APPLICATIONES MATHEMATICAE
31,1 (2004), pp. 97–106

B. Kopociński (Wrocław)

RANDOM SPLIT OF THE INTERVAL $[0,1]$

Abstract. We define two splitting procedures of the interval $[0,1]$, one using uniformly distributed points on the chosen piece and the other splitting a piece in half. We also define two procedures for choosing the piece to be split; one chooses a piece with a probability proportional to its length and the other chooses each piece with equal probability. We analyse the probability distribution of the lengths of the pieces arising from these procedures.

1. Introduction. In his collection of problems [3] Hugo Steinhaus formulates the following problem regarding the population growth of a rod-shaped bacterium. Initially, one part breaks off from the original bacillus and becomes an independent bacillus. Then, at each successive generation one part breaks off from the longest bacillus. The length of this offspring bacillus is equal to the length of the shortest bacillus at the moment. Steinhaus states that if the initial split is into incommensurable pieces, then at most three different bacillus lengths exist in any given generation and the fractions of small, medium sized and large individuals oscillate over time.

The question arises as to how Steinhaus's principle of growth may be modified by adding a random component, so that nevertheless the number of different lengths is "almost finite" (countable) and the probability distributions of bacillus lengths, after a suitable standardisation, oscillates over time.

Formally, assume that the interval $[0,1]$ is split into two pieces by using some procedure. Next, a piece is chosen and split into two parts using the same procedure. These actions are repeated ad infinitum.

We define two splitting procedures, one using uniformly distributed points on the chosen piece (called *uniform split*) and the other splitting a piece in half. We also define two procedures for choosing the piece to be

2000 *Mathematics Subject Classification*: Primary 60F05; Secondary 60J10.

Key words and phrases: Steinhaus Problem, uniform split, splitting in half, random choice, probability distribution, moments, limiting theorem.

split; one chooses a piece with a probability proportional to its length (called *proportional choice*) and the other chooses each piece with equal probability (called *random choice*). We analyse the probability distribution of the lengths of the pieces arising from these procedures.

2. Notation. Let n denote the index of a split, i.e. the number of points chosen as split points so far. The first point is either selected uniformly from the interval $[0,1]$ or can be taken to be $1/2$. We denote the vector of the lengths of the consecutive pieces after n splits by $\mathbf{D}_n = (D_{1,n}, D_{2,n}, \ldots, D_{n+1,n})$. For simplicity we use the notation $D_n = D_{1,n}$. Also, we set $Z_n = -\log_2 D_n$ and $T_n = D_n^{-1}$, $n \geq 1$.

Let R denote random variables uniformly distributed on $[0,1]$. We set $E_n = -\log_2 R_n$, and $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution. I_n, $n \geq 0$, denotes the binary random variable $P(I_n = 1) = (n+1)^{-1}$, $P(I_n = 0) = 1 - P(I_n = 1)$.

3. Results. Among the four procedures of choosing and splitting considered in this paper the most interesting case is proportional choice together with splitting in half. It answers our generalisation of Steinhaus's problem. For completeness, we also analyze the remaining cases. Except for one case, the random variables $D_{j,n}$, $1 \leq j \leq n+1$, are not identically distributed. We concentrate on $D_{1,n}$, but in the last subsection we consider the distribution of a mixture of coordinates of the vector $\mathbf{D}_n$ defined in two ways. The assumptions used are stated in the titles of the subsections.

3.1. *Random choice, uniform split*

THEOREM 1. *Under Assumptions* 3.1 *the random variable* Z_n *is asymptotically normal with expected value* $\log_2 n$ *and variance* $2(\log 2)^{-1} \log_2 n$.

Proof. Let the event $I_n = 1$ denote that the first piece of length D_n is to be partitioned and the random variable R_n be uniformly distributed on $[0,1]$ split into two pieces, with the length of the resulting first piece being $D_n R_n$. Let R_0, R_n, I_n, $n \geq 1$, be mutually independent. The following recursive formulas hold:

$$D_1 = R_0, \qquad D_{n+1} = \begin{cases} D_n & \text{if } I_n = 0, \\ D_n R_n & \text{if } I_n = 1,\ n \geq 1. \end{cases}$$

Hence we have

$$\begin{aligned} D_1 &= R_0, & Z_1 &= E_0, \\ D_{n+1} &= D_n R_n^{I_n}, & Z_{n+1} &= Z_n + I_n E_n, \qquad n \geq 1. \end{aligned}$$

Therefore,

$$D_{n+1} = \prod_{j=0}^{n} R_j^{I_j}, \qquad Z_{n+1} = \sum_{j=0}^{n} I_j E_j, \qquad n \geq 1,$$

where $I_0 = 1$ with probability 1.

