A reinforced urn process indexed by the vertices of a recombinant binary tree

Djilali Ait Aoudia, Paolo Bulla and Pietro Muliere

November 28, 2008

Abstract

According to Muliere, Secchi and Walker (2005) a reinforced random process indexed by k-ary tree, we introduce a class of discrete-time stochastic processes generated by interacting systeme of reinforced urns, we show that such processes are converge in distribution and asymptotically partially exchangeable.

Keywords: Urn model; reinforcement; Partial exchangeability.

AMS 2000 *subject classifications* : 60F15, 60G09, 60G07, 60E05

1 Introduction

According to Muliere, Secchi and Walker (2005) a reinforced random process indexed by a k-ary tree can be described as a stochastic process representing the outcomes of drawings in a system of urns, whose compositions are determined by the interaction of the geometrical structure (the tree) with a Pólya-like reinforcing rule.

An effective comprehension of this model can requirer a brief review of tree and Pólya urns.

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. In 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 ϕ, the *root*, is identified and the tree is considered as a directed graph where the vertices go in the direction away from ϕ. Given a vertex $\tau \in T$, there is a unique path $\pi(\phi, \tau)$ from ϕ to τ. The number of edges in $\pi(\phi, \tau)$ is the *level number* of τ and denoted $|\tau|$. Notice that $|\phi| = 0$. For all the vertices in $\tau \in T$ but the root, there exists a vertex $\sigma = \overleftarrow{\tau}$ called the *parent* of level $|\tau| - 1$ and with an edge to τ. Alternatively τ is said to be a *child* of σ and two 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 sampling scheme is run: (i) all the urns have the same initial composition of the two colors' balls and (ii) 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 that the compositions of the children urns are reinforced by the result of the parent's draw , 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 (1938) the collection of the sequences is said *partially exchangeable.* Section 2 provides more details about this process summarizing the main results of Muliere, Secchi and Walker (2005).

Retaining these basic ideas, Section 3 substitutes the original tree with a different structure with the purpose of introducing a closer dependence. The new frame G is actually a *recombinant binary tree* in which every vertex has two parents and two children except the root (without parents) and the vertices along the left and right extreme branches have two children and just one parent. Furthermore, vertices at the side of each other have a child in common.

The concern here is to study how the properties of the random sequences are offected with this new geometrical structure where the updated compositions may reflect a sort of merging between the outcomes of the parents. Section 4 tackles the problem of the asymptotic behaviour of this process and Section 4 provides a Poisson approximation for the total number of balls of color 1 at a given level.

2 A dichotomous reinforced process indexed by a k-ary tree

Let us recall Muliere, Secchi and Walker (2005) process and its main properties.

Given a probability space $(\Omega, \mathcal{F}, P)$ and a k-ary tree T, $X = \{X_\tau, \tau \in T\}$ is defined recursively. Let a and b be two real positive numbers so that the random variable X_ϕ indexed by the root is distributes Bernoulli(p_ϕ) with $p_\phi = \frac{a}{a+b}$

$$X_\phi \sim Bernoulli\,(p_\phi)\,.$$

For $n \geq 0$, let $\mathcal{F}_n = \sigma\left\{X_\tau :\, \tau \in T\, and\; |\tau| < n\right\}$ be the σ-field generated by the random variables corresponding to the vertices of T with level less or equal to n. Given $\mathcal{F}_n$, the k^{n+1} random variables X_τ, with $|\tau| = n$, are conditionally independent and Bernoulli distributed with parameter

$$p_\tau = \frac{a + \sum_{i=0}^{n-1} X_{\sigma_i}}{a+b+n}$$

if $\pi(\phi, \tau) = (0 = \sigma_0, \sigma_1, \dots, \sigma_{n-1}, \tau)$. Random variables indexed by siblings have the same conditional distribution.

