BERNARD FRIEDMAN'S URN[1]

By David A. Freedman

University of California, Berkeley

1. Introduction. Bernard Friedman (1949) proposed this urn model. An urn contains W_n white balls and B_n black balls at time n. One ball is drawn at random and then replaced, while α balls of the same color as the ball drawn and β balls of the opposite color are added to the urn. Friedman obtained elegant and almost explicit expressions for the generating functions of the W_n. This paper describes the asymptotic behavior of W_n, as n becomes large, under the conditions $\alpha \geqq 0$, $\beta \geqq 0$. It does not seem possible to obtain these results from his. Many of them were announced in (Freedman, 1963).

Acknowledgment. I would like to thank Professor David Blackwell for help which borders on collaboration.

Summary. The case $\beta = 0$ is the famous Pólya (1931) Urn; a discussion of its elementary properties can be found in (Feller, 1960, Chapter IV) and (Fréchet, 1943). These facts about the Pólya Urn are a classical part of the oral tradition, although some have yet to appear in print (see Blackwell and Kendall, 1964). The fractions $(W_n + B_n)^{-1}W_n$ converge with probability 1 to a limiting random variable Z, which has a beta distribution with parameters W_0/α, B_0/α. Given Z, the successive differences $W_{n+1} - W_n : n \geqq 0$ are conditionally independent and identically distributed, being α with probability Z and 0 with probability $1 - Z$. Proofs are in Section 2.

If $\beta > 0$, the situation is radically different. No matter how large α is in comparison with β, the fractions $(W_n + B_n)^{-1}W_n$ converge to $\frac{1}{2}$ with probability 1. This seemingly paradoxical result can be sharpened in several ways. Abbreviate ρ for $(\alpha + \beta)^{-1}(\alpha - \beta)$. If $\rho > \frac{1}{2}$, it is proved in Section 3 that $(W_n + B_n)^{-\rho} \cdot (W_n - B_n)$ converges with probability 1 to a nondegenerate limiting random variable. This result in turn fails for $\rho \leqq \frac{1}{2}$. If $0 < \rho \leqq \frac{1}{2}$, the sequence $(W_n + B_n)^{-\rho}(W_n - B_n)$ has plus infinity for superior limit and minus infinity for inferior limit, with probability 1. If $\rho < 0$, the sequence $(W_n - B_n)$ has plus infinity for superior limit and minus infinity for inferior limit, with probability 1. In both cases, the tail σ-field of $(W_n, B_n) : n \geqq 0$ is trivial.

If $\rho < \frac{1}{2}$, it is proved in Section 5 that the distribution of $n^{-\frac{1}{2}}(W_n - B_n)$ converges to normal with mean 0 and variance $(\alpha - \beta)^2/(1 - 2\rho)$. When $\rho = \frac{1}{2}$, the last fraction is not defined; but the distribution of $(n \log n)^{-\frac{1}{2}}(W_n - B_n)$ converges to normal with mean 0 and variance $(\alpha - \beta)^2$. The asymptotic normality of $W_n - B_n$ for $\rho \leqq \frac{1}{2}$ was observed by Bernstein (1940). I am grateful to J. A. McFadden for calling this paper to my attention.

Received 30 March 1964.

[1] Prepared with the partial support of the National Science Foundation, Grant GP-10.

Consider taking $\alpha = 0$ and $\beta = 1$, so $\rho < 0$. Since $(W_n + B_n)^{-1}W_n$ converges to $\frac{1}{2}$, therefore $W_n - B_n$ is asymptotically like the sum of n independent random variables, each equal to $+1$ with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. It is tempting to conclude that the distribution of $n^{-\frac{1}{2}}(W_n - B_n)$ converges to normal with mean 0 and variance 1. From the preceding paragraph, however, the asymptotic variance is $\frac{1}{3}$. There is an even more startling difference between the asymptotic behavior of $(W_n - B_n) : n \geqq 0$ and that of a coin-tossing game. Let $X_n : n \geqq 1$ be independent and ± 1 with probability $\frac{1}{2}$ each. Let $S_n = X_1 + \cdots + X_n$, $S_{j/n,n} = n^{-\frac{1}{2}}S_j$. Define $S_{t,n}$ for $0 \leqq t \leqq 1$ and nt not integral by linear interpolation. By the celebrated Invariance Principle of Donsker (1951), the law of $\{S_{t,n} : 0 \leqq t \leqq 1\}$ converges in a strong way to the law of a Brownian motion. However, for $\rho < \frac{1}{2}$ suppose we define $Z_{j/n,n} = n^{-\frac{1}{2}}(W_j - B_j)$ and $\{Z_{t,n} : 0 \leqq t \leqq 1\}$ by linear interpolation. The law of $\{Z_{t,n} : 0 \leqq t \leqq 1\}$ converges in the sense of the Invariance Principle to the law of a process $\{Z_t : 0 \leqq t \leqq 1\}$. Now Z_t is normal with mean 0 and variance $(1 - 2\rho)^{-1}(\alpha - \beta)^2 t$. But $\{Z_t : 0 \leqq t \leqq 1\}$, though Gaussian, does not have independent increments. On the other hand, $\{t^{-\rho}Z_t : 0 \leqq t \leqq 1\}$ is a nonhomogeneous Brownian motion. If $\rho = \frac{1}{2}$, it is necessary to put $Z_{j/n,n} = (n \log n)^{-\frac{1}{2}}(W_j - B_j)$. In the limit, $Z_t = t^{\frac{1}{2}}Z_1$, where Z_1 is normal with mean 0 and variance $(\alpha - \beta)^2$. These results were obtained independently by K. Ito and myself. Details will be given in a future joint paper.

D. Ornstein has obtained this very intuitive proof that $(W_n + B_n)^{-1}W_n$ converges to $\frac{1}{2}$ with probability 1 for $\beta > 0$. Suppose first $\alpha > \beta$. If $0 \leqq x \leqq 1$ and $P[\limsup (W_n + B_n)^{-1}W_n \leqq x] = 1$, by an easy variation of the Strong Law, with probability 1, in N trials there will be at most $Nx + o(N)$ drawings of a white ball; so at least $N(1 - x) - o(N)$ drawings of black. Therefore, with probability 1, $\limsup (W_n + B_n)^{-1}B_n$ is bounded above by

$$\lim_{N\to\infty} \{\alpha[Nx + o(N)] + \beta[N(1 - x) - o(N)]\}/N(\alpha + \beta)$$

or $(\alpha + \beta)^{-1}[\beta + (\alpha - \beta)x]$. Starting with $x = 1$ and iterating, $P[\limsup (W_n + B_n)^{-1} \leqq \frac{1}{2}] = 1$ follows. Interchange white and black to complete the proof for $\alpha > \beta$. If $\alpha < \beta$, and $P[\limsup (W_n + B_n)^{-1}W_n \leqq x] = 1$, then a similar argument shows $P[\limsup (W_n + B_n)^{-1}B_n \leqq (\alpha + \beta)^{-1} \cdot (\alpha + (\beta - \alpha)x)] = 1$. The argument proceeds as before, except both colors must be considered simultaneously.

