Journal of Theoretical Probability, Vol. 3, No. 4, 1990

A Time-Dependent Version of Pólya's Urn

Robin Pemantle[1]

Received May 31, 1989

A process is defined that consists of drawing balls from an urn and replacing them along with extra balls. The process generalizes the well-known *Pólya urn process*. It is shown that the proportion of red balls in the urn converges to a random limit that may have a nonzero probability of being 0 or 1, but is nonatomic elsewhere.

KEY WORDS: Pólya's urn process; nonatomic.

The Pólya urn process was introduced in Eggenberger and Pólya (1923). To run this process, let an urn contain R red balls and B black balls at time $n=1$. A ball is drawn at random from the urn and replaced in the urn along with another ball of the same color, so that at time $n=2$ there are $R+B+1$ balls; $R+1$ of them are red with $R/(R+B)$ and R of them are red otherwise. The draw and replacement are repeated ad infinitum, with the probability of drawing a red ball always equal to the proportion of balls in the urn that are red at that time. It is a well-known fact that the proportions of red balls converge almost surely to a limit that is random and has beta distribution with parameters R and B; see Feller[(4)] for a discussion of this.

This article considers a Pólya urn with the single change that the number of extra balls added of the color drawn is a function of time. A possible interpretation of this model is the American presidential primary election. Assume an initial amount of popular support for each candidate that dictates that candidate's chance of winning the first primary and then assume that the support increases proportionally to the size of the states won by the candidate in each primary. Then Proposition 6 below describes how the

[1] Department of Mathematics, White Hall, Cornell University, Ithaca, New York 14853.

0894-9840/90/1000-0627$06.00/0

order in which the primaries are held influences each candidate's chances and the final margin of victory. Many other variants of the Pólya urn have been studied: Friedman[5] allows balls of both colors to be added each time; Athreya[2] allows the number of balls added each time to be i.i.d. random variables; Hill, Lane, and Sudderth[6] allow the probability of drawing a red ball to vary nonlinearly with the proportion of red balls in the urn; and Arthur, Ermol'ev, and Kaniovskii[1] generalize Hill, Lane, and Sudderth's scheme to allow balls of finitely many different colors. Other discrete time random processes such as reinforced random walk can be embedded in urn models (see Pemantle[7]).

Let $F: \mathbf{Z}^{\geq 0} \to (0, \infty)$ be any function. Let $\{v_1, v_2, ...\}$ be the successive proportions of red balls in an urn that begins with R red balls and B black balls and evolves as follows: at discrete times $n = 1, 2, ...$, a ball is drawn and replaced in the urn along with $F(n)$ balls of the same color. (Allow F to take nonintegral values by defining the probability of drawing a red ball still to be proportional to the total mass of red balls in the urn.) The usual Pólya urn scheme is the case where $F(n) = 1$ for all n. It is shown that v_n must converge for any F and that the limit has no atoms except possibly at 0 and 1. Necessary and sufficient conditions for the limit to concentrate entirely on the two point set $\{0, 1\}$ are given. The proofs of Theorems 1-3 are based on variance calculations for a discrete-time Martingale. The methods are thus fairly elementary although the approach used to prove Theorem 3 is useful in a much more general stochastic approximation setting[8] and I have not seen it used before.

The following formal definition of the process is completely routine and may be skipped. Let Ω be $[0, 1]^{\mathbf{Z}^{\geq 0}}$ with the product uniform measure. All probabilities will be with respect to this space and all functions will be functions of ω where ω is a generic point in Ω, but the notation will supress the role of ω when no ambiguity arises. Let z_n be the nth coordinate of ω so that $\{z_n: n = 0, 1, 2, ...\}$ is a set of independent uniformly distributed variables on $[0, 1]$, and let $\mathcal{F}_n$ be the σ-algebra generated by $\{z_i: i \leq n\}$. Let $S_1(0) = R$, $S_2(0) = B$, and recursively define

$$\begin{aligned} S_1(n+1) &= S_1(n) + F(n)\, \mathbf{1}(z_n \leq S_1(n)/[S_1(n) + S_2(n)]) \\ S_2(n+1) &= S_2(n) + F(n)\, \mathbf{1}(z_n > S_1(n)/[S_1(n) + S_2(n)]) \end{aligned} \tag{1}$$

