Urn Models and Martingale Theory

Parth Karnawat
Supervisor: Prof. Arup Bose, Prof. K. Maulik

Abstract

In urn models, the asymptotic behaviour of proportions of balls of each colour are quantities of interest. In particular, a complete set of their linear combinations have been studied extensively. In this report we discuss a few such results. The proofs of these results are a good example of applications of martingale theory. Lastly, we use Berry-Esseen Theorem for martingales to obtain a rate of convergence result in Friedman's urn.

Contents

Note: Sections 1 to 4 and 10(plus a few theorems in Appendix) have been done during the last semester and have been included for continuity purposes.

1 Introduction

Consider the following urn model. We start with balls of d colours and let X_0 denote the vector of counts for the balls of different colours. We have a scheme of drawing balls and after each draw we put balls in the urn.

Let the vector $\boldsymbol{X}_n$ denote the number of balls in the urn after the n-th draw. Let $T_n = \boldsymbol{X}_n \cdot \mathbf{1}$, denoting the total number of balls in the urn, and $\boldsymbol{C}_n = \boldsymbol{X}_n / T_n$ the vector of proportions of each colour after the n-th turn.

At the n-th turn, draw a ball from the urn at random (i.e. probability of drawing a ball of i-th colour $= C_{n-1,i}$). If the drawn ball is of the i-th colour add r_{ij} balls of the j-th colour to the urn. The matrix $\boldsymbol{R} = (r_{ij})_{1 \leq i,j \leq K}$ is called the *replacement* matrix.

Hence $\boldsymbol{X}_n$ is a Markov chain such that

$$\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + \boldsymbol{\delta}_{n+1} \boldsymbol{R}$$

where $\mathbb{P}[\boldsymbol{\delta}_{n+1} = \boldsymbol{e}_i | \boldsymbol{X}_n] = C_{n-1,i}$. By abuse of notation, we can assume entries of $\boldsymbol{X}_n$, $\boldsymbol{R}$ to be non negative reals. Further assume that $\boldsymbol{R}$ has a *constant row sum.* This ensure a steady linear growth of the size of the urn.

Let $\boldsymbol{Y}_n = \boldsymbol{X}_n - \boldsymbol{X}_{n-1}, \mathcal{F}^n = \sigma(\boldsymbol{X}_0, \cdots, \boldsymbol{X}_n), \mathcal{F}^{(n)} = \sigma(\boldsymbol{X}_{n+1}, \cdots)$ and $\mathcal{F}^{(\infty)} = \cap_{n=0}^{\infty} \mathcal{F}^{(n)}$.

Here is the order notation we will use

1. $x_n \approx y_n$ if $\lim_{n\to\infty} x_n/y_n = 1$.
2. $x_n \sim y_n$ if $\lim_{n\to\infty} |x_n/y_n| > 0$.
3. $x_n = O(y_n)$ if $\lim_{n\to\infty} x_n/y_n = 0$.

2 Polya's Urn

Polya's urn is a very special urn model, where the replacement matrix $\boldsymbol{R} = \alpha \boldsymbol{I}$ (i.e the number of balls of the other colours added is zero). Most of it's good properties follow from the fact that Y_n's are exchangable(Refer [4]). It is well known that the proportion's of a Polya's urn converge almost surely to a random vector which has a Dirichlet disrtribution, as we prove now.

Theorem 1. *C_n converges almost surely to a random variable, say Z. Z follows a Dirichlet distribution with parameter X_0/α. Further, given Z, $\boldsymbol{Y}_i$'s are iid with $\mathbb{P}[\boldsymbol{Y}_i = \alpha \boldsymbol{e}_k] = Z_k$*

Proof. Note that $C_{n,i}$ is a bounded martingale for each i as,

$$\begin{aligned}\mathbb{E}[C_{n+1,i}|\mathcal{F}^n] &= \mathbb{E}\left[\frac{X_{n+1,i}}{T_{n+1}}\middle|\mathcal{F}^n\right] \\ &= \frac{\mathbb{E}[X_{n,i}+Y_{n+1,i}|\mathcal{F}^n]}{T_{n+1}} \\ &= \frac{X_{n,i}+\alpha C_{n,i}}{T_{n+1}} \\ &= \frac{X_{n,i}}{T_{n+1}}\cdot(1+\frac{\alpha}{T_n}) \\ &= C_{n,i} \quad .\end{aligned}$$

Thus by the martingale convergence theorem, $\boldsymbol{C}_n$ converges a.e. to some random vector say $\boldsymbol{Z}$.

Let $\boldsymbol{D} \sim Dir(X_0/\alpha)$, then we will show that

$$\mathbb{E}[\prod_{i=1}^{d} Z_i^{n_i}] = \mathbb{E}[\prod_{i=1}^{d} D_i^{n_i}]. \tag{1}$$

And hence for any polynomial p in $\boldsymbol{Z}$, we have $\mathbb{E}[p(\boldsymbol{Z})] = \mathbb{E}[p(\boldsymbol{D})]$. Now as Z is supported on a compact subset of $\mathbb{R}^d$, an application of Weierstrass's theorem gives us $\mathbb{E}[f(\boldsymbol{Z})] = \mathbb{E}[f(\boldsymbol{D})]$ for any bounded continuous function f. This proves that $\boldsymbol{Z}$ and $\boldsymbol{D}$ must have the same distribution.

Fix $k \in \mathbb{N}, \boldsymbol{\epsilon}_i$'s, $1 \le i \le k$, such that $\sum_{i=1}^{k} \boldsymbol{\epsilon}_i = \alpha(n_1, \cdots, n_d)$ and each $\boldsymbol{\epsilon}_i = \alpha \boldsymbol{e}_j$ for some j. Then **exchangibiliy** of Y_i's gives us that for $n \ge k-1$,

$$\begin{aligned}\mathbb{P}[\boldsymbol{Y}_1 = \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{Y}_k = \boldsymbol{\epsilon}_k|\mathcal{F}^{(n)}] &= \binom{n+1-k}{J_1-n_1, \cdots, J_d-n_d}\Big/\binom{n+1}{J_1, \cdots, J_d} \qquad (2)\\ &= \frac{(n-k+1)!}{\prod_{i=1}^{d}(J_i-n_i)!}\cdot\frac{\prod_{i=1}^{d}(J_i)!}{(n+1)!}.\end{aligned}$$

where $\boldsymbol{J} = \alpha^{-1}(\boldsymbol{X}_{n+1} - \boldsymbol{X}_0)$.

$$\begin{aligned}\mathbb{P}[\boldsymbol{Y}_1 = \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{Y}_k = \boldsymbol{\epsilon}_k|\mathcal{F}^{(n)}] &= \frac{(n-k+1)!}{(n+1)!}\cdot\prod_{i=1}^{d}\frac{(J_i)!}{(J_i-n_i)!} \\ &\approx n^{-k}\cdot\prod_{i=1}^{d}(J_i)^{n_i} \qquad (3)\\ &\approx \prod_{i=1}^{d}(J_i/n)^{n_i}. \qquad (4)\end{aligned}$$

Note that (3) follows from the fact $\mathbb{P}[J_i \to \infty] = 1, \forall i$ and by Stirling's Formula we have $\frac{\Gamma(n+A)}{\Gamma(n+B)} \approx n^{(A-B)}$.

R.H.S. of eqn (4) $\to \prod_{i=1}^{d} Z_i^{n_i}$ as $n \to \infty$, and L.H.S $\to \mathbb{P}[Y_0 = \epsilon_0, \cdots, Y_k = \epsilon_k | \mathcal{F}^{(\infty)}]$, by backward martingale convergence theorem. Thus we have,

$$\mathbb{P}[\boldsymbol{Y}_1 = \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{Y}_k = \boldsymbol{\epsilon}_k | \mathcal{F}^{(\infty)}] = \prod_{i=1}^{d} Z_i^{n_i}. \tag{5}$$

From this it follows that for any A measurable on finitely many $\boldsymbol{X}_n$'s (i.e $A \in \mathcal{F}_n$) $\mathbb{P}[A|\mathcal{F}^{(\infty)}]$ is a function of $\boldsymbol{Z}$. Then looking at the good sets for which this happens and applying Monotone Class Theorem (this is justified by DCT for conditional expectation), we get it to be true for any $A \in \sigma(\cup_{n=0}^{\infty} \mathcal{F}^n)$ measurable, in particular for $A \in \mathcal{F}^{(\infty)}$. Hence we have $\mathcal{F}^{(\infty)} \subset \sigma(\boldsymbol{Z})$. Clearly $\boldsymbol{Z} \in A$ is a tail event so $\sigma(\boldsymbol{Z}) \subset \mathcal{F}^{(\infty)}$. Hence σ-field generated by $\boldsymbol{Z}$ is indeed the tail σ-field $\mathcal{F}^{(\infty)}$ and given $\boldsymbol{Z}$, $\boldsymbol{Y}_i$'s are iid with $\mathbb{P}[\boldsymbol{Y}_i = \alpha \boldsymbol{e}_k] = Z_k$.

Now, taking expectation on both side of eqn (5), we have

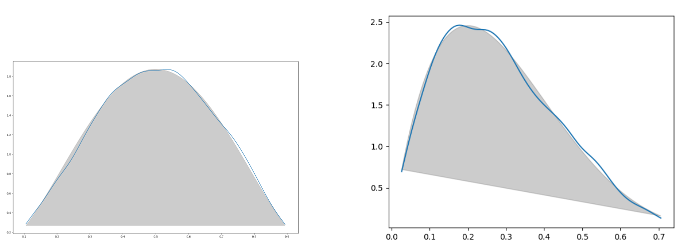
$$\mathbb{P}[\boldsymbol{Y}_1 = \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{Y}_k = \boldsymbol{\epsilon}_k] = \mathbb{E}[\prod_{i=1}^{d} Z_i^{n_i}]. \tag{6}$$

But direct caluclations of probabilities gives us that,

$$\mathbb{P}[\boldsymbol{Y}_1 = \boldsymbol{\epsilon}_1, \cdots, \boldsymbol{Y}_k = \boldsymbol{\epsilon}_k] = \frac{\prod_{i=1}^{d} \prod_{j=0}^{n_i - 1} (X_{0,i} + j\alpha)}{\prod_{j=0}^{k-1} T_j}. \tag{7}$$

Note that L.H.S. of eqn(7) is indeed $\mathbb{E}[\prod_{i=1}^{d} D_i^{n_i}]$, thus (6) and (7) give us eqn (1). $\square$

Here are some simulations of the above result, with $\alpha = 2, n = 2000$ draws and 10000 replications with the initial composition mentioned in the caption.



(a) (6,6) (b) (4,10)

Figure 1: Kernel Density Estimate(KDE)

3 Friedman's Urn

In this section we will consider Friedman's urn, an urn with two colours(i.e. $d = 2$) and $R = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$(Refer [5] for more details).

Notation For this section let

$$\delta = \alpha - \beta \quad , \sigma = \alpha + \beta > 0 \quad \text{and} \rho = \delta/\sigma.$$

As we saw in the previous section for $\beta = 0, C_{n,1}$ converges to a beta distribution. In contrast, if $\beta > 0$ we have that $C_{n,1} \to 1/2$ almost surely, for any values of $\boldsymbol{X}_0$ and ρ. Here is a proof due to Ornstein(Example 4.5.6 in [1]).

Theorem 2. *In Freedman's Urn, $C_{n,1}$, the proportion of balls of the first colour, converges to 1/2 almost surely when $\beta > 0$.*

Proof. Let B_n denote the event of drawing the ball of the 1-st colour in the n-th draw and $D_n = \sum_{k=1}^{n} I_{B_n}$ be the total number of balls of the 1-st colour drawn in the first n draws. Then from the *second Borel Cantelli lemma* for adapted events(Thm(28)), we have that $\frac{D_n}{\sum_{k=1}^{n} C_{k-1,1}} \to 1$ almost surely. Now consider the case $\alpha > \beta$. Assume $\limsup_n C_{n,1} \leq x$. This implies that $\limsup_n D_n/n \leq x$.

$$\begin{aligned} C_{n,1} &= \frac{X_{0,1} + \alpha D_n + \beta(n - D_n)}{T_0 + (\alpha+\beta)n} \\ \implies \limsup_n C_{n,1} &\leq \frac{\beta + (\alpha - \beta)x}{\alpha + \beta}. \end{aligned} \tag{8}$$

Note that R.H.S of eqn(8) is a linear function of x, has slope < 1 and a fixed point at 1/2. Hence starting with the trivial value $x = 1$ and iterating we get $\limsup_n C_{n,1} \leq 1/2$. Similarly, interchanging roles of 1-st colour and the 2-nd colour, we get that $\liminf_n C_{n,1} \leq 1/2$. For the case $\beta > \alpha$, we need to start with two assumptions, $\limsup_n C_{n,1} \leq x$ and $\limsup C_{n,2} \leq y$ and the result follows from the same argument. □

Thus we have that $X_{n,1} - X_{n,2} = o(n)$, this can be further strengthened. For the case $\rho > 1/2$ we have that $(X_{n,1} + X_{n,2})^{-\rho}(X_{n,1} - X_{n,2})$ converges to a non-degenerate random variable. This fails for $\rho \leq 1/2$, infact the limsup and liminf of the sequence $(X_{n,1} + X_{n,2})^{-\rho}(X_{n,1} - X_{n,2})$ are $+\infty$ and $-\infty$, respectively. Although it can be shown that for $\rho = 1/2, (n \ln n)^{-1/2}(X_{n,1} - X_{n,2}) \to \mathcal{N}(0, (\alpha - \beta)^2)$ and for $\rho < 1/2, n^{-1/2}(X_{n,1} - X_{n,2}) \to \mathcal{N}(0, (\alpha - \beta)^2/1 - 2\rho)$ These results are due to Freedman(1964) and Bernstein(1940).

3.1 $\rho > 1/2$

Observe that,

$$\begin{aligned} \mathbb{E}[X_{n+1,1} - X_{n+1,2}|\mathcal{F}^n] &= \mathbb{E}[(X_{n,1} - X_{n,2}) + (Y_{n+1,1} - Y_{n+1,2})|\mathcal{F}_n] \\ &= (X_{n,1} - X_{n,2}) + \mathbb{E}[Y_{n+1,1} - Y_{n+1,2}|\mathcal{F}_n] \\ &= (X_{n,1} - X_{n,2}) + (\delta C_{n,1} + (-\delta)C_{n,2}) \\ &= (X_{n,1} - X_{n,2}) + \frac{\delta(X_{n,1} - X_{n,2})}{T_n} \\ &= a_n(1)(X_{n,1} - X_{n,2}) \end{aligned} \tag{9}$$

where $a_n(k) = (1 + \frac{k\delta}{T_n})$. Thus we have that $Z_n = \frac{X_n}{\prod_{i=0}^{n-1} a_i(1)}$ is a martingale related to our quantity of concern, ie. $(X_{n,1} - X_{n,2})$. We will show that $\prod_{i=0}^{n-1} a_i(1) \sim n^\rho$ and that Z_n are $\mathcal{L}^{2k}$ bounded martingales for all $k \in \mathbb{N}$. Hence by the the theorem on $\mathcal{L}^p$ convergence of martingales, we would have the following theorem.

Theorem 3. *There exists a non degenrate random variable Z, such that $(X_{n,1}+X_{n,2})^{-\rho}(X_{n,1} - X_{n,2}) \to Z$ almost surely and in the p-th mean, $0 < r < \infty$.*

Lemma 1. *If $a_n = (1 + \frac{a}{b+cn})$, where $a, b, c > 0$, then $\prod_{i=0}^{n} a_i \approx \frac{\Gamma(b/c)}{\Gamma((a+b)/c)} n^{a/c}$.*

Proof. Note that since $\Gamma(x+1) = x\Gamma(x), \forall x > 0$, we have

$$\begin{aligned}
\prod_{i=0}^{n} a_i &= \prod_{i=0}^{n} \frac{((a+b)/c) + n}{(b/c) + n} \\
&= \frac{\Gamma(b/c)}{\Gamma((a+b)/c)} \cdot \frac{[\prod_{i=0}^{n}((a+b)/c) + n]\Gamma((a+b)/c)}{[\prod_{i=0}^{n}(b/c) + n]\Gamma(b/c)} \\
&= \frac{\Gamma(b/c)}{\Gamma((a+b)/c)} \cdot \frac{\Gamma(((a+b)/c) + n + 1)}{\Gamma((b/c) + n + 1)}
\end{aligned}$$

It follow directly from the Stirlings Formula that $\Gamma(n+A) \approx \Gamma(n)n^A$ and hence $\frac{\Gamma(((a+b)/c)+n+1)}{\Gamma((b/c)+n+1)} \approx n^{a/c}$, proving the result. $\square$

It follows from the above theorem that $\prod_{i=0}^{n-1} a_i(1) \sim n^\rho$. Observe that

$$\begin{aligned}
&\mathbb{E}[(X_{n+1,1} - X_{n+1,2})^{2K+2} | \mathcal{F}^n] \\
&= C_{n,1}(X_{n,1} - X_{n,2} + \delta)^{2K+2} + C_{n,2}(X_{n,1} - X_{n,2} - \delta)^{2K+2} \\
&= C_{n,1}\Bigg(\sum_{i=0}^{2K+2} \binom{2K+2}{i} \delta^i (X_{n,1} - X_{n,2})^{2K+2-i} \Bigg) \\
&\quad + C_{n,2}\Bigg(\sum_{i=0}^{2K+2} \binom{2K+2}{i} (-\delta^i)(X_{n,1} - X_{n,2})^{2K+2-i} \Bigg) \\
&= a_n(2K+2)(X_{n,1} - X_{n,2})^{2K+2} + \delta^{2K+2} \\
&\quad + \sum_{i=1}^{K} \Bigg[\binom{2K+2}{2i} \delta^{2i} + \binom{2K+2}{2i+1} (\delta^{2i+1}/T_n) \Bigg] (X_{n,1} - X_{n,2})^{2K+2-2i}
\end{aligned}$$

Now let $x_n(k) = \mathbb{E}[(X_{n,1} - X_{n,2})^k]$, taking expectation on both sides of the above equation we have the following recursive relation on $x_n(k)$.

$$x_{n+1}(2K+2) = a_n(2K+2)x_n(2K+2) + b_n(2K+2) \tag{10}$$

where $b_n(2K+2) = \sum_{i=1}^{K} \Bigg[\binom{2K+2}{2i} \delta^{2i} + \binom{2K+2}{2i+1} (\delta^{2i+1}/T_n) \Bigg] x_n(2K+2-2i) + \delta^{2K+2}$. A similar computation for odd powers will give us,

$$x_{n+1}(2K+1) = a_n(2K+1)x_n(2K+1) + b_n(2K+1) \tag{11}$$

where $b_n(2K+1) = \sum_{i=1}^{K}\left[\binom{2K+1}{2i}\delta^{2i} + \binom{2K+1}{2i+1}(\delta^{2i+1}/T_n)\right]x_n(2K+2-2i)$.