As sketched in the introduction the basic interpretation is that, at each vertex of the tree, there is a two-color (0 and 1) urn; the parameters a and b fix the initial composition of the urn in ϕ, while the random variables X_τ represents the outcome of the drawings. For a given vertex, the composition of its k children is the same of the parent plus 1 ball of the color drawn from the parent urn.

Before stating the proposition about the sequences of variables describing the drawing's outcomes, some additional definition are needed.

Let us define E be the *space of the ends* of T, that is the set of all infinite sequences $\varepsilon = (\phi = \varepsilon_0, \varepsilon_1, \varepsilon_2, \dots)$ of vertices of a k-ary tree with the property that $\forall i \geq 0$

$$|\varepsilon_i| = i \quad \text{and} \quad \varepsilon_i = \overleftarrow{\varepsilon_{i+1}}.$$

ε can be interpreted as a path on the tree T connecting the root with a point at a infinite level. Let $\overline{T} = T \cup E$ and define, for $\xi \neq \eta$, $\xi \wedge \eta$ as the *confluent* of η and ξ to be the vertex with the highest level belonging both to $\pi(0, \xi)$ and $\pi(0, \eta)$: A distance in $\overline{T}$ is defined as follows

$$d(\xi, \eta) = \begin{cases} \exp(-|\xi \wedge \eta|) & \xi \neq \eta \\ 0 & \text{otherwise} \end{cases}$$

for $\xi, \eta \in \overline{T}$. With this metric the space $\overline{T}$ is compact and totally unconnected; moreover, T is a discrete subspace of $\overline{T}$ and E is a compact subspace of $\overline{T}$.

An interesting matter of enquiry is the asymptotic behaviour of the process

$$p = \{p_\tau : \tau \in T\}$$

representing the proportion of balls of color 1 contained in the urns indexed by the different vertices of T.

Given a generic end $\varepsilon = (\phi, \varepsilon_1, \varepsilon_2, \dots)$ the sequence of random variables $\{X_\phi, X_{\varepsilon_1}, X_{\varepsilon_2}, \dots\}$ is a 0-1 Pólya sequence; hence the sequence is exchangeable, that is equivalent to saying that $X_\phi, X_{\varepsilon_1}, X_{\varepsilon_2}, \dots$ are conditionally independent $Bernoulli(p_\varepsilon)$ where

$$p_\varepsilon = \lim_{n\to\infty} p_{\varepsilon_n} \quad a.s.$$

Furthermore the almost sure limit p_ε is a random variable with distribution $Beta(a, b)$.

The geometrical structure of the k-ary tree induces a particular kind of dependence between the different sequences. Sequences along two different ends ε and η have the same marginal distribution, but they are dependent. If $n = |\varepsilon \wedge \eta|$, the subsequences $X_{\varepsilon_{n+1}}, X_{\varepsilon_{n+2}}, \dots$ and $X_{\eta_{n+1}}, X_{\eta_{n+2}}, \dots$ are conditionally independent given $\mathcal{F}_{n+1}$. The larger $n = |\varepsilon \wedge \eta|$ is, the greater the dependence. The following proposition summarizes some results.

Proposition 1. *[Muliere, Secchi and Walker (2005)] Let $\varepsilon, \eta \in E$, $\varepsilon \neq \eta$ and $n = |\varepsilon \wedge \eta|$.*

1. *For all* $x_1, x_2 \in [0,1]$

$$\begin{aligned} &P\left[p_\varepsilon \le x_1, p_\eta \le x_2\right] \\ &= \sum_{j=0}^{n+1} \binom{n+1}{j} \frac{B(a+j, b+n+1-j)}{B(a,b)} \prod_{i=1}^{2} \Psi(x_i | a+j, b+n+1-j) \end{aligned}$$

where $\Psi(x|c,d)$ *is the cumulative distribution function of a beta random variable with parameters* (c,d)*.*

2. *For* $x \in [0,1]$*, on a set of probability one,*

$$P\left[p_\eta \le x | p_\varepsilon\right] = \sum_{j=0}^{n+1} \binom{n+1}{j} p_\varepsilon^j (1-p_\varepsilon)^{(n+1-j)} \Psi(x|a+j, b+n+1-j).$$