where $\mathbf{1}$ denotes the indicator function of a set. So $S_1(n)$ and $S_2(n)$ represent the numbers of red and black balls in the urn after n draws. For convenience we let

$$\delta_n = F(n) \Big/ \left[R + B + \sum_{i=0}^{n-1} F(i) \right]$$

denote the fractional additions. Let $v_n = S_1(n)/[S_1(n)+S_2(n)]$ denote the proportion of red balls at time n. The following results will be proved.

Theorem 1. For any function F as above, the random variables v_n converge almost surely to some random variable v.

Theorem 2. The limit v satisfies $\text{prob}(v=0) = 1 - \text{prob}(v=1) = B/(R+B)$ if and only if $\sum_{i=1}^{\infty} \delta_n^2 = \infty$.

This theorem applies, for example, when $F(n) = 2^n$. Roughly, the hypothesis means that F grows faster than polynomially, but one needs to look more closely if the growth is irregular since the function

$$F(n) = \begin{cases} n & \text{if } \quad n \text{ is a power of 2} \\ 2^{\cdot\cdot n} & \text{otherwise} \end{cases}$$

satisfies the hypothesis.

Theorem 3. The distribution of v has no atoms on $(0, 1)$.

Theorem 4. If F is bounded by some constant M, then the distribution of v has no atoms at all, i.e., $\text{prob}(v=0) = \text{prob}(v=1) = 0$.

Remark. It is possible for the distribution of v to have atoms at 0 and 1 of total weight less than 1; then the remainder of the time v is in $(0, 1)$ and this part of the distribution is nonatomic. An example where this occurs is where $R = B = 1$ and $F(n) = n$. In this case the probability that all draws are of the same color is $\frac{2}{3} \times \frac{6}{7} \times \frac{15}{16} \times \cdots > 0$, but according to Theorem 2 the distribution is not entirely concentrated on $\{0, 1\}$. I do not know an interpolation between Theorem 2 and Theorem 4 giving a necessary condition for the probability that $v \in \{0, 1\}$ to be nonzero.

Proof of Theorem 1. $\{v_n : n = 1, 2, \ldots\}$ is a Martingale. To see this, calculate

$$\begin{aligned} \mathbf{E}(v_{n+1} \mid \mathscr{F}_n) &= v_n(S_1(n) + F(n))/[S_1(n) + F(n) + S_2(n)] \\ &\quad + (1 - v_n)\, S_1(n)/[S_1(n) + S_2(n) + F(n)] \\ &= [S_1(n) + v_n F(n)]/[S_1(n) + S_2(n) + F(n)] \\ &= v_n \end{aligned}$$

Now since $\{v_n\}$ is bounded, it converges almost surely to some v. □

Proof of Theorem 2. We calculate the expected value of v^2. Since $0 \leqslant v \leqslant 1$ and $\mathbf{E}(v) = R/(R+B)$, this is at most $R/(R+B)$ with equality if and only if $\text{prob}(v=0) = 1 - \text{prob}(v=1) = B/(R+B)$. Necessary and sufficient conditions for this will follow from the simple recurrence relation (2) below for the values of $R/(R+B) - \mathbf{E}(v_n^2)$, which are denoted W_n.

Since v_n converges almost surely to v and the variables are bounded by 1, we know that $\mathbf{E}(v_n^2)$ converges to $\mathbf{E}(v^2)$. Let V_n denote $\mathbf{E}(v_n^2)$. For a fixed F, v_n takes on only finitely many values and V_{n+1} can be recursively calculated as follows. If $v_n(\omega) = x = S_1(n)/[S_1(n) + S_2(n)]$ then

$$\begin{aligned} v_{n+1}(\omega) &= S_1(n)/[S_1(n) + S_2(n) + F(n)] \\ &= x/(1+\delta_n) \end{aligned}$$

with probability $1-x$, and

$$\begin{aligned} v_{n+1}(\omega) &= [S_1(n) + F(n)]/[S_1(n) + S_2(n) + F(n)] \\ &= (x+\delta_n)/(1+\delta_n) \end{aligned}$$