Notation. α and β are nonnegative real numbers (not necessarily natural). Abbreviate

$$\delta = \alpha - \beta, \tag{1.1}$$

$$\sigma = \alpha + \beta, \tag{1.2}$$

so $\rho = \delta/\sigma$. The process $(W_n, B_n) : n \geqq 0$ is a Markov chain on the probability triple $(\Omega, \mathfrak{F}, P)$ with pairs of nonnegative real numbers for states, (W_0, B_0) degenerate with $W_0 + B_0 > 0$, and the following stationary transition mechanism: from (j, k) move to $(j + \alpha, k + \beta)$ with probability $j/(j+k)$ or to $(j + \beta, k + \alpha)$ with probability $k/(j + k)$. The process $(W_n, B_n) : n \geqq 0$ is called a *Friedman Urn with parameters* α, β. Abbreviate

$$s = W_0 + B_0 \tag{1.3}$$

so $W_n = s + \sigma n$; and

$$\imath_n(j) = 1 + [j\delta/(s + \sigma n)], \tag{1.4}$$

$$x_n(k) = E[(W_n - B_n)^k]. \tag{1.5}$$

Let $\mathfrak{F}_n$ be the σ-field spanned by $(W_j, B_j) : 0 \leqq j \leqq n$, and $\mathfrak{F}^{(n)}$ the σ-field spanned by $(W_j, B_j) : j \geqq n + 1$. The *tail σ-field* of $(W_n, B_n) : n \geqq 0$ is $\bigcup_{n=0}^{\infty} \mathfrak{F}^{(n)}$. Plainly, it coincides with the σ-field of events measurable on $W_{n+1} - W_n : n \geqq 0$ and invariant under finite permutations of them.

If x_n and y_n are real numbers for $n \geqq 0$, then $x_n \approx y_n$ means $\lim_{n\to\infty} x_n/y_n = 1$; while $x_n \sim y_n$ means $\lim_{n\to\infty} x_n/y_n$ exists, is finite, and not zero. $x_n = o(y_n)$ means $\lim_{n\to\infty} x_n/y_n = 0$; while $x_n = O(y_n)$ means $\lim\sup_{n\to\infty} |x_n|/|y_n| < \infty$. As usual, an empty sum is 0 and an empty product is 1. The end of a proof is marked <>.

2. The Pólya Urn. This section is purely expository. We consider a Friedman urn process (W_n, B_n); $n \geqq 0$ on the probability triple $(\Omega, \mathfrak{F}, P)$, with parameter $\alpha > 0$ and parameter $\beta = 0$. Abbreviate $X_n = W_{n+1} - W_n$ for $n \geqq 0$.

LEMMA 2.1. *The process $(W_n + B_n)^{-1}W_n : n \geqq 0$ is a martingale, and converges with probability 1 to a limiting random variable Z.*

PROOF. The first assertion is easy to verify directly. The second follows from the first by appealing to the forward martingale convergence theorem (Doob, 1953, Theorem 4.1(i), p. 319). <>

THEOREM 2.1. *The tail σ-field of $(W_n, B_n) : n \geqq 0$ is equivalent to the σ-field spanned by Z. Given this σ-field, the $X_n : n \geqq 0$ are independent and identically distributed, being α with probability Z and 0 with probability $1 - Z$.*

PROOF. It is easy to check inductively that the $X_n : n \geqq 0$ are exchangeable. This observation goes back to Pólya (1931). See also (Fréchet, 1943) and (Feller, 1960). If $\epsilon_0 \cdots \epsilon_k$ are 0 or α, it follows that

$$P(X_j = \epsilon_j, 0 \leqq j \leqq k \,|\, \mathfrak{F}^{(n)}) = \binom{n - k + 1}{\alpha^{-1}[W_{n+1} - W_0 - \sum_{j=0}^{k} \epsilon_j]} \Big/ \binom{n + 1}{\alpha^{-1}[W_{n+1} - W_0]} \tag{2.1}$$

for $n \geqq k - 1$. As n tends to ∞, the left side of (2.1) converges to $P(X_j = \epsilon_j, 0 \leqq j \leqq k \,|\, \mathfrak{F}^{(\infty)})$ with probability 1, by the backward martingale convergence theorem (Doob, 1953, Theorem 4.2, p. 428). The right side of (2.1) converges to

$$Z^{\alpha^{-1}\Sigma_{j=0}^{k}\epsilon_j}(1 - Z)^{k+1-\alpha^{-1}\Sigma_{j=0}^{k}\epsilon_j}$$

with probability 1, by Lemma 2.1 and Stirling's formula. In particular, if A is measurable on a finite number of X_n's, then $P(A \mid \mathfrak{F}^{(\infty)})$ is equal to a function of Z with probability 1. By a routine argument, the last statement holds for any A measurable on $(W_n, B_n) : n \geqq 0$; in particular, for any $A \varepsilon \mathfrak{F}^{(\infty)}$. <>

Pólya (1931, p. 150) proved that the distribution of $(W_n + B_n)^{-1}W_n$ converges to beta with parameters W_0/α, B_0/α. This follows from Lemma 2.1 and the next theorem.

THEOREM 2.2. *The distribution of Z is beta with parameters W_0/α, B_0/α.*

PROOF. $P(X_j = \alpha, 0 \leqq j \leqq k)$ is $E(Z^{k+1})$ by Theorem 2.1 and $\prod_{j=0}^{k} (W_0 + \alpha j)/(W_0 + B_0 + \alpha j)$ by the definition of the urn process. <>

Of course, $P(W_n - B_n = W_0 - B_0$ for infinitely many $n \mid Z)$ is 0 or 1 according as $Z \neq \frac{1}{2}$ or $Z = \frac{1}{2}$. Since $P(Z = \frac{1}{2}) = 0$, therefore $P(W_n - B_n = W_0 - B_0$ for finitely many $n) = 1$. However, the expected number of n with $W_n - B_n = W_0 - B_0$ is infinite. A related fact is: if τ is the least n with $X_n \neq X_0$, then $P(\tau < \infty) = 1$ but $E(\tau) < \infty$ if and only if $\alpha < W_0$ and $\alpha < B_0$.

REMARK 2.1. The second part of Lemma 2.1 is: $n^{-1}\sum_{j=0}^{n-1} X_j$ converges with probability 1. This holds for any $\mathfrak{L}^1$ exchangeable process, either by the Birkhoff ergodic theorem (Doob, 1953, Theorem 2.1, p. 465), or by a direct martingale argument (Loève, 1963, III, p. 400).

REMARK 2.2. Of course, Theorem 2.2 is a special case of de Finetti's (1931) theorem: if $X_n : n \geqq 0$ is an exchangeable process of 0's and 1's, there is a random variable Z such that, given Z, the $X_n : n \geqq 0$ are conditionally independent and identically distributed, being equal to 1 with probability Z and 0 with probability $1 - Z$. Conversely, however, Remark 2.1 and the technique used to prove Theorem 2.2 give an immediate proof of de Finetti's theorem. Almost the same argument proves the generalization of de Finetti's theorem for an exchangeable process of random variables with finite range. An easy martingale passage to the limit gives the theorem for random variables whose range is compact metric. Standard functional analysis then gives the most general results known (Hewitt and Savage, 1955). For a different treatment, see (Freedman, 1962). Of course, de Finetti's theorem for $0 - 1$ valued processes is equivalent to the Hausdorff moment problem.

REMARK 2.3. Even the first part of Lemma 2.1 has a natural generalization. To begin with, $(W_n + B_n)^{-1}W_n$ is $P(X_{n+1} = 1 \mid X_0 \cdots X_n)$. If now $X_n : n \geqq 0$ is an arbitrary exchangeable process of 0's and 1's, with $Z = \lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1} X_j$, then $P(X_{n+1} = 1 \mid X_0 \cdots X_n) = E(Z \mid X_0 \cdots X_n)$ is clearly a martingale converging to Z. More abstractly, if $X_n : n = -1, 0, 1, \cdots$ is a completely general exchangeable process and h is a bounded, measurable function, then

$$E[h(X_{n+i} : i \geqq 1) \mid X_0 \cdots X_n] = E[h(X_{-i} : i \geqq 1) \mid X_0 \cdots X_n]$$

is a martingale.