Lemma 2. *If $x_{n+1} = a_n x_n + b_n, n \geq 0$, where x_n, a_n, b_n are sequences of real numbers, then $x_{n+1} = x_0 \prod_{i=0}^{n} a_i + \sum_{j=0}^{n-1} b_j(\prod_{i=j+1}^{n} a_i) + b_n, \forall n \geq 0$. Moreover if a_n is a sequence as in lemma 1, and $b_n = \mathcal{O}(n^d), d < \frac{a}{c} - 1$, then $\lim_{n\to\infty} \frac{x_n}{\prod_{i=0}^{n} a_i} = x_0 + \sum_{j=0}^{\infty} b_j \prod_{i=0}^{j} a_i^{-1}$, the series converging absolutely.*

Proof. The proof of the first part follows froms indutction. Now note that there exists constants K and K' such that $\sup_n \frac{|b_n|}{n^d} \leq K$ and $sup_n \frac{\prod_{i=0}^{n} a_i^{-1}}{n^{-a/c}} \leq K'$, because $b_n = \mathcal{O}(n^d)$ and $\prod_{i=0}^{n-1} a_i^{-1} \sim n^{a/c}$(Theorem 4). Therefore $\sum_{j=0}^{M} |b_j| \prod_{i=0} a_j^{-1} < KK' \sum_{j=0}^{M} 1/n^{a/c-d}$. Since $a/c - d > 1$ we have the series is absolutely convergent and the result follows. $\square$

Lemma 3. *For all non-negative integers k, $\lim_{n\to\infty} \mathbb{E}[(\frac{X_{n,1}-X_{n,2}}{n^{-\rho}})^k] = \mu(k)$, where $0 \geq \mu(k) < \infty$. Moreover, if k is even then $\mu(k) > 0$.*

Proof. Note for $k = 0$ the lemma is true trivially, for $k = 1$ we have from eqn(11) that $x_{n+1}(1) = a_n(1)x_n(1) = x_0 \prod_{i=0}^{n} a_i(1)$. By lemma 1, $\prod_{i=0}^{n} a_i(1) \sim n^\rho$ and hence the result follows for $k = 1$. We will now use equations (10) and (11) to complete the proof by induction. Suppose the result is true for $k \leq 2K+1$, this implies that $b_n(2K+2)$ and $b_n(2K+3)$ are $\mathcal{O}(n^{2K\rho})$ and $\mathcal{O}(n^{(2K+1)\rho})$ respectively. Note that $2K\rho < \delta(2K+2)/\sigma - 1$, thus we can apply lemma 2 to equation (10) to get $\lim_{n\to\infty} x_n(2K+2) \prod_{i=0}^{n} a_i(2K+2)^{-1} = x_0(2K+2) + \sum_{j=0}^{\infty} b_j(2K+2) \prod_{i=0}^{j} a_i(2K+2)^{-1}$, which is *positive* and finite. But $\prod_{i=0}^{n} a_i(2K+2) \sim n^{\rho(2K+2)}$, by lemma 1, thus the induction hypothes is true for $k = 2K+2$. The same argument using equation (11) proves the result for $k = 2K+3$. Hence by induction we are done. $\square$

Proof of Theorem 3. Note that from lemma 1 and lemma 3 we have that $\sup_n \mathbb{E}[Z_n^2] < \infty$. Thus by the martingale convergence theorem we have Z_n converges almost surely, hence $n^{-\rho}(X_{n,1} - X_{n,2})$ converges almost surely to a random variable say Z. Moreover $\sup_n \mathbb{E}[(n^{-\rho}(X_{n,1} - X_{n,2}))^{2k}] < \infty$, by lemma 3, hence the sequence $(n^{-\rho}(X_{n,1} - X_{n,2}))^p$ is uniformly integrable for all $0 < p < \infty$ and hence converges in the p-th mean. $\square$

3.2 $\rho = 1/2$

Lemma 4. *If a_n a sequence as in lemma 1 ,x_n as in lemma 2 and $b_n \approx B(n \ln n)^d$, $B \neq 0$ and $d = a/c - 1$, then $x_n \approx (d+1)^{-1} B(n \ln n)^{d+1}$.*

Proof. Note that $\sum_{j=1}^{n}[(\ln j)^d/j] \approx \int_1^n [(\ln x)^d/x]dx = (\ln n)^{d+1}/(d+1)$, for $d > -1$. This is true because from L'Hopitals rule we have $(\ln x)^d/x$ goes to zero as $x \to \infty$ and that it is an eventually decreasing function, using this fact basic analysis gives the required result(If f is monotone and area under f is infinite, then $\sum_1^N f(n) \approx \int_1^N f$, for eg. $\sum_1^N 1/n \approx \log N$).

$$\begin{aligned}\frac{x_{n+1}}{(n\log n)^{d+1}} &= (\prod_{v=0}^{n} a_v)(x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(n\log n)^{d+1} \\ &\approx n^{a/c}(x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(n\log n)^{d+1} \\ &\approx (x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(\log n)^{d+1}\end{aligned}$$

where the second equality follows from lemma 1. Hence it is enough to show that $\frac{B}{d+1}(\log n)^{d+1} \approx \sum_{j=0}^{n} b_j \prod_0^j a_v^{-1}$, given that $b_n \approx B(n\ln)^d$.

$$\begin{aligned}\sum_{j=0}^{n} \frac{b_j}{\prod_0^j a_v} &\approx \sum_{j=0}^{n} \frac{B(j\log j)^d}{j^{d+1}} = B\sum_{j=0}^{n} \frac{(\log j)^d}{j} \\ &\approx \frac{B}{d+1}(\log n)^{d+1}\end{aligned}$$

□

Lemma 5. *If a_n and x_n as in lemma 4 with $a/c > 1/2$ and $b_n \approx Bn^{d-1/2}(\log n)^{d-1}$ with $B \neq 0$ and $d = a/c - 1/2$. Then, $x_n \approx d^{-1}Bn^{d+\frac{1}{2}}(\log n)^d$.*

Proof. Follows from an exact same argument as that in the proof of lemma 4 with appropriate changes in order of b_n. □

Lemma 6. *If $X_{0,1} = X_{0,2}$, then $\mathbb{E}[(X_{n,1} - X_{n,2})^{2k-1}] = 0$, for all $k \in \mathbb{N}$. If $X_{0,1} \neq X_{0,2}$, then $\mathbb{E}[(X_{n,1} = X_{n,2})^{2k-1}] \sim n^{k-1/2}(\ln n)^{k-1}$, for all $k \in \mathbb{N}$.*

Proof. If $X_{0,1} = X_{0,2}$, then $(X_{n,1} - X_{n,2})$ has a symmetric distribution and hence $\mathbb{E}[(X_{n,1} - X_{n,2})^{2k-1}] = 0$, for all $k \in \mathbb{N}$. Now suppose $X_{0,1} \neq X_{0,2}$ then for k=1 the claim is that $\mathbb{E}[X_{n,1} = X_{n,2}] \sim n^{-1/2}$. Note that this follows directly from Eqn(11)(with $K = 0$) and lemma 1. Now suppose the lemma is true for $1 \leq k \leq K$, then in Eqn(11) we have that $b_n(2K+1) = \sum_{i=1}^{K}\left[\binom{2K+1}{2i}\delta^{2i} + \binom{2K+1}{2i+1}(\delta^{2i+1}/T_n)\right]x_n(2K+2-2i) \sim n^{K-1/2}(\log n)^{K-1}$, by the induction hypothesis. Hence lemma 5 applies and we have $x_n(2(K+1)-1) \sim n^{K+\frac{1}{2}}(\log n)^K$, and induction gives us the result. □

Lemma 7. *For each non negative integer k, $\lim_{n\to\infty}(n\log n)^{-k}\mathbb{E}[(X_{n,1}-X_{n,2})^{2k}] = \mu(2k)$ exists and is finite, with $\mu(0) = 1$ and $\mu(2k+2) = \frac{\delta^2}{k+1}\binom{2k+2}{2}\mu(2k)$.*

Proof. Again we prove this using induction and the recursive equation (10). For k=0 the lemma is trivially true. Now suppose it is true for $0 \leq k \leq K$, then in (10) $b_n(2K+2) \approx \delta^2\binom{2K+2}{2}\mu(2K)$. Then from lemma 4 we have $x_n(2K+2) \approx \frac{\delta^2}{K+1}\binom{2K+2}{2}\mu(2K)(n\log n)^{K+1}$. Hence by induction the result follows. □

Theorem 4. *As $n \to \infty, (n \log n)^{-1/2}(X_{n,1} - X_{n,2})$ converges in distribution to the normal distribution with mean 0 and variance $(\alpha - \beta)^2$.*

Proof. We know that the moments of a normal distribution determine it uniquely, lemma 7 tells us that $\mu(2k)$ is the $2k$-th moment of a normal distribution with mean 0 and variance $(\alpha - \beta)^2$. Furthermore lemma 6 tells us that $\mathbb{E}[(X_{n,1} - X_{n,2})^{2k-1}] = o((n \ln n)^{k-\frac{1}{2}})$ and $\mathbb{E}[(X_{n,1} - X_{n,2})^{2k}] \approx \mu(2k)(n \ln n)^k$, hence by the moment convergence crietrion(of Frechet and Shohat) we have the required result. □

Theorem 5. *With probability 1,* $\limsup_{n\to\infty} n^{-1/2}(X_{n,1} - X_{n,2}) = \infty$ *and* $\liminf_{n\to\infty} n^{-1/2}(X_{n,1} - X_{n,2}) = -\infty$.

Proof. Note that when n increases by 1, there is a change of $\delta n(X_{n,1} - X_{n,2})$. Hence the martingale Z_n in the discussion of $\rho = 1/2$ case, is such that $|Z_{n+1} - Z_n| \le \delta\left[1 + \frac{|(X_{n,1}-X_{n,2})|}{T_n}\right]/[\prod_{v=0}^{n} a_v(1)]$. Now we know form lemma 1 $[\prod_{v=0}^{n} a_v(1)] \sim n^\rho$ and $\frac{|(X_{n,1}-X_{n,2})|}{T_n} \le 1$ and so ess. $\sup |Z_{n+1} - Z_n| \le \mathcal{O}(n^{-\rho}) = \mathcal{O}(1)$. Hence we have that Z_n converges to a finite limit almost everywhere on $[\limsup Z_n < \infty]$. Hence we have $(n \log n)^{-1/2}(X_{n,1} - X_{n,2}) = o(Z_n)$ converges to 0 almost everywhere on $[\limsup Z_n < \infty]$. Now note that from Theorem 4 we have the following,

$$\mathbb{P}[\limsup Z_n < \infty] \le \lim_{\epsilon\to 0} \lim_{n\to\infty} \mathbb{P}[(n \log n)^{-1/2}(X_{n,1} - X_{n,2}) < \epsilon] = 0$$

Similarly, we can show $\mathbb{P}[\liminf Z_n > -\infty] = 0$, proving the required result. □

3.3 $\rho < 1/2$

The proofs of this section are step by step similar to that of the previous section, they use the following lemma instead of lemma 5, hence we state the results without writing the proof explicity. Also when $\rho = 0$ we have $\alpha = \beta$ and the urn process is degenerate, hence we may assume $\rho \neq 0$.

Lemma 8. *If a_n a sequence as in lemma 1 ,x_n as in lemma 2 and $b_n \approx Bn^d$, $B \neq 0$ and $d > a/c - 1$, then $x_n \approx \frac{B}{d-(a/c)+1} n^{d+1}$.*

Proof. We repeat the argument as in the proof of lemma 5.

$$\begin{aligned}
\frac{x_{n+1}}{n^{d+1}} &= (\prod_{v=0}^{n} a_v)(x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(n)^{d+1} \\
&\approx n^{a/c}(x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(n)^{d+1} \\
&\approx (x_0 + \sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1})/(n)^{d-(a/c)+1}
\end{aligned}$$

But note that $\sum_{j=1}^{n} j^m \approx n^{m+1}/(m+1)$, hence . $\sum_{j=0}^{n} b_j \prod_{v=0}^{j} a_v^{-1} \approx \sum \frac{Bj^d}{j^{a/c}} \approx B \sum j^{d-(a/c)} \approx \frac{B}{d-(a/c)+1} n^{d-(a/c)+1}$, completing the proof. $\square$

Lemma 9. *For a natural number k, if $X_{0,1} = X_{0,2}$ then $\mathbb{E}[(X_{0,1} - X_{0,2})^{2k-1}] = 0$. If $X_{0,1} \neq X_{0,2}$, but $(T_0+\delta)/\sigma$ is a negative integer, then $\mathbb{E}[(X_{0,1}-X_{0,2})^{2k-1}] = \mathcal{O}(n^{k+\rho-1})$. If $X_{0,1} \neq X_{0,2}$, but $(T_0+\delta)/\sigma$ is not negative integer, then $\mathbb{E}[(X_{0,1}-X_{0,2})^{2k-1}] \sim n^{k+\rho-1}$.*

Lemma 10. *For each nonnegative integer k, $\lim_{n\to\infty} n^{-k}\mathbb{E}[(X_{0,1} - X_{0,2})^{2k}] = \mu(2k)$ exists and is finite, with $\mu(0) = 1$ and $\mu(2k+2) = \frac{\delta^2}{(1-2\rho)(k+1)}\binom{2k+2}{2}\mu(2k)$.*

Theorem 6. *As $n \to \infty$, the $n^{-1/2}(X_{0,1} - X_{0,2})$ converges in distribution to a normal distribution with mean 0 and variance $(1-2\rho)^{-1}(\alpha - \beta)^2$.*

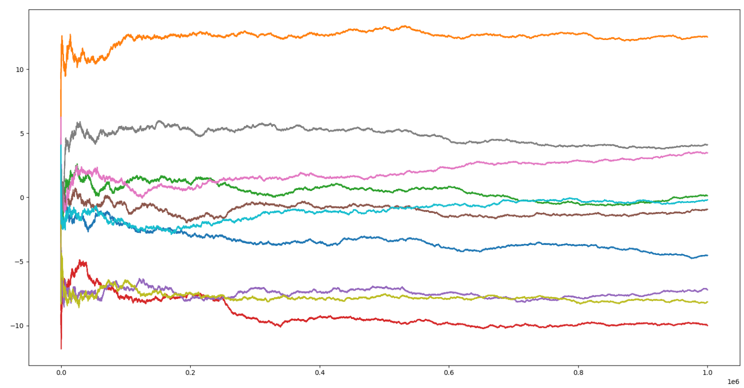
Theorem 7. *If $0 < \rho < 1/2$, then with probability 1 we have $\limsup_{n\to\infty} n^{-\rho}(X_{n,1} - X_{n,2}) = \infty$ and $\liminf_{n\to\infty} n^{-\rho}(X_{n,1} - X_{n,2}) = -\infty$.*

3.4 Simulations

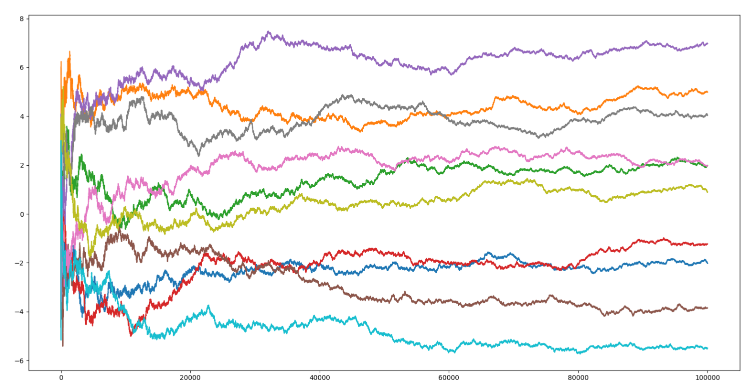
In this section we simulate the Friedmans Urn for different values of ρ and confirm the results we proved for them and draw some observations. Let N denote the number of draws simulated, and M denote the number of samples. Each colour following in the figures represents a sample.

3.4.1 $\rho > 1/2$

Let $U_n = n^{-\rho}(X_{n,1} - X_{n,2})$ 1. $\rho = 0.6, \alpha = 4, \beta = 1, (X_{0,1}, X_{0,2}) = (10, 10)$.

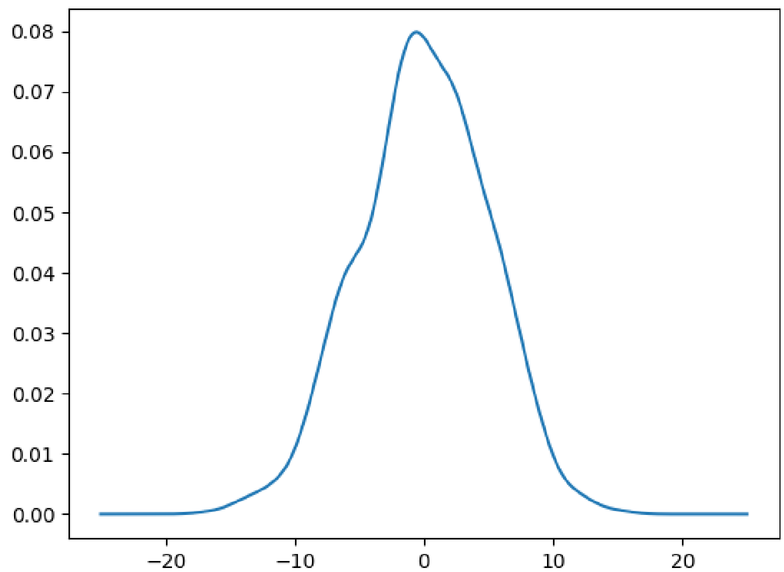


(a) N is $1\boldsymbol{M}$

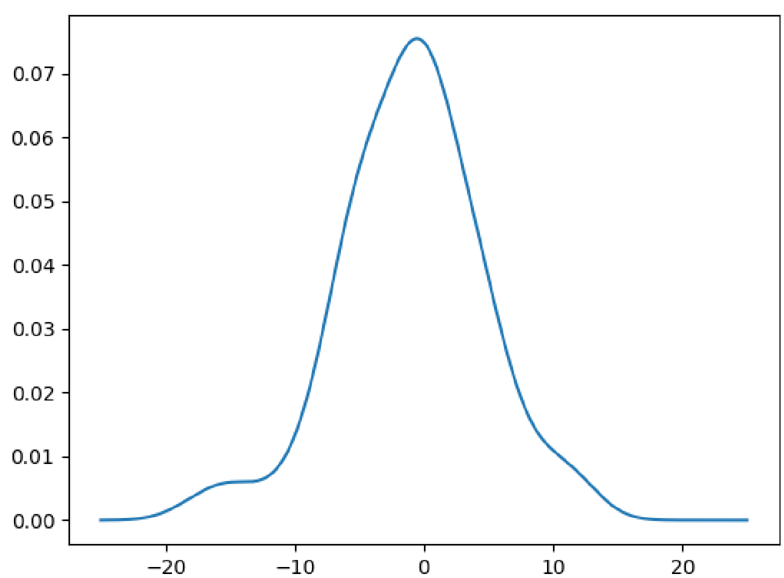


(b) N is $1\boldsymbol{M}$

Figure 2: Plot of U_n vs n



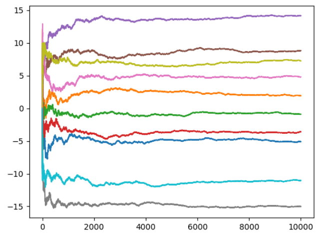
(a) 10k,1k



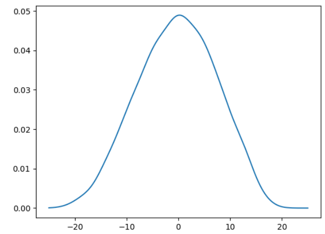
(b) 100k,100

Figure 3: KDE of U_N

Now $\rho = 0.8, \alpha = 9, \beta = 1, (X_{0,1}, X_{0,2}) = (1, 1)$.



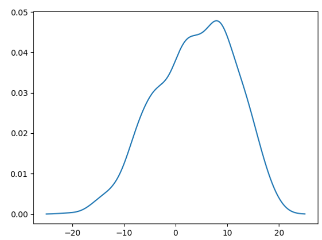
(a) Plot of U_n



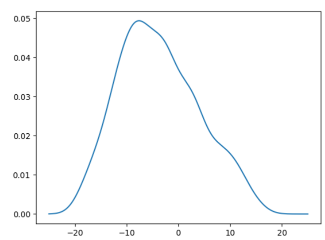
(b) N is 50k

Figure 4: KDE (10k,1k)

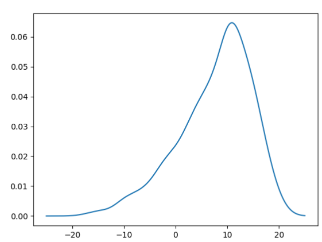
Observe that for larger values of ρ the convergence is faster. Now we plot KDE with different values of $(X_{0,1}, X_{0,2})$, as mentioned in their captions.



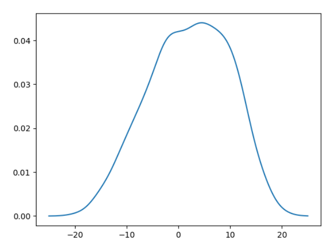
(a) (10,5)



(b) (5,10)



(c) (10,2)



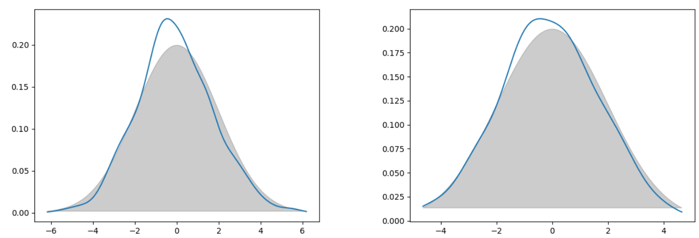
(d) (10,7)

Figure 5: KDE(10k,1k)

Note that more assymetric the initial distribution the more the skew in $Z = \lim_{n\to\infty} U_n$.

3.4.2 $\rho = 1/2$

Now let $U_n = (n \log n)^{-1/2}(X_{n,1}, X_{n,2})$ $\alpha = 3, \beta = 1, (X_{0,1}, X_{0,2}) = (10, 10)$.

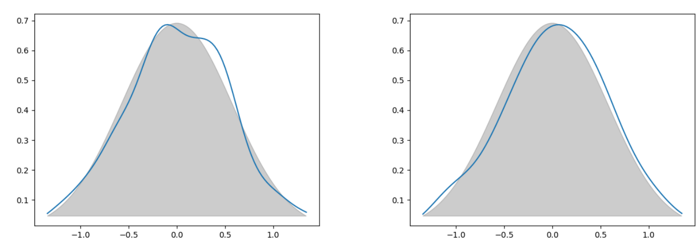


(a) (20k,1k) (b) (50k,400)

Figure 6: KDE

3.4.3 $\rho < 1/2$

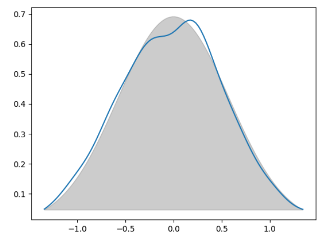
Now let $U_n = n^{-1/2}(X_{n,1}, X_{n,2})$ $1.\rho = -1\alpha = 0, \beta = 1, (X_{0,1}, X_{0,2}) = (1, 1)$.



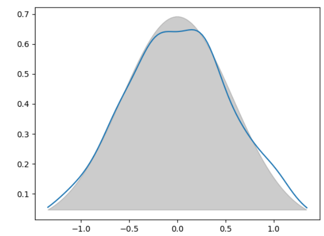
(a) (10k,1k) (b) (100k,100)

Figure 7: KDE

Now we plot the kernel density estimates for different initial compostitions of the urn, as mentioned in the captions.



