PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 113, Number 1, September 1991

WHEN ARE TOUCHPOINTS LIMITS FOR GENERALIZED PÓLYA URNS?

ROBIN PEMANTLE

(Communicated by Lawrence F. Gray)

ABSTRACT. Hill, Lane, and Sudderth (1980) consider a Pólya-like urn scheme in which $X_0, X_1, \ldots,$ are the successive proportions of red balls in an urn to which at the n th stage a red ball is added with probability $f(X_n)$ and a black ball is added with probability $1-f(X_n)$. For continuous f they show that X_n converges almost surely to a random limit X which is a fixed point for f and ask whether the point p can be a limit if p is a touchpoint, i.e. $p = f(p)$ but $f(x) > x$ for $x \neq p$ in a neighborhood of p. The answer is that it depends on whether the limit of $(f(x)-x)/(p-x)$ is greater or less than $1/2$ as x approaches p from the side where $(f(x)-x)/(p-x)$ is positive.

Hill, Lane, and Sudderth (1980), hereafter referred to as [HLS], consider the following urn scheme. Let $f:[0, 1] \to [0, 1]$ be any function and let an urn begin with l balls of which a proportion $X_{l-1} \in (0, 1)$ are red and the remainder black. Add a new ball to the urn, whose color is red with probability $f(X_{l-1})$ and black otherwise. Let X_l be the new proportion of red balls and iterate the procedure, producing a sequence of proportions $X_{l-1}, X_l, X_{l+1}, \ldots$. In the case where f is continuous, they show that X_n converges almost surely to some random variable X. Furthermore, $f(X) = X$ almost surely [HLS, Theorem 2.1 and Corollary 3.1]. Categorize points $p \in (0, 1)$ for which $p = f(p)$ by calling them *upcrossings* if $(y-p)(f(y)-y)$ is positive for all y in some neighborhood of p, and *downcrossings* if $(y-p)(f(y)-y)$ is negative for all y in some neighborhood of p. The terminology comes from the way the graph $y = f(x)$ crosses the graph $y = x$. The next results of [HLS] are that $\text{prob}(X_n \to p) > 0$ if p is a downcrossing and f maps $(0, 1)$ into itself, while $\text{prob}(X_n \to p) = 0$ if p is an upcrossing. The only other kind of isolated point, p, in the set $\{x : x = f(x)\}$ is a *touchpoint* where $f(y) > y$ for all $y \neq p$ in a neighborhood of p, or else $f(y) < y$ for all $y \neq p$ in a neighborhood of p. They ask whether touchpoints can be in the support of the limiting random variable X.

Received by the editors February 20, 1990 and, in revised form, May 30, 1990.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 60J05.
Key words and phrases. Pólya urn, touchpoint.
This research supported by a National Science Foundation postdoctoral fellowship.

This note answers their question both ways for continuous f, giving a condition on f near p implying $\text{prob}(X_n \to p) > 0$ and another condition that implies $\text{prob}(X_n \to p) = 0$. These conditions almost meet, in the sense that they cover all cases where $(f(x)-x)/(p-x)$ has a limit as $x \uparrow p$ except for the case where the limit is equal to $1/2$. By symmetry between red and black balls, there is no loss of generality in considering only touchpoints of the first kind, where $f(y) > y$ for $y \neq p$ in a neighborhood of p. Therefore, the proofs will be given only for the touchpoints of the first kind. Furthermore, whether X_n converges to p with positive probability depends only on the germ of f at p [HLS, Lemma 4.1], so the arguments below will assume without loss of generality that $f(y) > y$ for all $y \neq p, 1$, as well as assuming that f maps $(0, 1)$ into itself.