3. The case $\rho > \frac{1}{2}$. Let $(W_n, B_n): n \geqq 0$ be a Friedman urn defined on $(\Omega, \mathfrak{F}, P)$, with parameters $\alpha > 0$ and $\beta \geqq 0$. Throughout this section, unless specifically noted otherwise, we make the

ASSUMPTION. $(\alpha - \beta)/(\alpha + \beta) > \frac{1}{2}$.

This section is therefore a legitimate generalization of Section 2.

LEMMA 3.1. *For each nonnegative integer* k, $\lim_{n\to\infty} n^{-\rho k}E[(W_n - B_n)^k] = \mu(k)$, *with* $0 \leqq \mu(k) < \infty$. *If k is even, then* $\mu(k) > 0$.

PROOF. The result is trivial for $k = 0$, and $\mu(0) = 1$. An inductive proof will be given for even k. Recall notations (1.1) to (1.5). Now

$$\begin{aligned} &E\{(W_{n+1} - B_{n+1})^{2K+2} \mid \mathfrak{F}_n\} \\ &= [W_n/(s+\sigma_n)](W_n - B_n + \delta)^{2K+2} + [B_n/(s+\sigma n)](W_n - B_n - \delta)^{2K+2} \\ &= [W_n/(s+\sigma n)]\textstyle\sum_{j=0}^{2K+2}\binom{2K+2}{j}\delta^j(W_n - B_n)^{2k+2-j} \\ &\quad + [B_n/(s+\sigma n)]\textstyle\sum_{j=0}^{2K+2}\binom{2K+2}{j}(-\delta)^j(W_n - B_n)^{2K+2-j} \\ &= a_n(2K+2)(W_n - B_n)^{2K+2} \\ &\quad + \textstyle\sum_{j=1}^{K}[\binom{2K+2}{2j}\delta^{2j} + [\binom{2K+2}{2j+1}/(s+\sigma n)]\delta^{2j+1}](W_n - B_n)^{2K+2-2j} + \delta^{2K+2}, \end{aligned} \tag{3.1}$$

so

$$x_{n+1}(2K+2) = a_n(2K+2)x_n(2K+2) + b_n(2K+2), \tag{3.2}$$

with

$$\begin{aligned} &b_n(2K+2) \\ &= \textstyle\sum_{j=1}^{K}[\binom{2K+2}{2j}\delta^{2j} + [\binom{2K+2}{2j+1}/(s+\sigma n)]\delta^{2j+1}]x_n(2K+2-2j) + \delta^{2K+2}. \end{aligned} \tag{3.3}$$

Suppose the theorem is true for even $k \leqq 2K$. Then $0 \leqq b_n(2K+2) = O(n^{2K\rho})$, and by Lemma 6.4,

$$\begin{aligned} &\lim_{n\to\infty} x_n(2K+2)\textstyle\prod_{\nu=0}^{n} a_\nu(2K+2)^{-1} \\ &\qquad = x_0(2K+2) + \textstyle\sum_{j=0}^{\infty} b_j(2K+2)\prod_{\nu=0}^{j} a_\nu(2K+2)^{-1}, \end{aligned}$$

which is positive and finite. By Lemma 6.3, $\prod_{\nu=0}^{n} a_\nu(2K+2) \sim n^{(2K+2)\rho}$, and the theorem holds for $k = 2K+2$. By induction, it holds for even k.

Next,

$$E\{(W_{n+1} - B_{n+1}) \mid \mathfrak{F}_n\} = a_n(1)(W_n - B_n), \tag{3.4}$$

so

$$x_{n+1}(1) = a_n(1)x_n(1) = x_0(1)\textstyle\prod_{\nu=0}^{n} a_\nu(1). \tag{3.5}$$

Since $\prod_{\nu=0}^{n} a_\nu(1) \sim n^\rho$ by Lemma 6.3, the theorem holds when $k = 1$. The proof for odd k can be completed by induction. <>

THEOREM 3.1. *If* $0 < r < \infty$, *then* $\lim_{n\to\infty} n^{-\rho}(W_n - B_n) = Z$ *with probability 1 and in* rth *mean.*

PROOF. Define

$$Z_n = (W_n - B_n)\textstyle\prod_{\nu=0}^{n-1} a_\nu(1)^{-1}. \tag{3.6}$$

Equation (3.4) implies $Z_n : n \geqq 0$ is a martingale. Lemma 3.1 implies $\sup_n E(Z_n^2) < \infty$, so Z_n converges to a finite limit with probability 1 by (Doob, 1953, Theorem 4.1 (iii), p. 319). Since $\prod_{\nu=0}^{n-1} a_\nu(1) \sim n^\rho$ by Lemma 6.3, $\lim_{n\to\infty} n^{-\rho}(W_n - B_n) = Z$ exists and is finite with probability 1. Lemma 3.1 implies $\sup_n n^{-2k\rho}E[(W_n - B_n)^{2k}] < \infty$ for each natural number k, making $n^{-r\rho}(W_n - B_n)^r : 0 \leqq n < \infty$ uniformly integrable for each finite, positive r. < >

COROLLARY 3.1. *With probability* 1, *if* $\beta > 0$ *then* $\lim_{n\to\infty} (W_n + B_n)^{-1}W_n = \frac{1}{2}$.

A related fact was observed by Professor Jerome Sacks. Let τ be the least n with $W_{n+1} - W_n \neq W_1 - W_0$. If $\beta > 0$, for any α (the assumption $\rho > \frac{1}{2}$ is temporarily dropped), $E(\tau) < \infty$.

COROLLARY 3.2. *For each nonnegative integer* k, *the* kth *moment of* Z *is* $\mu(k)$.

COROLLARY 3.3. *With positive probability*, $W_n - B_n = W_0 - B_0$ *for finitely many* n.

Of course, $W_n - B_n = W_0 - B_0$ for infinitely many n with positive probability implies $P(Z = 0) > 0$. We do not know whether the first statement, or even the second, is true. In fact, we know almost nothing about the distribution of Z. We do not know whether Z spans the tail σ-field of $(W_n, B_n) : n \geqq 0$; nor do we have any information about the conditional distribution of $(W_n, B_n) : n \geqq 0$ given Z. Equation (3.2) is, of course, useless for evaluating the even moments of Z. More information can be obtained by choosing coefficients $c_{k,n}$ so that $\sum_{k=0}^{K} c_{2k,n}(W_n - B_n)^{2k}$ is a martingale. Similar remarks apply for odd moments.

Savkevitch (1940) proved that, for $W_0 = B_0$, the limiting distribution of $n^{-\rho}(W_n - B_n)$ is not normal. Combining this with Theorem 3.1: if $W_0 = B_0$ then Z is symmetric about 0 but is not normal because $E(Z^4) < 3[E(Z^2)]^2$.

4. The case $\rho = \frac{1}{2}$. Let (W_n, B_n): $n \geqq 0$ be a Friedman urn defined on $(\Omega, \mathfrak{F}, P)$ with parameters $\alpha, \beta > 0$. Throughout this section, unless specifically noted otherwise, we make the

ASSUMPTION. $(\alpha - \beta)/(\alpha + \beta) = \frac{1}{2}$.

We use notations (1.1) to (1.5).

LEMMA 4.1. *If* $W_0 = B_0$, *then* $E[(W_n - B_n)^{2k-1}] = 0$ *for each natural number* k. *If* $W_0 \neq B_0$, *then* $E[(W_n - B_n)^{2k-1}] \sim n^{k-\frac{1}{2}}(\log n)^{k-1}$, *for each natural number* k.