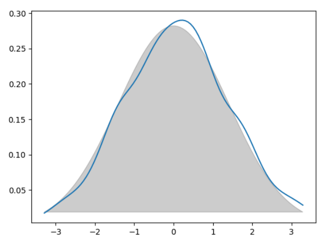
(a) (10,5)



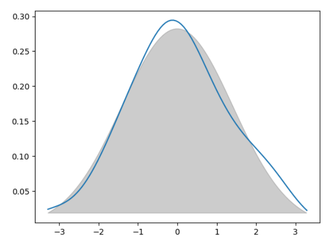
(b) (5,10)

Figure 8: KDE

$2. \rho = -0.5 \alpha = 1, \beta = 3, (X_{0,1}, X_{0,2}) = (10, 10)$



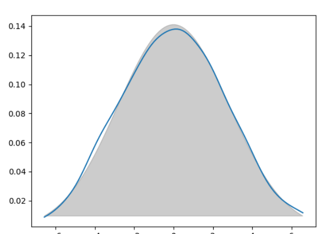
(a) (50k,1k)



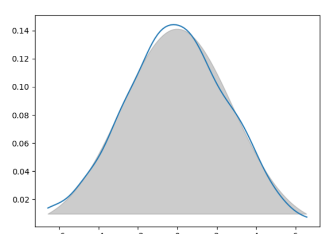
(b) (100k,100)

Figure 9: KDE

$2. \rho = 0.25 \alpha = 5, \beta = 3, (X_{0,1}, X_{0,2}) = (10, 10)$



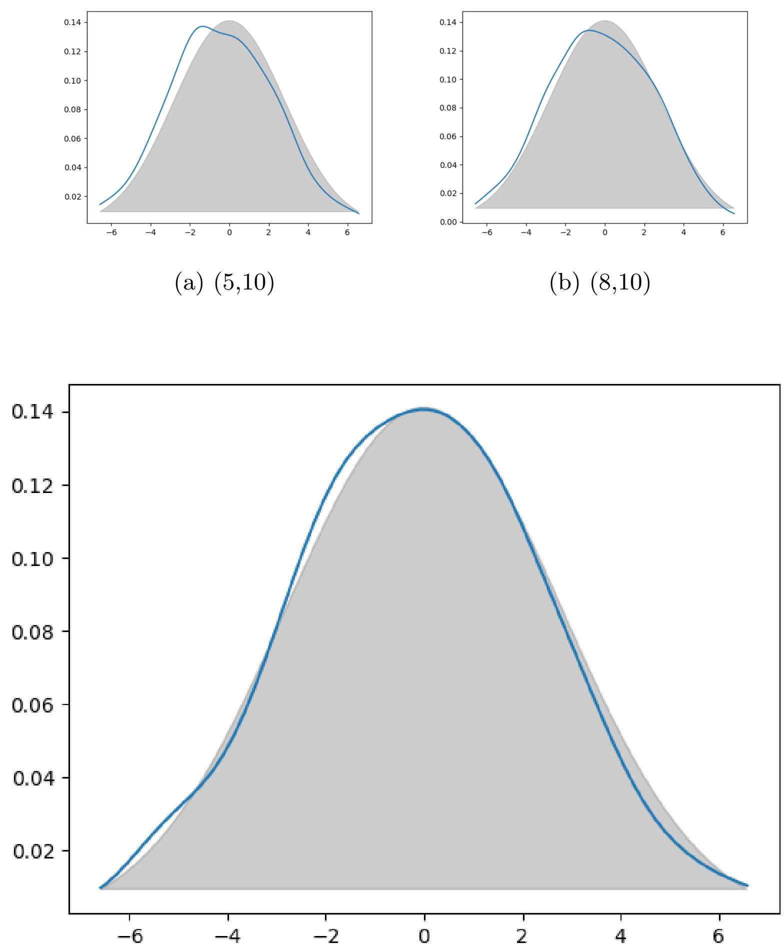
(a) (50k,1k)



(b) (5k,1k)

Figure 10: KDE

The KDE seem to converge faster for a ρ with a smaller absolute value, and thus we use this case to analyze what change in the intial composition would does, keeping other parameters same.



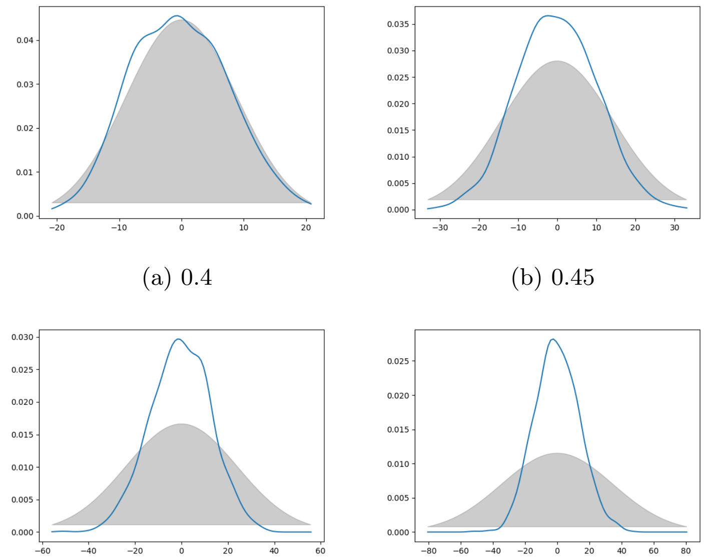
(a) (5,10) (b) (8,10)

(c) (9,10)

Figure 11: KDE

As it is visible, more the assymetry in the intial compsition of the urn, there is a skew in the KDE, which implies the rate of convergence is slower more the assymetry in the initial composition.

Now we try to see what happens as ρ approaches 0.5, we plot the kernel density estimates(using $N = 10k$ and $M = 1k$), with symmetric inital condition(10,10) and different ρ values.



(a) 0.4 (b) 0.45

(c) 0.48 (d) 0.49

Figure 12: KDE(10k,1k)

4 Almost Sure Convergence Of Proportions

The results in this section are due to Gouet (1997), for more details refer [6].

4.1 Terminology and technical lemmas

Colour of balls in the urn can be relabelled such that R can be written in the following normal form,

$$\boldsymbol{R} = \begin{bmatrix} R_{11} & 0 & \dots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & R_{rr} & 0 \\ 0 & \dots & 0 & Q \end{bmatrix}$$

with

$$Q = \begin{bmatrix} Q_{11} & \dots & Q_{1q} \\ \vdots & \ddots & \vdots \\ 0 & \dots & Q_{qq} \end{bmatrix}$$

where $r, q \geq 0$ and the diagonal submatrices $R_{11}, \dots, R_{rr}, Q_{11}, Q_{qq}$ are are irreducible and their corresponding sets of colours(connected components) are denoted by $C_1, \dots, C_r, D_1, \dots, D_q$. Moreover for all $i < q$ atleast one Q_{ij} is non zero for some $j \neq i$. Note that $\boldsymbol{R}$ is a constant row sum matrix, denote the row sum of $\boldsymbol{R}$ by s. Now I state some well know results for non negative irreducible matrices.

Theorem 8 (Perron-Frobenius)**.** *Let B be an irreducible non negative matrix. Then, there exists a dominant eigenvalue τ of B such that:*

1. *τ is real and has multiplicity 1 and the associated left and right eigenvectors are positive and unique upto multiplicative constants.*
2. *$\tau > Re(\lambda)$, where λ is any eigenvalue of B.*
3. *$\min_i \sum_j b_{ij} \leq \tau \leq \max_i \sum_j b_{ij}$, and*
4. *if there exists a non negative vector x and a real number σ such that $Bx \leq \sigma x$, then $\sigma \geq \tau$, and $\sigma = \tau$ if and only if $Bx = \sigma x$.*

As a corrolary of theorem 8 we have,

Theorem 9. *Let B be a non negative irreducible matrix. Then,*

1. *$\sum_j b_{ij} = 1$, for all i implies $\tau = 1$, and*
2. *$\sum_j b_{ij} \leq 1$ with atleast one strict inequality implies $\tau < 1$.*

From Theorem 8 we can conclude that given a non negative irreducible matrix there exists a unique positive left eigenvector such that the sum of its components is 1, call it the **dominant eigenvector** of B. In our case, we have that the dominant eigenvalue of $R_{11}, \dots, R_{rr}$ and Q_{qq} is s, and that of $Q_{11}, \dots, Q_{q-1,q-1} < s$, since for all $i < q$ atleast one Q_{ij} is non zero for some j.

Lemma 11. *Let λ and $y_n, n = 0, 1, \dots$ be complex numbers such that $Re(\lambda) < 1$ and $\delta_n = y_{n+1} - \bar{y}_n\lambda \to 0$, where $\bar{y}_n = (n+1)^{-1}\sum_0^n y_i$. Then, $y_n \to 0$.*

Proof. Note that $y_{n+1} = (n+2)\bar{y}_{n+1} - (n+1)\bar{y}_n$, and hence we have,

$$\begin{aligned}
\delta_n &= (n+2)\bar{y}_{n+1} - (n+1)\bar{y}_n - \bar{y}_n\lambda \\
\bar{y}_{n+1} &= \left(1 + \frac{\lambda - 1}{n+2}\right)\bar{y}_n + \frac{\delta_n}{n+2} \\
&= y_0 \prod_{i=0}^{n}\left(1 + \frac{\lambda - 1}{i+2}\right) + \mu_n \qquad (12)
\end{aligned}$$

where $\mu_n = \sum_{i=0}^{n-1}\left[\frac{\delta_i}{i+2} \cdot \prod_{j=i+1}^{n}\left(1 + \frac{\lambda-1}{j+2}\right)\right] + \frac{\delta_n}{n+2}$. As we proved earlier in lemma 1, $\prod_{i=0}^{n}\left(1 + \frac{\lambda-1}{i+2}\right) \sim n^{\lambda-1} \to 0$ as $Re(\lambda) < 1$. Now we show that $\mu_n \to 0$, to see this observe that,

$$\begin{aligned}
\mu_n &= \prod_{j=1}^{n}\left(1 + \frac{\lambda - 1}{j+2}\right) \cdot \sum_{i=0}^{n-1} \frac{\delta_i}{i+2} \prod_{j=1}^{i}\left(1 + \frac{\lambda - 1}{j+2}\right)^{-1} + \frac{\delta_n}{n+2} \\
&\sim n^{\lambda-1}\sum_{i=0}^{n-1}\delta_i i^{-\lambda} + \frac{\delta_n}{n+2} \to 0. \qquad (13)
\end{aligned}$$

Now it is well know that $\sum_{i=1}^{n} i^{-\lambda} \sim n^{1-\lambda}$ (Can be shown using Abels partial summation formula), also $\delta_n \to 0$. From these two facts we have eqn(13), hence by eqn(12) $y_n \to 0$. $\square$

Lemma 12. *Let $K(\lambda)$ be a $k \times k$ Matrix(Jordan Block) of the form,*

$$K(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ 0 & \dots & & 0 & \lambda \end{bmatrix}$$

where $Re(\lambda) < 1$ and $z_n = (z_{n,1}, \cdots, z_{n,k})$ be such that $\delta_n = z_{n+1} - \bar{z}_n K \to 0$. Then $z_n \to 0$.

Proof. Note that $\delta_{n,1} = z_{n+1,1} - \bar{z}_{n,1}\lambda \to 0$, hence by lemma 11 $z_{n,1} \to 0$. Convergence of other components of z_n follows from induction since $\delta_{n,k} = z_{n+1,k} - \bar{z}_{n,k}\lambda - \bar{z}_{n,k-1} \to 0$. $\square$

Lemma 13. *Suppose B is a non negative irreducible $p \times p$ matrix and $\sum_j b_{ij} \leq 1, \forall i$. Let $u_n = (u_{n,1}, \dots, u_{n,p})$ be a sequence of vectors such that $\delta_n = u_{n+1} - \bar{u}_n B \to 0$. Then,*

1. *if $\sum_j b_{ij} = 1, u_n \to u$, where u_n are non-negative normalized(sum of components=1) and u is the dominant eigenvector of B, and*
2. *if there exists i such that $\sum_j b_{ij} < 1, u_n \to 0$.*

Proof. Consider the Jordan canonical form of $B = TJT^{-1}$, where,

$$J = \begin{bmatrix} K_1 & & \dots & 0 \\ & K_2 & & \\ \vdots & & \ddots & \\ 0 & & \dots & K_q \end{bmatrix}$$

and $K_i = K(\lambda_i)$ a jordan block, $\lambda_1, \dots, \lambda_q$ are the eigenvalues of B and $\tau = \lambda_1$ is dominant. $\sigma_n = \delta_n T = u_{n+1}T - \bar{u}_n BT$ and $v_n = u_n T$. Then, $\sigma_n = v_{n+1} - \bar{(v)}_n J \to 0$ and thus $v_{n+1}^i - \bar{v}_n^i K_i \to 0$, where v_n^i denotes the subvector of v_n corresponding to K_i. Then assertion 2 follows directly from lemma 12 since in this case, from Theorem 8 and 9 it follows that $Re(\lambda_i) \leq \tau < 1$.

Now to show assertion 1, observe that Theorem 8 and 9 say that $\tau = 1$ is simple and $Re(\lambda_i) < 1$ for $i = 2, \dots, q$. Therefore, from lemma 12 we can conclude that $v_n^i \to 0$ for $i = 2, \dots, q$. It remains to show that $v_n^1 = v_{n,1}$ converges, as convergence of v_n implies u_n converges to unique normalized solution of $u = uT$, the dominant eigenvector. Clearly $v_{n,1}$ can't diverge as u_n are bounded, now suppose $v_{n,1}$ has two different limit points a, b then u_n has $A = (a, 0, \dots, 0)T^{-1}$ and $B = (b, 0, \dots, 0)T^{-1}$ as limit points. This is impossible since u_n normalized implies $A \cdot 1 = a(r_1 \cdot 1) \neq a(r_1 \cdot 1) = B \cdot 1$, where r_1 is the first row of T^{-1}. $\square$

4.2 Proof

For a class of consecutive colours $C = \{i, i+1, \ldots, j\}$ let $\boldsymbol{X}_{n,C} = (X_{n,i}, \ldots, X_{n,j})$, the subvector of $\boldsymbol{X}$ corresponding to the set of colours C, similarly define $\boldsymbol{C}_{n,C}$. Let $T_{n,C}$ denote the total number of balls with colours in C, that is $T_{n,C} = \sum_{i \in C} X_{n,i}$. Also, let $\mathbf{1}_C$ dentote the $p \times 1$ column vector with 1 in coordinates $i, i \in C$ and zero otherwise.

Recall $\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + I_{n+1}\boldsymbol{R} = \boldsymbol{X}_0 + \sum_{j=1}^{n+1} I_j \boldsymbol{R}$, where I_n is the vector with 1 in the coordinate of colour draw at the n-th turn and zero in the rest. From the borel cantelli lemma for adapted sequences we have that,

$$\sum_{j=1}^{n} I_{j,i} = (1 + o_1(n)) \sum_{j=1}^{n} C_{j-1,i} \tag{14}$$

as $\mathbb{P}[I_{n,i} = 1 | \mathcal{F}^{n-1}] = C_{n-1,i}$. In matrix notation we have,

$$\sum_{j=1}^{n} I_j = \sum_{j=1}^{n} C_{j-1}(I + o(n)) \tag{15}$$

where I is the identity matrix and $o(n)$ the diagonal matrix with $o_i(n)$ its diagonal entries. Multiplying by R we get,

$$\begin{aligned}
X_n - X_0 &= \sum_{j=1}^{n} I_j R \\
X_n - X_0 &= \sum_{j=1}^{n} C_{j-1}(I + o(n))R \\
C_n - \frac{X_0}{T_n} &= \frac{\sum_{j=1}^{n} C_{j-1}R}{T_n} + \frac{\sum_{j=1}^{n} C_{j-1}o(n)R}{T_n} \\
C_n - \frac{X_0}{T_n} &= \frac{\bar{C}_{n-1}R}{T_n/n} + \frac{\bar{C}_{n-1} \cdot o(n) \cdot R}{T_n/n}
\end{aligned} \tag{16}$$

Note that $\bar{C}_n$ is bounded and $\frac{T_n}{n} \to \frac{1}{s}$, and thus taking limits in eqn(16) gives us that,

$$C_n - \frac{\bar{C}_{n-1}\boldsymbol{R}}{s} \to 0 \tag{17}$$

Theorem 10. *Suppose $\boldsymbol{R}$ is irreducible with dominant eigenvector u. Then, $\boldsymbol{C}_n \to u$.*

Proof. Follows by part 1 of lemma 13 and eqn(17), $C_n \to u$. $\square$

Theorem 11. *Suppose $\boldsymbol{R}$ is reducible without isolated blocks($r = 0, q \geq 2$),i.e.*

$$Q = \begin{bmatrix} Q_{11} & \ldots & Q_{1q} \\ \vdots & \ddots & \vdots \\ 0 & \ldots & Q_{qq} \end{bmatrix}$$

where for all $i < q$ atleast one Q_{ij} is non zero for some $j \neq i$. Q_{ii} are irreducible for all i and the associated set of colours is D_i. Then,

1. *$C_{n,D_i} \to 0$, for $i < q$, and*
2. *$C_{n,D_q} \to v_q$, the dominant eigenvector of Q_{qq}.*

Proof. Look at the subvector corresponding to D_1 in eqn(17), we get $C_{n,D_1} - \bar{C}_{n-1,D_1}\frac{\boldsymbol{Q}_{11}}{s} \to 0$. Since D_1 is not isolated, Q_{11}/s is irreducible and atleast one row sum is strictly less than 1. Therefore it follows from lemma 13 part 2 that, $C_{n,D_1} \to 0$. Now suppose that $C_{n,D_l} \to 0$, for $l = 1, \ldots, j-1 < q-1$. Then again from eqn(17) we have that,

$$C_{n,D_j} - \bar{C}_{n-1,D_j}\frac{Q_{jj}}{s} - \sum_{l=1}^{j-1} \bar{C}_{n-1,D_l}\frac{Q_{ll}}{s} \to 0$$

then the induction hypothesis implies $C_{n,D_j} - \bar{C}_{n-1,D_j}\frac{\boldsymbol{Q}_{jj}}{s} \to 0$ and again we conlude from lemma 13 that $C_{n,D_j} \to 0, j = 1, \ldots, q-1$. Finally, we have

$$C_{n,D_q} - \bar{C}_{n-1,D_q}\frac{Q_{jj}}{s} - \sum_{l=1}^{q-1} \bar{C}_{n-1,D_l}\frac{Q_{ll}}{s} \to 0$$

which implies $C_{n,D_q} - \bar{C}_{n-1,D_q}\frac{\boldsymbol{Q}_{qq}}{s} \to 0$. Therefore the conclusion follows from lemma 13 part 1. □

Theorem 12. *Consider the urn model with replacement matrix $\boldsymbol{R}$ and $R_1, \ldots, R_r$ its isolated irreducible blocks and Q the reducible block, and the corresponding set of colours $C_1, \ldots, C_r$ and E_q. Let $u_i, i = 1, \ldots, r$ and v_q denote $1 \times p$ row vectors such that u_{i,C_i} and v_{q,D_q} are dominant eigenvectors of R_{ii} and Q_{qq} respectively, while the rest of their components are 0 and $(Y_1, \ldots, Y_r, Z_q)$ is a $1 \times r+1$ random vector with Dirichlet distribution with parameters $(\theta_1, \ldots, \theta_{r+1})$, where $\theta_j = T_{0,C_j}/s, j = 1, \ldots, r$ and $\theta_{r+1} = T_0/s - \sum_{j=1}^{r} \theta_j$. Then,*

$$\boldsymbol{C}_n \to \sum_{i=1}^{r} Y_i u_i + Z_q v_q$$

Here we can view the Dirichlet random vector as giving the proportion of the super colours $C_1, \ldots, C_r$ and $E_q (= \cup_1^q D_j)$, and the dominant eigen vectors, the (non-random) proportions of colours within supercolours. Here we consider a Dirichlet vector of length 1 is almost surely equal to 1.

Proof. If we consider the colours in C_i to be a single colour, for each i and also E_q, call them supercolours. Then the urn model for super colours is a Polyas urn model, hence as we proved before we have $(C_{n,C_1}, \ldots, C_{n,C_r}, C_{n,E_q}) \to Z$, a random vector with Dirichlet density as stated in the theorem. Now suppose R is irreducible($r = 1, q = 0$) then the result follows from Theorem 10, and if R is reducible with no isolated blocks then the result follows from Theorem 11.

Now suppose R is reducible and containing only irreducible blocks($q = 0, r > 1$), then define $N = N_{n,i} = \max k \leq n | \sum_{j \in C_i} I_{k,j} = 1$, the time of last draw of a ball with colour in C_i upto time n. Then clearly, N is well defined and grows to ∞ almost surely as $n \to \infty$ for all i. If we look at the urn process at times N, then we have from Theorem 10 that, $C_{N,C_i} \to u_{ii}$, where u_{ii} is the dominant eigenvector of R_{ii}. Now observe that,

$$C_{n,C_i} = \frac{X_{N,C_i}}{T_n} = \frac{X_{N,C_i}}{T_{N,C_i}} \frac{T_{N,C_i}}{T_n} \tag{18}$$

Now note that $\frac{T_{N,C_i}}{T_n} \to Y_i$, hence the result follows.