Let $\mathcal{F}_n$ be the σ-algebra generated by $\{X_i : i \leq n\}$, and let $\mathcal{F}_\tau$ be defined similarly for any stopping time τ. The key to the proof of both conditions will be the decomposition of the submartingale $\{X_n, \mathcal{F}_n\}$ into a martingale and an increasing process. Write $X_{n+1} = X_n + A_n + Y_n$, where

$$A_n = \mathbf{E}(X_{n+1} - X_n | \mathcal{F}_n)$$

is $\mathcal{F}_n$-measurable and $Y_n = X_{n+1} - X_n - A_n$, so $\mathbf{E}(Y_n|\mathcal{F}_n) = 0$. Then calculate the following conditional probabilities given $\mathcal{F}_n$:

$$X_{n+1} = \begin{cases} \frac{nX_n+1}{n+1} = X_n + \frac{1-X_n}{n+1} & \text{with probability } f(X_n), \\ \frac{nX_n}{n+1} = X_n - \frac{X_n}{n+1} & \text{with probability } 1 - f(X_n). \end{cases}$$

This gives $A_n = (f(X_n) - X_n)/(n+1)$, which is nonnegative by assumption; and hence

$$Y_n = \begin{cases} \frac{1-f(X_n)}{n+1} & \text{with probability } f(X_n), \\ -\frac{f(X_n)}{n+1} & \text{with probability } 1 - f(X_n). \end{cases}$$

Also, Y_n is a mean zero random variable given $\mathcal{F}_n$, with the conditional distribution of Y_n given $\mathcal{F}_n$ satisfying $\min(f(X_n), 1 - f(X_n))^2(n+1)^{-2} = \inf Y_n^2 \leq \mathbf{E}(Y_n^2|\mathcal{F}_n) \leq \sup_\omega Y_n^2 \leq (n+1)^{-2}$, where the inf is over ω in the $\mathcal{F}_n$-measurable set for which X_n has the given value. Defining

$$Z_{n,m} = \sum_{i=n}^{m-1} Y_i$$

yields for each fixed n a martingale $\{Z_{n,m}, \mathcal{F}_m\}$ with an L^2-bound $\mathbf{E}Z_{n,\infty}^2 \leq \sum_{i=n}^{\infty} (i+1)^{-2} \leq 1/n$. If f is bounded away from 0 and 1 near p, then a lower L^2-bound is gotten by stopping the process X_n when it exits an interval on which $\min(f(X_n), 1-f(X_n)) > b$. If τ is any stopping time bounded above by the exit time of the interval, then the above lower bound on $\mathbf{E}Y_m^2$ gives

$$\text{(1)} \qquad \mathbf{E}(Z_{n,\infty}^2|\mathcal{F}_n) \geq \mathbf{E}(Z_{n,\tau}^2|\mathcal{F}_n) \geq \text{prob}(\tau = \infty|\mathcal{F}_n) b^2 (n+1)^{-1}.$$

The idea will be that if $f(x) - x$ is less than $(p - x)/2$, then the increasing part A pushes X toward p so slowly that by the time X gets close to p, the increments of Z are very very small, and Z cannot push X above p. So, in fact, one gets convergence to p from below. On the other hand, if $f(x) - x$ is greater than $(p - x)/2$, then the increasing part pushes X toward p fast enough so that the increments of Z are big enough compared with $p - X$, so that, eventually, the addition of Z puts X over p. A result along the lines of Pemantle [P1, P2] then implies that X_n cannot converge to p.

Remark. It will be shown that convergence to a touchpoint near which $f(x) > x$ is always from the left. Thus the behavior of the function to the right of the touchpoint is irrelevant.

Theorem 1. *Let f be continuous in a neighborhood of a touchpoint p and suppose that f maps $(0, 1)$ into itself. Further suppose that $x < f(x) \leq x + k(p - x)$ for some $k < 1/2$ and all x in some left neighborhood, $(p - \varepsilon, p)$, of p. Then* $\text{prob}(X_n \to p) > 0$. *[Similarly, if $x > f(x) \geq x - k(x - p)$ for some $k < 1/2$ and all x in a right neighborhood, $(p, p + \varepsilon)$ of p, then also* $\text{prob}(X_n \to p) > 0$.]

Corollary 2. *If f is differentiable at a touchpoint p and continuous in a neighborhood of p, then* $\text{prob}(X_n \to p) > 0$ *under the same nontriviality assumption* $f((0, 1)) \subseteq (0, 1)$.