PROOF. If $W_0 = B_0$, then $W_n - B_n$ has a symmetric distribution. This proves the first claim. The second requires an induction. It is true for $k = 1$ by Equation (3.5) and Lemma 6.3. Suppose it is true for $1 \leqq k \leqq K$. The argument for Equation (3.1) implies

$$(4.1) \qquad x_{n+1}(2K + 1) = a_n(2K + 1)x_n(2K + 1) + b_n(2K + 1),$$

with

$$(4.2) \quad b_n(2K + 1) = \sum_{j=1}^{K} [\binom{2K+1}{2j}\delta^{2j} + [\binom{2K+1}{2j+1}/(s + \sigma n)]\delta^{2j+1}]x_n(2K + 1 - 2j).$$

By inductive assumption, $b_n(2K + 1) \sim n^{K-\frac{1}{2}}(\log n)^{K-1}$. Apply Lemma 6.5a. < >

LEMMA 4.2. *For each nonnegative integer* k,

$$\lim_{n\to\infty} (n \log n)^{-k} E[(W_n - B_n)^{2k}] = \mu(2k)$$

is finite, with $\mu(0) = 1$ *and*

$$\mu(2k+2) = \delta^2[1/(k+1)]\binom{2k+2}{2}\mu(2k). \tag{4.3}$$

PROOF. This is trivial when $k = 0$. Suppose it true for $0 \leqq k \leqq K$. It is then true when $k = K + 1$ by Equations (3.2), (3.3), and Lemma 6.5. <>

THEOREM 4.1. *As* n *tends to* ∞, *the distribution of* $(n \log n)^{-\frac{1}{2}}(W_n - B_n)$ *converges to normal with mean* 0 *and variance* $(\alpha - \beta)^2$.

PROOF. The moments of a normal distribution determine it uniquely, as is well known. Equation (4.3) implies $\mu(2k)$ is the $2k$th moment of a normal distribution with mean 0 and variance $(\alpha - \beta)^2$. Lemma 4.1 implies $E[(W_n - B_n)^{2k-1}] = o[(n \log n)^{k-\frac{1}{2}}]$, while Lemma 4.2 implies $E[(W_n - B_n)^{2k}] \approx \mu(2k)(n \log n)^k$. Now use the moment convergence criterion of Fréchet and Shohat (Loève, 1963, Theorem C, p. 185).

THEOREM 4.2. *If* $\epsilon > 0$, *then* $P\{|W_n - B_n| = O[n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\epsilon}]\} = 1$.

PROOF. Consider the martingale $Z_n : n \geqq 0$ defined by Equation (3.6). By (Doob, 1953, Theorem 1.1 (iii), p. 295), $Z_n^{2k} : n \geqq 0$ is an expectation-increasing martingale. Let

$$B_i = [\max_{0 \leqq n \leqq i} |Z_n| \geqq (\log \tfrac{1}{2}i)^{\frac{1}{2}+\epsilon}].$$

By Kolmogorov's inequality (Doob, 1953, Theorem 3.2, p. 314),

$$P(B_i) \leqq E(Z_i^{2k})/(\log \tfrac{1}{2}i)^{k+2k\epsilon}.$$

In particular, $P(B_{2^{j+1}}) = O(i^{-2k\epsilon})$ by Lemma 4.2.

Now $[|Z_n| \geqq (\log n)^{\frac{1}{2}+\epsilon}$ for infinitely many $n] \supset \{W_n - B_n \neq O[n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\epsilon}]\}$. And $[|Z_n| \geqq (\log n)^{\frac{1}{2}+\epsilon}$ for infinitely many $n] \subset [A_j$ for infinitely many $j]$, where $A_j = [|Z_n| \geqq (\log n)^{\frac{1}{2}+\epsilon}$ for some n with $2^j \leqq n < 2^{j+1}] \subset B_{2^{j+1}}$. If $k > (2\epsilon)^{-1}$, then $\sum P(A_j) < \infty$, so $P[A_j$ for infinitely many $j] = 0$ by the Borel-Cantelli lemma. <>

COROLLARY 4.1. *With probability* 1, $\lim_{n\to\infty} (W_n + B_n)^{-1}W_n = \frac{1}{2}$.

THEOREM 4.3. *With probability* 1,

$$\limsup_{n\to\infty} n^{-\frac{1}{2}}(W_n - B_n) = \infty$$

and

$$\liminf_{n\to\infty} n^{-\frac{1}{2}}(W_n - B_n) = -\infty.$$

PROOF. When n increases by 1, there is a change of δ in $(W_n - B_n)$. Hence, with Z_n defined by (3.6),

$$|Z_{n+1} - Z_n| \leqq \delta[1 + |W_n - B_n|/(s + \sigma n)]/[\textstyle\prod_{\nu=0}^n a_\nu(1)].$$

Now $\prod_{\nu=0}^n a_\nu(1) \sim n^\rho$ and $|W_n - B_n| \leqq W_n + B_n = s + \sigma n$, so ess. sup $|Z_{n+1} - Z_n| \leqq O(n^{-\rho}) = O(1)$. By (Doob, 1953, Theorem 1 (iv), p. 320), Z_n converges to a finite limit almost everywhere on [lim sup $Z_n < \infty$]. Consequently, $(n \log n)^{-\frac{1}{2}}(W_n - B_n) = o(Z_n)$ converges to 0 almost everywhere on [lim sup $Z_n < \infty$], and

$$P(\limsup Z_n < \infty) \leqq \lim_{\epsilon\to 0}\lim_{n\to\infty} P[|W_n - B_n| \leqq \epsilon(n \log n)^{\frac{1}{2}}] = 0$$

by Theorem 4.1. Similarly, $P(\liminf Z_n > -\infty) = 0$. <>

Since $W_n - B_n$ changes by δ at each move, $W_n - B_n$ occupies each of its possible positions, namely $(W_0 - B_0)$ plus an integer times δ, infinitely often with probability 1. See Theorem 5.5 for further results. Thus $n^{-\frac{1}{2}}(W_n - B_n)$ converges with probability 0 and $(n \log n)^{-\frac{1}{2}}(W_n - B_n)$ converges with probability 0.

Let I be a countable set, $\{p(i, j) : i \varepsilon I, j \varepsilon I\}$ a stochastic matrix, $\{p(n; i, j) : i \varepsilon I, j \varepsilon I\}$ the nth power of this matrix. Let $I(n, i) = \{j : j \varepsilon I, p(n; i, j) > 0\}$. Let $\mathfrak{X}$ be the space of sequences of elements of I; define ξ_n for $n \geqq 0$ as the function on $\mathfrak{X}$ sending a sequence into its nth term. Let T map $\mathfrak{X}$ into itself so that $\xi_n \circ T = \xi_{n+1}$.

Let $\mathfrak{A}^{(n)}$ be the σ-field spanned by $\xi_j : j \geqq n$; the tail σ-field $\mathfrak{A}^{(\infty)}$ of $\mathfrak{X}$ is $\bigcap_{n=0}^{\infty} \mathfrak{A}^{(n)}$ Plainly, $T^{-1} \mathfrak{A}^{(\infty)} = \mathfrak{A}^{(\infty)}$. For $i \varepsilon I$, let P_i be the unique probability on $(\mathfrak{X}, \mathfrak{A})$ making $\{\xi_n : n \geqq 0\}$ a Markov chain with initial state i and stationary transition probabilities $\{p(i, j) : i \varepsilon I, j \varepsilon I\}$.

The state $i \varepsilon I$ is called *merging* if two independent Markov chains with the state space I, stationary transition mechanism $p(\cdot, \cdot)$, and initial state i, meet infinitely often with probability 1.

LEMMA 4.3. *Let the state $i \varepsilon I$ be merging. Suppose j and k are in $I(n, i)$ for some n. Then two independent Markov chains with the state space I, stationary transition mechanism $p(\cdot, \cdot)$, and initial states j and k respectively, meet infinitely often with probability* 1.