3. *For* $s \ge 0$ *and* $i_0, \ldots, i_s \in \{0,1\}$*,*

$$\begin{aligned} &P\left[X_{\eta_0} = i_0, \ldots, X_{\eta_s} = i_s | \, p_\varepsilon\right] \\ &= \begin{cases} p_\varepsilon^{\xi_s} (1-p_\varepsilon)^{s+1-\xi_s} & s \le n \\ p_\varepsilon^{\xi_n} (1-p_\varepsilon)^{n+1-\xi_n} \frac{B(a+\xi_s, b+s+a-\xi_s)}{B(a+\xi_n, b+n+a-\xi_s))} & s > n \end{cases} \end{aligned}$$

where $\xi_s = \sum_{r=0}^{s} i_r$*.*

Remark 1. As a consequence, of the correlation between p_ε and p_η is

$$Corr(p_\varepsilon, p_\eta) = \frac{1 - \log d(\varepsilon, \eta)}{a + b + 1 - \log d(\varepsilon, \eta)}$$

where d is the distance on $\overline{T}$ defined above.

3 A reinforced urn process indexed by the vertices of a recombinant binary tree

In this section we propose a new stochastic process indexed by the vertices of a recombining binary tree G. As seen in the Introduction, G is obtained by starting from the simplest k-ary tree (the binary) so that each vertex splits into two children, and 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 along the left and right extreme branches (just one parent).

The process $X = \{X_\gamma,\, \gamma \in G\}$ is defined quite similarly to the process in the previous section. Given two positive real numbers a and b and $p_\phi = \frac{a}{a+b}$, let

$$X_\phi \sim Bernoulli\,(p_\phi)$$

and $\mathcal{F}_n = \sigma\,\{X_\gamma : \gamma \in G\, and\; |\gamma| \leq n-1\}$. Furthermore, for a given vertex τ, define the set of its ancestors $A(\tau) := \{\gamma \in G : \exists\, \pi(\phi, \tau) \text{ s.t. } \gamma \in \pi(\phi, \tau)\}$. Assume that, given $\mathcal{F}_n$ the random variables X_τ with $|\tau| = n$ are conditionally independent such that $X_\tau | \mathcal{F}_n \sim Bernoulli(p_\tau)$

$$p_\tau = \frac{a + \sum_{\gamma \in A(\tau)} X_\gamma}{a + b + \# A(\tau)}. \tag{1}$$

To ease the exposition every vertex $\tau \in G$ can be labeled by a couple (i, n) for $n = 0, 1, \ldots$ and $i = 0, \ldots, n$ where n is the level of the node and i increases from left to right. So the root ϕ is $(0, 0)$, the two nodes at level 1 $(0, 1)$ and $(1, 1)$ and so on. More generally, if (i, n) is a node at level n, its children are $(i, n+1)$ (left) and $(i+1, n+1)$ (right). Figure 1 displays the graph G.

Relying on the usual interpretation of the Pólya urn as a Bayesian learning and forecasting process, the new geometrical structure introduces a more

