^k \mathbb{E}[\rho^k_{(j,n)}]\right) \end{aligned}$$

by the monotome convergence's theorem

$$
\begin{aligned}
\lim_{n\longrightarrow\infty}\psi_{(j,n)}(t) &= \lim_{n\longrightarrow\infty}\sum_{k=0}^{\infty}\frac{(it)^k\mathbb{E}\left[\rho_{(j,n)}^k\right]}{k!} \\
&= \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\lim_{n\longrightarrow\infty}\mathbb{E}\left[\rho_{(j,n)}^k\right] \\
&= \sum_{k=0}^{\infty}\frac{(it)^k}{k!}\mathbb{E}\left[\Theta^k\right], \qquad (corollary\ 3.2.5) \\
&= \mathbb{E}\left[\sum_{k=0}^{\infty}\frac{(it\Theta)^k}{k!}\right] \\
&= \mathbb{E}\left[e^{it\Theta}\right]
\end{aligned}
$$

This shows that the distribution of the limit of $\rho_{(j,n}$ is beta with parameters (a,b).

3.3 Asymptotic exchangeability

Hill, Lane and Sudderth prove that the only exchangeable processes $\{V_n, n \geqslant 0\}$ produced by their urn scheme are Pòlya's process, the i.i.d. Bernoulli process and the deterministic one, by consequence for all i fixed the sequence $\left\{X_{\upsilon(i,n)}\right\}$ is not exchangeable

Definition 3.3.1

An infinity sequence $(V_1, V_2, \cdots)$ *is called asymptotically exchangeable if*

$$(V_{j+1}, V_{j+2}, \cdots) \stackrel{d}{=} (Z_1, Z_2, \cdots) \quad as \quad j \longrightarrow \infty \tag{3.3.1}$$

with $(Z_n, n \geqslant 1)$*is an infinity exchangeable sequence.*

Lemma 3.3.2 *(Aldous 1983)*

Let $(Z_n, n \geqslant 1)$*is an infinity exchangeable sequence directed by* α*.*

(a) Let α_n *be a regular conditional distribution for* Z_{n+1} *given* $(Z_1, \cdots, Z_n)$*.*

Then

$$\alpha_n \underset{n \to \infty}{\longrightarrow} \alpha \ a.s. \tag{3.3.2}$$

(b) Let $(X_n, n \geqslant 1)$ *be an infinite sequence, let* α_n *be a regular conditinal distribution for* X_{n+1} *given* $(X_1, X_2, \cdots, X_n)$*, and suppose* $\alpha_n \underset{n \to \infty}{\longrightarrow} \alpha \ a.s.$

Then

$$(X_{j+1}, X_{j+2}, \cdots) \stackrel{d}{=} (Z_1, Z_2, \cdots) \quad as \quad j \longrightarrow \infty \tag{3.3.3}$$

As a direct consequence the following Theorem is got.

Theorem 3.3.3

1) *The sequence* $\left\{X_{\upsilon(i,\varphi_i(n))}\right\}$ *is asymptotically exchangeable.*

2)

$$\lim_{n \mapsto \infty} \mathbb{P}\left[X_{\upsilon(i,\varphi_i(n)+1)} = x_1, \ldots, X_{\upsilon(i,\varphi_i(n)+k)} = x_k\right] = \mathbb{E}[M_i^{s(k)}(1-M_i)^{k-s(k)}]$$

For every $k \geqslant 1$, $x_j \in \{0,1\}$, $j = 1, \cdots, k$

with

$$s(k) = \textstyle\sum_{j=1}^{k} x_j \qquad\qquad M_i = \mathbb{E}[\Theta \mid \mathcal{G}_{\infty}^{(i)}]$$

$$\mathcal{G}_{\infty}^{(i)} = \bigvee_{n=1}^{\infty} \sigma(X_{\upsilon(i,\varphi_i(j))}, 0 \leq j \leq n) \qquad \Theta \sim beta(a,b)$$

Proof:

1)We all fixed i,and for $\geqslant 0$, the conditional distribution of $X_{\upsilon(i,\varphi_i(n)+1)}$ given $\Im_{\varphi_i(n)} = \sigma\left(X_\tau, |\sigma| < \varphi_i(n)\right)$ is a $Bernouilli\left(p_{\upsilon(i,\varphi_i(n))}\right)$; its converges with probability one to a $Bernouilli\,(\Theta)$ as n grows to infinity

with $\Theta \sim beta(a,b)$.The resulat now follows from Aldous lemma

2)Now

$$\mathbb{P}\left(X_{\upsilon(i,\varphi_i(n))} = 1 \mid \Im_{\varphi_i(n)}\right) = p_{\upsilon(i,\varphi_i(n))} \xrightarrow{a.s} \Theta$$

but

$$\mathbb{P}\left(X_{\upsilon(i,\varphi_i(n))} = 1 \mid \mathcal{G}_n^i\right) = \mathbb{E}\left(X_{\upsilon(i,\varphi_i(n))} = 1 \mid \mathcal{G}_n^i\right)$$

So that the Hunt's lemma (see hunt (1966) or Meyer (1969)) entails

$$\mathbb{E}\left(X_{\upsilon(i,\varphi_i(n))} = 1 \mid \mathcal{G}_n^i\right) \xrightarrow{a.s} \mathbb{E}\left(\Theta \mid \mathcal{G}_\infty^i\right) = M_i.$$

with

$$\mathcal{G}_{\infty}^{(i)} = \bigvee_{n=1}^{\infty} \mathcal{G}_n^i, \qquad \mathcal{G}_n^i = \sigma(X_{\upsilon(i,\varphi_i(j))}, 0 \leq j \leq n), \qquad \Theta \sim beta(a,b)$$

By (1) and monotone convergence theorem's, we have

$$\begin{aligned}
&\lim_{n\mapsto\infty} \mathbb{P}\left[X_{\upsilon(i,\varphi_i(n)+1)} = x_1, \ldots, X_{\upsilon(\varphi_i(n)+k)} = x_k\right] \\
&= \lim_{n\mapsto\infty} \int_\Omega \mathbb{P}\left[X_{\upsilon(i,\varphi_i(n)+1)} = x_1, \ldots, X_{\upsilon(i,\varphi_i(n)+k)} = x_k \mid M_i\right] d\mathbb{P}_{M_i} \\
&= \int_\Omega \lim_{n\mapsto\infty} \mathbb{P}\left[X_{\upsilon(i,\varphi_i(n)+1)} = x_1, \ldots, X_{\upsilon(i,\varphi_i(n)+k)} = x_k \mid M_i\right] d\mathbb{P}_{M_i} \\
&= \int_\Omega \lim_{n\mapsto\infty} \prod_{j=1}^{k} \mathbb{P}\left[X_{\upsilon(i,\varphi_i(n)+1)} = x_j \mid M_i\right] d\mathbb{P}_{M_i} \\
&= \int_\Omega \left[M_i^{\sum_{i=1}^k x_i} (1 - M_i)^{k - \sum_{i=1}^k x_i}\right] d\mathbb{P}_{M_i} \\
&= \mathbb{E}[M_i^{s(k)} (1 - M_i)^{k - s(k)}]
\end{aligned}$$

For every $k \geqslant 1$, $x_j \in \{0,1\}$, $j = 1, \cdots, k$ and $s(k) = \sum_{j=1}^k x_j$

□

In this section we note an interesting property of the proportion of balls of color 1 at level n in the tree T.