PROOF. Clear. <>

LEMMA 4.4. *Suppose $i \varepsilon I$ is merging, and $j \varepsilon I(n, i)$, $k \varepsilon I(n, i)$, for some n. Then $P_j = P_k$ on $\mathfrak{A}^{(\infty)}$.*

PROOF. (After Doeblin, 1938.) Let $\{X_n : n \geqq 0\}$ and $\{Y_n : n \geqq 0\}$ be independent Markov chains with the state space I, stationary transition mechanism $p(\cdot, \cdot)$, and initial states $j \neq k$ respectively. Let

$$\begin{aligned}\tilde{X}_n &= X_n \quad \text{provided } X_m \neq Y_m \quad \text{for } 1 \leqq m \leqq n\\ &= Y_n \quad \text{otherwise.}\end{aligned}$$

Then $\{\tilde{X}_n : n \geqq 0\}$ is distributed like $\{X_n : n \geqq 0\}$; so for $A \varepsilon \mathfrak{A}^{(\infty)}$, by Lemma 4.3,

$$\begin{aligned}P_j(A) &= \text{Prob. }\{[X_n : n \geqq 0] \varepsilon A\}\\ &= \text{Prob. }\{[\tilde{X}_n : n \geqq 0] \varepsilon A\}\\ &= \text{Prob. }\{[Y_n : n \geqq 0] \varepsilon A\}\\ &= P_k(A). \text{ <>}\end{aligned}$$

LEMMA 4.5. *If $i \varepsilon I$ is merging, then P_i takes only the values* 0 *and* 1 *on* $\mathfrak{A}^{(\infty)}$.

PROOF, Let $j_\nu \varepsilon I(\nu, i)$ and $k_\nu \varepsilon I(\nu, i)$ for $1 \leqq \nu \leqq n$. If $A \varepsilon \mathfrak{A}^{(\infty)}$, then

$$P_i(A \mid \xi_\nu = j_\nu, 1 \leqq \nu \leqq n) = P_{j_n}(T^{-n}A) = P_{k_n}(T^{-n}A)$$
$$= P_i(A \mid \xi_\nu = k_\nu, 1 \leqq \nu \leqq n),$$

where the first and third equality result from the Markov property, and the second is implied by Lemma 4.4. Under P_i, the set A is independent of all sets of the form $[\xi_\nu = j_\nu, 1 \leqq \nu \leqq n]$, so of all $\mathfrak{A}^{(0)}$ sets, and in particular of itself. < >

LEMMA 4.6. *Let $(W_n, B_n) : n \geqq 0$ and $(W_n^*, B_n^*) : n \geqq 0$ be independent Friedman urns on $(\Omega, \mathfrak{F}, P)$, with common parameters $\alpha = 3\beta > 0$, and $W_0 = W_0^*$, $B_0 = B_0^*$. Then $W_n = W_n^*$ for infinitely many n, with probability* 1.

PROOF. Let $\tilde{\mathfrak{F}}_n$ be the σ-field spanned by $W_j, B_j, W_j^*, B_j^* : 1 \leqq j \leqq n$, and $\tilde{Z}_n = [(W_n - B_n) - (W_n^* - B_n^*)]/\prod_{\nu=0}^{n-1} a_\nu(1)$. By Equation (3.4), $\{\tilde{Z}_n, \tilde{\mathfrak{F}}_n : 0 \leqq n < \infty\}$ is a martingale. We claim that with probability 1:

$$\limsup_{n\to\infty} \tilde{Z}_n = \infty, \tag{4.4}$$

$$\liminf_{n\to\infty} \tilde{Z}_n = -\infty. \tag{4.5}$$

Suppose by way of contradiction that (4.4) fails. As in the proof of Theorem 4.3, this entails: $[(W_n - B_n) - (W_n^* - B_n^*)]\,(n \log n)^{-\frac{1}{2}}$ converges to 0 with positive probability. This contradicts Theorem 4.1. Since $(W_n - B_n) - (W_n^* - B_n^*)$ changes by 0 or $\pm\delta$ as n increases by 1, the result follows. < >

THEOREM 4.4. *The tail σ-field of $(W_n - B_n) : n \geqq 0$ is trivial.*

PROOF. Apply Lemmas 4.5 and 4.6. < >

In particular, $(W_n - B_n)(n \log n)^{-\frac{1}{2}}$ cannot converge in probability. However, we do not have a guess as to the true asymptotic order of magnitude of $W_n - B_n$.

5. The case $\rho < \frac{1}{2}$.

Let $(W_n, B_n): n \geqq 0$ be a Friedman urn on $(\Omega, \mathfrak{F}, P)$ with parameters $\alpha \geqq 0$, $\beta > 0$. Throughout this section, unless specifically stated otherwise, we make the

ASSUMPTION. $(\alpha - \beta)/(\alpha + \beta) < \frac{1}{2}$.

If $\rho = 0$, then $W_n = W_0 + n\alpha$; the only interest, therefore, is in $\rho \neq 0$.

The proofs of Lemmas 5.1, 5.2, and Theorems 5.1, 5.2, 5.3, 5.4, and Corollary 5.1 are omitted. They parallel the arguments for the corresponding results in Section 4, appealing to Lemma 6.6 instead of Lemma 6.5. We use notations (1.1) to (1.5).

LEMMA 5.1. *For natural number k: If $W_0 = B_0$, then $E[(W_n - B_n)^{2k-1}] = 0$. If $W_0 \neq B_0$ but $(s + \delta)/\sigma$ is a negative integer, then $E[(W_n - B_n)^{2k-1}] = O(n^{k+\rho-1})$. If $W_0 \neq B_0$ and $(s + \delta)/\sigma$ is not a negative integer, then $E[(W_n - B_n)^{2k-1}] \sim n^{k+\rho-1}$.*

LEMMA 5.2. *For each nonnegative integer k,*

$$\lim_{n\to\infty} n^{-k} E[(W_n - B_n)^{2k}] = \mu(2k)$$

is finite, with $\mu(0) = 1$ and

$$\mu(2k + 2) = (1 - 2\rho)^{-1}\delta^2[1/(k + 1)]\binom{2k+2}{2}\mu(2k).$$

THEOREM 5.1. *As n tends to ∞, the distribution of $n^{-\frac{1}{2}}(W_n - B_n)$ converges to normal with mean 0 and variance $(1 - 2\rho)^{-1}(\alpha - \beta)^2$.*

THEOREM 5.2. *If $\epsilon > 0$, then $P\{|W_n - B_n| = O[n^{\frac{1}{2}}(\log n)^{\epsilon}]\} = 1$ for $0 < \rho < \frac{1}{2}$.*

COROLLARY 5.1. *With probability 1, $\lim_{n\to\infty} (W_n + B_n)^{-1}W_n = \frac{1}{2}$, when $0 < \rho < \frac{1}{2}$.*

We guess, but cannot prove, that

$$(5.1) \qquad P\{\limsup_{n\to\infty} [(W_n - B_n)/(2n \log\log n)^{\frac{1}{2}}] = \hat{\sigma}\} = 1$$

and

$$(5.2) \qquad P\{\liminf_{n\to\infty} [(W_n - B_n)/(2n \log\log n)^{\frac{1}{2}}] = -\hat{\sigma}\} = 1,$$

where $\hat{\sigma}^2 = (1 - 2\rho)^{-1}(\alpha - \beta)^2$, for any $\rho < \frac{1}{2}$.