with probability x. So

$$\begin{aligned} V_{n+1} &= \mathbf{E}v_{n+1}^2 \\ &= \mathbf{E}[(1-v_n)\, v_n^2/(1+\delta_n)^2 + v_n(v_n+\delta_n)^2/(1+\delta_n)^2] \\ &= [1/(1+\delta_n)^2]\, \mathbf{E}[(1-v_n)\, v_n^2 + v_n(v_n+\delta_n)^2] \\ &= [1/(1+\delta_n)^2]\, \mathbf{E}[v_n^2 + 2\delta_n v_n^2 + v_n \delta_n^2] \\ &= \mathbf{E}(v_n^2) + [\delta_n/(1+\delta_n)]^2\, (\mathbf{E}v_n - \mathbf{E}v_n^2) \\ &= V_n + [\delta_n/(1+\delta_n)]^2\, [R/(R+B) - V_n] \end{aligned}$$

To see better how the value of V_{n+1} relates to the value of V_n, we let W_k denote $R/(R+B) - V_k$. Then

$$W_{n+1} = W_n[1 - \delta_n^2/(1+\delta_n)^2] \tag{2}$$

Thus the value V_n converges to $R/(R+B)$ if and only if W_n converges to 0, which happens whenever the product of the values $[1 - \delta_n^2/(1+\delta_n)^2]$ converges to 0. This happens whenever $\sum_{n=1}^{\infty} \delta_n^2/(1+\delta_n)^2$ diverges, which in turn happens exactly when $\sum_{n=1}^{\infty} \delta_n^2$ diverges, and Theorem 2 is proved. □

Proof of Theorem 3. Fix $p \in (0, 1)$. If $\text{prob}(v_n \to p) > 0$ then there is some n and some event $\mathscr{A} \in \mathscr{F}_n$ such that $\text{prob}(v_n \to p \mid \mathscr{A})$ is arbitrarily close to 1. In fact, n can be taken to be as large as desired. Define

$$\alpha_n = \sum_{i=n}^{\infty} \delta_i^2$$

The quantity α_n can be thought of as the "remaining variance," since the expected square increments of the Martingale $\{v_i\}$ are bounded between constant multiples of δ_i^2 when v_i is near p. According to Theorem 2 there is no loss of generality in assuming α_n to be finite. Also assume without loss of generality that $p \leqslant 1/2$ since the case $p > 1/2$ is identical but with red balls and black balls interchanged.

Since $\alpha_n \to 0$ there is an N for which $n \geqslant N$ implies $\alpha_n < p/10$. Choose $c > 0$ small enough so that

$$9c^2 \leqslant 81p^2/800 \tag{3}$$

The essence of the proof is in the following two claims, holding for any $n > N$.

$$\text{Claim 1: prob}(\sup_{k \geqslant n} |v_k - p| > c\sqrt{\alpha_n} \mid \mathscr{F}_n) \geqslant 9p/10 \tag{4}$$

$$\text{Claim 2: prob}(\inf_{k \geqslant n} |v_k - p| \geqslant c\sqrt{\alpha_n}/2 \mid \mathscr{F}_n, \mathscr{B}) \geqslant c^2/16$$

$$\text{where} \quad \mathscr{B} \text{ is the event } |v_n - p| \geqslant c\sqrt{\alpha_n} \tag{5}$$

Putting these two claims together, we see that for any value of $n > N$, the probability given $\mathscr{F}_n$ is at least $9p/10 \cdot c^2/16$ that some v_{n+k} will be at least $c\sqrt{\alpha_n}$ away from p and that no subsequent v_{n+k+l} will ever return to the interval $[p - c\sqrt{\alpha_n}/2, p + c\sqrt{\alpha_n}/2]$. This contradicts the existence of the event $\mathscr{A}$ above, and the theorem follows.