For $n \geqslant 0$ le use define

p_n: the proportion of balls of color 1 at level n.

Y_n: the total of balls of color 1 at level n.

$\overline{Y_n}$: the total of balls of color 0 at level n.

C_n: the total number of balls at level n.

Lemma 3.3.4

For all $n \geqslant 0$

$$C_n = (n+1)(a+b) + \frac{n(n^2+6n+5)}{6} \tag{3.3.4}$$

Proof

We have that

$$Y_{n+1} = Y_n + 2\sum_{|\sigma|=n} X_\sigma + \sum_{|\sigma|<n} X_\sigma + a \tag{3.3.5}$$

$$\overline{Y}_{n+1} = \overline{Y}_n + 2\sum_{|\sigma|=n} (1-X_\sigma) + \sum_{|\sigma|<n} (1-X_\sigma) + b \tag{3.3.6}$$

Since $C_0 = a+b$, so

$$\begin{aligned}
C_{n+1} &= Y_{n+1} + \overline{Y}_{n+1} && (3.3.7)\\
&= C_n + a + b + 2\sum_{|\sigma|=n} 1 + \sum_{|\sigma|<n} 1 && (3.3.8)\\
&= C_n + a + b + 2(n+1) + \sum_{i=0}^{n-1}\sum_{|\sigma|=i} 1 && (3.3.9)\\
&= C_n + 2(n+1) + \sum_{i=0}^{n-1}(i+1) + a + b && (3.3.10)\\
&= C_n + 2(n+1) + \frac{n(n-1)}{2} + n + a + b && (3.3.11)\\
&= C_n + \frac{n^2+5n+4}{2} + a + b && (3.3.12)
\end{aligned}$$

We conclude that

$$C_n = (n+1)(a+b) + \frac{n(n^2+6n+5)}{6} \tag{3.3.13}$$

Lemma 3.3.5 *For all $n \geqslant 0$*

$$E\left[p_n\right] = \frac{a}{a+b} \tag{3.3.14}$$

Proof:

We have that

$$\begin{aligned} p_{n+1} &= \frac{Y_{n+1}}{C_{n+1}} \\ &= \frac{Y_n + 2\sum_{|\sigma|=n} X_\sigma + \sum_{|\sigma|<n} X_\sigma + a}{C_{n+1}} \\ &= \frac{C_n p_n + 2\sum_{|\sigma|=n} X_\sigma + \sum_{|\sigma|<n} X_\sigma + a}{C_{n+1}} \end{aligned}$$

So

$$\begin{aligned} C_{n+1}\mathbb{E}\left(p_{n+1}\right) &= C_n\mathbb{E}\left(p_n\right) + 2\sum_{|\sigma|=n}\mathbb{E}\left(X_\sigma\right) + \sum_{|\sigma|<n}\mathbb{E}\left(X_\sigma\right) + a \\ &= C_n\mathbb{E}\left(p_n\right) + 2\sum_{|\sigma|=n}\left(\frac{a}{a+b}\right) + \sum_{|\sigma|<n}\left(\frac{a}{a+b}\right) + a \\ &= C_n\mathbb{E}\left(p_n\right) + \frac{a}{a+b}\left(2(n+1) + \frac{n(n-1)}{2} + n\right) + a \\ &= C_n\mathbb{E}\left(p_n\right) + \frac{a}{a+b}\left(\frac{n^2+5n+4}{2} + a + b\right) \\ &= C_n\mathbb{E}\left(p_n\right) + (C_{n+1} - C_n)\frac{a}{a+b} \end{aligned}$$

Therefore

$$\begin{aligned} C_{n+1}\left(E\left[p_{n+1}\right] - \frac{a}{a+b}\right) &= C_n\left(\mathbb{E}\left[p_n\right] - \frac{a}{a+b}\right) \\ &= C_n\left(\mathbb{E}\left(p_n\right) - \frac{a}{a+b}\right) \\ &= \cdots = C_0\left(\mathbb{E}\left(p_0\right) - \frac{a}{a+b}\right) \\ &= (a+b)\left(\frac{a}{a+b} - \frac{a}{a+b}\right) = 0 \end{aligned}$$

This show that for all $n \geqslant 0$

$$\mathbb{E}\left[p_n\right] = \frac{a}{a+b} \tag{3.3.15}$$

The proportion of balls of color 1 generated by the urn associated with the vertices of the tree at level $n \geqslant 0$ is

$$K_n = \frac{2\sum_{|\sigma|=n} X_\sigma + \sum_{|\sigma|<n} X_\sigma + a}{C(n+1) - C(n)} \tag{3.3.16}$$

The total number of balls of color 1 in the urnsassociated with vertices of the tree T at level $n+1$ is $C(n+1)p_{n+1} = C(n)p_n + (C_{n+1} - Cn)\, K_n$

Since $\left\{X_{\upsilon(0,n)}\right\}_n$ is a Polya sequence of random varibles 0 or 1 unless $a = 0$ or $b = 0$,

$$\mathbb{P}\left(K_n \neq K_{n+1} \text{ for in infinitely many } n\right) = 1 \quad a.s \tag{3.3.17}$$

and thus $\{K_n\}$ does not converge.

The next result is obvious.

Lemma 3.3.6

$\mathbb{E}\left(K_n\right) = \frac{a}{a+b}$

Proof :

Since

$C(n+1)p_{n+1} = C(n)p_n + (C_{n+1} - Cn)\, K_n$ and $\mathbb{E}\left(p_n\right) \frac{a}{a+b}$, then $\mathbb{E}\left(K_n\right) = \frac{a}{a+b}$.

3.4 Connection with Dirichlet reinforced process

There is an interesting connection between the Dirichlet distribution and the reinforced stochastic process $X = \{X_\tau, \sigma \in T\}$ of random variable indexed by the vertices of tree T and with value in a Polish space S endowed with its Borel sigma-field $\mathcal{G}$.

As before, we assume that the random variable of the process X are difined on a rich enougn probability space $(\Omega, \Im, P)$, and we specify the law of X recursively on the levels of the tree T

Let G_0 be aprobability distribution on S and $c > 0$ a constant.

Set X_θ to be a random variable with value in S and probability distribution G_0 .

For $n \geqslant 0$ let $\Im_n \subseteq \Im$ be the sigma-field generated by the random variables X_σ with $|\sigma| \leq n$. Given $\Im_n$, assume that the $n+1$ random variables X_τ has values in S and probability distribution

$$G_\tau = \frac{cG_0 + \sum_{\sigma \in A(\tau)} \delta_{X_\sigma}}{c + \#A(\tau)} \tag{3.4.1}$$

If $A(\tau) := \{\gamma \in T : \exists\pi(\phi,\gamma)\ s.t.\ \gamma \in \pi(\phi,\gamma)\}$; for $s \in S$, δ_s indicate the point mass at s.