Proof. Since $f(x) - x$ does not change sign at p, the derivative of $f(x) - x$ must be zero at p and Theorem 1 applies. □

Proof of Theorem 1. Replacing f by a function agreeing with f on a neighborhood of p, there is no loss of generality in assuming that f is continuous and that $f(x) > x$ for all $x \in [0, 1)\backslash\{p\}$. Thus it will suffice to prove that with positive probability there is an N for which $n > N$ implies $X_n < p$, since X_n converges to a fixed point of f [HLS, Corollary 3.1], which must then be p. Pick a k for which the hypothesis is satisfied and pick k_1 with $k < k_1 < 1/2$. Pick a constant γ just barely greater than 1 so that $\gamma k_1 < 1/2$. The function $g(r) = re^{(1-r)/2k_1\gamma}$ has value 1 at $r = 1$ and derivative $g'(1) = 1 - 1/2k_1\gamma < 0$, so there is an $r \in (0, 1)$ for which $g(r) > 1$. Fix such an r. Define

$$T(n) = e^{n(1-r)/\gamma k_1}, \qquad \text{so } g(r)^n = r^n T(n)^{1/2} > 1.$$

Choose M big enough so that $\gamma r^M < \varepsilon$ and define

$$\tau_M = \inf\{j > T(M): X_{j-1} < p - r^M < X_j\}$$

if such a j exists, and $\tau_M = -\infty$ otherwise. By the nontriviality assumption that f maps $(0, 1)$ into itself, $\text{prob}(\tau_M > T(M)) > 0$. For each $n \geq M$, define $\tau_{n+1} = \inf\{j \geq \tau_n: X_j > p - r^{n+1}\}$. Note that if $X_j \geq p$ for some $j > T(M)$, then $\tau_n \leq j$ for all $n \geq M$. The theorem will be proved by showing that $\text{prob}(\tau_n > T(n) \text{ for all } n \geq M) > 0$, which will imply that with nonzero probability, X_n is eventually less than p, proving the theorem. Begin by assuming that $\tau_n > T(n)$ and calculate $\text{prob}(\tau_{n+1} > T(n+1) | \tau_n > T(N))$ as follows. Let $\mathscr{B}$ be the event $\{\inf_{j>\tau_n} X_j \geq p - \gamma r^n\}$ and estimate

$$\begin{aligned}
\text{prob}(\mathscr{B}^c | \tau_n > T(N)) &= \text{prob}\left(\inf_{j>\tau_n} X_j < p - \gamma r^n | \tau_n > T(N)\right) \\
&\leq \text{prob}\left(\inf_{j>\tau_n} Z_{\tau_n, j} < -(\gamma - 1) r^n | \tau_n > T(N)\right) \\
&\leq \mathbf{E}(Z^2_{\tau_n, \infty} | \tau_n > T(N)) / ((\gamma - 1) r^n)^2 \\
&\leq e^{-n(1-r)/k_1\gamma} (\gamma - 1)^{-2} r^{-2n} \\
&= (\gamma - 1)^{-2} [g(r)]^{-2n}.
\end{aligned}$$

Next, note that if $\mathscr{B}$ holds, then

$$\begin{aligned}
\sum_{T(n)<j<T(n+1)} A_j &= \sum_{T(n)<j<T(n+1)} (f(X_j) - X_j)/(j+1) \\
&< (\ln\lceil T(n+1)\rceil - \ln\lceil T(n)\rceil)(k\gamma r^n) \\
&\leq (k\gamma r^n)[(1-r)/\gamma k_1 + 1/T(n)] \\
&= (k/k_1)(r^n - r^{n+1}) + k\gamma r^n / T(n).
\end{aligned}$$

But then if $\mathscr{B}$ holds and $\tau_{n+1} = L \leq T(n+1)$, it must be the case that

$$\begin{aligned}
Z_{\tau_n, L} &= X_L - X_{\tau_n} - \sum_{j=\tau_n}^{L-1} A_j \\
&\geq X_L - X_{\tau_n} - \sum_{T(n)<j<T(n+1)} A_j \\
&\geq r^n - r^{n+1} - \xi_n - (k/k_1)(r^n - r^{n+1}) - k\gamma r^n / T(n) \\
&= r^n (1-r)(1 - (k/k_1)) - \xi_n - k\gamma r^n / T(n) \\
&= r^n (1-r)(1 - (k/k_1)) - \tilde{\xi}_n.
\end{aligned}$$