Proof of Claim 1. Let $\tau = \inf\{k \geqslant n: |v_k - p| > c\sqrt{\alpha_n}\}$. We need to show that $\text{prob}(\tau < \infty \mid \mathscr{F}_n) \geqslant 9p/10$. We will calculate the variance of $v_{t \wedge \tau}$. On the one hand, this is limited by the fact that $v_{t \wedge \tau}$ is never very far from p. (If the increment on which the stopping time is reached may be very large, then a different argument is used; see Case 1 below.) On the other hand, the variance always grows by at least a constant multiple of δ_i^2 until τ is reached, and c is chosen to be much smaller than this constant. These two facts together will imply that the stopping time is reached often enough for Claim 1 to be true.

Case 1. $\delta_i > 2c\sqrt{\alpha_n}/(1 - p - c\sqrt{\alpha_n})$ for some $i \geqslant n$. Basically what happens in this case is that there is a good enough chance of stopping on the $i+1$st draw:

$$\begin{aligned}
&\tau > i \text{ and draw } i \text{ is red} \\
&\Rightarrow v_i \geqslant p - c\sqrt{\alpha_n} \text{ and draw } i \text{ is red} \\
&\Rightarrow v_{i+1} = v_i + (1 - v_i)\,\delta_i/(1 + \delta_i) \geqslant p + c\sqrt{\alpha_n} \\
&\Rightarrow \tau = i + 1
\end{aligned}$$

The probability of a red draw is always at least $p - c\sqrt{\alpha_i}$ until τ is reached, which is at least $9p/10$ by choice of c and N. This easily implies that $\text{prob}(\tau < \infty \mid \mathscr{F}_n) \geq 9p/10$.

Case 2. No δ_i is that big. Then the increment on which τ is reached cannot be bigger than $2c\sqrt{\alpha_n}$. Pick any $i \geq n$ and use the fact that $v_{i \wedge \tau} - p$ is a Martingale to get

$$\begin{aligned} &\mathbf{E}((v_{(i+1)\wedge\tau} - p)^2 \mid \mathscr{F}_n) \\ &\quad = \mathbf{E}((v_{i\wedge\tau} - p)^2 \mid \mathscr{F}_n) + \mathbf{E}(\mathbf{1}_{\tau > i}(v_{i+1} - v_i)^2 \mid \mathscr{F}_n) \end{aligned} \tag{7}$$

But

$$(v_{i+1} - v_i)^2 = \begin{cases} v_i^2[\delta_i/(1+\delta_i)]^2 & \text{with probability } 1 - v_i \\ (1-v_i)^2\,[\delta_i/(1+\delta_i)]^2 & \text{with probability } v_i \end{cases} \tag{8}$$

Now since $1 + \delta_i < 2$ by the assumption that $\alpha_n < p/10$ and since $\tau > i \Rightarrow \min\{v_i, 1 - v_i\} \geq p - c\sqrt{\alpha_n} \geq 9p/10$, it follows that

$$(v_{i+1} - v_i)^2 \geq 81p^2\delta_i^2\mathbf{1}_{\tau > i}/400$$

So the right-hand side of Eq. (7) is at least

$$\mathbf{E}((v_{i\wedge\tau} - p)^2 \mid \mathscr{F}_n) + 81\delta_i^2 \,\text{prob}(\tau = \infty \mid \mathscr{F}_n)/400$$

Now summing over i and dropping the positive term $(v_n - p)^2$ gives

$$\mathbf{E}((v_{(n+M)\wedge\tau} - p)^2 \mid \mathscr{F}_n) \geq \left(81p^2 \sum_{i=n}^{n+M-1} \delta_i^2/400\right) \text{prob}(\tau = \infty \mid \mathscr{F}_n)$$

But Eq. (6) implies that

$$\mathbf{E}((v_{(n+M)\wedge\tau} - p)^2 \mid \mathscr{F}_n) \leq 9c^2\alpha_n = 9c^2 \sum_{i=n}^{\infty} \delta_i^2$$

so letting $M \to \infty$ gives

$$\text{prob}(\tau = \infty \mid \mathscr{F}_n) \leq 9c^2/(81p^2/400) \leq 1/2$$

by the choice of c in (3) above. So Claim 1 is proved. □

Proof of Claim 2. The idea this time is that the remaining variance is not enough to give a high probability of getting back to within $c\sqrt{\alpha_n}/2$ of p. The inequality we use is a one-sided Tschebysheff inequality saying that if v_n has a probability greater than $1 - \varepsilon$ of reentering the interval, then since it is a Martingale, the other ε of the time its average is on the order of ε^{-1} in the other direction, and this gives a contribution to the variance that gets impossibly large as ε goes to 0.

