PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 56, April 1976

ON ASYMPTOTIC BEHAVIOR FOR THE HAWKINS RANDOM SIEVE

C. C. HEYDE

ABSTRACT. This paper is concerned with the Hawkins random sieve which is a probabilistic analogue of the sieve of Eratosthenes. Analogues of the prime number theorem and Mertens' theorem have previously been obtained for this sieve by classical probabilistic methods. In the present paper, sharper results akin to the Riemann hypothesis are obtained by a more elegant martingale approach.

The following random sieve procedure introduced by Hawkins [1], [2] is a stochastic analogue of the sieve of Eratosthenes: Let $A_1 = \{2,3,4,5,6,\dots\}$.

Stage 1. Put $X_1 = \min A_1$. From the set $A_1 \backslash \{X_1\}$ each number in turn is (independently of the others) deleted with probability X_1^{-1} or not deleted with probability $1 - X_1^{-1}$. The set of elements of $A_1 \backslash \{X_1\}$ which remain is denoted by A_2.

Stage n. Put $X_n = \min A_n$. From the set $A_n \backslash \{X_n\}$ each number in turn is (independently of the others) deleted with probability X_n^{-1} or not deleted with probability $1 - X_n^{-1}$. The set of elements of $A_n \backslash \{X_n\}$ which remain is denoted by A_{n+1}.

Define

$$Y_n = \prod_{1 \leqslant k \leqslant n} (1 - X_k^{-1})^{-1}.$$

Wunderlich [5], [6] has obtained the results

$$\lim_{n\to\infty} (n \log n)^{-1} X_n = 1 \text{ a.s.,} \quad \text{and} \quad \lim_{n\to\infty} (\log n)^{-1} Y_n = 1 \text{ a.s.}$$

which are analogues of the prime number theorem and Mertens' theorem respectively. These results have been obtained by classical probabilistic methods involving detailed estimation of moments. In this paper we shall obtain sharper results, akin to the Riemann hypothesis, using a more elegant martingale approach. These results are given in the following theorem and are different in character from those obtained for a diffusion analogue by Williams [4].

THEOREM. (i) $\lim_{n\to\infty} (\log \log n)^{-1}(n^{-1}X_n - \log n) = 1$ *a.s.*,
(ii) $\lim_{n\to\infty} (\log \log n)^{-1}(Y_n - \log n) = 1$ *a.s.*

Received by the editors March 11, 1975.

AMS (MOS) subject classifications (1970). Primary 60F15, 10H30; Secondary 60J05.

Key words and phrases. Random sieve, prime numbers, Riemann hypothesis, martingales, strong laws.

PROOF. We first note, following Williams [4], that the process $\{(X_n, Y_n), n \geqslant 1\}$ is Markovian with $X_1 = Y_1 = 2$, and

$$P(X_{n+1} - X_n = j | X_n, Y_n) = Y_n^{-1}(1 - Y_n^{-1})^{j-1}, \qquad j \geqslant 1,$$

so that in particular

$$E(X_{n+1} | X_n, Y_n) = X_n + Y_n \tag{1}$$

and

$$E[(X_{n+1} - X_n - Y_n)^2 | X_n, Y_n] = Y_n^2 - Y_n. \tag{2}$$

Next note that

$$\begin{aligned} Y_n - \sum_{j=1}^{n} Y_j X_j^{-1} &= Y_n(1 - X_n^{-1}) - \sum_{j=1}^{n-1} Y_j X_j^{-1} \\ &= Y_{n-1} - \sum_{j=1}^{n-1} Y_j X_j^{-1}, \end{aligned}$$

and hence, by continued reduction,

$$Y_n - \sum_{j=1}^{n} Y_j X_j^{-1} = Y_1 - Y_1 X_1^{-1} = 1, \qquad \forall n \geqslant 1. \tag{3}$$

Now observe that the process

$$\left\{ \sum_{k=2}^{n} [X_k Y_k^{-1} - E(X_k Y_k^{-1} | X_{k-1}, Y_{k-1})], n > 2 \right\}$$

is a zero-mean martingale and, using (1),

$$\begin{aligned} &\sum_{k=2}^{n} [X_k Y_k^{-1} - E(X_k Y_k^{-1} | X_{k-1}, Y_{k-1})] \\ &\quad = \sum_{k=2}^{n} [X_k Y_k^{-1} - Y_{k-1}^{-1} E(X_k - 1 | X_{k-1}, Y_{k-1})] \\ &\quad = \sum_{k=2}^{n} [X_k Y_k^{-1} - X_{k-1} Y_{k-1}^{-1} - 1 + Y_{k-1}^{-1}] \\ &\quad = X_n Y_n^{-1} - n + \sum_{k=2}^{n} Y_{k-1}^{-1}, \end{aligned} \tag{4}$$

while, using (2),

$$\begin{aligned} E[(X_k Y_k^{-1} &- X_{k-1} Y_{k-1}^{-1} - 1 + Y_{k-1}^{-1})^2 | X_{k-1}, Y_{k-1}] \\ &= Y_{k-1}^{-2} E[(X_k - X_{k-1} - Y_{k-1})^2 | X_{k-1}, Y_{k-1}] \\ &= 1 - Y_{k-1}^{-1}. \end{aligned}$$

Hence, if $\{c_n\}$ is a sequence of positive constants such that $c_n \uparrow \infty$ and $\sum c_n^{-2} < \infty$, then

$$c_n^{-1} \sum_{k=2}^{n} [X_k Y_k^{-1} - E(X_k Y_k^{-1} | X_{k-1}, Y_{k-1})] \to 0 \quad \text{a.s.}$$

as $n \to \infty$ (Neveu [3, Proposition IV-6-2]) and, in particular, using (4),

$$n^{-3/4}\left(X_n Y_n^{-1} - n + \sum_{k=2}^{n} Y_{k-1}^{-1}\right) \to 0 \quad \text{a.s.} \tag{5}$$

as $n \to \infty$.