Using the notation of section 1, we can write $X_{\nu(0,0)}$ to be a random variable with value in S and probability distribution G_0 and for $n \geqslant 0$ and $0 \leq i \leq n$. Given $\Im_n$, assume that the $n+1$ random variables $X_{\upsilon(i,n)}$ has values in S and probability distribution

$$G_{\upsilon(i,n)} = \frac{cG_0 + \sum_{k=(i-j)^+,\cdots,i\wedge(n-j),j=1,\cdots,n} \delta_{X_{\upsilon(k,n-j)}}}{c + i(n-i) + n} \tag{3.4.2}$$

In this last section, we adress the following question:

For all $i \geqslant 0$ the sequence $G_{\upsilon(i,n)}$ weakly converge to a Dirichlet process G_υ with parameter cG_0on a set of probability one?

Chapter 4

On sum of Product of independent and identically distributed random variables with Bernoulli distribution

Abstract :
We develop new discrete distributions that describe the behavior of a sum of independent Bernoulli random variables and we are interested a multidimensional case.
Keywords: Probability generating function, Reccurence, Poisson distribution, Beta mixture, pólya urn schemes.

4.1 Introduction:

Since the middle of the 90s, a great amount of research has been done to study the distribution of the number of runs of length 2 : $S_n = \sum_{k=1}^{n} X_k X_{k+1}$ and $S = \lim_{n \longrightarrow \infty} S_n$ in sequences of independent Bernoulli randon variables $\{X_n\}_{n=1}^{\infty}$ with parameters p_n. For the particular case $p_n = \frac{1}{n}$, S follows a

Poisson distribution with mean 1, and this result is due to Diaconis (1991), this result is generalised by A.Joffe, E.Marchand, F.Perron and P.Popadiuk (1999, 2004) for $p_n = \frac{1}{b+n}$, b a non-negative integer, and by Holst (2006) for $p_n = \frac{a}{a+b+n}$.

In particular for the independent and identically distributed case where $p_n = p$, the distribution of S_n is know to be type II binomial of order $k = 2$. Ling (1988) introduced type II binomial distribution of order k as the number of (possibly overlapping) run of length k in a infinite sequence of independent and identically distributed Bernoulli random variables. In contrast, type I binomial distribution of order k (Hirano, 1986, Philippou and Mahri, 1986) involve only the counting of non-overlapping runs of length k. Properties type II binomial distributions of order k have been obtained by Ling (1988), Hirano et al (1991), Godbole (1992), and Han and Aki (2000).

The chapter is orgnised as follows. In a first part we find the distribution of the sum of product of exchangeable random variables with Bernoulli distribution an application for a Polya urn model .

In a second part we study an extention ;with results fo the numbers of runs of length 2. The problem, which permits us to work with non-independ Bernoulli components, leads to many particular interesting cases.

4.2 On sum of product of exchangeables random variables with Bernoulli distribution:

In this section we find the exact distribution of an arbitrary of an infinite sum of overlapping products of a sequence of exchangeable Bernoulli random variables.

Let $\{X_k\}_{k=1}^{n}$ be a sequence of independent Bernoulli random variables with parmeters $p \in [0, 1]$, in this case, the distribution of $M_n = \sum_{i=1}^{n} X_i X_{i+1}$ is know to be type II binomial of order $k = 2$ (e.g, Joff e al, 2004, Hirano, 1986) . In particular Holst gave an implicit expression for the probability generating function of type II binomial distribution and we have the next theorem.

Theorem 4.2.1 *(L.Holst, 2006)*
Let $\{X_n\}_{n\geq 1}$ *be independent and identically distributed random variables with* $X_n \sim Ber(p)$ *and let*

$$M_n = \sum_{i=1}^{n} X_i X_{i+1}$$

Then,

$$\mathbb{E}\binom{M_n}{r} = \sum_{k=1}^{r} \binom{r-1}{r-k} \binom{n-r}{k} p^{k+r}$$

4.2.1 Application to exchangeable random variables:

Theorem('4.2.1) can be applied to obtain the following result.

Theorem 4.2.2 *Let* $\{X_n\}_{n\geq 1}$ *be the sequence of indicators of the events of an infinite exchangeable sequence.*
Let

$$S_n = \sum_{i=1}^{n} X_i X_{i+1} \tag{4.2.1}$$

Then there is a random variable $\Theta \in [0, 1]$ *such that:*

$$\mathbb{P}\left(S_n = i\right) = \sum_{r=i}^{\infty} (-1)^{r-i} \binom{r}{i} \sum_{k=1}^{r} \binom{r-1}{r-k} \binom{n-r}{k} \mathbb{E}\left(\Theta^{k+r}\right)$$

Proof :
i)The sequence $\{X_n\}_{n\geq 1}$ is exchangeable and by de Finetti's Representation Theorem there exists a random variable $\Theta \in [0, 1]$,such that :

$$X_n \mid \Theta \overset{i.i.d}{\sim} Ber\left(\Theta\right).$$

Therefore

$$\begin{aligned}
\mathbb{E}\left[\binom{S_n}{r}\right] &= \int_0^1 \mathbb{E}\left[\binom{S_n}{r} \mid \Theta = x\right] f_\Theta(x)dx \\
&= \int_0^1 \sum_{k=1}^{r} \binom{r-1}{r-k}\binom{n-r}{k} x^{k+r} f_\Theta(x)dx \qquad \text{(by (1.1))} \\
&= \sum_{k=1}^{r} \binom{r-1}{r-k}\binom{n-r}{k} \int_0^1 x^{k+r} f_\Theta(x)dx \\
&= \sum_{k=1}^{r} \binom{r-1}{r-k}\binom{n-r}{k} \mathbb{E}\left(\Theta^{k+r}\right)
\end{aligned}$$

And we have

$$\begin{aligned}
\mathbb{E}\left(t^{S_n}\right) &= \mathbb{E}\left((1+(t-1))^{S_n}\right) \\
&= \sum_{r=0}^{\infty} \mathbb{E}\binom{S_n}{r}(t-1)^r \\
&= \sum_{i=0}^{\infty} t^i \sum_{r=i}^{\infty} (-1)^{r-i}\binom{r}{i}\mathbb{E}\binom{S_n}{r} \\
&= \sum_{i=0}^{\infty} t^i \mathbb{P}\left(S_n = i\right)
\end{aligned}$$

We conclude that

$$\begin{aligned}
\mathbb{P}\left(S_n = i\right) &= \sum_{r=i}^{\infty} (-1)^{r-i}\binom{r}{i}\mathbb{E}\binom{S_n}{r} \\
&= \sum_{r=i}^{\infty} (-1)^{r-i}\binom{r}{i} \sum_{k=1}^{r} \binom{r-1}{r-k}\binom{n-r}{k}\mathbb{E}\left(\Theta^{k+r}\right)
\end{aligned}$$

4.2.2 Example : Eggenberger and Pólya Urn

Consider an urn initially containing $b > 0$ black balls and $a > 0$ white balls.At time $n = 1, 2, \cdots$, a ball is sampled from the urn and replaced into it along with $\alpha > 0$ balls of the same color.This generates an infinite sequence (X_n) of Bernoulli random variables, where X_n is 1 or 0 according to the color black or white respectively of ball extracted from the urn at time n. And let Y_n be the number of white balls in the urn after n draws.