Let $\mathscr{B}$ be the event $|v_n - p| \geqslant c\sqrt{\alpha_n}$ as in (5) above. Define a new stopping time by $\tau = \inf\{k \geqslant n;\ |v_k - p| \leqslant c\sqrt{\alpha_n}/2\}$. From (7) again, calculate

$$\mathrm{Var}(v_{(n+M)\wedge\tau} \mid \mathscr{F}_n) = \sum_{i=n}^{n+M-1} \mathbf{E}(\mathbf{1}_{\tau > i}(v_{i+1} - v_i)^2 \mid \mathscr{F}_n) \leqslant \sum_{i=n}^{\infty} \delta_i^2 \tag{9}$$

according to the values for $(v_{i+1} - v_i)^2$ given in (8). So $\{v_{(n+i)\wedge\tau}\}$ is an L^2-bounded Martingale with variance $\mathbf{E}((v_\tau - v_n)^2 \mid \mathscr{F}_n)$ at most $\sum_{i=n}^{\infty} \delta_i^2 = \alpha_n$. On the other hand,

$$\begin{aligned}
&\mathbf{E}((v_\tau - v_n)^2 \mid \mathscr{F}_n, \mathscr{B}) \\
&\quad\geqslant \mathrm{prob}(\tau < \infty \mid \mathscr{F}_n, \mathscr{B})(c\sqrt{\alpha_n}/2)^2 \\
&\qquad + \mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B})\, \mathbf{E}((v_\infty - v_n)^2 \mid \mathscr{F}_n, \mathscr{B}, \tau = \infty) \\
&\quad\geqslant \mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B})\, \mathbf{E}(v_\infty - v_n \mid \mathscr{F}_n, \mathscr{B}, \tau = \infty)^2 \\
&\quad\geqslant \mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B}) \left[\frac{c\sqrt{\alpha_n}}{2} \frac{\mathrm{prob}(\tau < \infty \mid \mathscr{F}_n, \mathscr{B})}{\mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B})}\right]^2 \\
&\quad= \frac{c^2\alpha_n}{4} \frac{\mathrm{prob}(\tau < \infty \mid \mathscr{F}_n, \mathscr{B})^2}{\mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B})}
\end{aligned} \tag{10}$$

where the penultimate term is calculated from the fact that $|\mathbf{E}(v_\infty - v_n \mid \mathscr{F}_n, \mathscr{B}, \tau < \infty)| > c\sqrt{\alpha_n}/2$ while $\mathbf{E}(v_\infty - v_n)$ must be zero. Combining the two inequalities (9) and (10) gives

$$\alpha_n \geqslant (c^2/4)\,\alpha_n \frac{\mathrm{prob}(\tau < \infty \mid \mathscr{F}_n, \mathscr{B})^2}{\mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B})}$$

It follows easily from this that $\mathrm{prob}(\tau = \infty \mid \mathscr{F}_n, \mathscr{B}) \geqslant \min\{1/2, c^2/16\} = c^2/16$ and Claim 2 is proved, along with Theorem 3. □

Proof of Theorem 4. Suppose $\mathrm{prob}(v = 0) = p > 0$. For any $\varepsilon > 0$ let $h_\varepsilon\colon [0, 1] \to [0, 2]$ be defined by $h_\varepsilon(x) = \max\{0, 2 - 2x/\varepsilon\}$. Then $h_\varepsilon > \mathbf{1}_{[0,\varepsilon]}$, so

$$\liminf_{n\to\infty} \mathbf{E}h_\varepsilon(v_n) \geqslant \liminf_{n\to\infty} \mathrm{prob}(v_n \in [0, \varepsilon]) \geqslant p$$