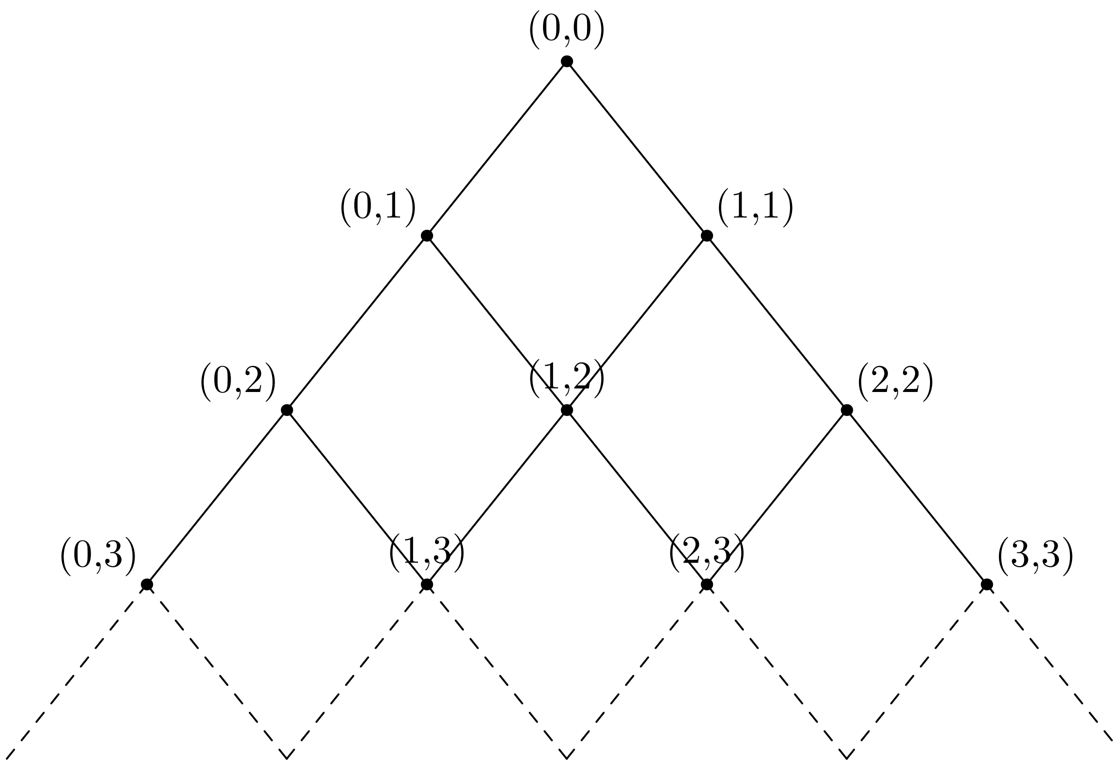

Figure 1: the recombining binary tree G.

complex prediction rule where, in order to fix the predictive distribution of the color of the ball extracted in the node τ, all the outcomes in the urns in the set of the ancestors must be considered and not only the nodes of particular path from ϕ to τ.

In determining these predictive distributions the number of balls added as reinforcement to the initial composition is a key element. For a given vertex $\tau = (i, n) \in G$ this number corresponds to the number of ancestors, that is the cardinality of the set $A(i, n)$. The next lemma computes this cardinality.

Lemma 1. *For* $n = 0, 1, \dots$ *and* $i = 0, \dots, n$

$$\#A(i, n) = i(n - i) + n.$$

Proof. By construction

$$A(i, n) = \left\{(k, n - j) : k = (i - j)^+, \dots, i \wedge (n - j),\ j = 1, \dots, n\right\}.$$

Trivially we have $\#A(0, n) = n$ and, by recurrence,

$$A(i, n) = \bigcup_{k=0}^{n-i} \{(i, i + k)\} \cup A(i - 1, n - 1).$$

Therefore it follows that

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

$\square$

Using the couples (i,n) as indices we can rewrite $X_{(0,0)} \sim Bernoulli\left(p_{(0,0)}\right)$, $p_{(0,0)} = \frac{a}{a+b}$, and for $n \geq 0$ and $0 \leq i \leq n$:

$$X_{(i,n)}|\mathcal{F}_n \sim Bernoulli\left(p_{(i,n)}\right)$$

with

$$p_{(i,n)} = \frac{a + \sum_{k=(i-j)^+,\ldots,i\wedge(n-j),\, j=1,\ldots,n} X_{(k,n-j)}}{a+b+i(n-i)+n}. \tag{2}$$

Exception, unlike the process introduced by Muliere, Secchi and Walker (2005) a sequence of X_γ's obtained by starting from the root and moving down from a parent to one of its children is not exchangeable and the corresponding sequence of p_γ's is not a martingale.