The same argument proves the result for the general case when R has both isolated and non isolated components. In this case consider $E_q == \cup_1^q D_j$ as the last supercolour. $\square$

5 The Reducible Case

In this section we assume $\boldsymbol{R}$ to be a stocastic matrix and $\boldsymbol{X}_0$ to be a probability vector. Denote $\prod_{j=0}^{n-1}(1 + \frac{\lambda}{j+1})$ by $a_n(\lambda)$. We have shown before that $a_n(\lambda) \sim n^\lambda/\Gamma(\lambda+1)$ when λ is not a negative integer. We will obtain results on the asymptotic behaviour of a *complete* set of linear combinations of $\boldsymbol{X}_n$, for the following urn models. All the results in this section are due to Prof. A. Bose, Prof. A. Dasgupta and Prof. K. Maulik (2009), for more details refer [7].

When a 2×2 replacement matrix is reducible, but not the identity matrix, then with appropriate labelling of colours we have, for some $0 < s < 1$

$$\boldsymbol{R} = \begin{bmatrix} s & 1-s \\ 0 & 1 \end{bmatrix} \tag{19}$$

In the 3-colour urn model, suppose we have one dominant colour. We further assume that the first diagonal block is a *multiple* of an irreducible stochastic matrix. That is,

$$\boldsymbol{R} = \begin{bmatrix} \begin{matrix} & sQ & \end{matrix} & \begin{matrix} 1-s \\ 1-s \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 1 \end{bmatrix} \tag{20}$$

where Q is an irreducible stochastic matrix. Note that this assumption would be crucial for the results which will follow.

Now suppose we have two dominant colour. Then after a relabelling of colours, we have for $0 < s < 1$

$$\boldsymbol{R} = \begin{bmatrix} s & (1-s)\boldsymbol{p} \\ \boldsymbol{0} & P \end{bmatrix} \tag{21}$$

where $\boldsymbol{p}$ is a row probability vector and P an irreducible stocastic matrix.

The results, which we shall be proving for three colour urn models with two dominant colours, can be easily extended to the four colour model with the following replacement matrix, $0 < s < 1$

$$\boldsymbol{R} = \begin{bmatrix} sQ & E \\ 0 & P \end{bmatrix} \tag{22}$$

here we assume that Q is an irreducible stochastic matrix.

5.1 Two-colour urn models

Before moving on to the reducible case in Eqn(19), I will first mention some results for the irreducible case.

Theorem 13. *In an urn model with $d = 2$, $\boldsymbol{R}$ irreducible we have the following,*

1. *$\boldsymbol{C}_n \to \boldsymbol{\pi}_R$ almost surely, where $\boldsymbol{\pi}_R$ is the stationary distribution of $\boldsymbol{R}$.*
2. *Let λ denote the non-principal eigenvalue of $\boldsymbol{R}$ and $\boldsymbol{\xi}$ be a left λ-eigenvector,*
   (a) *If $\lambda < \frac{1}{2}$, then $\boldsymbol{X_n}\boldsymbol{\xi}/\sqrt{n} \implies N(0, \frac{\lambda^2}{1-2\lambda}\boldsymbol{\pi}_R\boldsymbol{\xi}^2)$*
   (b) *If $\lambda = \frac{1}{2}$, then $\boldsymbol{X_n}\boldsymbol{\xi}/\sqrt{n \ln n} \implies N(0, \lambda^2\boldsymbol{\pi}_R\boldsymbol{\xi}^2)$*
   (c) *If $\lambda > \frac{1}{2}$, then $\boldsymbol{X_n}\boldsymbol{\xi}/a_n(\lambda)$ is an L^2-bounded martingale and converges almost surely, as well as in L^2, to a non-degenerate random variable, whose distribution depends on $\boldsymbol{X}_0$.*

When $\boldsymbol{R}$ is as in Eqn(19) we will see that we have almost sure limit for all values of λ, in contrast to the irreducible case.

Theorem 14. *In a two colour urn model, with $\boldsymbol{R}$ as in Eqn(19), we have:*

1. *$C_n \to (0, 1)$ almost surely.*
2. *$C_n\boldsymbol{\xi}/a_n(s) = W_n/a_n(s)$ is an L^2 bounded martingale. Further, W_n/n^s converges to a non-degenerate, positive random variable almost surely, as well as in L^2.*

Proof. The first statement follows from the result of Gouet[6], since the last irreducible block corresponds to the dominant colours.

Next, we will show that $V_n = W_n/a_n(s)$ is L^2 bounded and that variance of V_n increases to that of the limit, hence the limit is non degenerate. Using the fact that $a_n(s) \sim n^s/\Gamma(s+1)$, the second statement follows.

The proof of a.s. positivity of the limit requires methods from branching process theory and we shall skip that proof. A proof is available in [10].

From the evolution equation of W_n if follows that, for $\chi_n =$ the indicator of a white ball drawn in n-th trial, we have

$$V_{n+1} - V_n = \frac{s}{a_{n+1}(s)}\Big(\chi_{n+1} - \frac{W_n}{n+1}\Big).$$

This shows that V_n is indeed a martingale. Further, $V_{n+1} = V_n + (V_{n+1} - V_n)$ and since martingale increments are uncorrelated we have,

$$E[V_{n+1}^2|\mathcal{F}_n] = V_n^2 + \frac{s^2}{a_{n+1}^2(s)}\left[\frac{W_n}{n+1} - \frac{W_n^2}{(n+1)^2}\right].$$

Now choose N large enough such that $\forall n \geq N$, we have

$$\begin{aligned} E[V_{n+1}^2|\mathcal{F}_n] &\leq V_n^2 + \frac{V_n}{(n+1)a_n(s)} \\ &\leq V_n^2 + \Gamma(s+1)\frac{1+V_n^2}{(n+1)^{s+1}} \\ &\leq V_n^2\left[1 + \frac{\Gamma(s+1)}{(n+1)^{s+1}}\right] + \frac{\Gamma(s+1)}{(n+1)^{s+1}}. \end{aligned}$$

Here the second inequality follows from $V_n \leq (1+V_n^2)/2$ and $a_n(\lambda) \sim n^\lambda/\Gamma(\lambda+1)$.

Now taking further expectations and iterating we get $\forall n \geq N$,

$$E[V_{n+1}^2] + 1 \leq (E[V_N^2] + 1)\prod_{i=N}^{n}\left[1 + \frac{\Gamma(s+1)}{(i+1)^{s+1}}\right].$$

It follows from the fact $1 + x \leq e^x$, that

$$E[V_{n+1}^2] + 1 \leq (E[V_N^2] + 1)\exp\left(\Gamma(s+1)\sum_{i=N}^{\infty} i^{-(s+1)}\right) < \infty.$$

□

5.2 Three colour urn model

5.2.1 One dominant colour

In this section we consider the replacement matrix as in Eqn(20). Suppose λ is the non principal eigenvalue of Q and the corresponding eigenvector is ξ. Then $(1,1,1)', (1,1,0)'$ and $(\xi', 0)'$ are the eigenvectors of $\boldsymbol{R}$ corresponding to the eigenvalues 1, s and $s\lambda$. Let us say that the three colours are white, black and green, with green being the dominant colour. Denote by S_n the subvector corresponding to the non-dominant colours, ie. $S_n = (W_n, B_n)$.

Theorem 15. *In a three colour urn model, with replacement matrix as in Eqn(20), we have the following:*

1. $C_n \to (0,0,1)$ *almost surely.*
2. $S_n \cdot 1/(n+1)^s$ *converges almost surely and in* L^2, *to a non-degenerate positive random variable* U.
3. $S_n/(n+1)^s \to \pi_Q U$.

4. *If $\lambda < 1/2$, then $S_n\xi/n^{s/2} \implies N(0, \frac{s^2\lambda^2}{s(1-2\lambda)}U\pi_Q\xi^2)$.*

5. *If $\lambda = 1/2$, then $S_n\xi/\sqrt{n^s \ln n} \implies N(0, s^2\lambda^2 U\pi_Q\xi^2)$.*

6. *If $\lambda > 1/2$, then $S_n\xi/a_n(s\lambda)$ is an L^2 bounded martingale and it converges almost surely and in L^2, to V which is a non-degenerate random variable.*

Moreover, the distributions of U and V depend on the initial value S_0.

Proof. Again, the first part follows directly form Gouet's result, as green is the only dominant colour. Now, obsereve that $(S_n \cdot 1, G_n)$ follows a two colour urn model with replacement matrix as in Eqn(19). Therefore, by Part 2 of Thm(14) the second part follows.

Let τ_k denote the time of k-th draw of a ball of white or black colour. Note that τ_k increases to ∞ as $n \to \infty$, we showed this before using the second Borel-Cantelli lemma for martingales. Then, $S_{\tau_k} \cdot 1/S_{\tau_k}$, are proportions from a two colour urn with *irreducible* replacement matrix. Thus by Part 1. of Thm(13), we have $S_{\tau_k} \cdot 1/S_{\tau_k} \to \pi_Q$. Further note that, for $\tau_k \leq n < \tau_{k+1}$ we have $S_n = S_{\tau_k}$ and $S_n \cdot 1/(n+1)^s \to U$, by the second part. Combining these two facts we have proved Part 3.

Let

$$\chi_n = (1,0), (0,1) \text{ or } (0,0),$$

accordingly as white, black or green ball are drawn in the n-th draw.

Consider the martingale $T_n = S_n/a_n(s\lambda)$. Then the martingale difference is,

$$T_{n+1} - T_n = \frac{s\lambda}{a_{n+1}(s\lambda)}\Big(\chi_n\xi - \frac{S_n\xi}{n+1}\Big).$$

Then again using the fact that martingale differences are uncorrelated we have,

$$\begin{aligned} E[T_{n+1}^2] &= E[T_n^2] + \Big(\frac{s\lambda}{a_{n+1}(s\lambda)}\Big)^2 E\Big[\frac{S_n\xi^2}{(n+1)} - \Big(\frac{S_n\xi}{n+1}\Big)^2\Big] \\ &= E[T_n^2]\Big[1 - \Big(\frac{s\lambda}{(n+1)(1+s\lambda/(n+1)^2)}\Big)^2\Big] \\ &\quad + \Big(\frac{s\lambda}{a_{n+1}(s\lambda)}\Big)^2 \frac{1}{(n+1)^{1-s}} E\Big[\frac{S_n\xi^2}{(n+1)^s}\Big]. \end{aligned}$$

Note that the first term is bounded by $E[T_n^2]$. From Part 2 we have that $S_n 1/(n+1)^s$ is L^2 bounded, which implies that $S_n\xi^2/(n+1)^s$ is L^1 bounded. Now using $a_n(s) \sim n^s/\Gamma(s+1)$, we have a large N such that $\forall n \geq N$,

$$E[T_{n+1}^2] \leq E[T_n^2] + \frac{C}{n^{1+s(2\lambda-1)}},$$

for some constant C. Iterating the above inequality and noting that $n^{-(1+s(2\lambda-1))}$ is summable, we have that T_n is an L^2 bounded martingale. Again since variance of T_n increases to that of its limit, V is non-degenerate. This proves the sixth part.

Suppose $\lambda < 1/2$, and let $Z_n = S_n\xi/n^{s/2}$. The evolution equation for $S_n\xi$ is as follows

$$S_{n+1}\xi = S_n\xi + s\chi_{n+1}Q\xi = S_n\xi + s\lambda\chi_{n+1}\xi.$$

Now we decompose Z_n into a sum of a conditional expectation and a martingale difference.

$$Z_{n+1} = E[Z_{n+1}|\mathcal{F}_n] + (Z_{n+1} - E[Z_{n+1}|\mathcal{F}_n]). \tag{23}$$

But,

$$\begin{aligned}
E[Z_{n+1}|\mathcal{F}_n] &= \frac{1}{(n+1)^{s/2}} E[S_{n+1}\xi|\mathcal{F}_n] \\
&= \frac{1}{(n+1)^{s/2}} \Big[S_n\xi + \frac{s\lambda}{n+1} S_n\xi\Big] \\
&= Z_n(1+1/n)^{-s/2}\Big(1 + \frac{s\lambda}{n+1}\Big) \\
&= Z_n\Big(1 - \frac{s}{2n} + O\Big(\frac{1}{n^2}\Big)\Big)\Big(1 + \frac{s\lambda}{n+1}\Big) \\
&= Z_n\Big(1 - \frac{s(1/2-\lambda)}{n}\Big) + Z_nO(n^{-2}).
\end{aligned} \tag{24}$$

On the other hand the martingale difference is,

$$M_{n+1} = Z_{n+1} - E[Z_{n+1}|\mathcal{F}_n] = \frac{s\lambda}{(n+1)^{s/2}}\Big(\chi_{n+1} - \frac{S_n}{n+1}\Big)\xi. \tag{25}$$

Combining Eqn(23) and Eqn(24) we get the following

$$Z_{n+1} = Z_n\Big(1 - \frac{s(1/2-\lambda)}{n}\Big) + Z_nO(n^{-2}) + M_{n+1}.$$

Iterating the above equation we get the following,

$$\begin{aligned}
Z_{n+1} = Z_1 \prod_{i=1}^{n}\Big(1 - \frac{s(1/2-\lambda)}{i}\Big) &+ \sum_{j=1}^{n} Z_jO(j^{-2}) \prod_{i=j+1}^{n}\Big(1 - \frac{s(1/2-\lambda)}{i}\Big) \\
&+ \sum_{j=1}^{n} M_{j+1} \prod_{i=j+1}^{n}\Big(1 - \frac{s(1/2-\lambda)}{i}\Big).
\end{aligned} \tag{26}$$

Observe that the first term converges to 0, because

$$a_n(-s(1/2-\lambda)) \sim n^{-s(1/2-\lambda)} \to 0.$$

And in the second term note that the product term can be bounded by 1. Since coordinates of $S_n/n+1$ are bounded by 1, we have that $|Z_n|/n^{1-s/2}$ is bounded. Hence the second term is bounded by a multiple of $\sum_{i=1}^{\infty} i^{-(1+s/2)} < \infty$. Moreover, note that each term of the summation converges to zero as the product term converges to 0 as $n \to \infty$. Hence the second term also converges to zero.

It remains to analyze the third term,

$$B_{n+1} = \sum_{j=1}^{n} M_{j+1} \prod_{i=j+1}^{n} \Big(1 - \frac{s(1/2-\lambda)}{i}\Big). \tag{27}$$

We will show using the Lyapunov version of the Martingale Central Limit Theorem that B_n converges weakly to a mixture of normal distributions. Thus we want to show for some $k > 2$,

$$\sum_{j=1}^{n} E[|M_{j+1}|^k|\mathcal{F}_j] \prod_{i=j+1}^{n} \Big(1 - \frac{s(1/2-\lambda)}{i}\Big)^k \to 0.$$

Observe that M_{n+1} is bounded by a constant multiple of $(n+1)^{-s/2}$, since all co-ordinates of χ_n and $S_n/(n+1)$ are bounded by 1. Thus the above sum is bounded by a constant multiple of

$$\sum_{j=1}^{n} j^{-ks/2} \prod_{i=j+1}^{n} \Big(1 - \frac{s(1/2-\lambda)}{i}\Big)^k.$$

If we choose $k > 2/s$ then the above sum converges to zero by DCT. Therefore, by Corrollary 3.1 of [8], we have B_n converges to a mixture of normal distributions. We shall now explicitly compute the limiting conditional variance.

Upon computing the conditional variance of M_{n+1} we get,

$$E[M_{n+1}^2|\mathcal{F}_n] = \frac{(s\lambda)^2}{(n+1)^s}\left[\frac{S_n\xi^2}{n+1} - \Big(\frac{S_n\xi}{n+1}\Big)^2\right].$$

From Part (3) of the theorem it follows that

$$E[M_{n+1}^2|\mathcal{F}_n] \sim \frac{(s\lambda)^2}{n+1} U\pi_Q\xi^2.$$

Now writing

$$\prod_{i=j+1}^{n} \Big(1 - \frac{s(1/2-\lambda)}{i}\Big)^k \text{ as } a_n(-s(1/2-\lambda))/a_j(-s(1/2-\lambda)),$$

and using $a_n(s) \sim n^s/\Gamma(s+1)$, we have with probability 1

$$\sum_{j=1}^{n} E[M_{j+1}^2|\mathcal{F}_j] \prod_{i=j+1}^{n} \Big(1 - \frac{s(1/2-\lambda)}{i}\Big)^2 \sim \frac{(s\lambda)^2 U\pi_Q\xi^2}{n^{s(1-2\lambda)}} \sum_{j=1}^{n} \frac{1}{j^{1-2(1-2\lambda)}}.$$

Thus as $n \to \infty$ the above sum converges to $(s\lambda)^2 U\pi_Q\xi^2/s(1-2\lambda)$, the required conditional variance. This finishes the proof of the fourth part.

The proof of the fifth part is similar to that of the fourth. The following approximations are key:

$$(1+1/n)^{-s/2} = 1 - \frac{s}{2n} + O(n^{-2}) \text{ and } \frac{\ln n}{\ln(n+1)} = \frac{\ln n}{\ln n + 1/n + O(n^{-2})}.$$

Using both of them gives us,

$$\Big(\frac{n}{n+1}\Big)^{s/2} \sqrt{\frac{\ln n}{\ln(n+1)}} = 1 - \frac{s}{2n} - \frac{1}{2n\ln n} + O(n^{-2}).$$

Upon computing the decomposition of $Z_n = S_n\xi/\sqrt{n^s \ln n}$ and using the above approximation we get the following:

$$E[Z_{n+1}|\mathcal{F}_n] = Z_n[1-(2n\ln n)^{-1} + O(n^{-2})],$$

$$M_{n+1} = \frac{s}{2\sqrt{(n+1)^s \ln(n+1)}}\Big(\chi_{n+1} - \frac{S_n}{n+1}\Big)\xi,$$

$$Z_{n+1} = Z_n[1-(2n\ln n)^{-1}] + Z_n O(n^{-2}) + M_{n+1}.$$

Iterating the above equation we get an equation similar to Eqn(26), and a similar argument shows that the first two terms converge to zero. Then we can verify the Conditional Lyapunov condition to conclude the result using a Martingale central limit theorem. □

5.2.2 Two dominant colours

In this section we consider the replacement matrix as in Eqn(21). Let the three colours be white, black and green, the latter two being the dominant colours. Suppose λ is the non-principal eigenvalue of P and the corresponding eigenvector is ξ. Then $1, s$ and λ are the eigenvalues of $\boldsymbol{R}$ (since it is block upper triangular). The vector $(1,1,1)'$ and $(1,0,0)'$ are always the eigenvectors of $\boldsymbol{R}$ corresponding to the eigenvalues 1 and s. The asymptotic behaviour of these linear combinations follows from the result of the two colour case (by collapsing the two dominant colours into one). We state this result now for the sake of completeness.

Theorem 16. *In a three colour urn model, with replacement matrix as in Eqn(21), we have the following:*

1. $C_n \mathbf{1} = X_n \mathbf{1}/(n+1) = 1$.
2. $C_n \to (0, \pi_P)$ *almost surely.*
3. $W_n/n^s \to V$ *almost surely and in* L^2*, where distribution of* V *depends on the initial composition* $(W_0, B_0 + G_0)$.

Now suppose $\boldsymbol{R}$ has a complete set of eigenvectors, ie. it is diagonalizable. This is the case when $\lambda \neq s$, where $\nu = ((1-s)\boldsymbol{p}\xi/(\lambda - s), \xi')'$ is the λ-eigenvector. If $\lambda = s$, then note that $\boldsymbol{R}$ is diagonalizable iff $\boldsymbol{p}\xi = 0$. In this case let $\nu = (0, \xi')'$, an s-eigenvector independent of $(1,0,0)'$. Also note that, in this case $\boldsymbol{p}$ is orthogonal to ξ and is a probability vector and hence $\boldsymbol{p} = \pi_P$. The following theorem describes the limiting behaviour of the linear combination corresponding to ν.

Theorem 17. *In a three colour urn model, with a **diagonalizable** replacement matrix as in Eqn(21), we have the following:*

1. *If $\lambda < 1/2$, then $X_n\nu/\sqrt{n} \implies N(0, \frac{\lambda^2}{1-2\lambda}\pi_P\xi^2)$.*
2. *If $\lambda = 1/2$, then $X_n\nu/\sqrt{n \ln n} \implies N(0, \lambda^2\pi_P\xi^2)$.*
3. *If $\lambda > 1/2$, then $X_n\nu/a_n(\lambda)$ is an L^2 bounded martingale and $X_n\nu/n^\lambda$ converges almost surely to a non-degenrate random variable.*

Proof. We omit the proofs of the first and the second parts since they are similar to those of part four and five of Thm(15) (decomposition into a conditional expectation and a martingale difference works).