The term ξ_n comes from the fact that X_{τ_n} may overshoot the stopping point $p - r^n$, and $\tilde{\xi}_n$ denotes the sum of ξ_n and the $k\gamma r^n / T(n)$ term. Then ξ_n is bounded by $X_{\tau_n} - X_{\tau_n - 1} < \tau_n^{-1} < T(n)^{-1}$ by assumption. Since $T(n)^{-1}$ is of order less than r^{2n}, the $\tilde{\xi}_n$ contribution vanishes asymptotically in the sense that

$$\frac{r^n (1-r)(1 - (k/k_1)) - \tilde{\xi}_n}{r^n (1-r)(1 - (k/k_1))} \to 1.$$

Now $\mathbf{E}(Z^2_{\tau_n,\infty}|\tau_n > T(N)) < T(n)^{-1}$, so

$$\begin{aligned}
\operatorname{prob}(\tau_{n+1} \le T(n+1)|\tau_n > T(N)) \\
&\le \operatorname{prob}(\mathscr{B}^c|\tau_n > T(N)) \\
&\quad + \operatorname{prob}\left(\mathscr{B} \text{ and } \sup_L Z_{\tau_n,L} \ge r^n(1-r)\left(1-\frac{k}{k_1}\right) - \tilde{\xi}_n|\tau_n > T(N)\right) \\
&\le (1-\gamma)^{-2}[g(r)]^{-2n} + T(n)^{-1}\Big/\left[r^n(1-r)\left(1-\frac{k}{k_1}\right)-\tilde{\xi}_n\right]^2 \\
&\le (1-\gamma)^{-2}[g(r)]^{-2n} + [(1-r)(1-(k/k_1))]^{-2}[g(r)]^{-2n} \\
&\quad \times \left[\frac{r^n(1-r)(1-(k/k_1))-\tilde{\xi}_n}{r^n(1-r)\left(1-\frac{k}{k_1}\right)}\right]^2 .
\end{aligned}$$

Because the last term of the numerator vanishes asymptotically, the sum of these probabilities converges. Then $\operatorname{prob}(\tau_n > T(n)$ for all $n > M) = \operatorname{prob}(\tau_M > T(M))\prod_{n\ge M}(1-\operatorname{prob}(\tau_{n+1} \le T(n+1)|\tau_n > T(N))) > 0$ since each factor is positive and $\sum \operatorname{prob}(\tau_{n+1} \le T(n)|\tau_n > T(N))$ is finite. In this case, X_n must converge to p from below. □

Theorem 3. *Suppose that $f(x) \ge x + k(p-x)$ for some $k > 1/2$ and all x in some left neighborhood, $(p-\varepsilon, p)$, of p. Then* $\operatorname{prob}(X_n \to p) = 0$. [*Similarly, if $f(x) \le x - k(x-p)$ for some $k > 1/2$ and all x in a right neighborhood, $(p, p+\varepsilon)$ of p, then also* $\operatorname{prob}(X_n \to p) = 0$.]

Remark. No continuity assumptions are needed this time.

Proof. Again there is no loss of generality in assuming that $f(x) \ge x$ for all x; similarly, assume $f(x) \ge \min(1, x + k|p-x|)$ on $[0, p]$. Furthermore, Lemma 2.2 of [HLS] says that replacing f by a pointwise smaller function gives a process which can be defined on the same probability space so as always to be smaller. Thus replacing f by the minimum of 1 and $x + k|p-x|$ on $[0, p]$ and by x on $[p, 1]$ gives a process which converges to p whenever the original process does, so it suffices to prove the theorem for this choice of f. The importance of assuming this lies only in getting f bounded away from 0 and 1 near p (without assuming continuity) so that there will be a lower L^2-bound on Z.

The following argument is self-contained, but the reader may wish to look at Pemantle [P2, Lemmas 1 and 2] to see the template from which this proof was constructed.