Nevertheless the expected value of the proportion of balls of color 1 is constant.

Lemma 2. *For $n \geq 0$ and $0 \leq l \leq n$*

$$E\left[p_{(l,n)}\right] = \frac{a}{a+b}.$$

Proof. We have $E\left[p_{(0,0)}\right] = \frac{a}{a+b}$, so we proceed by induction assuming, for $i \leq m$ and $m \leq n$, that $E\left[p_{(i,m)}\right] = \frac{a}{a+b}$.

Notice that $E\left[p_{(n+1,n+1)}\right] = \frac{a}{a+b}$ because the sequence $\left\{p_{(n,n)}\right\}_{n\geq 1}$ is a $\mathcal{F}_n$-martingale.

Hence for $i \le n$ fixed, we have

$$
\begin{aligned}
E\left[p_{(i,n+1)}\right] &= \left[\frac{a + \sum_{k=(i-j)^+,\dots,i\wedge(n+1-j),\, j=1,\dots,n+1} E\left[E\left(X_{(k,n+1-j)}\middle|\,\mathcal{F}_n\right)\right]}{a+b+i(n+1-i)+n+1}\right] \\
&= \left[\frac{a + \sum_{k=(i-j)^+,\dots,i\wedge(n+1-j),\, j=1,\dots,n+1} E\left[p_{(k,n+1-j)}\right]}{a+b+i(n+1-i)+n+1}\right] \\
&= \left[\frac{a + \sum_{k=(i-j)^+,\dots,i\wedge(n+1-j),\, j=1,\dots,n+1} \frac{a}{a+b}}{a+b+i(n+1-i)+n+1}\right] \\
&= \frac{a + \frac{a}{a+b}(i(n+1-i)+n+1)}{a+b+i(n+1-i)+n+1} = \frac{a}{a+b}.
\end{aligned}
$$

□

As a consequence, we obtain the probability distribution of $X_{(l,n)}$.

Corollary 1. *For $n \ge 0$ and $0 \le l \le n$,*

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

Proof. For $n \ge 0$ and $0 \le l \le n$, let compute

$$
\begin{aligned}
E\left[e^{itX_{(l,n)}}\right] &= E\left[E\left(e^{itX_{(l,n)}}\middle|\,\mathcal{F}_n\right)\right] \\
&= E\left[p_{(l,n)}e^{it} + 1 - p_{(l,n)}\right] \\
&= \frac{a}{a+b}e^{it} + 1 - \frac{a}{a+b} = E\left[e^{itU}\right]
\end{aligned}
$$

with $U \sim Bernoulli\left(\frac{a}{a+b}\right)$. □

The processes $Y = \{Y_\gamma, \gamma \in G\}$ and $p = \{p_\gamma, \gamma \in G\}$ derived by X and defined like in the previous section enjoy the property of symmetry in that the random variables associated to vertices symmetric with respect to the vertical axis have the same law.

Lemma 3. *For $n \geq 0$ and $0 \leq i \leq n$:*

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

2. $Y_{(i,n)} \stackrel{L}{=} Y_{(n-i,n)}$.

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 $|\tau| = n$, define $P^1(\tau)$ as the set of the parents of τ that is the ancestors of τ of level $n-1$ and, more generally, for $j \in 1, \ldots, n$, $P^j(\tau)$ 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 $k = 1, \ldots, n-1$ the recursive relation $\bigcup_{\eta \in P^k(\tau)} A(\eta) = \bigcup_{j=k+1}^{n} P^k(\tau)$.

Now let $s : G \to G$, $s((i,n)) = (n-i,n)$ be the function relating a node to its symmetric with respect to the vertical symmetry axis of G. Notice that $A(n-i,n) = \{s(\tau), \tau \in A(i,n)\}$.