THEOREM 5.3. *If $0 < \rho < \frac{1}{2}$, then*

$$P[\limsup_{n\to\infty} n^{-\rho}(W_n - B_n) = \infty] = P[\liminf_{n\to\infty} n^{-\rho}(W_n - B_n) = -\infty] = 1.$$

THEOREM 5.4. *If $0 < \rho < \frac{1}{2}$, the tail σ-field of $(W_n, B_n): n \geqq 0$ is trivial.*

LEMMA 5.3. *Let $(\Omega, \mathfrak{F}, P)$ be a probability triple, $\mathfrak{A}_n : n \geqq 0$ a nondecreasing sequence of sub-σ-fields of $\mathfrak{F}$, $A_n \in \mathfrak{A}_n$ for $n \geqq 0$. The set of $\omega \in \Omega$ which are members of an infinite number of the A_n's differs by a P-null set from the set where $\sum P(A_{n+1} \mid \mathfrak{A}_n)$ diverges.*

PROOF. This is an obvious modification of Levy's conditional form of the Borel-Cantelli lemmas (Doob, 1953, Corollary 2, p. 324). <>

THEOREM 5.5. *If $0 < \rho \leqq \frac{1}{2}$, then*

$$P\{\liminf_{n\to\infty} [|W_n - B_n|/(2n \log\log n)^{\frac{1}{2}}] \geqq \alpha - \beta\} = 1.$$

If τ is the least $n > 0$ with $W_n - B_n = W_0 - B_0$, then $E(\tau) = \infty$, although $P(\tau < \infty) = 1$ by Theorems 4.3 and 5.3.

PROOF. Realize the process $(W_n, B_n): n \geqq 0$ on a probability triple $(\Omega, \mathfrak{F}, P)$ which supports a sequence $U_n : n \geqq 0$ of independent random variables having common distribution uniform over [0, 1], independent of $(W_n, B_n): n \geqq 0$. We will define a new process $T_n : n \geqq 0$ on $(\Omega, \mathfrak{F}, P)$ with the properties:

$$(5.3) \qquad T_0 = |W_0 - B_0|;$$

$$(5.4) \qquad T_n \leqq |W_n - B_n|;$$

(5.5) $\{T_n : n \geqq 0\}$ is distributed like $\{T_0 + X_1 + \cdots + X_n : n \geqq 0\}$, where the X_i's are independent and $\pm\delta$ with probability $\frac{1}{2}$ each.

Let $|W_0 - B_0|$ have remainder δ_0 on division by δ, and put $\delta_1 = \delta - \delta_0$. Let R_n be the indicator function of the event

$$[U_n \leqq (s + \sigma n)/(s + \sigma n + |W_n - B_n|)].$$

Define T_0 by (5.3). For $n \geqq 1$, we define T_n inductively. If $|W_n - B_n| \geqq \delta$, then $T_{n+1} = T_n - \delta$ on $|W_{n+1} - B_{n+1}| = |W_n - B_n| - \delta$, while $T_{n+1} = T_n +$

$\delta R_n - \delta(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = |W_n - B_n| + \delta$. Suppose $\delta_0 \neq 0$. If $|W_n - B_n| = \delta_0$, then $T_{n+1} = T_n - \delta$ on $|W_{n+1} - B_{n+1}| = \delta_1$ and $T_{n+1} = T_n + \delta R_n - \delta(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = \delta_0 + \delta$. If $|W_n - B_n| = \delta_1$, then $T_{n+1} = T_n - \delta$ on $|W_{n+1} - B_{n+1}| = \delta_0$ and $T_{n+1} = T_n + \delta R_n - \delta(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = \delta_1 + \delta$. If $\delta_0 = 0$, and $|W_n - B_n| = 0$, then $T_{n+1} = T_n + \delta 1_{[U_n \leqq \frac{1}{2}]} - \delta 1_{[U_n > \frac{1}{2}]}$. Plainly, (5.4) is satisfied. The conditional distribution of T_{n+1} given $W_j, B_j : 1 \leqq j \leqq n$ and $U_j : 0 \leqq j \leqq n - 1$ is: $T_{n+1} = T_n \pm \delta$ with probability $\frac{1}{2}$ each. Hence $\{T_n : n \geqq 0\}$ is a Markov chain with the transition rule just given, and therefore satisfies (5.5). The waiting time for $\{T_n : n \geqq 0\}$ to achieve a value less than T_0 is known to have infinite mean (Feller, 1960, p. 256). This proves: if $|W_0 - B_0| \geqq \delta$, the mean waiting time for $|W_n - B_n|$ to reach $|W_0 - B_0| - \delta$ in infinite. Now the last result follows; for there is positive probability that $|W_1 - B_1| = |W_0 - B_0| + \delta$, and then the previous remark applies. The Law of Iterated Logarithm (Feller, 1960, pp. 191 ff) implies:

$$P(\limsup_{n\to\infty} [T_n/(2n \log\log n)^{\frac{1}{2}}] = \delta) = 1,$$

which proves the first result. < >

Theorems 4.3 and 5.3 prove that $W_n - B_n$ enters each of its possible states infinitely often with probability 1. Theorem 5.5 proves that the return times are stochastically larger than the return times for a coin-tossing game, and in particular have infinite mean. Is the mean waiting time for $|W_n - B_n|$ to exceed $|W_0 - B_0|$ finite?

THEOREM 5.6. *If* $\rho < 0$, *then*

(i) $P\{\limsup_{n\to\infty} [|W_n - B_n|/(2n \log\log n)^{\frac{1}{2}}] \leqq |\delta|\} = 1$, *and*

(ii) $W_n - B_n - (W_0 - B_0)$ *visits each multiple of* δ *infinitely often with probability* 1.

PROOF. Realize $(W_n, B_n) : n \geqq 0$ on a triple $(\Omega, \mathfrak{F}, P)$ which supports a sequence $\{U_n : n \geqq 0\}$ of independent random variables having common distribution uniform over $[0, 1]$, independent of $(W_n, B_n) : n \geqq 0$.

We will construct a new process $\{T_n : n \geqq 0\}$ on $(\Omega, \mathfrak{F}, P)$ satisfying:

$$T_0 = |W_0 - B_0| + |\delta|; \tag{5.6}$$

$$T_n \geqq |W_n - B_n|; \tag{5.7}$$

(5.8) $\{T_n - T_0 : n \geqq 0\}$ is distributed like $\{|X_1 + \cdots + X_n| : n \geqq 0\}$, where the X_i's are independent and $\pm\delta$ with probability $\frac{1}{2}$ each.

Let $|W_0 - B_0|$ have remainder δ_0 on division by $|\delta|$, and put $\delta_1 = |\delta| - \delta_0$. Let R_n be the indicator function of the event

$$[U_n \leqq (s + \sigma n)/(s + \sigma n + |W_n - B_n|)].$$

Define T_0 by (5.6). For $n \geqq 1$, define T_n inductively. If $T_n = T_0$, then $T_{n+1} = T_n + |\delta|$. For $T_n > T_0$, proceed thus. If $|W_n - B_n| \geqq |\delta|$, then $T_{n+1} = T_n + |\delta|$ on $|W_{n+1} - B_{n+1}| = |W_n - B_n| + |\delta|$, while $T_{n+1} = T_n + |\delta| R_n - |\delta|(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = |W_n - B_n| - |\delta|$. Suppose $\delta_0 \neq 0$.

If $|W_n - B_n| = \delta_0$, then $T_{n+1} = T_n + |\delta|$ on $|W_{n+1} - B_{n+1}| = \delta_0 + |\delta|$, and $T_{n+1} = T_n + |\delta|R_n - |\delta|(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = \delta_1$. If $|W_n - B_n| = \delta_1$, then $T_{n+1} = T_n + |\delta|$ on $|W_{n+1} - B_{n+1}| = \delta_1 + |\delta|$ and $T_{n+1} = T_n + |\delta|R_n - |\delta|(1 - R_n)$ on $|W_{n+1} - B_{n+1}| = \delta_0$. If $\delta_0 = 0$, and $|W_n - B_n| = 0$, then

$$T_{n+1} = T_n + |\delta|1_{[U_n \leqq \frac{1}{2}]} - |\delta|1_{[U_n > \frac{1}{2}]}.$$

Plainly, T_n satisfies (5.7) and is of the form $|W_0 - B_0| + k|\delta|$ with $k \geqq 1$.