Lemma 4. *There are constants $a, c > 0$ and a neighborhood $\mathscr{N}$ of p such that for any n*

$$\operatorname{prob}(Z_{n,\infty} > cn^{-1/2} \textit{ or } X_{n+j} \notin \mathscr{N} \textit{ for some } j|\mathscr{F}_n) > a.$$

Proof. Pick $b > 0$ and $\mathcal{N}$ a neighborhood of p such that $f(\mathcal{N}) \subseteq [b, 1-b]$. Assume that $X_n \in \mathcal{N}$ or else the result is trivially true. For $k > 0$, let $\tau \leq \infty$ be the first time X_j exits $\mathcal{N}$ or $Z_{n,j}$ exits $(-kn^{-1/2}, kn^{-1/2})$. Then equation (1) gives $\mathbf{E}(Z_{n,\tau}^2|\mathcal{F}_n) \geq \text{prob}(\tau = \infty|\mathcal{F}_n)b^2(n+1)^{-1}$. On the other hand, $\mathbf{E}(Z_{n,\tau}^2|\mathcal{F}_n) \leq \mathbf{E}(X_\tau - X_n)^2 \leq k^2/n$, since Z is just the martingale part of X. Putting these together gives $\text{prob}(\tau = \infty|\mathcal{F}_n) \leq k^2(n+1)/b^2 n$, and choosing k small enough makes this at most $1/3$. Let

$$q = \text{prob}(\tau < \infty, \ X_\tau \notin \mathcal{N}|\mathcal{F}_n),$$

so that the conditional probability of $Z_{n,j}$ exiting $(-kn^{-1/2}, kn^{-1/2})$ given $\mathcal{F}_n$ is at least $2/3 - q$. Any martingale $\mathcal{M}$ started at zero that exits an interval $(-L, L)$ with probability at least r and has increments bounded by $L/2$ satisfies $\text{prob}(\sup \mathcal{M} \geq L/2) \geq (3r-1)/4$; stopping $\mathcal{M}$ upon exiting $(-L, L/2)$ and letting $s = \text{prob}(\sup \mathcal{M} > L/2)$ gives $0 = \mathbf{E}\mathcal{M} \leq sL + (r-s)(-L) + (1-r)(L/2) = 2L(s - (3r-1)/4)$. Thus $Z_{n,j} \geq k/2\sqrt{n}$ for some j with probability at least $(1-3q)/4$.

Now for any j, condition on the event $Z_{n,j} \geq k/2\sqrt{n}$; then the conditional probability of the event $Z_{n,\infty} < k/4\sqrt{n}$ can be bounded away from 1 using the following one-sided Tschebysheff estimate:

Lemma 5. *If $\mathcal{M}$ has mean zero and $L < 0$, then*

$$\text{prob}(\mathcal{M} \leq L) \leq \mathbf{E}\mathcal{M}^2/(\mathbf{E}\mathcal{M}^2 + L^2).$$

Proof. Write w for $\text{prob}(\mathcal{M} \leq L)$. From

$$0 = \mathbf{E}\mathcal{M}^2 = w\mathbf{E}(\mathcal{M}|\mathcal{M} \leq L) + (1-w)\mathbf{E}(\mathcal{M}|\mathcal{M} > L)$$

and $\mathbf{E}(\mathcal{M}|M \leq L) \leq L$, it is immediate that

$$\mathbf{E}(\mathcal{M}|\mathcal{M} > L) \geq -L\frac{w}{1-w}.$$

Then

$$\begin{aligned}
\mathbf{E}\mathcal{M}^2 &= w\mathbf{E}(M^2|M \leq L) + (1-w)\mathbf{E}(\mathcal{M}^2|\mathcal{M} > L) \\
&\geq wL^2 + (1-w)(\mathbf{E}(\mathcal{M}|\mathcal{M} > L))^2 \\
&\geq wL^2 + (1-w)L^2(w^2/(1-w)^2) \\
&= L^2 w/(1-w),
\end{aligned}$$

from which the desired conclusion follows. □

Apply this to the process $Z_{j,i}$ stopped at the entrance time τ of the interval $(-\infty, -k/4\sqrt{n})$ to get

$$\begin{aligned}
\text{prob}(Z_{n,\infty} \leq k/4\sqrt{n}|\mathcal{F}_j) &\leq \text{prob}(Z_{j,\tau} \leq -k/4\sqrt{n}|\mathcal{F}_j) \\
&\leq \mathbf{E}Z_{j,\tau}^2/(\mathbf{E}Z_{j,\tau}^2 + k^2/16n) \\
&\leq \mathbf{E}Z_{n,\infty}^2/(\mathbf{E}Z_{n,\infty}^2 + k^2/16n) \\
&\leq 16/(k^2 + 16).
\end{aligned}$$