Let χ_n be the random row vector taking values $(1,0,0)'$, $(0,1,0)'$, $(0,0,1)'$ depending on whether a white, black or green ball is drawn at the n-th draw. Let $Z_n = X_n\nu/a_n(\lambda)$. From the evolution equation of $X_n\nu$ we get the following

$$Z_{n+1} - Z_n = \frac{\lambda}{a_{n+1}(\lambda)}\Big(\chi_{n+1} - \frac{X_n}{n+1}\Big)\nu.$$

Thus we can see Z_n is indeed a martingale, Moreover we have the following

$$E[(Z_{n+1} - Z_n)^2|\mathcal{F}_n] = \frac{\lambda^2}{a_{n+1}^2(\lambda)}\left[\frac{X_n\nu^2}{n+1} - \Big(\frac{X_n\nu}{n+1}\Big)^2\right].$$

Now note that $X_n/(n+1)$ is bounded by 1 in each co-ordinate. Using this and the fact that $a_n(\lambda) \sim n^\lambda/\Gamma(\lambda+1)$, we have that the above equation is bounded by a constant multiple of $n^{-2\lambda}$. Now since martingale differences are uncorrelated we have $E[Z_{n+1}^2] = \sum_{i=1}^n E[(Z_{i+1} - Z_i)^2]$, where $Z_0 = 0$. Hence $E[Z_{n+1}^2]$ is bounded by a constant multiple of $\sum_{i=1}^\infty i^{-2\lambda} < \infty$. Thus, $\{Z_n\}$ is indeed an L^2 bounded martingale and part three follows. $\square$

When $\boldsymbol{R}$ is not diagonalizable, we can not have a complete set of eigenvectors. Therefore we work with a complete set of vectors obtained from the Jordan decomposition of $\boldsymbol{R}$, such that two of them are the eigenvectors of $\boldsymbol{R}$. So consider the Jordan decomposition $\boldsymbol{R}T = TJ$ where

$$J = \begin{bmatrix} s & 1 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Further T is chosen such that the first and the third columns t_1 and t_3 are $(1,0,0)'$ and $(1,1,1)'$ respctively. Also the subvector of the last two coordinates of t_2 is an eigenvector of P corresponding to s ($= \lambda$ in this case). We shall denote it by ξ. As we shall see, the behaviour of $X_n t_2$ is substantially different from that of the diagonalizable case.

Theorem 18. *In a three colour urn model, with a **non-diagonalizable** replacement matrix as in Eqn(21), we have the following:*

1. *If $s < 1/2$, then $X_n t_2/\sqrt{n} \implies N(0, \frac{s^2}{1-2s}\pi_P \xi^2)$.*
2. *If $s \geq 1/2$, then $X_n t_2/n^s \ln n$ converges to V almost surely, as well as in L^2. Here V is the almost sure limiting random variable obtained in Part (3) of Thm(17).*

Proof. Again we split $Z_n = X_n t_2/\sqrt{n}$ into a conditional expectations and a Martingale difference. From the Jordan decomposition of $\boldsymbol{R}$, we have the following evolution equation,

$$X_{n+1}t_2 = X_n t_2 + (t_1 + st_2)\chi_{n+1}$$

where χ_n is the vector indicating the colour of the ball drawn. Hence the conditional expectation becomes

$$\begin{aligned} E[Z_{n+1}|\mathcal{F}_n] &= \frac{X_n t_2}{\sqrt{n+1}}\Big(1 + \frac{s}{n+1}\Big) + \frac{1}{(n+1)^{3/2}} X_n t_1 \\ &= Z_n\big(1 - \frac{1/2 - s}{n+1}\big) + Z_n O(n^{-2}) + \frac{1}{(n+1)^{3/2}} W_n. \end{aligned}$$

Let us denote $\boldsymbol{st} = t_1 + st_2$. Then the martingale difference term is

$$M_{n+1} = \frac{\boldsymbol{st}}{\sqrt{n+1}}\Big(\chi_{n+1} - \frac{X_n}{n+1}\Big).$$

Thus we have the following recursion on Z_n

$$Z_{n+1} = Z_n\Big(1 - \frac{1/2 - s}{n+1}\Big) + Z_n O(n^{-2}) + \frac{1}{(n+1)^{3/2}} W_n + M_{n+1}.$$

Upon iterating we get,

$$\begin{aligned} X_{n+1} = X_1 \prod_{i=1}^{n}\Big(1 - \frac{1/2 - s}{i}\Big) + \sum_{j=1}^{n} X_j O(j^{-2}) \prod_{i=j+1}^{n}\Big(1 - \frac{1/2 - s}{i}\Big) \\ + \sum_{j=1}^{n} \frac{W_j}{(j+1)^{3/2}} \prod_{i=j+1}^{n}\Big(1 - \frac{1/2 - s}{i}\Big) + \sum_{j=1}^{n} M_{j+1} \prod_{i=j+1}^{n}\Big(1 - \frac{1/2 - s}{i}\Big). \end{aligned}$$

Except the additional third term, this decomposition is similar to Eqn(26), so we will only show that this term converges to 0 almost surely. Note that we saw

$W_n/n^s \to V$ almost surely, and since $a_n(s) \sim n^s/\Gamma(s+1)$ we get that the third term is of the order

$$\frac{1}{n^{1/2-s}} \sum_{j=1}^{n} \frac{V}{j^{3/2-s}} \frac{1}{j^{-(1/2-s)}} \sim \frac{V \ln n}{n^{1/2-s}} \to 0$$

almost surely, as required. To estimate the conditional variance note that

$$\begin{aligned} E[M_{n+1}^2|\mathcal{F}_n] &= \frac{s^2}{n+1}\Big(\frac{X_n \boldsymbol{t}^2}{n+1} - \Big(\frac{X_n \boldsymbol{t}}{n+1}\Big)^2\Big) \\ &\sim \frac{s^2(0,\pi_P)\boldsymbol{t}^2}{n+1} \\ &\sim \frac{s^2}{n+1}\pi_P \xi^2. \end{aligned}$$

□

5.3 Four colour urn model

In this section we will extend the result obtained for three colour urn models with two dominant colours to four colour urn models, whoose replacement matrix is as in Eqn(22). Let the four colours be white, black, green and yellow.

Denote the non-principal eigenvalues of P and Q by β and α respectively. Then it follows that 1, β, s and $s\lambda$ are the eigenvalues of $\boldsymbol{R}$ (since it is block triangular). Moreover, if ξ is a λ-eignevector of Q, then $\nu_1 = (1,1,0,0)'$, $\nu_2 = (\xi', 0, 0)'$ and $\nu_4 = (1,1,1,1)'$ are eigenvectors of $\boldsymbol{R}$ corresponding to $s, s\lambda$ and 1 respectively.

Theorem 19. *In a four colour urn model, with replacement matrix as in Eqn(22), we have the following:*

1. $X_n\nu_4/(n+1) = 1$.
2. $C_n = X_n/(n+1) \to (0,0,\pi_P)$ *almost surely.*
3. $(W_n, B_n)/n^s \to \pi_q U$ *almost surely.*
4. $X_n\nu_1/n^s \to U$ *almost surely and in* L^2.
5. *If* $\lambda < 1/2$, *then* $X_n\nu_2/n^{s/2} \implies N(0, \frac{s^2\lambda^2}{s(1-2\lambda)} U\pi_Q\xi^2)$.
6. *If* $\lambda = 1/2$, *then* $X_n\nu_2/\sqrt{n^s \ln n} \implies N(0, s^2\lambda^2 U\pi_Q\xi^2)$.
7. *If* $\lambda > 1/2$, *then* $X_n\nu_2/a_n(s\lambda)$ *is an* L^2 *bounded martingale which converges almost surely and in* L^2, $X_n\nu_2/n^{s\lambda} \to V$, *where* V *is a non-degenerate random variable.*

Proof. If we collapse the two dominant colours to one colour, then $(W_n, B_n, G_n + Y_n)$ follows the urn model with replacement matrix as in Eqn(20). Then these results follow directly from Thm(15). □

Now suppose $\boldsymbol{R}$ is diagonalizable. Then there is another β-eigenvector, ν_3, of $\boldsymbol{R}$ which is independent of $\nu_i, i = 1, 2, 4$. Following are the weak/strong laws for linear combination corresponding to ν_3.

Theorem 20. *Suppose we have a four colour urn model, with a replacement matrix as in Eqn(22). Moreover assume that all eigenvalues of $\boldsymbol{R}$ are **distinct**, Then we have the following:*

1. *If $\beta < 1/2$, then $X_n\nu_3/\sqrt{n} \implies N(0, \frac{\beta^2}{1-2\beta}\pi_P\xi^2)$.*
2. *If $\beta = 1/2$, then $X_n\nu_3/\sqrt{n\ln n} \implies N(0, \beta^2\pi_P\xi^2)$.*
3. *If $\beta > 1/2$, then $X_n\nu_3/a_n(\beta)$ is an L^2 bounded martingale and $X_n\nu_3/n^\beta$ converges almost surely to a non-degenerate random variable.*

Proof. The proof is similar to that of Thm(17) and hence we omit it. □

Now suppose $\boldsymbol{R}$ is not diagonalizable. Then one of its eigenvalues must repeat. In particular we must have β equal s or $s\lambda$, Then denote the other one by α. Again we consider the Jordan decomposition of $\boldsymbol{R}$, $\boldsymbol{R}T = TJ$, where J is of the form

$$J = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 1 & 0 \\ 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and the first two columns of T, say t_1 and t_2, are the eigenvectors of $\boldsymbol{R}$ corresponding to the eigenvalues α and β. The fourth column of T, say t_4, is equal to ν_4. However, the third column of T, say t_3, is not an eigenvector of $\boldsymbol{R}$. Although the two-dimensional subvector ν formed by the lower half of t_3 is a β-eigenvector of P. Let

$$W = \begin{cases} U & \text{when } \beta = s \\ V & \text{when } \beta = s\lambda \end{cases} \tag{28}$$

where U and V are the random variables as in Thm(19).

Further, if $\beta = s$ then we can take $t_2 = \nu_1$ and $t_1 = \nu_2$ and if $\beta = s\lambda$ then we can take $t_2 = \nu_2$ and $t_1 = \nu_1$.

Then it follows from Thm(19) that $X_nt_2/(n+1)^\beta \to W$ almost surely and in L^2.

Theorem 21. *In a four colour urn model, with a **non-diagonalizable** replacement matrix as in Eqn(22), we have the following:*

1. *If $\beta < 1/2$, then $X_nt_3/\sqrt{n} \implies N(0, \frac{\beta^2}{1-2\beta}\pi_P\xi^2)$.*
2. *If $\beta \geq 1/2$, then $X_nt_3/n^\beta \ln n$ converges to W almost surely, as well as in L^2, where W is as in Eqn(28).*

Proof. The proof is similar to that of Thm(18) and hence I would omit it. □

6 Berry-Esseen Theorems

The following important theorem was dicovered independently by A.C. Berry (1941) and C. G. Esseen (1942)(with radically different proofs). We borrow the proof from [9], Section 5, Chapter XVI.

Theorem 22. *Let X_k be iid random variable, with distribution function F such that $E[X_k] = 0$, $E[X_k^2] = \sigma^2 > 0$ and $E[|X_k|^3] = \rho < \infty$, Let F_n stand for the normalized sum $\frac{(X_1 + \cdots X_n)}{\sigma\sqrt{n}}$. Then*

$$|F_n(x) - \Phi(x)| \leq \frac{3\rho}{\sigma^3\sqrt{n}} \ \forall \ x, n$$

(where Φ is the distribution function of standard normal distribution.)

Proof. We shall use the following lemma,

Lemma 14 (Smoothing Inequality)**.** *Let F be a probability distribution with vanishing expectation and characteristic function φ. Suppose $F - G$ vanish at $\pm\infty$ and that G has a derivative g, such that $|g| \leq m$. Finally, suppose g has a C^1 Fourier transform γ such that $\gamma(0) = 1$ and $\gamma\prime(0) = 0$. Then $\forall x$ and $T > 0$,*

$$|F(x) - G(x)| \leq \frac{1}{\pi}\int_{-T}^{T}\left|\frac{\varphi(\xi) - \gamma(\xi)}{\xi}\right|d\xi + \frac{24m}{\pi T}.$$

Proof. Refer to [9], Lemma 2, XVI.4 . □

For our case, $G = \Phi$ and $F = F_n$, We choose

$$T = \frac{4}{3}\frac{\sigma^3}{\rho}\sqrt{n} \leq \frac{4}{3}\sqrt{n} \quad \text{as} \quad \sigma^3 \leq \rho.$$

Also for the standard normal density we have $m < \frac{2}{5}$. And thus we have,

$$\pi|F_n(x) - \Phi(x)| \leq \int_{-T}^{T}\left|\frac{\varphi^n(\frac{\xi}{\sigma\sqrt{n}}) - e^{-\frac{-1}{2}\xi^2}}{\xi}\right|d\xi + \frac{9 \cdot 6}{T}. \tag{29}$$

Now we estimate the integrand in the integral of (29).
Note for $\alpha, \beta \in \mathbb{C}$ and $|\alpha| \leq \gamma$, $|\beta| \leq \gamma$, $|\alpha^n - \beta^n| \leq n|\alpha - \beta|\gamma^{n-1}$,

$$\left|\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{(-\frac{1}{2n}\xi^2)n}\right| \leq n\left|\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{\frac{\xi^2}{2n}}\right|\gamma^{n-1}.$$

Lemma 15.
$$\left|e^{it} - 1 - \frac{it}{1!} - \cdots - \frac{(it)^{n-1}}{(n-1)!}\right| \leq \frac{t^n}{n!}.$$

Proof. Let

$$\rho_n(t) = \left(e^{it} - 1 - \frac{it}{1!} - \cdots - \frac{(it)^{n-1}}{(n-1)!}\right)$$

then

$$\rho_1(t) = e^{it} - 1 = i\int_0^t e^{ix}dx$$

$$|\rho_1(t)| \leq \left|\int_0^t e^{ix}dx\right| \leq t.$$

Furthermore for $n > 1$,

$$\rho_n(t) = i\int_0^t \rho_{n-1}(x)dx$$

and the rest follows by induction. □

$$\left|\varphi(t) - 1 + \frac{1}{2}\sigma^2 t^2\right| = \left|\int_{-\infty}^{\infty} e^{itx} - 1 - itx + \frac{1}{2}t^2x^2 dF\right| \leq \int_{-\infty}^{\infty} \frac{|t|^3|x|^3}{3!}dF(x) = \frac{\rho}{6}|t|^3$$

$$\implies |\varphi(t)| \leq 1 - \frac{1}{2}\sigma^2t^2 + \frac{1}{6}\rho|t|^3 \text{ if } \frac{1}{2}\sigma^2t^2 \leq 1.$$

For $|\xi| \leq T$ i.e. $|\xi| \leq \frac{4}{3}\frac{\rho}{\sigma^3}\sqrt{n} \implies |\xi|^2 \leq 2n$,

$$\left|\varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right)\right| \leq 1 - \frac{\xi^2}{2n} + \frac{\rho}{6\sigma^3 n^{\frac{3}{2}}}|\xi|^3 \leq 1 - \frac{5}{18n}\xi^2 \leq e^{-\frac{5}{18n}\xi^2}.$$

For $\sqrt{n} \leq 3$ the statement of Thm(22) is trivially true. So without loss of generality we assume $\sqrt{n} > 3$. Then,

$$\left|\varphi\left(\frac{\xi}{\sigma\sqrt{n}}\right)\right|^{n-1} \leq e^{-\frac{5}{18}\frac{n-1}{n}\xi^2} \leq e^{-\frac{\xi^2}{4}}.$$

Now note that as $n \geq 10$,

$$(e^{-\frac{\xi^2}{2n}})^{n-1} \leq e^{-(\frac{n-1}{n})\xi^2} \leq e^{-\frac{\xi^2}{4}}.$$

Hence $e^{-\frac{\xi^2}{4(n-1)}}$ serves as a γ. Now note that

$$n\left|\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - e^{-\frac{\xi^2}{2n}}\right| \leq n\left|\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right) - 1 + \frac{\xi^2}{2n}\right| + n\left|1 - \frac{\xi^2}{2n} - e^{-\frac{\xi^2}{2n}}\right|.$$

Now note that $e^{-x} - 1 + x \leq \frac{x^2}{2}$, $\forall x > 0$ and since

$$(1-x) \leq e^{-x} \implies (1 - e^{-x}) \leq x \implies \int_0^t (1 - e^{-x})dx \leq \frac{t^2}{2}$$

we have

$$n\left|\varphi^n\left(\frac{\xi}{\sigma\sqrt{n}}\right)-e^{-\frac{\xi^2}{2n}}\right|\le n\frac{\rho}{6}\left(\frac{|\xi|}{\sigma\sqrt{n}}\right)^3+\frac{1}{2}\left(\frac{|\xi|^2}{2n}\right)^2=\frac{\rho|\xi|^3}{6\sigma^3\sqrt{n}}+\frac{1}{8n}\xi^4.$$

the integrand in Eqn(29) is bounded by

$$\begin{aligned}&\le\left(\frac{\rho}{6\sigma^3\sqrt{n}}|\xi|^3+\frac{1}{8n}\xi^4\right)\frac{e^{-\frac{\xi^2}{4}}}{|\xi|}\\&=\frac{1}{T}\left(\frac{2\xi^2}{9}+\frac{|\xi|^3}{6\sqrt{n}}\right)e^{-\frac{\xi^2}{4}}.\end{aligned}$$

But as $\sqrt{n}>3$ we have

$$\leqslant\frac{1}{T}\left(\frac{2\xi^2}{9}+\frac{|\xi|^3}{18}\right)e^{-\frac{\xi^2}{4}}.$$

Note that this is integrable over $\mathbb{R}$ (Moreover after change of variables the integrals are Gamma integrals and can be computed explicitly. We omit the details). Thus,

$$\pi|F_n(x)-\Phi(x)|\le\frac{1}{T}\left(\frac{8}{9}\sqrt{\pi}+\frac{8}{9}\right)+\frac{4\cdot 6}{T}.$$

$$T=\frac{4}{3}\frac{\sigma^3}{\rho}\sqrt{n}\implies|F_n(x)-\Phi(x)|\le\frac{3\rho}{\sigma^3\sqrt{n}}.$$

□

We now state and prove a Berry-Esseen Theorem for martingales (Theorem 3.7 of [8]).

Theorem 23 (Berry-Esseen for martingales)**.** *Let $\{S_i=\sum_1^i Z_j,\mathcal{F}_i,1\leqslant i\leqslant n\}$ be a zero-mean martingale with $\mathcal{F}_i$ equal to the σ-field generated by $Z_1,\dots,Z_i$. Let*

$$V_i^2=\sum_1^i E(Z_j^2\mid\mathcal{F}_{j-1}),\quad 1\leqslant i\leqslant n,$$

and suppose that

$$\max_{i\leqslant n}|Z_i|\leqslant n^{-1/2}M\qquad a.s.\tag{30}$$

and

$$P(|V_n^2-1|>9M^2Dn^{-1/2}(\log n)^2)\leqslant Cn^{-1/4}\log n\tag{31}$$

for constants $M,C,$ and $D(\geqslant e)$. Then for $n\geqslant 2$,

$$\sup_{-\infty<x<\infty}|P(S_n\leqslant x)-\Phi(x)|\leqslant(2+C+7MD^{1/2})n^{-1/4}\log n.$$

Proof. We begin with a lemma, whose proof is due to Hyede and Brown (1970).

Lemma 16. *Let $W(t), t \geqslant 0$, be a standard Brownian motion and let T be a nonnegative random variable. Then for all real x and all $\varepsilon > 0$,*

$$|P(W(T) \leqslant x) - \Phi(\sigma^{-1/2}x)| \leqslant \left(\frac{2\varepsilon}{\sigma}\right)^{1/2} + P(|T-\sigma| > \varepsilon).$$

Here Φ denotes the standard normal distribution function.