By the definition of the process in equation (1), we have, for fixed $\tau \in G$ and given x and $x_\eta, \eta \in A(\tau)$,

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

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

so that

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

Let consider a given node (i,n) and fix $\{x_\tau, \tau \in A(i,n)\}$. We have

$$\begin{aligned} P\left[X_\tau = x_\tau, \tau \in A(i,n)\right] &= \prod_{j=1}^{n-1} \prod_{\tau \in P^j(i,n)} P\left[X_\tau = x_\tau \middle| X_\eta = x_\eta, \eta \in A(\tau)\right] P\left[X_\phi = x_\phi\right] \\ &= \prod_{j=1}^{n-1} \prod_{\tau \in P^j(i,n)} P\left[X_{s(\tau)} = x_\tau \middle| X_{s(\eta)} = x_\eta, \eta \in A(\tau)\right] P\left[X_{s(\phi)} = x_\phi\right] \\ &= P\left[X_{s(\tau)} = x_\tau, \tau \in A(i,n)\right] \end{aligned}$$

(the x_η's are the different x_τ's corresponding to the different subsets of the above set).

So the random vectors $(X_\tau, \tau \in A(i,n))$ and $(X_\tau, \tau \in A(n-i,n))$ have the same probability distribution. As $Y_{(i,n)} = a + \sum_{\tau \in A(i,n)} X_\tau$ and $Y_{(n-i,n)} = a + \sum_{\tau \in A(n-i,n)} X_\tau$, it turns out that $Y_{(i,n)}$ and $Y_{(n-i,n)}$ have the same law too. Analogously we obtain the result for $p_{(i,n)}$ and $p_{(n-i,n)}$. $\square$

This lemma is very useful with respect to the study of the convergence of the process in that it allows to focus just on the sequences of the type $\left\{p_{(i,n)}\right\}_n$ omitting to consider the symmetric $\left\{p_{(n-i,n)}\right\}_n$.

4 Some asymptotic properties of the process

Failing the martingale properties of the sequence of proportions $\left\{p_{(i,n)}\right\}_n$, the asymptotic behaviour of the new stochastic process is more difficult to describe. This section gives some results.

For $i \geq 0$, we define the processes $\left\{\rho_{(i,n)}\right\}_n$ such that

$$\rho_{(i,n)} = \begin{cases} p_{(n,n)} & \text{if} \quad n < i \\ \\ p_{(i,n)} & \text{if} \quad n \geq i \end{cases} \tag{3}$$

We note that for a fixed $i \geq 0$ and $n < i$ the processes $\left\{\rho_{(i,n)}\right\}_n$ correspond to Pólya urn processes.

Theorem 1. *For fixed $i \geq 0$, there exists $\varphi_i : \mathbb{N} \to \mathbb{N}$, such that $\varphi_i(n) \nearrow \infty$ and the process $\left\{\rho_{(i,\varphi_i(n))}\right\}_{n\geq 0}$ converges a.s. to the random variable $\Theta \sim Beta(a,b)$.*

Proof. We know that for $i = 0$, the process $\left\{\rho_{(0,n)}\right\}_n$ is a Pólya sequence, 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 $\left\{\rho_{(i,n)}\right\}_{n\geq 0}$ is bounded, there exists $\varphi_i : \mathbb{N} \to \mathbb{N}$, such that $\varphi_i(n) \nearrow \infty$, and the process $\left\{\rho_{\varphi_i(n)}\right\}_{n\geq 0}$ converges almost surely to the random variable $L_i \in [0,1]$.

Now, we show that the limit is always Θ with $\Theta \sim Beta(a,b)$.

For $i = 1$, applying Theorem 2.13 in Hall and Heyde (1980) we get

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

We have

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

Therefore

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

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

In the general case, by similar computation we have that

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

and

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

So we conclude by induction that

$$\begin{cases} L_i = \frac{L_i + iL_{i-1}}{i+1} & \text{a.s.} \\ L_0 = \Theta & \text{a.s.} \end{cases} \tag{4}$$