We define

$$S_n = \sharp\{k \geqslant 1 : Y_{k+1} = Y_k + 1,\ k \leq n-1\} = \sum_{i=1}^{n} X_i X_{i+1}$$

We knew that the sequence $\{X_n\}_{n\geq 1}$ generated by a Pólya urn is exchangeable and its de Finetti measure is Beta with parameters$\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)$. Therfore

$$\begin{aligned}\mathbb{P}\left(S_n = i\right) &= \sum_{r=i}^{\infty}(-1)^{r-i}\binom{r}{i}\sum_{k=1}^{r}\binom{r-1}{r-k}\binom{n-r}{k}E\left(\Theta^{k+r}\right) \qquad \text{with } \Theta \sim Beta\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right) \\ &= \sum_{r=i}^{\infty}(-1)^{r-i}\binom{r}{i}\sum_{k=1}^{r}\binom{r-1}{r-k}\binom{n-r}{k}\frac{B\left(\frac{a}{\alpha}+k+r, \frac{b}{\alpha}\right)}{B\left(\frac{a}{\alpha}, \frac{b}{\alpha}\right)}\end{aligned}$$

with $B(a,b)$ is the beta function evaluated in (a,b),that is
$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$,with Γ the usual gamma function .

4.3 On Sums of Products of Bernoulli variables, The multidimensional case :

Let $\overline{X}_k = (X_{k,1}, X_{k,2}, \cdots, X_{k,p})$, $k = 1, 2, \cdots,$ be $(1 \times p)$ random vector such that $\{X_{k,i}\}_{k\geq 1}$ be independent Bernoulli random variables with parameters $p_{k,i}$ in $(0,1)$.
Let

$$S_n = \sum_{k=1}^{n}\overline{X}_k\overline{X}_{k+1}^T = \sum_{k=1}^{n}\sum_{j=1}^{p}X_{k,j}X_{k+1,j} \tag{4.3.1}$$

Remark 4.3.1
i)

$$S_n = \sum_{k=1}^{n}\sum_{j=1}^{p}X_{k,j}X_{k+1,j} = \sum_{j=1}^{p}Z_{j,n} \qquad \textit{with} \qquad Z_{j,n} = \sum_{k=1}^{n}X_{k,j}X_{k+1,j}$$

ii) If

$$\sum_{k=1}^{\infty}p_{k,j}p_{k+1,j} < \infty \qquad 1 \leq j \leq p.$$

The Borel Cantelli lemma shows that

$$S = \sum_{k=1}^{\infty}\overline{X}_k\overline{X}_{k+1}^T$$

exist almost surely.

iii)

$$\{\overline{X}_k,\ independent\ for\ all\ k\} \Rightarrow \{X_{k,j}\ independent\ for\ all\ k,\ j\} \Rightarrow \left\{Z_j^n\ independent\ for\ all\ j\right\}$$

Here, we allow the components $X_{k,1}, \cdots, X_{k,p}$ to be dependent .As a consequence, the random vriaables $Z_j^n,\ j = 1, \cdots, p$ are not necessarily and not in general independent .
The particular case of independence of the composent of $\overline{X}_k$ for $k = 1, 2, \cdots$ can be to deducted by studing the convolution of $Z_j^n, j = 1, 2, \cdots, p$

4.3.1 Bidimensional case (p=2):

For $p = 2$,

$$S_n = S_{n-1} + X_{n,1}X_{n+1,1} + X_{n,2}X_{n+1,2} \tag{4.3.2}$$

therefore

$$\begin{cases} \mathcal{L}\left(S_n \mid \overline{X}_{n+1} = (1,1)\right) = \mathcal{L}\left(S_{n-1} + X_{n,1} + X_{n,2}\right) \\ \mathcal{L}\left(S_n \mid \overline{X}_{n+1} = (0,1)\right) = \mathcal{L}\left(S_{n-1} + X_{n,2}\right) \\ \mathcal{L}\left(S_n \mid \overline{X}_{n+1} = (1,0)\right) = \mathcal{L}\left(S_{n-1} + X_{n,1}\right) \\ \mathcal{L}\left(S_n \mid \overline{X}_{n+1} = (0,0)\right) = \mathcal{L}\left(S_{n-1}\right) \end{cases} \tag{4.3.3}$$

As, well, define the random variables $\{W_{i,n}\}_n, i = 1, 2, 3.$ such that

$$W_{n,1} := S_{n-1} + X_{n,1} + X_{n,2} \tag{4.3.4}$$

$$W_{n,2} := S_{n-1} + X_{n,1} \tag{4.3.5}$$

$$W_{n,3} := S_{n-1} + X_{n,2} \tag{4.3.6}$$

And let $\varphi_{i,n}(t) = \mathbb{E}\left(t^{W_{n,i}}\right), i = 1, 2, 3.$ be their generating function, $\varphi_{n,0}(t) = \mathbb{E}\left(t^{S_n}\right)$ be the generating function of S_n.

Proposition 4.3.2

For $n \geq 2$.

$$\begin{pmatrix} \varphi_{0,n-1}(t) \\ \varphi_{1,n}(t) \\ \varphi_{2,n}(t) \\ \varphi_{3,n}(t) \end{pmatrix} = \begin{pmatrix} d_n & a_n & b_n & c_n \\ d_n & t^2 a_n & t b_n & t c_n \\ d_n & t a_n & t b_n & c_n \\ d_n & t a_n & b_n & t c_n \end{pmatrix} \begin{pmatrix} \varphi_{0,n-2}(t) \\ \varphi_{1,n-1}(t) \\ \varphi_{2,n-1}(t) \\ \varphi_{3,n-1}(t) \end{pmatrix} \tag{4.3.7}$$

$$a_n = \mathbb{P}\left(\overline{X}_n = (1,1)\right), b_n = \mathbb{P}\left(\overline{X}_n = (1,0)\right)$$

$$c_n = \mathbb{P}\left(\overline{X}_n = (0,1)\right), c_n = \mathbb{P}\left(\overline{X}_n = (0,0)\right)$$

Proof :

The following result follows since

$$\begin{cases} \mathcal{L}\left(W_{1,n} \mid \overline{X}_n = (1,1)\right) = \mathcal{L}\left(W_{1,n-1} + 2\right) \\ \mathcal{L}\left(W_{1,n} \mid \overline{X}_n = (1,0)\right) = \mathcal{L}\left(W_{2,n-1} + 1\right) \\ \mathcal{L}\left(W_{1,n} \mid \overline{X}_n = (0,1)\right) = \mathcal{L}\left(W_{3,n-1} + 1\right) \\ \mathcal{L}\left(W_{1,n} \mid \overline{X}_n = (0,0)\right) = \mathcal{L}\left(S_{n-2}\right) \end{cases} \tag{4.3.8}$$

$$\begin{cases} \mathcal{L}\left(W_{2,n} \mid \overline{X}_n = (1,1)\right) = \mathcal{L}\left(W_{1,n-1} + 1\right) \\ \mathcal{L}\left(W_{2,n} \mid \overline{X}_n = (1,0)\right) = \mathcal{L}\left(W_{2,n-1} + 1\right) \\ \mathcal{L}\left(W_{2,n} \mid \overline{X}_n = (0,1)\right) = \mathcal{L}\left(W_{3,n-1}\right) \\ \mathcal{L}\left(W_{2,n} \mid \overline{X}_n = (0,0)\right) = \mathcal{L}\left(S_{n-2}\right) \end{cases} \tag{4.3.9}$$

$$\begin{cases} \mathcal{L}\left(W_{3,n} \mid \overline{X}_n = (1,1)\right) = \mathcal{L}\left(W_{1,n-1} + 1\right) \\ \mathcal{L}\left(W_{3,n} \mid \overline{X}_n = (1,0)\right) = \mathcal{L}\left(W_{2,n-1}\right) \\ \mathcal{L}\left(W_{3,n} \mid \overline{X}_n = (0,1)\right) = \mathcal{L}\left(W_{3,n-1} + 1\right) \\ \mathcal{L}\left(W_{3,n} \mid \overline{X}_n = (0,0)\right) = \mathcal{L}\left(S_{n-2}\right) \end{cases} \tag{4.3.10}$$