We must have $Y_n \uparrow \infty$ a.s., for if not (5) gives $X_n = O(n)$ a.s. on the exceptional set $\{\lim_{n\to\infty} Y_n < \infty\}$ which provides a contradiction (since $Y_n \uparrow \infty$ on the set $\{\sum X_n^{-1} = \infty\}$) unless the exceptional set has zero probability. Thus, $n^{-1} \sum_{k=2}^{n} Y_{k-1}^{-1} \to 0$ a.s. as $n \to \infty$ and (5) gives

$$n^{-1} X_n Y_n^{-1} = n^{-1}\left[X_n Y_n^{-1} - n + \sum_{k=2}^{n} Y_{k-1}^{-1}\right] + 1 - n^{-1} \sum_{k=2}^{n} Y_{k-1}^{-1} \to 1 \text{ a.s.} \tag{6}$$

as $n \to \infty$.

Next, using (6) together with (3),

$$\begin{aligned}(\log n)^{-1} Y_n &= (\log n)^{-1}\left(Y_n - \sum_{j=1}^{n} Y_j X_j^{-1}\right) \\ &\quad +(\log n)^{-1} \sum_{j=1}^{n} (Y_j X_j^{-1} - j^{-1}) + (\log n)^{-1} \sum_{j=1}^{n} j^{-1} \to 1 \text{ a.s.}\end{aligned} \tag{7}$$

since

$$(\log n)^{-1} \sum_{j=1}^{n} (Y_j X_j^{-1} - j^{-1}) = (\log n)^{-1} \sum_{j=1}^{n} j^{-1}(j Y_j X_j^{-1} - 1) \to 0 \quad \text{a.s.}$$

via Toeplitz' lemma. An application of (7) in (6) then gives

$$(n \log n)^{-1} X_n \to 1 \quad \text{a.s.} \tag{8}$$

The results (7) and (8) are those of Wunderlich cited above.

Proceeding further, we have from (5) and (7) and with another application of Toeplitz' lemma that almost surely

$$\begin{aligned} X_n Y_n^{-1} - n &= -\sum_{k=1}^{n-1} Y_k^{-1} + o(n^{3/4}) \\ &= -\sum_{k=2}^{n-1} [Y_k^{-1} - (\log k)^{-1}] - \sum_{k=2}^{n-1} (\log k)^{-1} + o(n^{3/4}) \\ &= -\sum_{k=2}^{n-1} (\log k)^{-1}[Y_k^{-1} \log k - 1] - \sum_{k=2}^{n-1} (\log k)^{-1} + o(n^{3/4}) \\ &= -(1 + o(1)) \sum_{k=2}^{n-1} (\log k)^{-1} + o(n^{3/4}) \\ &= -(1 + o(1)) n (\log n)^{-1}, \end{aligned} \tag{9}$$

so that

$$n^{-1}X_n - Y_n = -Y_n(\log n)^{-1}(1 + o(1)) \quad \text{a.s.}$$

and hence, from (7),

$$n^{-1}X_n - Y_n = -1 + o(1) \quad \text{a.s.} \tag{10}$$

It therefore remains only to prove part (i) of the Theorem, for part (ii) will then follow using (10).

We have, using (3),

$$\begin{aligned}\lim_{n\to\infty} (\log\log n)^{-1}(Y_n - \log n) &= \lim_{n\to\infty} (\log\log n)^{-1}\left(\sum_{j=1}^{n} Y_j X_j^{-1} - \log n\right)\\ &= \lim_{n\to\infty} (\log\log n)^{-1} \sum_{j=2}^{n} j^{-1}(jY_jX_j^{-1} - 1),\end{aligned} \tag{11}$$

since $\sum_{j=1}^{n} j^{-1} - \log n \to \gamma$, Euler's constant, as $n \to \infty$. To deal with (11) we rewrite (9) in the form

$$nY_nX_n^{-1} = (1 - (\log n)^{-1} + R_n)^{-1}$$

where $|R_n| = o(\log n)^{-1}$ a.s., and expand to obtain

$$nY_nX_n^{-1} = 1 + (\log n)^{-1} + T_n,$$

where $|T_n| = o(\log n)^{-1}$ a.s. as $n \to \infty$. Then,

$$\sum_{j=2}^{n} j^{-1}(jY_jX_j^{-1} - 1) = \sum_{j=2}^{n} (\log j)^{-1} + \sum_{j=2}^{n} T_j = (1 + o(1))\log\log n \quad \text{a.s.}$$

and result (i) follows from (11). This completes the proof.

Note added in proof. Since this paper was written, the discrete analogue of the result of Williams [4] has appeared in a paper by W. Neudecker and D. Williams, Compositio Math. **29** (1974), 197–200.

References

1 D. Hawkins, *The random sieve*, Math. Mag. **31**(1957/58), 1–3. MR **20** #5761.

2. ———, *Random sieves*. II, J. Number Theory **6**(1974), 192–200.

3. J. Neveu, *Cacul des probabilités*, 2nd ed., Masson, Paris, 1970.

4. D. Williams, *A study of a diffusion process motivated by the sieve of Erastosthenes*, Bull. London Math. Soc. **6** (1974), 155–164.

5. M. C. Wunderlich, *The prime number theorem for random sequences*, J. Number Theory (to appear).

6. ———, *A probabilistic setting for prime number theory*, Acta Arith. **26**(1974), 59–81.

Commonwealth Scientific and Industrial Research Organization (CSIRO), Division of Mathematics and Statistics, P.O. Box 1965, Canberra City, 2601 Australia