Combining this with the previous result shows that the conditional probability of $Z_{n,\infty} > k/4\sqrt{n}$ given $\mathcal{F}_n$ is at least $(1-3q)k^2/64+4k^2)$. Recall that q is the conditional probability of the process exiting $\mathcal{N}$ given $\mathcal{F}_n$, so that the probability we are trying to bound below is at least the maximum of q and $(1-3q)k^2/(64+4k^2)$. For any value of q the maximum is at least $k^2/(64+7k^2)$, thus the statement of the lemma is proved with $c = k/4$ and $a = k^2/(64+7k^2)$. □

Let τ be any finite stopping time. Conditioning on $\mathcal{F}_\tau$ then gives a stopping time version of the previous lemma:

$$\text{prob}(Z_{\tau,\infty} > c\tau^{-1/2} \text{ or } X_{\tau+j} \notin \mathcal{N} \text{ for some } j|\mathcal{F}_\tau) > a. \tag{2}$$

A corollary of this is a sort of converse to the proof of Theorem 1, saying that if $X_n \to p$ then it does so from the left.

Corollary 6. *Let p be a touchpoint of the first kind, i.e.. $f(y) > y$ for all $y \neq p$ in a neighborhood of p. Then the probability of the event that either $X_n > p$ finitely often or X_n does not converge to p is 1.*

Proof. Suppose to the contrary that the probability that X_n converges to p and is greater than or equal to p infinitely often is nonzero. Then there are n, M, and some event $\mathcal{B} \in \mathcal{F}_n$ such that $n < M$ and conditional on $\mathcal{B}$, the probability of X_j converging to p and being greater than p some time before M but never leaving $\mathcal{N}$ after time n is at least $1-a/3$. Define τ to be the minimum of M and the least $j \geq n$ such that $X_j > p$. Then letting $\mathcal{C}$ be the event that X_j converges to p without leaving $\mathcal{N}$ after time n,

$$\begin{aligned}\text{prob}(\tau < M|\mathcal{B})\text{prob}(\mathcal{C}|\mathcal{B}, \tau < M) + \text{prob}(\tau = M|\mathcal{B})\text{prob}(\mathcal{C}|\mathcal{B}, \tau = M)\\ = \text{prob}(\mathcal{C}|\mathcal{B}) \geq 1 - a/3.\end{aligned}$$

So

$$\text{prob}(\mathcal{C}|\mathcal{B}, \tau < M) \geq 1 - a/3 - \text{prob}(\tau = M|\mathcal{B}) \geq 1 - 2a/3.$$

Now $\tau < M$ implies that $X_\tau > p$. But since A_n is an increasing process, it follows that $X_j \to p$ and $X_\tau > p$ together imply $Z_{\tau,\infty} < 0$. Thus

$$\text{prob}(Z_{\tau,\infty} < 0 \text{ and } X_{n+j} \in \mathcal{N} \text{ for all } j|\mathcal{B}, \tau < M) \geq 1 - 2a/3,$$

and hence

$$\text{prob}(Z_{\tau,\infty} > c\tau^{-1/2} \text{ or } X_{\tau+j} \notin \mathcal{N} \text{ for some } j|\mathcal{B}, \tau > M) \leq 2a/3.$$

But this contradicts (2), since the events $\mathcal{B}$ and $\tau < M$ are both in $\mathcal{F}_\tau$. □

Continuation of the proof of Theorem 3. It remains to show that under the hypothesis of the theorem, the probability is zero that X_n eventually resides in $(p-\varepsilon, p)$. If the probability were nonzero, then for any δ there would be an event $\mathscr{B}$ in some $\mathscr{F}_M$ for which $\text{prob}(X_{M+j} \in (p-\varepsilon, p)$ for all $j \geq 0|\mathscr{B}) > 1-\delta$. In fact, conditioning on X_M, $\mathscr{B}$ may be taken to determine X_M. So it suffices to show that the probability of the event $X_{M+j} \in (p-\varepsilon, p)$ for all $j \geq 0$ given X_M is bounded away from 1. For what follows condition on $\mathscr{F}_M$ and on $X_M \in (p-\varepsilon, p)$. Also choose M large enough so that for any $n > M$, $n^{-k/2k_1} < cn^{-1/2}$ where c is chosen as in Lemma 4, and choose ε small enough so that $(p-\varepsilon, p)$ is a subset of a neighborhood $\mathscr{N}$ to which Lemma 4 applies.