In fact $(4.3.2), (4.3.4), (4.3.5)$ and $(4.3.6)$ implied that

$$\begin{cases} \varphi_{0,n}(t) = a_{n+1}\varphi_{1,n}(t) + b_{n+1}\varphi_{2,n}(t) + c_{n+1}\varphi_{3,n}(t) + d_{n+1}\varphi_{0,n-1}(t) \\ \varphi_{1,n}(t) = t^2 a_n \varphi_{1,n-1}(t) + t b_n \varphi_{2,n-1}(t) + t c_n \varphi_{3,n-1}(t) + d_n \varphi_{0,n-2}(t) \\ \varphi_{2,n}(t) = t a_n \varphi_{1,n-1}(t) + t b_n \varphi_{2,n-1} + c_n \varphi_{3,n-1}(t)(t) + d_n \varphi_{0,n-2}(t) \\ \varphi_{3,n}(t) = t a_n \varphi_{1,n-1}(t) + b_n \varphi_{2,n-1} + t c_n \varphi_{3,n-1}(t) + d_n \varphi_{0,n-2}(t) \end{cases} \tag{4.3.11}$$

$$(4.3.11) \Leftrightarrow \begin{cases} \varphi_{0,n-1}(t) = a_n\varphi_{1,n-1}(t) + b_n\varphi_{2,n-1}(t) + c_n\varphi_{3,n-1}(t) + d_n\varphi_{0,n-2}(t) \\ \varphi_{1,n}(t) = t^2a_n\varphi_{1,n-1}(t) + tb_n\varphi_{2,n-1}(t) + tc_n\varphi_{3,n-1}(t) + d_n\varphi_{0,n-2}(t) \\ \varphi_{2,n}(t) = ta_n\varphi_{1,n-1}(t) + tb_n\varphi_{2,n-1} + c_n\varphi_{3,n-1}(t)(t) + d_n\varphi_{0,n-2}(t) \\ \varphi_{3,n}(t) = ta_n\varphi_{1,n-1}(t) + b_n\varphi_{2,n-1} + tc_n\varphi_{3,n-1}(t) + d_n\varphi_{0,n-2}(t) \end{cases} \qquad (4.3.12)$$

$$(4.3.12) \Leftrightarrow \begin{pmatrix} \varphi_{0,n-1}(t) \\ \varphi_{1,n}(t) \\ \varphi_{2,n}(t) \\ \varphi_{3,n}(t) \end{pmatrix} = \begin{pmatrix} d_n & a_n & b_n & c_n \\ d_n & t^2a_n & tb_n & tc_n \\ d_n & ta_n & tb_n & c_n \\ d_n & ta_n & b_n & tc_n \end{pmatrix} \begin{pmatrix} \varphi_{0,n-2}(t) \\ \varphi_{1,n-1}(t) \\ \varphi_{2,n-1}(t) \\ \varphi_{3,n-1}(t) \end{pmatrix}$$

By the proposition 4.3.2, we conclude that

$$\begin{aligned} \begin{pmatrix} \varphi_{0,n-1}(t) \\ \varphi_{1,n}(t) \\ \varphi_{2,n}(t) \\ \varphi_{3,n}(t) \end{pmatrix} &= \begin{pmatrix} d_n & a_n & b_n & c_n \\ d_n & t^2a_n & tb_n & tc_n \\ d_n & ta_n & tb_n & c_n \\ d_n & ta_n & b_n & tc_n \end{pmatrix} \begin{pmatrix} \varphi_{0,n-2}(t) \\ \varphi_{1,n-1}(t) \\ \varphi_{2,n-1}(t) \\ \varphi_{3,n-1}(t) \end{pmatrix} \\ &= \prod_{i=2}^{n} \begin{pmatrix} d_i & a_i & b_i & c_i \\ d_i & t^2a_i & tb_i & tc_i \\ d_i & ta_i & tb_i & c_i \\ d_i & ta_i & b_i & tc_i \end{pmatrix} \begin{pmatrix} \varphi_{0,0}(t) \\ \varphi_{1,1}(t) \\ \varphi_{2,1}(t) \\ \varphi_{3,1}(t) \end{pmatrix} \\ &= \prod_{i=2}^{n} \begin{pmatrix} d_i & a_i & b_i & c_i \\ d_i & t^2a_i & tb_i & tc_i \\ d_i & ta_i & tb_i & c_i \\ d_i & ta_i & tc_i & tc_i \end{pmatrix} \begin{pmatrix} 1 \\ a_1t^2 + (b_1 + c_1)t + d_1 \\ (a_1 + b_1)t + c_1 + d_1 \\ (a_1 + c_1)t + b_1 + d_1 \end{pmatrix} \end{aligned}$$

Lemma 4.3.3

If for all k,

$X_{k,1},\cdots,X_{k,p}$ *are independent , for all p, such that*

$$X_{k,j} \sim Bernoulli\,(p_{k,j}), \qquad with \qquad p_{k,j} = \frac{a_j}{a_j+b_j+k-1},\ j=1,2,\cdots p$$

and a_j, b_j be positive reals

Then

1)*S admits the following representation*

$$S \mid T \sim \ Poisson\,(T)$$

where

$$T \stackrel{d}{=} \sum_{i=1}^{p} a_i U_i$$

and $\{U_i\}_{1\le i\le p}$ are independent with beta(a_i,b_i) distribution

2)

$$\mathbb{E}\,(S) = \sum_{i=1}^{p} \frac{a_i^2}{a_i+b_i}$$

$$\mathbb{V}ar\,(S) = \sum_{i=1}^{p} \frac{a_i^2(3a_ib_i + a_i^2 + b_i^2 + b_i)}{(a_i+b_i)^2(a_i+b_i+1)}$$

3)

$$\mathbb{P}\,(S=j) = \int_{\mathbb{R}} \frac{x^j e^{-x}}{j!} f_T(x)dx$$

$$with \qquad f_T(x) = \bigotimes_{i=1}^{p} g_i(x) \qquad and \qquad g_i(x) = \frac{1}{B(a_i,b_i)}(\frac{1}{a_i})^{a_i+b_i-1}x^{a_i-1}(a_i-x)^{b_i-1}, \qquad 0\le x\le a_i.$$

with $\otimes$, is the convolution product.

Proof :

1) We have that

$$S_n = \sum_{k=1}^{n} \overline{X}_k \overline{X}_{k+1}^T = \sum_{j=1}^{p} Z_{j,n} \tag{4.3.13}$$

with $Z_{j,n} = \sum_{k=1}^{n} X_{k,j} X_{k+1,j}$ and $(Z_{j,n} n \geqslant 1)$ are independent for all $j = 1, \cdots, p$.