Proof. First we note that if $0 < \varepsilon < \frac{\sigma}{2}$,

$$P(W(T) \leqslant x) \leqslant P(W(T) \leqslant x; |T-\sigma| \leqslant \varepsilon) + P(|T-\sigma| > \varepsilon),$$

and

$$\begin{aligned}
&P(W(T) \leqslant x; |T-\sigma| \leqslant \varepsilon) \leqslant P\left(\inf_{|t-\sigma|\leqslant\varepsilon} W(t) \leqslant x\right) \\
&= P\left(W(\sigma-\varepsilon) + \inf_{|t-\sigma|\leqslant\varepsilon}[W(t) - W(\sigma-\varepsilon)] \leqslant x\right) \\
&= \int_{-\infty}^{0} P(W(\sigma-\varepsilon) \leqslant x-y) dP\left(\inf_{0\leqslant t\leqslant 2\varepsilon} W(t) \leqslant y\right) \\
&= (\pi\varepsilon)^{-1/2} \int_0^\infty \Phi\left((\sigma-\varepsilon)^{-1/2}(x+y)\right) \exp\left(-y^2/4\varepsilon\right) dy \\
&= \pi^{-1/2} \int_0^\infty \Phi\left((\sigma-\varepsilon)^{-1/2}\left(x+\varepsilon^{1/2}z\right)\right) e^{-z^2/4} dz.
\end{aligned}$$

Consequently,

$$\begin{aligned}
P(W(T) \leqslant x) - \Phi(\sigma^{-1/2}x) \leqslant &\pi^{-1/2} \int_0^\infty \Phi\left((\sigma-\varepsilon)^{-1/2}\left(x+\varepsilon^{1/2}z\right)\right) e^{-z^2/4} dz \\
&- \Phi(\sigma^{-1/2}x) + P(|T-\sigma| > \varepsilon) \\
\leqslant &\pi^{-1/2} \int_0^\infty \mid \Phi\left((\sigma-\varepsilon)^{-1/2}\left(x+\varepsilon^{1/2}z\right)\right) \\
&- \Phi(\sigma^{-1/2}x) \mid e^{-z^2/4} dz + P(|T-\sigma| > \varepsilon).
\end{aligned}$$

The term within the modulus sign in the integrand is bounded by

$$\left|\Phi\left((\sigma-\varepsilon)^{-1/2}\left(x+\varepsilon^{1/2}z\right)\right) - \Phi\left((\sigma-\varepsilon)^{-1/2}x\right)\right| + \left|\Phi\left((\sigma-\varepsilon)^{-1/2}x\right) - \Phi(\sigma^{-1/2}x)\right|.$$

The first term here is not greater than $(2\pi)^{-1/2}(\sigma-\varepsilon)^{-1/2}\varepsilon^{1/2}z$, and for $x \geqslant 0$ the second term is not greater than

$$\begin{aligned}
(2\pi)^{-1/2}(e^{\frac{-x^2}{2\sigma}} x\sigma^{-1/2}) \left[\left(\frac{\sigma}{\sigma-\varepsilon}\right)^{1/2} - 1\right] &\leqslant (2\pi)^{-1/2} e^{-1/2} \left[(1-\varepsilon\sigma^{-1})^{-1/2} - 1\right] \\
&< \pi^{-1/2} e^{-1/2} \varepsilon\sigma^{-1},
\end{aligned}$$

provided that $0 < \varepsilon < \frac{\sigma}{2}$. By symmetry the same bound applies when $x < 0$, and combining these estimates we deduce that for $0 < \varepsilon < \frac{\sigma}{2}$,

$$\begin{aligned}
P(W(T) \leqslant x) - \Phi(\sigma^{-1/2}x) - P(|T-\sigma| > \varepsilon) \\
&\leqslant \pi^{-1/2} e^{-1/2} \varepsilon \sigma^{-1} + \pi^{-1/2} \varepsilon^{1/2} \sigma^{-1/2} \int_0^\infty z e^{-z^2/4} dz \\
&\leqslant (\varepsilon\sigma^{-1})^{1/2} \left((2\pi e)^{-1/2} + 2\pi^{-1/2} \right) \\
&< (2\varepsilon\sigma^{-1})^{1/2}.
\end{aligned}$$

That is,

$$P(W(T) \leqslant x) - \Phi(\sigma^{-1/2}x) \leqslant (2\varepsilon\sigma^{-1})^{1/2} + P(|T-\sigma| > \varepsilon). \tag{32}$$

This bound holds trivially if $\varepsilon \geqslant \frac{\sigma}{2}$. Further, using a similar procedure we deduce that

$$\begin{aligned}
P(W(T) \leqslant x) &\geqslant P(W(T) \leqslant x; |T-\sigma| \leqslant \varepsilon) - P(|T-\sigma| > \varepsilon) \\
&\geqslant P\left(\sup_{|t-\sigma| < \varepsilon} W(t) \leqslant x \right) - P(|T-\sigma| > \varepsilon) \\
&= \pi^{-1/2} \int_0^\infty \Phi\left((\sigma-\varepsilon)^{-1/2} \left(x - \varepsilon^{1/2} z \right) \right) e^{-z^2/4} dz - P(|T-\sigma| > \varepsilon) \\
&\geqslant \Phi(x) - (2\varepsilon\sigma^{-1})^{1/2} - P(|T-\sigma| > \varepsilon). \qquad (33)
\end{aligned}$$

Now combining inequalities (32) and (33), the proof is complete. □

Returning to the proof of the theorem, we deduce from the Skorokhod representation theorem (Thm(30)) that there exists a standard Brownian motion W, nonnegative r.v.'s $T_i, 1 \leqslant i \leqslant n$, such that (without loss of generality)

$$S_i = W(T_i), \quad 1 \leqslant i \leqslant n.$$

Lemma[16] now asserts that for all n, x, and $\Delta > 0$,

$$\begin{aligned}
\left| P(S_n \leqslant x) - \Phi(\sigma^{-1/2}x) \right| &\leqslant 2(\Delta\sigma^{-1})^{1/2} + P\left(\left| T_n - V_n^2 \right| > \Delta \right) \\
&\quad + P\left(\left| V_n^2 - \sigma \right| > \Delta \right) \qquad (34)
\end{aligned}$$

Let $\tau_1 = T_1$ and $\tau_i = T_i - T_{i-1}, 2 \leqslant i \leqslant n$. Thm(30) implies that each τ_i is $\mathscr{G}_i$-measurable and $E(\tau_i \mid \mathscr{G}_{i-1}) = E\left(X_i^2 \mid \mathscr{F}_{i-1} \right)$ a.s., where $\mathscr{G}_i$ is the σ-field generated by $S_1, \ldots, S_i$ and $W(t)$ for $t \leqslant T_i$. Therefore

$$T_n - V_n^2 = \sum_1^n (\tau_i - E(\tau_i \mid \mathscr{G}_{i-1}))$$

is a sum of martingale differences. For any martingale with differences $Z_i, 1 \leqslant i \leqslant n$, and any $p \geqslant 2$ we have from Hölder's and Burkholder's inequalities (see

Theorem[31])

$$E\left|\sum_{1}^{n} Z_i\right|^p \leqslant \left(18pq^{1/2}\right)^p E\left|\sum_{1}^{n} Z_i^2\right|^{p/2} \leqslant \left(18pq^{1/2}\right)^p n^{p/2-1}\sum_{1}^{n} E\left|Z_i\right|^p$$
$$\leqslant \left(18pq^{1/2}\right)^p n^{p/2}\max_{i\leqslant n} E\left|Z_i\right|^p,$$

where $q = \left(1-p^{-1}\right)^{-1} \leqslant 2$ for $p \geqslant 2$.

Applying these inequalities to the martingale with differences $Z_i = \tau_i - E\left(\tau_i \mid \mathscr{G}_{i-1}\right)$ we deduce that

$$P\left(\left|T_n - V_n^2\right| > \Delta\right) \leqslant \Delta^{-p} E\left|\sum_{1}^{n} Z_i\right|^p$$
$$\leqslant \Delta^{-p}\left(18p2^{1/2}\right)^p n^{p/2}\max_{i\leqslant n} E\left|Z_i\right|^p. \tag{35}$$

Now, $\left|Z_i\right| \leqslant \max\left(\tau_i, E\left(\tau_i \mid \mathscr{G}_{i-1}\right)\right)$ and $E\left[E\left(\tau_i \mid \mathscr{G}_{i-1}\right)^p\right] \leqslant E\left(\tau_i^p\right)$, by Jensen's inequality. Hence by Theorem[30],

$$E\left|Z_i\right|^p \leqslant E\left[\tau_i^p + E\left(\tau_i \mid \mathscr{G}_{i-1}\right)^p\right]$$
$$\leqslant 2E\left(\tau_i^p\right) \leqslant 4\Gamma(p+1)E\left|X_i\right|^{2p}. \tag{36}$$

Stirling's expansion of $\Gamma(p+1)$ implies that for $p \geqslant 2$,

$$\Gamma(p+1) \leqslant (2\pi)^{1/2}p^{p+1/2}e^{-p+1/24},$$

and combining this with both Eqn(35) and Eqn(36) we see that

$$P\left(\left|T_n - V_n^2\right| > \Delta\right) \leqslant 10.5\left(9.4p^2\right)^p p^{1/2}n^{p/2}\Delta^{-p}\max_{i\leqslant n} E\left|X_i\right|^{2p}. \tag{37}$$

From the hypothesis of the theorem we have

$$\max_{i\leqslant n} E\left|X_i\right|^{2p} \leqslant n^{-p}M^{2p},$$

and this together with Eqn(34) and Eqn(37) implies that for all $\Delta > 0$,

$$\left|P\left(S_n \leqslant x\right) - \Phi(\sigma^{-1/2}x)\right| \leqslant 2(\Delta\sigma^{-1})^{1/2} + 10.5\left(9.4p^2M^2\right)^p p^{1/2}n^{-p/2}\Delta^{-p} + P\left(\left|V_n^2 - \sigma\right| > \Delta\right).$$

We now choose $\Delta = \Delta(n) \to 0$ and $p = p(n) \to \infty$ to minimize the sum of the first two terms above. Let

$$\Delta = 9.4M^2Dn^{-1/2}(\log n)^2, \quad p = \log n \quad \text{and } \sigma = 1.$$

If $n > e^2$ then $p > 2$, and

$$2\Delta^{1/2} + 10.5\left(9.4p^2M^2\right)^p p^{1/2}n^{-p/2}\Delta^{-p}$$
$$\leqslant 6.2MD^{1/2}n^{-1/4}\log n + 10.5D^{-\log n}(\log n)^{1/2}$$
$$\leqslant \left(7MD^{1/2} + 2\right)n^{-1/4}\log n.$$

Here we have assumed that $D \geqslant e$. Consequently,

$$|P(S_n \leqslant x) - \Phi(x)| \leqslant \left(7MD^{1/2} + 2\right) n^{-1/4} \log n + P\left(\left|V_n^2 - 1\right| > \Delta\right),$$

and combined with hypothesis (Eqn(31)) on V_n this proves Eqn(23) for $n > e^2$. The bound in Eqn(23) applies trivially if $2 \leqslant n < e^2$. Hence the proof is complete. □

Remark: Note that we proved Thm(23) just for the natural filtration, but it is true for arbitrary filtrations with slightly modified constants. For more details read the Remarks after Theorem 3.7 of [8].

7 Final Result

In this section we will apply Thm(23) to the Friedman's urn model. Upon making a further assumption of $\lambda < 1/4$, we will be able to obtain a rate of $n^{\lambda-1/4}$ for the weak convergence in Thm(6). We will see later why Thm(23) is inadequate to give any results for the case $\lambda \geq 1/4$.

Recall Friedman's urn model has the following replacement matrix,

$$\boldsymbol{R} = \begin{pmatrix} s & 1-s \\ 1-s & s \end{pmatrix}, \quad s \in [0,1].$$

Let λ denote the non-principal eigenvalue and ξ the corresponding eigenvector, i.e. $\lambda = 2s - 1$ and $\xi = (1, -1)'$. Let $X_n = (W_n, B_n)$ be the vector of the number of white and black balls after the n-th draw. Let $Y_n = \frac{X_n \cdot \xi}{\sqrt{n}}$.

Note that Y_n converges in distribution to $N(0, \frac{\lambda^2}{1-2\lambda})$, given $\lambda < 1/2$. To prove this we showed that $\lim_{n\to\infty} E[Y_n^{2k}] \to \mu(2k)$, where $\mu(k)$ is the k-th moment of $N(0, \frac{\lambda^2}{1-2\lambda})$.

We shall use this fact along with the Berry-Esseen theorem for martingales (Thm[23]) to get a rate of convergence for Y_n.

Theorem 24. *Let F_{Y_n} denote the distribution function of Y_n and Φ_{σ_n} the distribution function of a normal distribution with variance σ_n. Then,*

$$||F_{Y_n} - \Phi_{\sigma_n}||_\infty \leq Const. \cdot \frac{\ln n}{n^{1/4-\lambda}}$$

where $\sigma_n = \lambda^2 a_n^2 \sum_{k=1}^n \frac{1}{(k+1)a_k^2}$.

Proof. In order to use the Berry-Esseen theorem for martingales, we will prove weak convergence of Y_n again, using martingale techniques. To do this we split Y_n into a conditional expectation and a martingale difference, just like in Thm[15]. Upon similar computation we get

$$Y_{n+1} = a_n Y_1 + \sum_{i=1}^n \frac{a_n}{a_i} Y_j O(j^{-2}) + \sum_{i=1}^n \frac{a_n}{a_i} M_{i+1}, \tag{38}$$

where

$$\begin{aligned} M_{n+1} &= Y_{n+1} - E[Y_{n+1}|\mathcal{F}_n] \\ &= \frac{\lambda}{\sqrt{n+1}}\left[\chi_{n+1} - \frac{X_n}{n+1}\right]\xi, \end{aligned} \tag{39}$$

and

$$E[M_{n+1}^2|\mathcal{F}_n] = \frac{\lambda^2}{n+1}\left[1 - \frac{(W_n - B_n)^2}{(n+1)^2}\right]. \tag{40}$$

Moreover, the conditional Lyapunov condition is satisfied and conditional variance converges to $\frac{\lambda^2}{1-2\lambda}$ for the martingale array $\{(Z_{n+1,k}, \mathcal{F}_{n+1,k})|n \geq 1, k = 1, \ldots, n\}$. Here

$$Z_{n+1,k} = \frac{a_n}{a_k}M_{k+1} \quad \text{and} \quad \mathcal{F}_{n+1,k} = \mathcal{F}_{k+1}.$$

(Note: In fact the calculations for obtaining the above equations are exactly the same as that in the proof of Thm[15], $\lambda < 1/2$ case with $s = 1$. Hence we omit them.)

We shall now apply the Berry-Esseen theorem for martingales on the finite zero mean martingale $(S_k, \mathcal{F}_{k+1})_{1\leq k\leq n}$, with martingale differences $Z_{n+1,k}$. First note that, $\max_{k\leq n}\sqrt{n}|Z_{n+1,k}| = \max_{k\leq n}\sqrt{n}|M_{k+1}|a_n a_k^{-1}$. From Eqn(39) it follows that M_{n+1} is bounded by $2\lambda/\sqrt{n+1}$. Therefore, $\max_{k\leq n}\sqrt{n}|Z_{n+1,k}| \leq 2\lambda\sqrt{n}a_n \max_{k\leq n}(\sqrt{k+1}a_k)^{-1}$. Now since $a_n \sim n^{\lambda-1/2}$, we have for some constant $C_1 > 0$

$$\max_{k\leq n}\sqrt{n}|Z_{n+1,k}| \leq C_1 \cdot n^{\lambda}.$$

Hence we choose $M = C_1 n^{\lambda}$ in the hypothesis of Berry-Esseen theorem.

The conditional variance is

$$V_n^2 = \sigma_n - \lambda^2 a_n^2 \sum_{k=1}^{n}\frac{(W_k - B_k)^2}{(k+1)^3 a_k^2}.$$

where $\sigma_n = \lambda^2 a_n^2 \sum_{k=1}^{n}\frac{1}{(k+1)a_k^2}$. Then Eqn(23) gives us, $\forall p \geq 2$

$$\begin{aligned} \left|P\left(S_n \leqslant x\right) - \Phi(\sigma_n^{-1/2}x)\right| &\leqslant 2(\Delta\sigma_n^{-1})^{1/2} + P\left(\left|V_n^2 - \sigma_n\right| > \Delta\right) \\ &\quad + C_2\left(C_3 p^2 n^{2\lambda-\frac{1}{2}}\right)^p p^{1/2}\Delta^{-p}. \end{aligned} \tag{41}$$

Now we analyze the middle term,

$$P\left(\left|V_n^2 - \sigma_n\right| > \Delta\right) \leq \frac{E[\left|V_n^2 - \sigma_n\right|^r]}{\Delta^r}. \tag{42}$$

It can be shown that

$$E[\left|V_n^2 - \sigma_n\right|^r] \leq D_r a_n^{2r}, \forall r \geq 2, \tag{43}$$

where D_r is a constant. We will first show this for $r = 2$.

$$E[\left|V_n^2 - \sigma_n\right|^2] = \lambda^4 a_n^4 \left[\sum_{k=1}^{n} \frac{E(W_k - B_k)^4}{(k+1)^6 a_k^4} + 2 \sum_{i<j} \frac{E[(W_i - B_i)^2 (W_k - B_k)^2]}{(i+1)^3 a_i^2 (j+1)^3 a_j^2} \right].$$

Note that

$$\sum_{k=1}^{n} \frac{E(W_k - B_k)^4}{(k+1)^6 a_k^4} \sim \sum_{k=1}^{n} \frac{k^2}{k^6 k^{4(\lambda - 1/2)}} = \sum_{k=1}^{n} \frac{1}{n^{4\lambda + 2}} < \infty.$$

Form Holders inequality we have

$$E[(W_i - B_i)^2 (W_k - B_k)^2] \leq E[(W_i - B_i)^4]^{1/2} \cdot E[(W_j - B_j)^4]^{1/2}.$$

Therefore we have

$$\begin{aligned} \sum_{i<j} \frac{E[(W_i - B_i)^2 (W_k - B_k)^2]}{(i+1)^3 a_i^2 (j+1)^3 a_j^2} &\leq \sum_{i<j} \frac{E[(W_i - B_i)^4]^{1/2} E[(W_j - B_j)^4]^{1/2}}{(i+1)^3 a_i^2 (j+1)^3 a_j^2} \\ &\sim \sum_{i<j} \frac{i \cdot j}{i^3 i^{2(\lambda - 1/2)} j^3 j^{2(\lambda - 1/2)}} \\ &= \sum_{i<j} \frac{1}{i^{1+2\lambda} j^{1+2\lambda}} < \infty. \end{aligned}$$

This proves Eqn(43) for $r = 2$. Similarly, using the generalized holder's inequality and the fact that $E[(W_n - B_n)^{2k}] \sim n^k$, we can prove Eqn(43) for $r > 2$.

Now substituting Eqn(43) in Eqn(41) we get

$$\begin{aligned} \left| P\left(S_n \leqslant x\right) - \Phi(\sigma_n^{-1/2} x) \right| &\leqslant 2(\Delta \sigma_n^{-1})^{1/2} + D_r' n^{r(2\lambda - 1)} \Delta^{-r} \\ &\quad + C_2 \left(C_3 p^2 n^{2\lambda - \frac{1}{2}} \right)^p p^{1/2} \Delta^{-p}. \end{aligned} \tag{44}$$

Now we will choose Δ and p to minimize the sum of the first and the last term. Let $\Delta = C_3 (\ln n)^2 n^{2\lambda - 1/2} \cdot D$, where $D > e$ and let $p = \ln n$. Thus we get,

$$\begin{aligned} \left| P\left(S_n \leqslant x\right) - \Phi(\sigma_n^{-1/2} x) \right| &\leqslant 2\sqrt{\frac{C_3 D}{\sigma_n}} \frac{\ln n}{n^{1/4 - \lambda}} + D_r''((\ln n)^2 n^{1/2})^{-r} \\ &\quad + C_2 D^{-\ln n} (\ln n)^{1/2}. \end{aligned} \tag{45}$$

Since $\frac{\ln n}{n^{1/4 - \lambda}}$ is the dominating rate we have

$$\left| P\left(S_n \leqslant x\right) - \Phi(\sigma_n^{-1/2} x) \right| \leqslant \text{Const.} \cdot \frac{\ln n}{n^{1/4 - \lambda}}. \tag{46}$$

Lemma 17. *Suppose X, Y are the two rvs. Let $F_X(x)$ denote the cdf of X (i.e., $F_X(x) = P(X \leq x)$) and let Φ_σ denote the cdf of normal distribution with variance σ. Then,*

$$||F_{X+Y} - \Phi_\sigma||_\infty \leq ||F_X - \Phi_\sigma||_\infty + \frac{\epsilon}{\sqrt{2\pi\sigma}} + P(|Y| > \epsilon|).$$

Proof. First note that,

$$\begin{aligned}\{X+Y\le x\} &= \{X+Y\le x, Y\ge -h\}\sqcup\{X+Y\le x, Y<-h\}\\ &\subseteq \{X\le x-Y, -Y\le h\}\cup\{Y<-h\}\\ &\subseteq \{X\le x+h\}\cup\{Y<-h\}.\end{aligned}$$

Therefore,

$$P(X+Y\le x)\le P(X\le x+h)+P(Y<-h).$$

From this we get,

$$\begin{aligned}F_{X+Y}(x)-\Phi_\sigma(x) &\le F_X(x+h)-\Phi_\sigma(x+h)+\Phi_\sigma(x+h)-\Phi_\sigma(x)\\ &\le ||F_X-\Phi_\sigma||_\infty+\int_x^{x+h}\frac{1}{\sqrt{2\pi\sigma}}e^{-t^2/2\sigma^2}dt+P(Y<-h)\\ &\le ||F_X-\Phi_\sigma||_\infty+\frac{h}{\sqrt{2\pi\sigma}}+P(Y<-h).\end{aligned}$$

Now similarly we have

$$\{X\le x-h\}\cap\{Y\le h\}\subseteq\{X+Y\le x\}\implies P(X+Y\le x)\ge P(X\le x-h)-P(Y>h).$$

The above implication follows from the fact $P(A-B)\ge P(A)-P(B)$. Therefore,

$$\begin{aligned}&P(X+Y\le x)\ge P(X\le x-h)-P(Y>h)\\ &\implies -P(X+Y\le x)\le -P(X\le x-h)+P(Y>h)\\ &\implies -F_{X+Y}(x)+\Phi_\sigma(x)\le -F_X(x-h)+\Phi_\sigma(x-h)-\Phi_\sigma(x-h)+P(Y>h)\\ &\implies -[F_{X+Y}(x)-\Phi_\sigma(x)]\le -[F_X(x-h)-\Phi_\sigma(x-h)]+\int_{x-h}^{x}\frac{1}{\sqrt{2\pi\sigma}}e^{-t^2/2\sigma^2}dt+P(Y>h)\\ &\implies -[F_{X+Y}(x)-\Phi_\sigma(x)]\le |F_X-\Phi_\sigma||_\infty+\frac{h}{\sqrt{2\pi\sigma}}+P(Y>h).\end{aligned}$$