Then to prove Theorem 4 it suffices to show that $\lim_{\varepsilon\downarrow 0} \mathbf{E}h_\varepsilon(v_n) = 0$

uniformly in n. To do this, define a partial order on distributions of $[0, 1]$-valued random variables by writing

$$\mu \preccurlyeq \nu \Leftrightarrow \int h\, d\mu \leqslant \int h\, d\nu \qquad \text{for all continuous convex } h\colon [0,1] \to [0,2]$$

Note that

$$\begin{aligned} &\mu_i \preccurlyeq \nu_i \text{ and } \mu_i \to \mu,\ \nu_i \to \nu \text{ weakly} \\ &\Rightarrow \mu \preccurlyeq \nu \end{aligned} \tag{11}$$

Another characterization of this partial order is that $\mu \preccurlyeq \nu$ if and only if there are random variables y and z with laws μ and ν, respectively, such that $\mathbf{E}(z \mid y) = y$. Also say $y \preccurlyeq z$ if y and z have laws μ and ν, respectively, with $\mu \preccurlyeq \nu$ (this does not mean that y and z need to be the y and z of the previous sentence, or even be defined on the same probability space). As a notational convention, for any function $F^{(i)}\colon \mathbf{Z}^{\geqslant 0} \to (0, \infty)$, let $v_n^{(i)}$ be the successive proportions of red balls in an urn that is governed by Eq. (1) but with $F^{(i)}$ in place of F. Let all other superscripted variables refer to such an urn as well. The initial distribution is always taken to be R red balls and B black balls. The proof of Theorem 4 rests on the following two propositions.

Proposition 1.

(i) For any n, if $v_n^{(1)} \succcurlyeq v_n^{(2)}$ and $\delta_n^{(1)} \geqslant \delta_n^{(2)}$ then $v_{n+1}^{(1)} \succcurlyeq v_{n+1}^{(2)}$.

(ii) If $\delta_n^{(1)} \geqslant \delta_n^{(2)}$ for all n, then $v_{n+1}^{(1)} \succcurlyeq v_{n+1}^{(2)}$.

Proposition 2. Suppose that for some k, $F^{(1)}(k) = F^{(2)}(k+1) > F^{(2)}(k) = F^{(1)}(k+1)$ and $F^{(2)}(n) = F^{(1)}(n)$ for all $n \neq k,\ k+1$. Then $v_n^{(1)} \succcurlyeq v_n^{(2)}$ for all n.

Assuming these for the moment, the proof of Theorem 4 is finished as follows. Fix N and let $F^{(1)}(n) = F(n)$ for $n \leqslant N$ and $F^{(1)}(n) = 0$ for $n > N$. Then $v_n^{(1)} = v_n$ for $n \leqslant N$. Let $F^{(2)}$ take on the same values as $F^{(1)}$ but rearranged in descending order and let $F^{(3)}(n) = M$ for all n, where M is a bound for the values of F. So, for example, if the values of F are 1, 4, 3, 2, 9,... and $N = 3$, then the values of $F^{(1)}$ are 1, 4, 3, 0, 0,..., the values of $F^{(2)}$ are 4, 3, 1, 0, 0,... and the values of $F^{(3)}$ are M, M, M, M, M,... for some $M \geqslant 9$. Now Propositions 5 and 6 and Eq. (11) give

$$v_N = v_N^{(1)} = v^{(1)} \preccurlyeq v^{(2)} \preccurlyeq v^{(3)}$$

But the distribution of $v^{(3)}$ is known to be beta with parameters R/M

and B/M.[4] So $\mathbf{E}h_\varepsilon(v_N) \leqslant \int h_\varepsilon \, d\beta(R/M, B/M)$ which goes to zero as $\varepsilon \to 0$ independently of N. It remains now to prove Propositions 1 and 2.