Therefore $Z_j = \lim\limits_{n \longmapsto \infty} Z_{j,n}$ admits the following representation

$$Z_j \mid T_j \sim Poisson(T_j) \quad \text{where} \quad T_j \overset{d}{=} a_j U_j \quad \text{and} \quad U_j \sim beta(a_j, b_j)$$

We conclude that

$S = \lim\limits_{n \longmapsto \infty} S_n = \lim\limits_{n \longmapsto \infty} \sum_{j=1}^{p} Z_j^n$ admits the following representation

$$S \mid T \sim \; Poisson\,(T)$$

Where

$$T \overset{d}{=} \sum_{i=1}^{p} a_i U_i$$

and $\{U_i\}_{1 \leq i \leq p}$ are independent with $beta(a_i, b_i)$ distribution

2)Trivially

$$\begin{aligned}
\mathbb{E}\,(S) &= \mathbb{E}\,(\mathbb{E}\,(S \mid T)) \\
&= \mathbb{E}\,(T) \quad , \qquad\qquad S \mid T \sim Poisson(T) \\
&= \mathbb{E}\left(\sum_{i=1}^{p} a_i U_i\right) \qquad\qquad T \overset{d}{=} \sum_{i=1}^{p} a_i U_i \\
&= \sum_{i=1}^{p} a_i \mathbb{E}(U_i) \\
&= \sum_{i=1}^{p} \frac{a_i^2}{a_i + b_i}
\end{aligned}$$

And since

$$\mathbb{V}ar(S) = \mathbb{E}\,(\mathbb{V}ar(S \mid T)) + \mathbb{V}ar\,(\mathbb{E}(S \mid T)) \tag{4.3.14}$$

Therefore

$$\begin{aligned}
\mathbb{V}ar(S) &= \mathbb{E}\left(\mathbb{V}ar(S \mid T)\right) + \mathbb{V}ar\left(\mathbb{E}(S \mid T)\right) \\
&= \mathbb{E}\left(T\right) + \mathbb{V}ar\left(T\right) \qquad\qquad , S \mid T \sim Poisson(T) \\
&= \sum_{i=1}^{p} \frac{a_i^2}{a_i + b_i} + \sum_{i=1}^{p} a_i^2 Var(U_i) \qquad , \ T \overset{d}{=} \sum_{i=1}^{p} a_i U_i \\
&= \sum_{i=1}^{p} \frac{a_i^2}{a_i + b_i} + \sum_{i=1}^{p} a_i^2 \left(\frac{a_i b_i}{(a_i + b_i)^2 (a_i + b_i + 1)}\right) \\
&= \sum_{i=1}^{p} \frac{a_i^2 (3a_i b_i + a_i^2 + b_i^2 + b_i)}{(a_i + b_i)^2 ((a_i + b_i + 1)}
\end{aligned}$$

3)

$$\begin{aligned}
\mathbb{P}\left(S = j\right) &= \int_{\Omega} \mathbb{P}\left(S = j/T\right) d\mathbb{P}_T \\
&= \int_{-\infty}^{+\infty} \frac{x^j e^{-x}}{j!} f_T(x) dx \\
&= \int_{-\infty}^{+\infty} \frac{x^j e^{-x}}{j!} f_{\sum_{i=1}^{p} a_i U_i}(x) dx
\end{aligned}$$

Since $T \overset{d}{=} \sum_{i=1}^{p} a_i U_i$ and $\{U_i\}_{1 \leq i \leq p}$ are independent with $beta(a_i, b_i)$ distributions.

So

$$f_T(x) = \overset{p}{\underset{i=1}{\otimes}} g_i(x) \qquad \text{and} \qquad g_i(x) = \tfrac{1}{B(a_i, b_i)} (\tfrac{1}{a_i})^{a_i + b_i - 1} x^{a_i - 1} (a_i - x)^{b_i - 1}, \qquad 0 \leq x \leq a_i.$$

with $\otimes$, is the convolution product.

Lemma 4.3.4

If for all k,

$X_{k,1}, \cdots, X_{k,p}$ are independent sucht that

$$X_{k,j} \sim Bernoulli\,(p_j), \qquad \textit{with} \qquad 0 < p_j < 1, \ j = 1, 2, \cdots p$$

(ie: for j fixed $\{X_{k,j}\}_k$ are i.i.d with $X_{k,j} \sim Bernoulli\,(p_j)$)

Then,

S_n admits the following representation

$$\mathcal{L}\left(S_n\right) = \overset{p}{\underset{j=1}{\otimes}} B_2^{II}(n, p_j)$$

With $B_2^{II}(n, p_j)$ be the type II binomial distribution of order 2.

Lemma 4.3.5

Let

$$B = \begin{pmatrix} tp & 1-p \\ p & t(1-p) \end{pmatrix}, \qquad t \in [0,1] \quad , p \in [0,1].$$

Then for all $n \geqslant 0$

$$B^n = \begin{pmatrix} b_{1,1}^{(n)}(t,p) & b_{1,2}^{(n)}(t,p) \\ b_{2,1}^{(n)}(t,p) & b_{2,2}^{(n)}(t,p) \end{pmatrix} \tag{4.3.15}$$

with

$$b_{2,2}^{(n)}(t,p) = b_{1,1}^{(n)}(t,1-p)$$

And

$$b_{2,1}^{(n)}(t,p) = b_{1,2}^{(n)}(t,1-p)$$

Proof :

Since

$$B^{n+1} = \begin{pmatrix} b_{1,1}^{(n)}(t,p) & b_{1,2}^{(n)}(t,p) \\ b_{2,1}^{(n)}(t,p) & b_{2,2}^{(n)}(t,p) \end{pmatrix} \begin{pmatrix} tp & 1-p \\ p & t(1-p) \end{pmatrix} \tag{4.3.16}$$

$$= \begin{pmatrix} tpb_{1,1}^{(n)}(t,p) + pb_{1,2}^{(n)}(t,p) & (1-p)b_{1,1}^{(n)}(t,p) + t(1-p)b_{1,2}^{(n)}(t,p) \\ tpb_{2,1}^{(n)}(t,p) + pb_{2,2}^{(n)}(t,p) & (1-p)b_{1,2}^{(n)}(t,p) + t(1-p)b_{2,2}^{(n)}(t,p) \end{pmatrix} \tag{4.3.17}$$

Therefore we conclude by induction on $n\square$.