Combining both the inequlities we get that, $\forall\ x\in\mathbb{R}$ and $\forall\ h>0$,

$$\begin{aligned}&|F_{X+Y}(x)-\Phi_\sigma(x)|\le |F_X-\Phi_\sigma||_\infty+\frac{h}{\sqrt{2\pi\sigma}}+P(|Y|>h)\\ &\qquad\implies ||F_{X+Y}(x)-\Phi_\sigma(x)||_\infty\le |F_X-\Phi_\sigma||_\infty+\frac{h}{\sqrt{2\pi\sigma}}+P(|Y|>h).\end{aligned}$$

□

We now use the previous lemma to combine the rate in Eqn(46) along with Eqn(38). Note that the first two terms in Eqn(38) are bounded by a constant multiple of a_n (Since $|Y_n|\le\sqrt{n}+1/\sqrt{n}$). Suppose $W=a_nY_1+\sum_{i=1}^n\frac{a_n}{a_i}Y_jO(j^{-2})$, then $Y_{n+1}=S_n+W$. Now the previous lemma gives us

$$||F_{S_n+W}-\Phi_{\sigma_n}||_\infty\le||F_{S_n}-\Phi_{\sigma_n}||_\infty+\frac{\epsilon}{\sqrt{2\pi\sigma_n}}+P(|W|>\epsilon)$$

Now since W is bounded by a constant times a_n, we apply Markov inequality to get

$$P(|W| > \epsilon) \leq \frac{E[|W|^s]}{\epsilon^s} \leq \text{Const.} \frac{n^{s(\lambda-1/2)}}{\epsilon^s}$$

Thus we have

$$||F_{S_n+W} - \Phi_{\sigma_n}||_\infty \leq ||F_{S_n} - \Phi_{\sigma_n}||_\infty + \frac{\epsilon}{\sqrt{2\pi\sigma_n}} + \text{Const.} \frac{n^{s(\lambda-1/2)}}{\epsilon^s}$$

Equating the last two terms and solving for ϵ we get

$$\epsilon = \text{Const.} \sigma_n^{1/2(s+1)} n^{(\frac{s}{s+1})(\lambda-1/2)}$$

Combinining this with Eqn(46) and the fact $\sigma_n \to \frac{\lambda^2}{1-2\lambda}$ (non-zero), we get

$$||F_{S_n+W} - \Phi_{\sigma_n}||_\infty \leq \text{Const.} \quad \cdot \frac{\ln n}{n^{1/4-\lambda}} + \text{Const.} n^{(\frac{s}{s+1})(\lambda-1/2)}.$$

For a large value of s, $n^{\lambda-1/4}$ would be the dominating rate, hence we have

$$||F_{Y_n} - \Phi_{\sigma_n}||_\infty \leq Const. \cdot \frac{\ln n}{n^{1/4-\lambda}}$$

Finishing the proof of the result. □

Remark: Note that if $\lambda \geq 1/4$ then the third term in Eqn(41) is large as $n \to \infty$ and $\Delta \to 0$. This is why Thm(23) is inadequate to get any results in this case. If we could weaken the assumption in Eqn(30), which would be satisfied by the urn model, then we may be able to get rates for the case $\lambda \geq 1/4$.

8 Appendix

Theorem 25 (Sub-Martingale Convergence Theorem). *Suppose* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *is a submartingale and* $\sup_n \mathbb{E}[X_n^+] < \infty$, *then* $\exists X_\infty$, *an integrable random variable* $\ni X_n \to X_\infty$ *almost surely.*

Proof. Refer Ash Thm 6.4.3 □

Theorem 26 (Backward Sub-MCT). *Suppose* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *is a backward submartingale and* $\inf_n \mathbb{E}[X_n] > -\infty$, *then* $\exists X_\infty$, *an integrable random variable* $\ni X_n \to X_\infty$ *almost surely.*

Theorem 27 (DCT for conditional expectation). *If* $|X_n| \leq Z$, *an integrable random variable, and* $X_n \to X$ *almost surely, then* $\mathbb{E}[X_n|F] \to \mathbb{E}[X|F]$.

Theorem 28 (Second Borel–Cantelli lemma for adapted sequences). *Suppose* $B_n \in \mathcal{F}_n \forall n \geq 1$, *where* $(\mathcal{F}_n)_{n\geq 0}$ *is a filtration, then*

$$\frac{\sum_{i=1}^n I_{B_i}}{\sum_{i=1}^n \mathbb{P}(B_n|\mathcal{F}_{n-1})} \to 1$$

almost surely on $[\sum_{i=1}^\infty \mathbb{P}(B_n|\mathcal{F}_{n-1}) = \infty]$.

Theorem 29. *In a Polya's urn, $\mathbb{P}(X_{n,i} \leq K, \forall n) = 0$ and hence $\mathbb{P}(X_{n,i} \not\to \infty) = 0$.*

Proof. Note that it suffices to show this for a Polya's urn with two colours, as we can always consider a group of colours as one colour. Note that $\mathbb{E}[C_{n,1}] = C_{0,1}$, $\forall n$, this can be shown easily by induction. $K \geq \mathbb{E}[X_{n,1}] = T_n\mathbb{E}[C_{n,1}] = T_n C_{0,1}$, a contradiction as $T_n \to \infty$ and $C_{0,1} > 0$. $\square$

Theorem 30 (**Skorokhod Representation**). *Let $\{S_n = \sum_1^n X_i, \mathscr{F}_n, n \geqslant 1\}$ be a zero-mean, square integrable martingale. Then there exists a probability space supporting a (standard) Brownian motion W and a sequence of nonnegative variables $\tau_1, \tau_2, \ldots$ with the following properties. If $T_n = \sum_1^n \tau_i, S_n' = W(T_n), X_1' = S_1', X_n' = S_n' - S_{n-1}'$ for $n \geqslant 2$, and $\mathscr{G}_n$ is the σ-field generated by $S_1', \ldots, S_n'$ and by $W(t)$ for $0 \leqslant t \leqslant T_n$, then*

1. $\{S_n, n \geqslant 1\} \stackrel{\mathrm{d}}{=} \{S_n', n \geqslant i\}$,
2. *T_n is $\mathscr{G}_n$-measurable,*
3. *for each real number $r \geqslant 1$,*

$$E\left(\tau_n^r \mid \mathscr{G}_{n-1}\right) \leqslant C_r E\left(\left|X_n'\right|^{2r} \mid \mathscr{G}_{n-1}\right) = C_r E\left(\left|X_n'\right|^{2r} \mid X_1', \ldots, X_{n-1}'\right) \quad a.s.,$$

where $C_r = 2\left(8/\pi^2\right)^{r-1}\Gamma(r+1)$, and

4. $E\left(\tau_n \mid \mathscr{G}_{n-1}\right) = E\left(X_n'^2 \mid \mathscr{G}_{n-1}\right) \quad a.s.$

Proof. Refer to Theorem A.1 of [8]. $\square$

Theorem 31 (**Burkholder's Inequality**). *If $\{S_i, \mathscr{F}_i, 1 \leqslant i \leqslant n\}$ is a martingale and $1 < p < \infty$, then there exist constants C_1 and C_2 depending only on p such that*

$$C_1 E\left|\sum_{i=1}^n X_i^2\right|^{p/2} \leqslant E\left|S_n\right|^p \leqslant C_2 E\left|\sum_{i=1}^n X_i^2\right|^{p/2}.$$

(The proof shows that suitable constants are given by $C_1^{-1} = \left(18p^{1/2}q\right)^p$ and $C_2 = \left(18pq^{1/2}\right)^p$, where $p^{-1} + q^{-1} = 1$.)

Proof. Refer to Theorem 2.10 of [8]. $\square$

8.1 Martingale Central Limit Theorem

Notation: Let $(X_{ni}, \mathcal{F}_{ni})_{i\geq 1}$ be a martingale for each n. Now define the martingale difference array as $Y_{nk} = X_{nk} - X_{n,k-1}, \forall k \geq 2$. Let $X_{n0} = 0$ and $\mathcal{F}_{n0} = \{\phi, \Omega\}$. *Assume* that Y_{nk} has finite second moment and let $\sigma_{nk}^2 = \mathbb{E}[Y_{nk}^2|\mathcal{F}_{n,k-1}]$. Moreover, assume that $\sum_{k=1}^\infty Y_{nk}$ and $\sum_{k=1}^\infty \sigma_{nk}^2$ converge with probability 1.

Following is a martingale central limit theorem from [3].

Theorem 32. *Suppose that $\sum_{k=1}^{\infty} \sigma_{nk}^2 \to \sigma^2$ in probability, where σ is a positive constant, and that $\sum_{k=1}^{\infty} \mathbb{E}[Y_{nk}^2 I_{[|Y_{nk}|\geq \epsilon]}] \to 0, \forall \epsilon > 0$. Then $\sum_{k=1}^{\infty} Y_{nk} \implies \sigma N$, where N is the standard Normal distribution.*

Proof. We will first assume that $\sum_{k=1}^{\infty} \sigma_{nk}^2 \leq c$, and then later remove this assumption.

Let
$$S_0 = 0,\ S_m = \sum_{k=1}^{m} Y_{nk},\ S_\infty = \sum_{k=1}^{\infty} Y_{nk},$$
similarly let,
$$\Sigma_0 = 0, \Sigma_m = \sum_{k=1}^{m} \sigma_{nk}^2, \Sigma_\infty = \sum_{k=1}^{\infty} \sigma_{nk}^2.$$
Note that S and Σ both depend on n but for convinience of notation we omit writing their dependence.

We wish to show that $S_\infty \to \sigma N$ in distribution. To show this we will show that $\phi_{S_\infty}(t) \to \phi_{\sigma N}(t)$ as $n \to \infty$. The result then follows by Levy's continuity theorem.

$$\begin{aligned}
|\phi_{S_\infty}(t) - \phi_{\sigma N}(t)| &= |\mathbb{E}[e^{itS_\infty} - e^{\frac{-t^2\sigma^2}{2}}]| \\
&= |\mathbb{E}[e^{itS_\infty}(1 - e^{\frac{t^2\Sigma_\infty}{2}} \cdot e^{\frac{-t^2\sigma^2}{2}}) + (e^{itS_\infty} \cdot e^{\frac{t^2\Sigma_\infty}{2}} - 1)e^{\frac{-t^2\sigma^2}{2}}]| \\
&\leq \mathbb{E}[|1 - e^{\frac{t^2\Sigma_\infty}{2}} \cdot e^{\frac{-t^2\sigma^2}{2}}|] + |\mathbb{E}[(e^{itS_\infty} \cdot e^{\frac{t^2\Sigma_\infty}{2}} - 1)]|. \qquad (47)
\end{aligned}$$

Let the first term in (47) be A and the second term B.

$|1 - e^{\frac{t^2(\Sigma_\infty - \sigma^2)}{2}}|$, by our assumption $\Sigma_\infty = \sum_{k=1}^{\infty} \sigma_{nk}^2 \leq c$ is bounded and $\Sigma_\infty \to \sigma^2$ in probability. Hence we can apply extended DCT to get $A \to 0$ as $n \to \infty$.

Now note that B can be written as the telescoping sum (49), as follows.

$$\begin{aligned}
e^{itS_\infty} \cdot e^{\frac{t^2\Sigma_\infty}{2}} - 1 &= \lim_{m\to\infty} e^{itS_m} \cdot e^{\frac{t^2\Sigma_m}{2}} - 1 \qquad (48) \\
&= \lim_{m\to\infty} \sum_{k=1}^{m} e^{itS_{k-1}}(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})e^{\frac{t^2\Sigma_k^2}{2}}. \qquad (49)
\end{aligned}$$

Now as the m-th partial sum in (48) is uniformly bounded in m due to our initial assumption, $\Sigma_m \leq \sum_{k=1}^{\infty} \sigma_{nk}^2 \leq c$, we can apply DCT to interchange sum and the integral to get

$$
\begin{aligned}
B &= |\mathbb{E}[\lim_{m\to\infty}\sum_{k=1}^{m} e^{itS_{k-1}}(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})e^{\frac{t^2\Sigma_k^2}{2}}]| \\
&= |\lim_{m\to\infty}\sum_{k=1}^{m} \mathbb{E}[e^{itS_{k-1}}(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})e^{\frac{t^2\Sigma_k^2}{2}}]| \\
&= |\sum_{k=1}^{\infty} \mathbb{E}[\mathbb{E}[e^{itS_{k-1}}(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})e^{\frac{t^2\Sigma_k^2}{2}}|\mathcal{F}_{n,k-1}]]| \qquad (50) \\
&= |\sum_{k=1}^{\infty} \mathbb{E}[e^{itS_{k-1}}e^{\frac{t^2\Sigma_k^2}{2}}\mathbb{E}[(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})|\mathcal{F}_{n,k-1}]]| \qquad (51) \\
&\le \sum_{k=1}^{\infty} \mathbb{E}[|e^{itS_{k-1}}e^{\frac{t^2\Sigma_k^2}{2}}\mathbb{E}[(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})|\mathcal{F}_{n,k-1}]|] \\
&\le e^{\frac{t^2c^2}{2}}\sum_{k=1}^{\infty} \mathbb{E}[|\mathbb{E}[(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})|\mathcal{F}_{n,k-1}]|] \qquad (52)
\end{aligned}
$$

where (51) follows from the fact that S_{k-1}, Σ_k are $\mathcal{F}_{n,k-1}$ measurable. Hence it suffices to show that the sum in (52) goes to zero. To do this we use the Taylor's expansion and bound the error terms appropriately using the Linderberg condition. Note that

$$e^x = 1 + x + x^2/2 + R,$$

where $|R| \le \min(|x|^3/6, |x|^2)$ and

$$e^z = 1 + z + R'$$

, where $|R'| \le |z|^2 e^{|z|}$. Using this we have,

$$
\begin{aligned}
\mathbb{E}[e^{itY_{nk}}|\mathcal{F}_{n,k-1}] &= 1 + it\mathbb{E}[Y_{nk}|\mathcal{F}_{n,k-1}] - \mathbb{E}[Y_{nk}^2]t^2/2 + \mathbb{E}[\theta|\mathcal{F}_{n,k-1}] \\
&= 1 - \sigma_{nk}^2 t^2/2 + \mathbb{E}[\theta|\mathcal{F}_{n,k-1}], \\
\mathbb{E}[e^{\frac{-t^2\sigma_{nk}^2}{2}}|\mathcal{F}_{n,k-1}] &= e^{\frac{-t^2\sigma_{nk}^2}{2}} = 1 - \sigma_{nk}^2 t^2/2 + \theta', \\
|\mathbb{E}[e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}}|\mathcal{F}_{n,k-1}]| &= |\mathbb{E}[\theta|\mathcal{F}_{n,k-1}] - \theta'| \\
&\le \mathbb{E}[|\theta||\mathcal{F}_{n,k-1}] + |\theta'|, \\
\mathbb{E}[|\mathbb{E}[e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}}|\mathcal{F}_{n,k-1}]|] &\le \mathbb{E}[|\theta| + |\theta'|] \qquad (53) \\
|\theta| &\le \min(|tY_{nk}|^3, |tY_{nk}|^3) \le K_t|Y_{nk}|^2\min(|Y_{nk}|, 1) \\
&\le K_t Y_{nk}^2(I_{|Y_{nk}|\ge\epsilon} + \epsilon) \qquad (54) \\
|\theta'| &\le \frac{t^4\sigma_{nk}^4}{4}e^{\frac{t^2\sigma_{nk}^2}{2}} \le \frac{t^4}{4}e^{\frac{t^2c^2}{2}}\sigma_{nk}^4 \le K_t\sigma_{nk}^4 \qquad (55)
\end{aligned}
$$

where $K_t = max(t^3, t^2, t^4 e^{\frac{t^2c^2}{2}})$. Using (54) and (55) we can bound (53) to get

$$\sum_{k=1}^{\infty}\mathbb{E}[|\mathbb{E}[(e^{itY_{nk}} - e^{\frac{-t^2\sigma_{nk}^2}{2}})|\mathcal{F}_{n,k-1}]|] \le K_t\sum_{k=1}^{\infty}(\mathbb{E}[Y_{nk}^2 I_{|Y_{nk}|\ge\epsilon}] + \epsilon\mathbb{E}[\sigma_{nk}^2] + \mathbb{E}[\sigma_{nk}^4]). \qquad (56)$$

Now note that

$$\sigma_{nk}^2 \leq \mathbb{E}[\epsilon^2 + Y_{nk}^2 I_{|Y_{nk}|\geq\epsilon}|\mathcal{F}_{n,k-1}] \leq \epsilon^2 + \sum_{k=1}^{\infty} \mathbb{E}[Y_{nk}^2 I_{|Y_{nk}|\geq\epsilon}|\mathcal{F}_{n,k-1}].$$

Hence we have

$$\mathbb{E}[\sup_{n\geq 0}\{\sigma_{nk}^2\}] \leq \epsilon^2 + \sum_{k=1}^{\infty} \mathbb{E}[Y_{nk}^2 I_{|Y_{nk}|\geq\epsilon}].$$

Hence for n large enough, (56) is bounded by $K_t[\epsilon + 2c\epsilon + c\epsilon^2]$. Since ϵ is arbitrary, we have proved our result under the assumption, $\sum_{k=1}^{\infty} \sigma_{nk}^2 \leq c$.

To remove the assumption, take $c > \sigma^2$, and consider $A_{nk} = [\sum_{j=1}^{k} \sigma_{nj}^2 \leq c]$, and take $Z_{nk} = Y_{nk} I_{A_{nk}}$. Then note that $A_{nk} \in \mathcal{F}_{n,k-1}$ and A_{nk} decreases to $A_{n\infty} = [\sum_{j=1}^{\infty} \sigma_{nj}^2 \leq c]$. Thus we have,

$$\mathbb{E}[Z_{nk}|\mathcal{F}_{n,k-1}] = I_{A_{nk}} \mathbb{E}[Y_{nk}|\mathcal{F}_{n,k-1}] = 0$$

and

$$\tau_{nk}^2 = \mathbb{E}[Z_{nk}^2|\mathcal{F}_{n,k-1}] = I_{A_{nk}} \sigma_{nk}^2.$$

Observe that

$$\sum_{j=1}^{\infty} \tau_{nj}^2 = \sum_{j=1}^{k} \sigma_{nj}^2$$

on $A_{nk} - A_{n,k+1}$ and

$$\sum_{j=1}^{\infty} \tau_{nj}^2 = \sum_{j=1}^{\infty} \sigma_{nj}^2$$

on $A_{n\infty}$, therefore

$$\sum_{j=1}^{\infty} \tau_{nj}^2 \leq c.$$

Since $|Z_{nk}| \leq |Y_{nk}|$ we also have that

$$\sum_{k=1}^{\infty} \mathbb{E}[Z_{nk}^2 I_{[|Z_{nk}|\geq\epsilon]}] \leq \sum_{k=1}^{\infty} \mathbb{E}[Y_{nk}^2 I_{[|Y_{nk}|\geq\epsilon]}] \to 0, \forall \epsilon > 0.$$

And finally note that $\mathbb{P}(A_{n\infty}) \to 1$ as $n \to \infty$, because

$$\mathbb{P}(A_{n\infty}) \leq \mathbb{P}(|(\sum_{k=1}^{\infty} \sigma_{nk}^2) - \sigma^2| > c) \to 0.$$

Thus $\sum_{k=1}^{\infty} \tau_{nk}^2 \to \sigma^2$ in probability.

Hence we have $\sum_{k=1}^{\infty} Z_{nk} \implies \sigma N$, but $\sum_{k=1}^{\infty} Z_{nk} = \sum_{k=1}^{\infty} Y_{nk}$ on $A_{n\infty}$, and $\mathbb{P}(A_{n\infty}) \to 1$ as $n \to \infty$. Thus $\sum_{k=1}^{\infty} Y_{nk} \implies \sigma N$. $\square$

Following is another Martingale Central limit theorem from [8].

Def: If Y_n is a sequence of r.v.'s' on a probability space $(\Omega, \mathcal{F}, P)$ converging in distribution to an r.v.Y, we say that the convergence is *stable* if for all continuity points y of Y and all events$E \in \mathcal{F}$, the limit

$$\lim_{n\to\infty} P(Y_n \leq y \cap E) = Q_y(E)$$

exists, and if $Q_y(E) \to P(E)$ as $y \to \infty$. We designate the convergence by writing $Y_n \xrightarrow{d} Y$ (stably).

Theorem 33. *Suppose that $Y_n \xrightarrow{d} Y$, where all the Y_n are on the same space $(\Omega, \mathcal{F}, P)$. Then $Y_n \to Y$ (stably) if and only if there exists a variable Y' on an extension of $(\Omega, \mathcal{F}, P)$, with the same distribution as Y, such that for all real t,*

$$\exp(itY_n) \to Z(t) = \exp(itY') \quad (\textit{ weakly in } L^1) \quad \textit{as} \quad n \to \infty,$$

and $E[Z(t)I(E)]$ is a continuous function of t for all $E \in \mathcal{F}$.