Begin by setting up constants and stopping times: pick a $k < 3/4$ for which the hypothesis of the theorem is satisfied and pick k_1 so that $k > k_1 > 1/2$. For $n \geq M$ define

$$V_n = (k/k_1)\ln(n) + 2\ln(p - X_n) \quad \text{for } X_n < p \quad \text{and } -\infty \text{ otherwise.}$$

By assumption on X_M, $V_M > -\infty$. Let τ be the least $n \geq M$ such that $X_n \notin (p-\varepsilon, p)$ or $V_n < 0$. Observe that if $V_n > 0$ then $1/n < (p - X_n)^{2k_1/k} \leq (p - X_n)^{4/3}$, so $|X_{n+1} - X_n|$ is small compared to $p - X_n$, so $V_{\tau\wedge n}$ can never reach $-\infty$ and is in fact bounded below by $\min(-1, V_M)$. Now for $n < \tau$ calculate

$$\begin{aligned}
\mathbf{E}(\ln(p - X_{n+1})|\mathscr{F}_n) &\leq \ln \mathbf{E}(p - X_{n+1}|\mathscr{F}_n) \\
&= \ln(p - X_n - A_n) \\
&\leq \ln((p - X_n)(1 - k/(n+1))) \\
&= \ln(p - X_n) + \ln(1 - k/(n+1));
\end{aligned}$$

so

$$\begin{aligned}
\mathbf{E}(V_{n+1}|\mathscr{F}_n) &\leq V_n + (k/k_1)(\ln(n+1) - \ln(n)) + 2\ln(1 - k/(n+1)) \\
&= V_n + (k/k_1)(n^{-1} + o(n^{-1})) - 2k(n^{-1} + o(n^{-1})) \\
&= V_n - ((2 - 1/k_1)k + o(1))n^{-1} < V_n - Cn^{-1}
\end{aligned}$$

for large n and some $C > 0$. So $V_{n\wedge\tau}$ is a supermartingale for large n, bounded below by $\min(-1, V_M)$, and hence converges almost surely. Clearly it cannot converge without stopping, since the increments of the expectation sum to $-\infty$, therefore the stopping time is reached almost surely.

In other words, conditional upon any event in any $\mathscr{F}_M$, the probability is 1 that for some $n > M$, either X_n will leave $(p-\varepsilon, p)$ or $(k/k_1)\ln(n) < -2\ln(p - X_n)$. Let $\sigma \leq \infty$ be the least $n > M$ for which $(k/k_1)\ln(n) < -2\ln(p - X_n)$. We have just shown that the conditional probability of some X_n leaving $(p-\varepsilon, p)$ given $\sigma = \infty$ is one. On the other hand, the conditional probability of some X_{n+j} leaving $(p-\varepsilon, p)$ given $\sigma = n < \infty$ is at least a by Lemma 4 since $X_{n+j} \notin \mathscr{N}$ trivially implies $X_{n+j} \notin (p-\varepsilon, p)$, while $Z_{n,\infty} > cn^{1/2}$ implies $Z_{n,n+j} > cn^{1/2} > n^{-k/2k_1} > p - X_n$ for some j, which implies $X_{n+j} > p$. □

REFERENCES

[HLS] B. Hill, D. Lane, and W. Sudderth, *A strong law for some generalized urn processes*, Ann. Prob. **8** (1980), 214–226.

[P1] R. Pemantle, *Random processes with reinforcement*, Doctoral thesis, Massachusetts Institute of Technology, 1988.

[P2] ——, *Nonconvergence to unstable points in urn models and stochastic approximations*, Ann. Prob. **18** (1990), 698–712.

DEPARTMENT OF STATISTICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720

Current address: Department of Mathematics, Kidder Hall, Oregon State University, Corvallis, Oregon 97331-4005