Let us define, for all

$$\alpha_n(t,p) = p\ b_{1,1}^{(n-1)}(t,p) + (1-p)b_{2,1}^{(n-1)}(t,p), \qquad n \geqslant 0$$

Proposition 4.3.6

The generating function of S_n satisfies,

$$\varphi_{S_n}(t) = \beta_n(t,p) + \beta_n(t,1-p)$$

with

$$\beta_n(t,p) = \alpha_n(p,t)(pt+1-p)$$

Proof :

We suppose $a_n = d_n = 0$ and $b_n = 1 - c_n = p$,by propostion 1,we have for all $n \geq 2$.

$$\begin{pmatrix} \varphi_{0,n}(t) \\ \varphi_{1,n+1}(t) \\ \varphi_{2,n+1}(t) \\ \varphi_{3,n+1}(t) \end{pmatrix} = \begin{pmatrix} 0 & A \\ 0 & B \end{pmatrix}^{n+1} \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & AB^n \\ 0 & A^{n+1} \end{pmatrix} \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & AB^n \\ 0 & A^{n+1} \end{pmatrix} \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

With

$$A = \begin{pmatrix} p & 1-p \\ tp & t(1-p) \end{pmatrix} \qquad and \qquad B = \begin{pmatrix} tp & 1-p \\ p & t(1-p) \end{pmatrix}$$

By lemma 3 , we have

$$AB^n = \begin{pmatrix} p & 1-p \\ tp & t(1-p) \end{pmatrix} \begin{pmatrix} b_{1,1}^{(n)}(t,p) & b_{1,2}^{(n)}(t,p) \\ b_{2,1}^{(n)}(t,p) & b_{2,2}^{(n)}(t,p) \end{pmatrix}$$
$$= \begin{pmatrix} \alpha_n(t,p) & \alpha_n(t,1-p) \\ t\alpha_n(t,p) & t\alpha_n(t,1-p) \end{pmatrix}$$

Therefore, the generating function of S_n satisfied,

$$\varphi_{S_n}(t) = \beta_n(t,p) + \beta_n(t,1-p)$$

with

$$\beta_n(t,p) = \alpha_n(p,t)(pt+1-p)$$

Theorem 4.3.7

Whenver $a_n = d_n = 0$ *and* $b_n = 1 - c_n = p$ *for all* $n \geq 2$.

$$\varphi_{0,n-1}(t) = \frac{1}{\lambda_+ - \lambda_-}[(t{+}1)(2p^2t^2{-}pt^2)(\lambda_+^n{-}\lambda_-^n){+}(t{+}1)(\lambda_+{-}pt)\lambda_+^n{+}(1{-}2p)\lambda_-^n\,((pt+1-p)\lambda_+ + (t-pt+p)\lambda_-)]$$

Where

$$\lambda_+ = \left(t + \sqrt{t^2 - 4p(1-p)(t^2-1)}\right)/2 \qquad \textit{and} \qquad \lambda_- = t - \lambda_+$$

Proof :

We suppose $a_n = d_n = 0$ and $b_n = 1 - c_n = p$,by propostion 1,we have for all $n \geq 2$.

$$\begin{pmatrix} \varphi_{0,n-1}(t) \\ \varphi_{1,n}(t) \\ \varphi_{2,n}(t) \\ \varphi_{3,n}(t) \end{pmatrix} = \begin{pmatrix} 0 & 0 & p & 1-p \\ 0 & 0 & tp & t(1-p) \\ 0 & 0 & tp & 1-p \\ 0 & 0 & p & t(1-p) \end{pmatrix}^n \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & A \\ 0 & B \end{pmatrix}^n \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & AB^{n-1} \\ 0 & A^n \end{pmatrix} \begin{pmatrix} 1 \\ t \\ pt+1-p \\ p+t(1-p) \end{pmatrix}$$

With

$$A = \begin{pmatrix} p & 1-p \\ tp & t(1-p) \end{pmatrix} \qquad \textit{and} \qquad B = \begin{pmatrix} tp & 1-p \\ p & t(1-p) \end{pmatrix}$$

With λ_+ and λ_- as eigenvalue of B, the above systeme yield the stated solution.

Corollary 4.3.8

If $a_n = d_n = 0$ *and* $b_n = 1 - c_n = p$ *for all* $n \geq 2$, *then*

$$p \in \left\{1, \frac{1}{2}\right\} \qquad \Leftrightarrow \qquad S_n \overset{d}{=} Bin(n,p). \tag{4.3.18}$$

Proof :

i) $\Rightarrow$:

Trivially for $p = 1$, $S_n = n$.

By theorem 1 ,for $p = 1/2$ we have that

$$\varphi_{(0,n-1)}(t) = \mathbb{E}\left(t^{S_{n+1}}\right) = (\frac{t+1}{2})^{n+1} \tag{4.3.19}$$

ii) $\Leftarrow$: We suppose $S_n \overset{d}{=} Bin(n,p)$,then

$$\begin{aligned} \mathbb{E}\left(S_n\right) - np &= \sum_{k=1}^{n} \mathbb{E}\left(X_{1,k}X_{1,k+1}\right) + \sum_{k=1}^{n}\left(X_{2,k}X_{2,k+1}\right) - np \\ &= n(p^2 + (1-p)^2) - np \\ &= n(p-1)(p-1/2) = 0 \end{aligned}$$

Therefore

$$p \in \left\{1, \tfrac{1}{2}\right\}$$

Example (**Bernoulli urn**) :

Consider an urn initially containing α white and β black balls.One ball at a time is drawn at random from the urn ant its color is inspected.It is then returned to the urn (sampling with replacement). This generates a sequence or random variables $\{X_n\}$,(resp:$\{Y_n\}$) each equal to 0 or 1 (resp :1 or 0) according to the color white or black respectively.

If $\alpha = \beta$, then

$$S_n := \sum_{k=1}^{n} X_k \overset{\mathcal{L}}{=} \sum_{k=1}^{n} Y_k \overset{\mathcal{L}}{=} \sum_{k=1}^{n} X_k X_{k+1} \overset{L}{=} \sum_{k=1}^{n} Y_k Y_{k+1} \overset{\mathcal{L}}{=} Bin(n, 1/2)$$

Theorem 4.3.9 *Whenver* $b_n = c_n = 0$ *and* $a_n = 1 - b_n = p_n$ *for all* $n \geq 2$ *with* $p_n = \frac{a}{a+b+n-1}$, *and* a, b *be positive reals*

Then

S *almost the following representation*

$$S \mid T \sim \; Poisson\,(2T)$$

Where

$$T \overset{d}{=} aU \qquad and \qquad U \sim beta(a,b).$$

Proof :

Since $S_n = \sum_{k=1}^{n} (X_{k,1}X_{k+1,1} + X_{k,2}X_{k+1,2})$

With

$$a_n = \mathbb{P}\,(X_{n,1} = 1, X_{n,2} = 1)\,, b_n = \mathbb{P}\,(X_{n,1} = 1, X_{n,2} = 0)$$

$$c_n = \mathbb{P}\,(X_{n,1} = 0, X_{n,2} = 1)\,, c_n = \mathbb{P}\,(X_{n,1} = 0, X_{n,2} = 0)$$

If $b_n = c_n = 0$ and $a_n = 1 - b_n = p_n$

Then

$$X_{k,1} = X_{k,2} \quad a.s \qquad \text{for all} \quad k. \tag{4.3.20}$$

Therefore

$$S_n = \sum_{k=1}^{n} 2X_{k,1}X_{k+1,1} \quad a.s \tag{4.3.21}$$

with $X_{k,1} \sim \text{Bernouilli}\,(p_k)$. We conclude by Holst's theorem .

Bibliography

[1] Aldous.D. Exchangeability and related topics.Ecole d'été de probabilités de Saint-Flour,XIII-1983,Lecture note in Math.1117,Springer Berlin (1989).