□

Remark 2. We note that for all $\varphi_i : \mathbb{N} \longrightarrow \mathbb{N}$, such that $\varphi_i(n) \nearrow \infty$ and the process $\left\{\rho_{(i,\varphi_i(n))}\right\}_{n\geq 0}$ converges a.s, we have always $\lim_{n\to\infty} \rho_{(i,\varphi_i(n))} = \Theta$ with $\Theta \sim Beta(a,b)$, so by the monotone convergence's theorem, for all fixed k

$$\lim_{n\to\infty} \mathbb{E}\left[\rho^k_{(i,\varphi_i(n))}\right] = \mathbb{E}\left[\Theta^k\right] \tag{5}$$

Now,theorem 1 can be used to prove the convergence in law of the overall sequence $\left\{\rho_{(j,n)}\right\}_n$.

Theorem 2. *For fixed $j \geq 0$*

$$\rho_{(j,n)} \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} \Theta \sim Beta(a,b). \tag{6}$$

Proof. Let $\psi_{(j,n)}$ be the characteristic function of $\rho_{(j,n)}$. Then we have

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

Now, for fixed $i \geq 0$ and $k \geq 0$, we have

$$\mathbb{E}\left[\rho^k_{(i,n)}\right] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}\left[\Theta^k\right] \tag{7}$$

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

In fact, for fixed $i \geq 0$, $\rho_{(i,n)} \in [0,1]$, and therefore, for all $k \geq 0$, $\mathbb{E}[\rho^k_{(i,n)}] \in [0,1]$, so that $\lim_{n\to+\infty} \mathbb{E}[\rho^k_{(i,n)}]$ exists and is finite.

Now, we show that this limit is unique.

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

By Remark 2

$$\begin{cases} \ell_1 = \lim_{n \to +\infty} \mathbb{E}[\rho^k_{(i,n)}] = \mathbb{E}\left[\Theta^k\right] \\ \\ \ell_2 = \lim_{n \to +\infty} \mathbb{E}[\rho^k_{(i,n)}] = \mathbb{E}\left[\Theta^k\right] \end{cases} \tag{8}$$

We conclude that

$$\mathbb{E}\left[\rho^k_{(i,n)}\right] \underset{n \to +\infty}{\longrightarrow} \mathbb{E}\left[\Theta^k\right] \tag{9}$$

By the monotone convergence theorem

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

This shows that the distribution of the limiting distribution of $\rho_{(j,n)}$ is $Beta(a, b)$. $\square$

Theorem 1 implies also that the subsequences $\left\{X_{(i,\phi_i(n))}\right\}_n$ are asymptotically exchangeable. Before proving this fact, we recall the definition of asymptotic exchangeability.

Definition 1. *An infinite sequence* $(V_1, V_2, \cdots)$ *is called asymptotically exchangeable if*

$$(V_{j+1}, V_{j+2}, \cdots) \overset{d}{=} (Z_1, Z_2, \cdots) \textit{ as } j \longrightarrow \infty \tag{10}$$

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

Now we can state the following

Theorem 3. *The sequence* $\left\{X_{(i,\varphi_i(n))}\right\}_n$ *is asymptotically exchangeable.*

Proof. For a fixed $i \geq 0$, the conditional distribution of $X_{(i,\varphi_i(n)+1)}$ given $\mathcal{F}_{\varphi_i(n)} = \sigma\left(X_\tau, |\sigma| < \varphi_i(n)\right)$ is a *Bernouilli* $\left(p_{(i,\varphi_i(n))}\right)$; we have seen that, as $n \to \infty$, it converges with probability one to a *Bernouilli* (Θ) with $\Theta \sim Beta(a,b)$.

So

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

But

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

and Hunt's lemma (see Hunt (1966) or Meyer (1969)) entails

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

with $\mathcal{G}_\infty^{(i)} = \bigvee_{n=1}^{\infty} \mathcal{G}_n^i$, $\mathcal{G}_n^i = \sigma(X_{(i,\varphi_i(j))}, 0 \leq j \leq n)$ and $\Theta \sim beta(a,b)$. Finally by virtue of Lemma 8.2 in Aldous (1985) we get the asymptotic exchangeability of the subsequence of interest. □

5 Poisson approximation

Finally we provide a conditional Poisson approximation for the overall number of balls of color 1 at level n given all the history of the process up to level $n-1$.