Proof of Proposition 1. Part (ii) follows inductively from part (i). To show (i) begin with the following definition. For any [0, 1]-valued random variable y let $T_\delta(y)$ be a random variable whose distribution is given by

$$\begin{aligned}\text{prob}(T_\delta(y) = y - [\delta/(1+\delta)]\, y \mid y) \\ = 1 - \text{prob}(T_\delta(y) = y + [\delta/(1+\delta)](1-y) \mid y) = 1 - y\end{aligned}$$

In other words, $T_\delta(y)$ is the proportion of red balls in an urn after a fractional addition of δ to an initial proportion y. By conditioning on y it is easy to see that

$$\delta \geqslant \eta \Rightarrow T_\delta(y) \succcurlyeq T_\eta(y) \tag{12}$$

Then to show (i) it suffices to show

$$y \succcurlyeq z \Rightarrow T_\delta(y) \succcurlyeq T_\delta(z) \tag{13}$$

So let $y \succcurlyeq z$ be random variables, with y' and z', respectively, dnoting $T_\delta(y)$ and $T_\delta(z)$. A few reductions are helpful. First, assume without loss of generality that $\mathbf{E}(y \mid z) = z$. In other words, since the law of $T_\delta(y)$ depends only on the law of y, we now do assume that y and z are defined on the same probability space with $\mathbf{E}(y \mid z) = z$. Second, assume that z is deterministic. This is possible because $T_\delta(y)$ and $T_\delta(z)$ can both be defined separately for each value of z, and then $\mathbf{E}(h(T_\delta(y))) \geqslant \mathbf{E}(h(T_\delta(z)))$ will follow by conditioning on z. Third, assume that y takes on only two values, since otherwise it is possible to interpolate random variables $y = y_n \succcurlyeq y_{n-1} \succcurlyeq \cdots \succcurlyeq y_1 \succcurlyeq y_0 = z$ so that $\mathbf{E}(y_i \mid y_{i-1}) = y_{i-1}$ and y_i takes on only two values for each value of y_{i-1}. (Recall that y takes on only finitely many values). Thus without loss of generality let $a, r > 0$ be such that

$$\text{prob}(y = z - ar) = 1 - \text{prob}(y = z + (1-a)r) = 1 - a$$

Then the laws of y' and z' are given by

$$\text{prob}(z' = Q = z - [\delta/(1+\delta)]\, z) = 1 - z$$

$$\text{prob}(z' = R = z + [\delta/(1+\delta)](1-z)) = z$$

$$\text{prob}(y' = A = z - [\delta/(1+\delta)]\, z - ar/(1+\delta)) = (1-a)(1-z+ar)$$

$$\text{prob}(y' = B = z - [\delta/(1+\delta)]\, z + (1-a)r/(1+\delta)) = a[1 - z - (1-a)r]$$

$$\text{prob}(y' = C = z + [\delta/(1+\delta)](1-z) - ar/(1+\delta)) = (1-a)(z - ar)$$

$$\text{prob}(y' = D = z + [\delta/(1+\delta)](1-z) + (1-a)r/(1+\delta)) = a[z + (1-a)r]$$

From this it is easy to see that a variable $y'' \stackrel{\mathscr{D}}{=} y'$ can be constructed so that $\mathbf{E}(y'' \mid z') = z'$. To do this, let $y'' = B$ only when $z' = Q$ and let $y'' = C$ only when $z' = R$. Since the ratio of prob$(y'' = A)$ to prob$(y'' = B)$ is at least $1 - a$ to a, the remaining values of y'' when $z' = Q$ can be assigned so that $\mathbf{E}(y'' \mid z' = Q) = Q$ without exceeding the value for prob$(y'' = A)$ required above. Then it follows that $\mathbf{E}(y'' \mid z' = R) = R$ and hence $y' \succcurlyeq z'$. □

Proof of Proposition 2. The two urns agree up to time k, so by conditioning on v_k, assume $k = 1$. It suffices to show $v_2^{(1)} \succcurlyeq v_2^{(2)}$, since $v_1^{(1)} \succcurlyeq v_1^{(2)}$ by (12), and it then follows from Proposition 5 that $v_2^{(1)} \succcurlyeq v_2^{(2)} \Rightarrow v_i^{(1)} \succcurlyeq v_i^{(2)}$ for $i \geqslant 3$. To show that $v_2^{(1)} \succcurlyeq v_2^{(2)}$, construct two more urns. They both begin with R red balls and B black balls. From the first urn, two balls are drawn with replacement and then $F(1)$ balls of the first color and $F(2)$ balls of the second color are added to the urn. Let μ be the law of the resulting proportion of red balls. From the second urn, draw a single ball and replace it along with $F(1) + F(2)$ balls of the color drawn. Let ν be the law of the resulting proportion of red balls in the second urn. The majorization inequality implies that $\mu \preccurlyeq \nu$:

$$\begin{aligned}
\int h\, d\mu = {} & \frac{B}{R+B}\,\frac{B+F(1)}{R+B+F(1)}\, h\left[\frac{R}{R+B+F(1)+F(2)}\right] \\
& + \frac{B}{R+B}\,\frac{R}{R+B+F(1)}\, h\left[\frac{R+F(2)}{R+B+F(1)+F(2)}\right] \\
& + \frac{R}{R+B}\,\frac{B}{R+B+F(1)}\, h\left[\frac{R+F(1)}{R+B+F(1)+F(2)}\right] \\
& + \frac{R}{R+B}\,\frac{R+F(1)}{R+B+F(1)}\, h\left[\frac{R+F(1)+F(2)}{R+B+F(1)+F(2)}\right] \\
\geqslant {} & \frac{B}{R+B}\, h\left[\frac{R}{R+B+F(1)+F(2)}\right] \\
& + \frac{R}{R+B}\, h\left[\frac{R+F(1)+F(2)}{R+B+F(1)+F(2)}\right] \\
= {} & \int h\, d\nu
\end{aligned}$$

Now the variables $v_2^{(1)}$ and $v_2^{(2)}$ can be constructed as follows. To construct $v_2^{(1)}$, begin with R red balls and B black balls, draw a ball, and replace it along with $F(1)$ balls of the same color bearing a special mark. Now draw another ball. If it is unmarked, replace it with along with $F(2)$ balls of the same color. If it is marked, replace it along with $F(2)$ more balls of that color that are also marked. The probability of drawing a marked ball is $F(1)/R+B+F(1)$ and it is clear from the construction that the law of $v_2^{(1)}$ is given by ν when a marked ball is drawn and by μ when an unmarked ball is drawn. So the law of $v_2^{(1)}$ is $[F(1)/R+B+F(1)]\nu+\{(R+B)/[R+B+F(1)]\}\mu$. Similarly, the law of $v_2^{(2)}$ is $[F(2)/R+B+F(2)]\nu+\{(R+B)/[R+B+F(2)]\}\mu$. Then $v_2^{(1)} \succcurlyeq v_2^{(2)}$ follows from $\nu \succcurlyeq \mu$ and $F(1) \geqslant F(2)$. □

Knowing that the distribution of v is nonatomic on $(0, 1)$, it is logical to ask when the distribution is absolutely continuous with respect to Lebesgue measure. Nothing is known about this except when F is constant and the distribution of v is known to be a beta, or when $F(n)$ goes to zero faster than 2^{-n} and v is supported on a Cantor set.

REFERENCES

1. Arthur, B., Ermol'ev, Y., and Kaniovskii, Y. (1983). A generalized urn problem and its applications. *Cybernetics* **19**, 61–71.
2. Athreya, K. (1969). On a characteristic property of Pólya's urn. *Stud. Sci. Math. Hung.* **4**, 31–35.
3. Eggenberger, F., and Pólya, G. (1923). Über die Statistik Verketter Vorgänge. *Z. Angew. Math. Mech.* **3**, 279–289.
4. Feller, W. (1966). *Introduction to Probability Theory and Its Applications*, Vol. 2, Wiley, New York.
5. Friedman, B. (1949). A simple urn model. *Commun. Pure Appl. Math.* **2**, 59–70.
6. Hill, B., Lane, D., and Sudderth, W. (1980). A strong law for some generalized urn processes. *Ann. Prob.* **8**, 214–226.
7. Pemantle, R. (1988). Random processes with reinforcement. Doctoral thesis. Massachusetts Institute of Technology.
8. Pemantle, R. (1989). Nonconvergence to unstable points in urn models and stochastic approximations. To appear in *Ann. Prob.* **18**, 2.

Printed in Belgium