The conditional distribution of T_{n+1} given W_j, $B_j : 1 \leqq j \leqq n$ and $U_j : 0 \leqq j \leqq n - 1$ is: if $T_n = |W_0 - B_0| + |\delta|$, then $T_{n+1} = T_n + |\delta|$, while if $T_n > |W_0 - B_0| + |\delta|$ then $T_{n+1} = T_n \pm |\delta|$ with probability $\frac{1}{2}$ each. Hence $\{T_n : n \geqq 0\}$ is a Markov chain with the transition rule just given, and (5.8) holds.

In particular (Feller, 1960, pp. 191 ff), for each $\epsilon > 0$, with probability 1,

$$T_n \leqq (|\delta| + \epsilon)(2n \log \log n)^{\frac{1}{2}}$$

for all large n. This proves result (i).

Moreover, $T_n = T_0$ for infinitely many n with probability 1 (Feller, 1960, p. 288). Hence $|W_n - B_n| \leqq |W_0 - B_0| + |\delta|$ for infinitely many n with probability 1.

Let d be a natural number. Let $\tau_0 = 0$ and τ_{k+1} be the least $n > \tau_k + 3d$ with $|W_n - B_n| \leqq |W_0 - B_0| + |\delta|$. Let $\mathcal{A}_k = \{A : A \varepsilon \mathfrak{F} \text{ and } A \cap [\tau_k = j] \varepsilon \mathfrak{F}_j\}$. Let $A_{k+1} = [W_{j+1} = W_j + \alpha \text{ for } \tau_k \leqq j < \tau_k + d;\ W_{j+1} = W_j + \beta \text{ for } \tau_k + d \leqq j < \tau_k + 3d]$. Then $A_k \varepsilon \mathcal{A}_k$ and there is an $\epsilon > 0$ with $P(A_{k+1} \mid \mathcal{A}_k) \geqq \epsilon$, all k. Apply Lemma 5.3. <>

We do not know whether the mean return times are finite.

COROLLARY 5.2. *If* $\rho < 0$, *then* $(W_n + B_n)^{-1}W_n$ *converges to* $\frac{1}{2}$ *with probability* 1.

LEMMA 5.4. *Let* $(W_n, B_n) : n \geqq 0$ *and* $(W_n^*, B_n^*) : n \geqq 0$ *be two independent Friedman urns on* $(\Omega, \mathfrak{F}, P)$ *with common parameters* $0 \leqq \alpha < \beta$ *and* $W_0 = W_0^*$, $B_0 = B_0^*$. *Then* $W_n = W_n^*$ *for infinitely many n with probability* 1.

PROOF. $W_{n+1} - W_{n+1}^*$ is either $W_n - W_n^*$ or $W_n - W_n^* + |\delta|$ or $W_n - W_n^* - |\delta|$. Given W_j, $W_j^* : 1 \leqq j \leqq n$, the last two values are taken with conditional probabilities

$$[B_n/(s + \sigma n)] \cdot [W_n^*/(s + \sigma n)] \quad \text{and} \quad [W_n/(s + \sigma n)] \cdot [B_n^*/(s + \sigma n)]$$

respectively, whose ratio is

$$(B_n/B_n^*)/(W_n/W_n^*).$$

If $W_n > W_n^*$, then $B_n < B_n^*$, and this ratio is less than 1; if $W_n < W_n^*$, then $B_n > B_n^*$, and this ratio is greater than 1. It is therefore possible to define a process $T_n : n \geqq 0$ on $(\Omega, \mathfrak{F}, P)$ such that

(i) $T_0 = W_0 - W_0^* = 0$;
(ii) $T_{n+1} = T_n$, or $T_n - |\delta|$, or $T_n + |\delta|$;
(iii) $T_{n+1} = T_n$ if and only if $W_{n+1} - W_{n+1}^* = W_n - W_n^*$;

(iv) $P(T_{n+1} = T_n \pm |\delta| \mid W_1 \cdots W_n, W_1^* \cdots W_n^*, T_1 \cdots T_n, T_{n+1} \neq T_n) = \frac{1}{2}$;
(v) $|T_n| \geqq |W_n - W_n^*|$.

We omit the construction of T_n; see the proofs of Theorems 5.5 and 5.6. From (iii),

$$(5.9) \quad P(T_{n+1} \neq T_n \mid W_1 \cdots W_n, W_1^* \cdots W_n^*, T_1 \cdots T_n) = (B_n W_n^* + W_n B_n^*)/(s + \sigma n)^2.$$

Let us pause for a moment to study the function $f(x, y) = x(1 - y) + y(1 - x)$ on $0 \leqq x, y \leqq 1$. Since $\partial f/\partial x = 1 - 2y$,

$$\begin{aligned} f(x, y) &\geqq f(0, y) = y, && \text{for } y < \tfrac{1}{2}, \\ &= \tfrac{1}{2}, && \text{for } y = \tfrac{1}{2}, \\ &\geqq f(1, y) = 1 - y, && \text{for } y > \tfrac{1}{2}. \end{aligned}$$

In particular, $f(x, y) \geqq \min [\frac{1}{2}, y, 1 - y]$. Apply this inequality to the right side of (5.9); it is bounded below by

$$\min [\tfrac{1}{2}, W_n/(s + \sigma n), 1 - (W_n/(s + \sigma n))].$$

The sum over n of these quantities is plainly $+\infty$ with probability 1, so $T_{n+1} \neq T_n$ for infinitely many n with probability 1 by Lemma 5.3. Let τ_k be the kth n with $T_{n+1} \neq T_n$, and $X_k = T_{\tau_k+1} - T_{\tau_k}$. Then $X_k : k \geqq 0$ are independent, $\pm|\delta|$ with probability $\frac{1}{2}$ each. Now $\{X_1 + \cdots + X_n : n \geqq 0\}$ is a subsequence of $\{T_n : n \geqq 0\}$, so $T_n = 0$ for infinitely many n with probability 1. But $|W_n - W_n^*| \leqq T_n$. < >

THEOREM 5.7. *If $\rho < 0$, the tail σ-field of $(W_n, B_n): n \geqq 0$ is trivial.*

PROOF. Apply Lemmas 4.5 and 5.4. < >

These results exhibit a substantial difference between urn processes with $\rho > \frac{1}{2}$ and $\rho \leqq \frac{1}{2}$. What is the intuitive meaning of the $\frac{1}{2}$? Is there a qualitative difference between urns with $\rho > 0$ and $\rho < 0$?

Let $(W_n, B_n): n \geqq 0$ be a Friedman urn on $(\Omega, \mathfrak{F}, P)$, with parameters $\alpha = 0$ and $\beta = 1$. Blackwell asked whether the distribution D of $W_n - W_{n-1} : n \geqq 1$ is singular with respect to the distribution F of a sequence of independent random variables, taking the values 0 and 1 with probability $\frac{1}{2}$ each. Using the results of this section, it is not hard to see that the answer is yes.

Restrict D and F to the tail σ-field. Since both are then trivial, they must be singular or equal. The latter possibility is ruled out by Theorem 5.1, the central limit theorem of De Moivre, and the standard martingale argument of Lemma 5.5 below.