Proposition 2. *Let $H_n = \sum_{i=0}^{n} X_{(i,n)}$ be the number of balls of color 1 sampled at level n and W_n a Poisson random variable with parameter $\lambda_n = \sum_{i=0}^{n} p_{(i,n)}$ for $p_{(i,n)}$ defined as in (2).*

Then for any A,

$$|P(H_n \in A|\,\mathcal{F}_n) - P(L_n \in A)| \leq \sum_{i=0}^{n} p_{(i,n)}^2.$$

Proof. We recall the definition of total variation distance between two random variables is given by :

$$\begin{aligned} d_{TV}(X,Y) &= \sum_k |P(X=k) - P(Y=k)| \\ &= 2\sup_{A\subseteq\mathbb{N}} |P(X\in A) - P(Y\in A)| \end{aligned} \tag{11}$$

So we can compute,

$$\begin{aligned} d_{TV}(H_n,W_n) &= \sum_k |P(H_n=k) - P(W_n=k)| \leq \sum_{i=0}^{n}\sum_k \left|P\left(X_{(i,n)}=k\right) - P(W_{i,n}=k)\right| \\ &= \sum_{i=0}^{n}\left(\left|1-p_{(i,n)}-e^{p_{(i,n)}}\right| + \left|p_{v(i,n)} - p_{(i,n)}e^{-p_{(i,n)}}\right| + \sum_{k\geq 2} P(W_{i,n}=k)\right) \\ &= \sum_{i=0}^{n}\left(\left|1-p_{(i,n)}-e^{-p_{(i,n)}}\right| + \left|p_{(i,n)} - p_{v(i,n)}e^{-p_{(i,n)}}\right| + 1 - e^{-p_{(i,n)}}(1+p_{(i,n)})\right) \end{aligned} \tag{12}$$

where $W_{i,n}$'s are independent Poisson random variables with parameters $p_{(i,n)}$.

But, since, for $x \geq 0$, $1-x \leq e^{-x} \leq 1$, we obtain by (12

$$\begin{aligned} d_{TV}\left(H_n, W_n\right) &= \sum_{i=0}^{n}\left(-1+p_{(i,n)}+e^{-p_{(i,n)}}+p_{v(i,n)}-p_{v(i,n)}e^{-p_{(i,n)}}+1\right. \\ &\quad \left. -e^{-p_{(i,n)}}-e^{-p_{v(i,n)}}-p_{(i,n)}e^{-p_{(i,n)}}\right) \\ &= \sum_{i=0}^{n}\left(2p_{(i,n)}-2p_{(i,n)}e^{p_{(i,n)}}\right) = 2\sum_{i=0}^{n}\left(p_{v(i,n)}(1-e^{p_{(i,n)}})\right) \\ &\leq 2\sum_{i=0}^{n} p_{(i,n)}^2. \end{aligned}$$

Hence using (11) we conclude that:

$$\sup_{A \subseteq \mathbb{N}} \left|P\left(H_n \in A\right)-P\left(L_n \in A\right)\right| \leq \sum_{i=0}^{n} p_{(i,n)}^2.$$

$\square$

Acknowledgements

Thanks to Éric Marchand for useful comments.

References

Ait Aoudia, D. (2008). Reinforced Urn Processes and Binary Tree. PhD thesis, Department of Decision Science, Bocconi University.

Aldous, D. (1985). Exchangeability and related topics. Ecole d'été de Probabilités de Saint Flour XIII. Lecture Notes in Math. 1117, Springer.

de Finetti, B. (1938). VI Colloque Génève, Sur la condition d'équivalence partielle. Actualités Scientifiques et Industrielles, 738, Paris.

Hall, P., Heyde, C.C. (1980). Martingale limit theory and its application. Academic Press Inc.

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

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

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.