Lemma 18. *Let η^2 be an a.s. finite r.v. and suppose that*

$$\max_i |X_{ni}| \xrightarrow{p} 0, \tag{57}$$

$$\sum_i X_{ni}^2 \xrightarrow{p} \eta^2, \tag{58}$$

and

$$\textit{for all real } t, T_n(t) \to 1 \ (\textit{weakly in } L^1) \textit{ as } n \to \infty. \tag{59}$$

Then $S_{nk_n} \xrightarrow{d} Z$ (stably) where the r.v. Z has characteristic function $E(\exp(-\frac{1}{2}\eta^2 t^2))$.

Proof. Define $r(x)$ by

$$e^{ix} = (1 + ix)\exp(-\frac{1}{2}x^2 + r(x))$$

and note that $|r(x)| \leqslant |x|^3$ for $|x| \leqslant 1$. Let $I_n = \exp(itS_{nk_n})$ and

$$W_n = \exp(-\frac{1}{2}t^2 \sum_i X_{ni}^2 + \sum_i r(tX_{ni})).$$

Then

$$I_n = T_n \exp(-\eta^2 t^2/2) + T_n(W_n - \exp(-\eta^2 t^2/2)).$$

In view of Thm(33) it suffices to prove that for all $E \in \mathcal{F}$,

$$E(I_n I(E)) \to E(\exp(-\eta^2 t^2/2) I(E)). \tag{60}$$

Since $\exp(-\eta^2 t^2/2)I(E)$ is bounded, (3.14) ensures that

$$E(T_n \exp(-\eta^2 t^2/2) I(E)) \to E(\exp(-\eta^2 t^2/2) I(E)). \tag{61}$$

Moreover, any sequence of r.v. which converges weakly in L^1 is uniformly integrable, and so the sequence

$$T_n(W_n - \exp(-\eta^2 t^2/2)) = I_n - T_n \exp(-\eta^2 t^2/2)$$

is uniformly integrable. (The uniform integrability of I_n follows from the fact that $|I_n| = 1$.) Conditions (57) and (58) imply that when $\max_i |X_{ni}| \leqslant 1$,

$$\begin{aligned}|\sum_i r(X_{ni}t)| &\leqslant |t|^3 \sum_i |X_{ni}|^3 \\ &\leqslant |t|^3 (\max_i |X_{ni}|)(\sum_i X_{ni}^2) \xrightarrow{p} 0 \quad \text{ as } n \to \infty.\end{aligned}$$

It follows that $W_n - \exp(-\eta^2 t^2/2) \xrightarrow{p} 0$, and in view of the uniform integrability,

$$E[T_n(W_n - \exp(-\eta^2 t^2/2))I(E)] \to 0. \tag{62}$$

Conditions (61) and (62) imply (60). □

Theorem 34. *Let $\{S_{ni}, \mathcal{F}_{ni}, 1 \leqslant i \leqslant k_n, n \geqslant 1\}$ be a zero-mean, squareintegrable martingale array with differences X_{ni}, and let η^2 be an a.s. finite r.v. Suppose that*

$$\max_i |X_{ni}| \xrightarrow{\mathrm{p}} 0, \tag{63}$$

$$\sum_i X_{ni}^2 \xrightarrow{\mathrm{p}} \eta^2, \tag{64}$$

$$E\left(\max_i X_{ni}^2\right) \textit{ is bounded in } n, \tag{65}$$

and

$$\textit{the } \sigma\textit{-fields are nested } : \mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i} \textit{ for } 1 \leqslant i \leqslant k_n, n \geqslant 1. \tag{66}$$

Then $S_{nk_n} = \sum_i X_{ni} \xrightarrow{d} Z$ (stably), where the r.v. Z has characteristic function $E\exp(-\frac{1}{2}\eta^2 t^2)$.

Proof. Suppose first that η^2 is a.s. bounded, so that for some $C(> 1)$,

$$P(\eta^2 < C) = 1. \tag{67}$$

Let

$$X'_{ni} = X_{ni} I(\sum_{j=1}^{i-1} X_{nj}^2 \leqslant 2C) \text{ and } S'_{ni} = \sum_{j=1}^{i} X'_{nj}.$$

Then $\{S'_{ni}, \mathcal{F}_{ni}\}$ is a martingale array. Since

$$P(X'_{ni} \neq X_{ni} \text{ for some } i \leqslant k_n) \leqslant P(U_{nk_n}^2 > 2C) \to 0 \tag{68}$$

we have $P(S'_{nk_n} \neq S_{nk_n}) \to 0$, and so

$$E|\exp(itS'_{nk_n}) - \exp(itS_{nk_n})| \to 0.$$

Hence $S_{nk_n} \xrightarrow{\mathrm{d}} Z$ (stably) if and only if $S'_{nk_n} \xrightarrow{d} Z$ (stably). In view of (68), the martingale differences $\{X'_{ni}\}$ satisfy conditions (57) and (58) of Lemma(18); we must check (59). Let

$$T'_n = \prod_j (1 + itX'_{nj}),$$

and

$$J_n = \begin{cases} \min\{i \leqslant k_n \mid U^2_{ni} > 2C\} & \text{if } U^2_{nk_n} > 2C \\ k_n & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned} E|T'_n|^2 = E\prod_j (1 + t^2 X'^2_{nj}) &\leqslant E[\{\exp(t^2 \sum_{j=1}^{J_n - 1} X'^2_{nj})\}(1 + t^2 X^2_{nJ_n})] \\ &\leqslant \{\exp(2Ct^2)\}(1 + t^2 E X^2_{nJ_n}), \end{aligned}$$

which is bounded uniformly in n, by (65). Consequently $\{T'_n\}$ is uniformly integrable.

Let $m \geqslant 1$ be fixed and let $E \in \mathcal{F}_{mk_m}$; then by (66), $E \in \mathcal{F}_{nk_n}$ for all $n \geqslant m$. For such an n,

$$\begin{aligned} E[T'_n I(E)] &= E[I(E) \prod_{j=1}^{k_n} (1 + itX'_{nj})] \\ &= E[I(E) \prod_{j=1}^{k_m} (1 + itX'_{nj}) \prod_{j=k_m+1}^{k_n} E(1 + itX'_{nj} \mid \mathcal{F}_{n,j-1})] \\ &= E[I(E) \prod_{j=1}^{k_m} (1 + itX'_{nj})] = P(E) + R_n, \end{aligned}$$

where the remainder term R_n consists of at most $2^{k_m} - 1$ terms of the form

$$E[I(E)(it)^r X'_{n_{j_1}} X'_{nj_2} \cdots X'_{nj_r}],$$

where $1 \leqslant r \leqslant k_m$ and $1 \leqslant j_1 \leqslant j_2 \leqslant \cdots \leqslant j_r \leqslant k_m$.

Since

$$\begin{aligned} |X'_{nj_1} \cdots X'_{njj_r}|^2 &\leqslant (\sum_{j=1}^{J_n - 1} X'^2_{nj})^{r-1} (\max_i X'^2_{ni}) \\ &\leqslant (2C)^{r-1} (\max_i X'^2_{ni}), \end{aligned}$$

it follows that

$$|R_n| \leqslant (2^{k_m}-1)(2C)^{k_m/2}E(\max_i |X'_{ni}|).$$

But, for any $\varepsilon > 0$,

$$\begin{aligned} E(\max_i \mid X_{ni})). &\leqslant \varepsilon + E[\max_i |X_{ni}| I |X_{ni}| > \varepsilon)] \\ &= \varepsilon + E[(\max_i |X_{ni}|) I(\max_i |X_{ni}| > \varepsilon)] \\ &\leqslant \varepsilon + [E(\max_i X_{ni}^2) P(\max_i |X_{ni}| > \varepsilon)]^{1/2} \\ &\to \varepsilon \quad \text{ as } n \to \infty. \end{aligned}$$

It follows that $E(\max_i |X_{ni}|) \to 0$ and so $R_n \to 0$. Therefore

$$E[T'_n I(E)] \to P(E). \tag{69}$$

Let $\mathcal{F}_\infty = \bigvee_1^\infty \mathcal{F}_{nk_n}$ be the σ-field generated by $\bigcup_1^\infty \mathcal{F}_n$. For any $E' \in \mathcal{F}_\infty$ and any $\varepsilon > 0$ there exists an m and an $E \in \mathcal{F}_{mk_m}$ such that $P(E \Delta E') < \varepsilon$ (Δ denotes symmetric difference). Since $\{T'_n\}$ is uniformly integrable and

$$|E[T'_n I(E')] - E[T'_n I(E)]| \leqslant E[|T'_n| I(E \Delta E')],$$

$\sup_n |E[T'_n I(E')] - E[T'_n I(E)]|$ can be made arbitrarily small by choosing ε sufficiently small.

It now follows from (69) that for any $E' \in \mathcal{F}_\infty$,

$$E[T'_n I(E')] \to P(E').$$

This in turn implies that for any bounded $\mathcal{F}_\infty$-measurable r.v. $X, E[T'_n X] \to E(X)$. Finally, if $E \in \mathcal{F}$, then

$$E[T'_n I(E)] = E[T'_n E(I(E) \mid \mathcal{F}_\infty)] \to E[E(I(E) \mid \mathcal{F}_\infty)] = P(E).$$

This establishes (59) and completes the proof in the special case where (67) holds.

It remains only to remove the boundedness condition (67). If η^2 is not a.s. bounded, then given $\varepsilon > 0$, choose a continuity point C of η^2 such that $P(\eta^2 > .$ C$) > \varepsilon$. Let

$$\eta_C^2 = \eta^2 I(\eta^2 \leqslant C) + CI(\eta^2 > C),$$

$$X''_{ni} = X_{ni} I(\sum_{j=1}^{i-1} X_{nj}^2 \leqslant C) \quad \text{ and } \quad S''_{ni} = \sum_{j=1}^{i} X''_{nj}.$$

Then $\{S''_{ni}, \mathcal{F}_n\}$ is a martingale array and conditions (63),(65), and (66) are satisfied. Now,

$$\begin{aligned} &(\sum_i X_{ni}^2) I(\sum_i X_{ni}^2 \leqslant C) + CI(\sum_i X_{ni}^2 > C) \leqslant \sum_i X''^2_{ni} \\ &\leqslant (\sum_i X_{ni}^2) I(\sum_i X_{ni}^2 \leqslant C) + (C + \max_i X_{ni}^2) I(\sum_i X_{ni}^2 > C). \end{aligned}$$

Since C is a continuity point of the distribution of η^2,

$$I(\sum_i X_{ni}^2 \leqslant C) \xrightarrow{\text{p}} I(\eta^2 \leqslant C),$$

and so

$$\sum_i X_{ni}''^2 \xrightarrow{p} \eta_C^2.$$

As η_C^2 is a.s. bounded, the first part of the proof tells us that $S_n'' \xrightarrow{d} Z_C$ (stably), where the r.v. Z_C has characteristic function $E\exp(-\frac{1}{2}\eta_C^2 t^2)$. If $E \in \mathcal{F}$, then

$$\begin{aligned}
&|E[I(E)\exp(itS_{nk_n})] - E[I(E)\exp(-\frac{1}{2}\eta^2 t^2)]| \\
&\leqslant E|\exp(itS_{nk_n}) - \exp(itS''_{nk_n})| + |E[I(E)\exp(itS''_{nk_n})] - E[I(E)\exp(-\frac{1}{2}\eta_C^2 t^2)]| \\
&+ E|\exp(-\frac{1}{2}\eta_C^2 t^2) - \exp(-\frac{1}{2}\eta^2 t^2)|.
\end{aligned}$$

Since $S''_{nk_n} \to Z_C$ (stably), the second term on the right-hand side converges to zero as $n \to \infty$. The first and third terms are each less than 2ε in the limit since

$$\begin{aligned}
P({S_{nk_n}}' \neq S''_{nk_n}) &\leqslant P(X''_{ni} \neq X_{ni} \text{ for some } i) \\
&\leqslant P(U_{nk_n}^2 > C) \to P(\eta^2 > C) < \varepsilon.
\end{aligned}$$

Hence for all $\varepsilon > 0$,

$$\limsup_{n\to\infty} |E[I(E)\exp(itS_{nk_n})] - E[I(E)\exp(-\frac{1}{2}\eta^2 t^2)]| \leqslant 4\varepsilon.$$

It follows that the limit equals zero, and in view of Thm(33) this completes the proof.

□

9 Acknowledgement

I would like to thank both my supervisors, Arup Sir and Krishnau Sir, for their valuable time and guidance. I am grateful for the discussions which we had during project meetings. It was a wonderful opportunity to learn from them.

References

[1] Rick Durret; Probability: Theory and Examples, ISBN: 9781108473682, 1108473687.

[2] Robert Ash, Catherine Doleans-Dade; Probability and Measure Theory. ISBN: 978-0120652020.

[3] Patrick Billingsley; Probability and Measure, ISBN: 978-1118122372.

[4] William Feller; An Introduction to Probability Theory and Its Applications, Vol 1,ISBN: 978-0471257097.

[5] Freedman, D.A.; (1965). Bernard Friedman's urn. *Ann. Math. Statist.* **36** 956–970.

[6] Gouet, R.; (1997). Strong convergence of proportions in a multicolor P ´olya urn. *J. Appl. Probab.* **34** 426–435.

[7] A. Bose, A. Dasgupta, K. Maulik; (2009) Multicolor urn models with reducible replacement matrices. *ISI/BS* 279-295.

[8] Hall, P. and Heyde, C.C. (1980). Martingale Limit Theory and Its Application. *New York: Academic Press Inc.*

[9] William Feller; An Introduction to Probability Theory and Its Applications, Vol 2,ISBN: 9780471257097.

[10] Janson, S. ;(2006) Limit theorems for traingular urn schemes. *Probab. Theory Related Fields* **134** 417-452. MR2226887

10 Code

This is the code used to draw all simulations in this report.

```
import random #for generating random number
import math
import matplotlib.pyplot as plt
import numpy as np
from scipy.stats import beta

#for picking colours during runtime
from matplotlib import cm
from matplotlib.colors import ListedColormap, LinearSegmentedColormap

hsv = cm.get_cmap('hsv', 8)

from scipy.stats import gaussian_kde
from scipy.stats import norm

def calcProportion(X):#Normalizes a given vector
  sum = 0
  C = []
  for i in range(len(X)):
      sum+=X[i]
  for i in range(len(X)):
```

```
        C.append(X[i]/sum)
    return C
def drawBall(X):#draws a ball at random from urn with composition X
    sum = 0
    sumList = []
    for i in range(len(X)):
        sum += X[i]
        sumList.append(sum)
    #print(sum, sumList)

    r = random.uniform(0,sum)
    #print(r)
    for i in range(len(X)):
        if r<=sumList[i]:
            return i
def runExperiment(d,R,X,N):#runs the experiment till time N and returns ordere
    Y = list(X)
    drawList = []
    for n in range(N):
        draw = drawBall(Y)
        for i in range(d):
            Y[i] = Y[i]+R[draw][i]
        drawList.append(draw)
    return(drawList)
def sampleUrn(d,R,X,N,M):#makes a sample using runExperiment
    list = []
    for n in range(M):
        list.append(runExperiment(d,R,X,N))
    return(list)

def getCountN(d,draw):#gets count vector at time n
    counter = [0]*d
    for i in draw:
        counter[i]+=1
    return(counter)
def getXN(d,R,X,draw):#gets the vector of no. of balls at time n
    Y = list(X)
    for i in draw:
        for j in range(d):
            Y[j] = Y[j]+R[i][j]
    return(Y)

def getCountList(d,draw):#returns a list of count vector of colour drawn till
    counterList = []
    counter = [0]*d
    for i in draw:
```

```
        counter[i]+=1
        counterList.append(list(counter))
    return(counterList)
def getConcList(d,R,X,draw):#same as getConcList but via a different approach
    Y = list(X)
    concList = [calcProportion(Y)]
    for i in draw:
        for j in range(d):
            Y[j] += R[i][j]
        concList.append(calcProportion(Y))
    return(concList)
def getConcList1(d,R,X,draw):#returns a list of concentration vector wrt time
    concList = []
    initial = sum(X)
    #print(initial)
    rowSum = [sum(R[i]) for i in range(d)]
    #print(rowSum)
    counter = [0]*d
    conc = [0]*d
    for i in draw:
        counter[i] += 1
        #print(counter)
        sum1 = initial + sum([(counter[k]*rowSum[k]) for k in range(d)])
        #print(sum1)
        for j in range(d):
            sum2 = X[j] + sum([(counter[k]*R[k][j]) for k in range(d)])
            conc[j] = sum2/sum1
        concList.append(list(conc))
    return(concList)
def getDiffList(d,R,X,draw):#returns difference between white and black balls
    Y = list(X)
    diffList = [Y[0]-Y[1]]
    for i in draw:
        for j in range(d):
            Y[j] += R[i][j]
        diffList.append(Y[0]-Y[1])
    return(diffList)

def getConcSamp(d,R,X,sample):
    concSamp = []
    for i in range(len(sample)):
        concSamp.append(getConcList(d,R,X,sample[i]))
    return(concSamp)

def plotList(list):#plots any given list against time
    arr = np.array(list)
```

```
    plt.plot(arr)
    #plt.show()
def plotSample(sample):#plots sample with consistent colouring between samples
    d = len(sample[0][0])
    list = np.array(sample)
    for j in range(len(sample)):
        for i in range(d):
            ypoints = np.array(list[j,:,i])
            plt.plot(ypoints,label=("C"+str(i)), color=hsv(i/d))
    plt.xlabel("Draw")
    plt.ylabel("Proportion")
    #plt.show()

#KDE estimates for rate of convergence of difference
def sampleFriedUrn(alpha,beta,X,N,M):
    return(sampleUrn(2,[[alpha,beta],[beta,alpha]],X,N,M))
def friedman_diff(alpha,beta,X,sample):
    diffList = []
    R = [[alpha,beta],[beta,alpha]]
    for i in range(len(sample)):
        Y = getXN(2,R,X,sample[i])
        diffList.append(Y[0]-Y[1])
    return(diffList)
def rhoPlot1(alpha,beta,X,N,M):
    sample = sampleFriedUrn(alpha,beta,X,N,M)
    rho = (alpha-beta)/(alpha+beta)
    for j in range(M):
        diffList = getDiffList(2,[[alpha,beta],[beta,alpha]],X,sample[j])
        for i in range(1,N+1):
            diffList[i] /= (i)**rho
        plotList(diffList)
    plt.show()

def rhoCase1(alpha,beta,X,N,M):
    sample = sampleFriedUrn(alpha,beta,X,N,M)
    rho = (alpha-beta)/(alpha+beta)
    scale = N**rho
    diffList = friedman_diff(alpha,beta,X,sample)
    for i in range(M):
        diffList[i] /= scale

    x_grid = np.linspace(-25, 25, 1000)
    list = np.array(diffList)
    pdf = kde_func(list,x_grid)
    plt.plot(x_grid,pdf)
    plt.show()
```

```
def rhoCase2(alpha, beta, X, N, M):
    sample = sampleFriedUrn(alpha, beta, X, N, M)
    scale = (N*math.log(N))**(1/2)
    diffList = friedman_diff(alpha, beta, X, sample)
    for i in range(M):
        diffList[i] /= scale

    rv = norm(scale=(alpha-beta))
    x_grid = np.linspace(rv.ppf(0.01), rv.ppf(0.99), 1000)
    list = np.array(diffList)
    pdf = kde_func(list, x_grid)
    plt.plot(x_grid, pdf)
    plt.fill(x_grid, rv.pdf(x_grid), ec='gray', fc='gray', alpha=0.4)
    plt.show()
def rhoCase3(alpha, beta, X, N, M):
    sample = sampleFriedUrn(alpha, beta, X, N, M)
    rho = (alpha-beta)/(alpha+beta)
    scale = N**(1/2)
    diffList = friedman_diff(alpha, beta, X, sample)
    for i in range(M):
        diffList[i] /= scale
    stdDev = (abs(alpha-beta))/((1-2*rho)**(1/2))
    print(stdDev)
    rv = norm(scale=stdDev)
    x_grid = np.linspace(rv.ppf(0.01), rv.ppf(0.99), 100)
    list = np.array(diffList)
    pdf = kde_func(list, x_grid)
    plt.plot(x_grid, pdf)
    plt.fill(x_grid, rv.pdf(x_grid), ec='gray', fc='gray', alpha=0.4)
    plt.show()

def kde_func(list, x_grid):
    kde = gaussian_kde(list)
    return kde.evaluate(x_grid)

def beta_kde(alpha, X, N, M):
    sample = sampleUrn(2, [[alpha, 0], [0, alpha]], X, N, M)
    list = []
    for draw in sample:
        Y = getXN(2, [[alpha, 0], [0, alpha]], X, draw)
        list.append(Y[0]/(Y[0]+Y[1]))

    a = X[0]/alpha
    b = X[1]/alpha

    x_grid = np.linspace(beta.ppf(0.01, a, b), beta.ppf(0.99, a, b), 100)
```

```
    kList = np.array(list)
    pdf = kde_func(kList, x_grid)
    plt.plot(x_grid, pdf)
    plt.fill(x_grid, beta.pdf(x_grid, a, b), ec='gray', fc='gray', alpha=0.4)
    plt.show()


#rhoPlot1(8.5,1.5,[10,10],50000,10)
#rhoCase1(8.5,1.5,[5,10],50000,200)
#rhoCase2(3,1,[5,10],10000,1000)
#rhoCase3(7.45,2.55,[10,10],10000,1000)
#beta_kde(2,[4,10],2000,10000)
```