LEMMA 5.5. *Let P_1 and P_2 be probabilities on a σ-field Σ of subsets of a set $\mathfrak{X}$. Let $\Sigma^{(n)}: n \geqq 1$ be sub-σ-fields of Σ which shrink to the σ-field $\Sigma^{(\infty)}$. If P is a probability on Σ, let $P^{(n)}$ be its restriction to $\Sigma^{(n)}$, $1 \leqq n \leqq \infty$. If $P_1^{(\infty)} = P_2^{(\infty)}$, then $\|P_1^{(n)} - P_2^{(n)}\| \to 0$ as $n \to \infty$.*

PROOF. Let $P = \frac{1}{2}(P_1 + P_2)$. Then

$$\|P_1^{(n)} - P_2^{(n)}\| = \int_{\mathfrak{X}} |[dP_1^{(n)}/dP^{(n)}] - [dP_2^{(n)}/dP^{(n)}]|\, dP.$$

Since $dP_i^{(n)}/dP^{(n)}$ is the P-conditional expectation of dP_i/dP given $\Sigma^{(n)}$, it converges in $\mathfrak{L}^1(P)$-norm to the P-conditional expectation of dP_i/dP given Σ^∞. The latter function, being $dP_i^{(\infty)}/dP^{(\infty)}$, is equal to 1 with P-probability 1. < >

We still do not know, for example, whether the distribution of $W_n - W_{n-1}$: $n \geqq 1$ is equivalent to the distribution of $W_{n+1} - W_n : n \geqq 1$ if $W_0 = B_0 = 1$.

6. Some difference equations. In this section, we solve the difference equations that appear in Section 3, 4, 5 and estimate the order of magnitude of the solutions. Let x_n, a_n, b_n be real numbers for $n \geqq 0$ with

$$x_{n+1} = a_n x_n + b_n . \tag{6.1}$$

LEMMA 6.1. *If x_n, a_n, b_n are real numbers satisfying* (6.1) *for $n \geqq 0$, then*

$$x_{n+1} = x_0 \prod_{\nu=0}^{n} a_\nu + \sum_{j=0}^{n-1} b_j \prod_{\nu=j+1}^{n} a_\nu + b_n$$
$$= \Big(\prod_{\nu=0}^{n} a_\nu\Big)\Big(x_0 + \sum_{j=0}^{n} b_j \prod_{\nu=0}^{j} a_\nu^{-1}\Big),$$

the second form being valid when $a_\nu \neq 0$ for $0 \leqq \nu \leqq n$.

PROOF. Direct verification. < >

LEMMA 6.2. *If A and B are real numbers, then* $\Gamma(A + n)/\Gamma(B + n) \approx n^{A-B}$

PROOF. Use Stirling's formula. < >

Suppose $b > 0$, $c > 0$, a is real and

$$a_n = 1 + [a/(b + cn)] \quad \text{for} \quad n \geqq 0. \tag{6.2}$$

LEMMA 6.3. *If $\{a_n\}$ is defined by* (6.2), *with $b > 0$, $c > 0$, and $(a + b)/c$ not a negative integer, then*

$$\prod_{\nu=0}^{n} a_\nu \approx [\Gamma(b/c)/\Gamma((a + b)/c)]n^{a/c}.$$

PROOF. Since

$$\prod_{\nu=0}^{n} a_\nu = [\Gamma(b/c)/\Gamma((a + b)/c)]$$
$$\cdot\{\Gamma[((a + b)/c) + n + 1]/\Gamma[(b/c) + n + 1]\},$$

Lemma 6.2 applies. < >

LEMMA 6.4. *If $\{a_n\}$ is defined by* (6.2) *with $a > 0$; and $b_n = O(n^d)$ with $d < c^{-1}a - 1$; and $\{x_n\}$ satisfies* (6.1); *then* $\lim_{n\to\infty} x_n \prod_{\nu=0}^{n} a_\nu^{-1} = x_0 + \sum_{j=0}^{\infty} b_j \prod_{\nu=0}^{j} a_\nu^{-1}$, *the series converging absolutely.*

PROOF. Lemmas 6.1 and 6.3. < >

LEMMA 6.5. *If $\{a_n\}$ is defined by* (6.2) *with $a > 0$; and $b_n \approx B(n \log n)^d$ with $B \neq 0$ and $d = c^{-1}a - 1$; and $\{x_n\}$ satisfies* (6.1): *then $x_n \approx (d + 1)^{-1}B (n \log n)^{d+1}$.*

PROOF. We can replace x_n, a_n, b_n by x_{n+k}, a_{n+k}, b_{n+k} if necessary, replacing b by $b + ck$, and therefore assume without loss of generality that $(a + b)/c > 0$ and $a_n > 0$ for all n. Now

$$\sum_{j=1}^{n} [(\log j)^d/j] \approx \int_1^n [(\log x)^d/x]\, dx$$
$$= (\log n)^{d+1}/(d + 1) \qquad \text{for} \quad d > -1,$$

the first relation holding because $x \to (\log x)^d/x$ is ultimately decreasing. Apply Lemmas 6.1 and 6.3. < >

LEMMA 6.5a. *If* $\{a_n\}$ *is defined by* (6.2) *with* $c^{-1}a > \frac{1}{2}$; *and* $b_n \approx Bn^{d-\frac{1}{2}}(\log n)^{d-1}$ *with* $B \neq 0$ *and* $d = c^{-1}a - \frac{1}{2}$; *and* $\{x_n\}$ *satisfies* (6.1); *then* $x_n \approx d^{-1}\ B_n^{d+\frac{1}{2}}\ (\log n)^d$.

PROOF. As in Lemma 6.5. < >

LEMMA 6.6. *If* $\{a_n\}$ *is defined by* (6.2); *and* $b_n \approx Bn^d$ *with* $B \neq 0$ *and* $d > c^{-1}a - 1$; *and* $\{x_n\}$ *satisfies* (6.1) *then* $x_n \approx [B/(d - (a/c) + 1)]n^{d+1}$.

PROOF. As in Lemma 6.5, using the observation $\sum_{j=1}^{n} j^C \approx n^{C+1}/(C + 1)$, for $C > -1$. < >

REFERENCES

BERNSTEIN, SERGE (1940). Sur un problème du schéma des urnes à composition variable. *C. R. Dokl. Acad. Sci. URSS.* **28** 5–7.

BLACKWELL, DAVID and KENDALL, DAVID (1964). The Martin boundary for Pólya's Urn scheme. (Abstract, *Ann. Math. Statist.* **35** 460); paper to appear in *J. Appl. Prob.*

DE FINETTI, BRUNO (1931). Funzione caratteristica di un fenomeno aleatorio. *Atti Accad. Naz. Lincei. Memo. Cl. Sci. Fis. Mat. Nat. Sez.* 6, **4** 251–299.

DOEBLIN, W. (1938). Sur deux problèmes de M. Kolmogoroff concernant les chaines dénombrables. *Bull. Soc. Math. France* **66** 210–220.

DONSKER, MONROE D. (1951). An invariance principle for certain probability limit theorems *Mem. Amer. Math. Soc.* **6** 1–12.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

FELLER, WILLIAM (1960). *An Introduction to Probability Theory and Its Applications.* **1** (2nd ed.). Wiley, New York.

FRÉCHET, MAURICE (1943). *Les Probabilités Associées à un Système d'Évenements Compatibles et Dépendants*. Hermann, Paris.

FREEDMAN, DAVID A. (1962). Invariants under mixing which generalize de Finetti's theorem. *Ann. Math. Statist.* **33** 916–923.

FREEDMAN, DAVID (1963). L'Urne de Bernard Friedman. *C. R. Acad. Sci. Paris.* **257** 3809.

FRIEDMAN, BERNARD (1949). A simple urn model. *Comm. Pure Appl. Math.* **2** 59–70.

HEWITT, EDWIN and SAVAGE, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501.

LOÈVE, MICHEL (1963). *Probability Theory* (3rd ed.). Van Nostrand, Princeton.

PÓLYA, G. (1931). Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré.* **1** 117–161.

SAVKEVITCH, V. (1940). Sur le schéma des urnes à composition variable. *C. R. Dokl. Acad. Sci. URSS.* **28** 8–12.