[2] B. Arthur, Y. Ermoliev, and Y. Kaniovskii. A generalized urn problem and its applications. Cybernetic 19:61-71, (1983).

[3] K. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. Ann. Math. Statist., 39:1801-1817, (1968).

[4] K. Athreya. Some results on multitype continuous time Markov branching processes. Ann. Math. Statist., 38:347-357, 1968.

[5] P. Cartier, Functiones harmoniques sur un arbre, in: Symposia Mathematica, vol. 9, Academic Press, London, 1972, pp. 203-270.

[6] Blackwell,D and MacQueen,J.B. Ferguson distribution via Pólya urn schemes,Annal of Statisics,1 353-355 (1973).

[7] D.M. Cifarelli, P. Muliere, P. Secchi, Prior processes for Bayesian nonparametrics, Technical Report, 377/P, Politecnico di Milano, 1999.

[8] Chow . Y.S, Tei.cher H. Probability theory. Springer Berlin (1992).

[9] Eggenberger, F. and Pòlya, G. Über die Statitik Verketter Vorgänge Zeitschrift fur Angewandte Mathematik 3 279-289. (1923)

[10] De Finetti, B . funzione caratteristica di un fenomeno aleatorio Atti Accad. Naz Linei Mem. Cl. Sci. Fis. Mat Natur. Sez (6) 4 251-299 (1931).

[11] De Finetti ,B. La prévision ses lois logiques, ses souces subjectives. Ann. Inst. H. Poincaré 7 1-68 (1937)

[12] De Finetti ,B. Theory of Probability , New York. (1975).

[13] Feller,W. An Introduction to Probability Theory and its Applications, vol. I. John Wiley - Sons, New York, third edition, 1968.

[14] Feller,W An Introduction to Probability Theory and its Applications, vol. II. John Wiley - Sons, New York, second edition, 1971.

[15] Ferguson, T.S., 1973. A Bayesian analysis of some nonparametric problems. Ann. Statist. 1, 209-230.

[16] Freedman, D.A. Bernard Freidman's urn. Ann. Math. Staist 36. 956-970 (1965).

[17] Friedman, B. A simple urn model. Comm. Pure Appl. Math 2 59-70 (1949).

[18] Hall.P, C.C.Heyde. Martingale Limit Theory and Its Application. Academic Press INC (1980).

[19] Hill, B.M., Lane, D., Sudderth, W., 1980. A strong law for some generalized urn processes. Ann. Probab. 8 (2), 214-226.

[20] Hill, B, Lane, D., Sudderth,W., 1987. Exchangeable urn processes. Ann. Probab. 15 (4), 1586-1592.

[21] Holst .L. On the number of consecutive successes in Bernoulli trials. Departement of Mathematics , Royal Institute of Technology SE-100 44 Stockholm, Sweden. (2006).

[22] Hunt, G.A. Martingales et Processus de Markov. Masson. (1966).

[23] Johnson,N.L, and ,Kotz,S.Urn Model and Their Applications.New York:John Wiley & Sons (1977).

[24] Johnson,N.L, Kotz,S, and Kemp ,A W. Univariate Discrete Distributions,Second edition ,New York John: Wiley & Sons (1992).

[25] of Dependent Bernoulli and Its Relationship to a Matching Type Problem, Technical Report 2686, CRM , Université de Montreal (2000). Joff, A. , Marchand, E. , Perron, F. , and Popadiuk, P. On a Particular Sum

[26] Joff, A. , Marchand, E. , Perron, F. , and Popadiuk, P. On sums of products of Benoulli variables and random permutations. Journal of Theoretical Probabiity, 17 , 285-292 , (2004).

[27] Kotz. S, Mahmoud .H, and Robert. P. On generalized urn models. Sta. Prob Let, 49:163-173,(2000).

[28] Limic,V and P. Tarrès. Attracting edge and strongly edge reinforced random walks. Preprint, page 25, 2006.

[29] Mahmoud,H On rotations in fringe-balanced binary trees. Information Processing Letters, 65:41-46, 1998.

[30] Mahmoud, H P´olya urn models and connections to random trees: a review. J. Iranian. Stat. Soc., 2:53-114, 2003. Mahmoud,H Polya-type urn models with multiple drawings. J. Iranian Stat. Soc., 3:165-73, 2004.

[31] Mauldin,D W. Sudderth, and S. Williams. Poolya trees and random distributions. Ann. Statist., 20:1203-1221, 1992.

[32] May, C., Paganoni, A.M., Secchi, P., 2002. On a two-color generalized Polya urn. Quaderno di Dipartimento n. 505/P, Aprile 2002, Dipartimento di Matematica, Politecnico di Milano.

[33] Meyer, P.A. (1969). Un lemme de théorie des martingales. Séminaire de probabilité de Strasbourg, Tome 3: 08, p. 143, LMN 88.

[34] Mori , T.F. On the distribution of sums of overlapping products .Acta Scientiarum Mathematica (Szeged), 67. 833-422 (2001).

[35] Muliere, P., Secchi, P. and Walker, S. (2005). Partially exchangeable processes indexed by the vertices of a k-tree constructed via reinforcement. Stochastic Processes and their Applications, 115, 661-677.

[36] Neveu,J Arbres et Processus de Galton-Watson .Ann.Ins.H Poincaré 22, 199-207, 1986.

[37] Paganoni, A.M., Secchi, P., 2004. Interacting reinforced urn systems. Adv. in Appl. Probab. 36 (3), 791-804.

[38] Pemantle,R Phase transition of reinforced random walk and rwre on trees. Annals of Probability, 16:1229-1241, 1988.

[39] Pemantle,R Random processes with reinforcement. Doctoral Dissertation. M.I.T., 1988.

[40] Pemantle,R Nonconvergence to unstable points in urn models and stochastic approximations. Annals of Probability, 18:698-712, 1990.

[41] Pemantle,R A time-dependent version of polya's urn. Journal of Theoretical Probability, 3:627-637, 1990.

[42] Pemantle,R When are touchpoints limits for generalized Polya urns? Proceedings of the American Mathematical Society, 113:235-243, 1991.

[43] E. Regazzini, G. Petris, Some critical aspects on the use of exchangeability in Statistics, J. Ital. Statist. Soc. 1 (1992) 103-130.

[44] Robbins, H. and Siegmund, D. (1971). A convergence theorem for non negative almost supermartingales and some applications. In J.S. Rustagi, Optimizing Methods in Statistics, 233-257, New York. Academic Press.

[45] Sethuraman, J. and Sethuraman, S . On counts of Bernoulli strings and connections to rank orders and random permutations . In A festschrift for Herman Rubin. IMS Lecture Note Monograph Series, 45 . 140-152, Institute of Mathematical Statistics , Beachwood , Ohio.

[46] Walker, S., Muliere, P., 1997. Beta-Stacy processes and a generalization of the Polya urn scheme. Ann. Statist. 25 (4), 1762-1780.

[47] Walker, S.G., Muliere, P., 2003. Reinforcement and finite exchangeability. Studi Statistici, N. 71, Istituto di Metodi Quantitativi, Università Bocconi, Milano.

[48] W. Woess, Catene di Markov e Teoria del Potenziale nel Discreto, Quaderni dell'Unione Matematica Italiana, vol. 41, Pitagora Editrice, Bologna, 1996.