Let $c = (\log 2)^{-1}$. Since $\mathrm{E}(I_j) = (j+1)^{-1}$, $\mathrm{D}^2(I_j) = j(j+1)^{-2}$ for $j \geq 0$, it follows that

$$\mathrm{E}((I_j E_j)^r) = \frac{c^r r!}{j+1}, \qquad r \geq 1,$$

$$\mathrm{E}(Z_{n+1}) = \sum_{j=0}^{n} \mathrm{E}(I_j E_j) = \sum_{j=0}^{n} \frac{c}{j+1} \sim \log_2 n,$$

$$\sigma_{n+1}^2 = \mathrm{D}^2(Z_{n+1}) = \sum_{j=0}^{n} \mathrm{D}^2(I_j E_j) = \sum_{j=0}^{n} (\mathrm{E}((I_j E_j)^2) - (\mathrm{E}(I_j)\mathrm{E}(E_j))^2)$$

$$= c^2 \sum_{j=0}^{n} \left(\frac{2}{j+1} - \frac{1}{(j+1)^2} \right) \sim 2c \log_2 n,$$

$$\mathrm{E}\left(\left| I_j E_j - \frac{1}{j+1} \right|^3 \right) = \frac{j}{(j+1)^4} + \frac{1}{j+1} \mathrm{E}\left(\left| E_j - \frac{1}{j+1} \right|^3 \right)$$

$$\leq \frac{j}{(j+1)^4} + \frac{1}{j+1} \left(6c^3 + \frac{6c^2}{j+1} + \frac{3c}{(j+1)^2} + \frac{1}{(j+1)^3} \right),$$

$$\varrho_{n+1}^3 = \sum_{j=0}^{n} \mathrm{E}\left(\left| I_j E_j - \frac{1}{j+1} \right|^3 \right) \leq 19 \log_2 n \quad \text{for sufficiently large } n.$$

Since $\sqrt[3]{\varrho_n^3}/\sqrt{\sigma_n^2} \to 0$, Theorem 1 follows from Lyapunov's Theorem ([1, p. 211]). ∎

3.2. *Proportional choice, uniform split.* If $[0,1]$ includes n split points, then the random variable R_n uniformly distributed on $[0,1]$ determines a piece to be divided and simultaneously designates the break off point chosen uniformly along the piece. In this case the $D_{j,n}$, $1 \leq j \leq n+1$, are identically distributed.

THEOREM 2. *Under Assumptions* 3.2, *asymptotically the random variable* $Z_n - \log_2 n$ *has a Gompertz distribution*

$$\lim_{n\to\infty} P(Z_n - \log_2 n \leq x) = \exp(-\exp(-(\log 2)x), \qquad -\infty < x < \infty.$$

Proof. If R_n falls into $[0, D_n]$, then it is the first piece to be split, and the location of R_n splits it into pieces. Let R_n, $n \geq 0$, be mutually independent. The following recursive formulas hold:

$$D_1 = R_0, \qquad D_{n+1} \stackrel{\mathrm{d}}{=} \begin{cases} D_n & \text{if } R_n > D_n, \\ R_n & \text{if } R_n \leq D_n,\ n \geq 1. \end{cases}$$

This may be rewritten as $D_1 = R_0$, $D_{n+1} \overset{\mathrm{d}}{=} \min(D_n, R_n)$, $n \geq 1$. Hence, $D_{n+1} \overset{\mathrm{d}}{=} \min(R_0, \ldots, R_n)$ and the limiting theorem for minimum order statistics, $\lim_{n\to\infty} P(nD_n \leq x) = 1 - e^{-x}$, $x > 0$, holds (see [2, p. 22]). Theorem 2 is a disguised version of that theorem. ∎

3.3. *Random choice, splitting in half*

THEOREM 3. *Under Assumptions* 3.3 *the random variable* Z_n *is asymptotically normal with expected value* $\log n$ *and variance* $\log n$.

Proof. Suppose that the event $I_n = 1$ denotes that D_n is to be split. Let I_n, $n \geq 1$, be mutually independent. Then the following recursive formulas hold:

$$D_1 = 2^{-1}, \qquad D_{n+1} = 2^{-I_n} D_n, \qquad n \geq 1. \tag{1}$$

Hence, we have

$$D_{n+1} = 2^{-(I_0 + \ldots + I_n)}, \qquad Z_{n+1} = \sum_{j=0}^{n} I_j, \qquad n \geq 1.$$

where $I_0 = 1$ with probability 1.

We have

$$\mathrm{E}(Z_{n+1}) = \sum_{j=0}^{n} \mathrm{E}(I_j) = \sum_{j=0}^{n} \frac{1}{j+1} \sim \log n,$$

$$\sigma_{n+1}^2 = \mathrm{D}^2(Z_{n+1}) = \sum_{j=0}^{n} \mathrm{D}^2(I_j) = \sum_{j=0}^{n} \left(\frac{1}{j+1} - \frac{1}{(j+1)^2} \right) \sim \log n,$$

$$\mathrm{E}\left(\left| I_j - \frac{1}{j+1} \right|^3 \right) = \frac{j}{(j+1)^4} + \frac{j^3}{(j+1)^4},$$

$$\varrho_{n+1}^3 = \sum_{j=0}^{n} \mathrm{E}\left(\left| I_j - \frac{1}{j+1} \right|^3 \right) \sim \log n.$$

Because $\sqrt[3]{\varrho_n^3} / \sqrt{\sigma_n^2} \to 0$, Theorem 3 follows from Lyapunov's Theorem. ∎

Using (1) we obtain:

PROPOSITION 4. *Under Assumptions* 3.3 *the probabilities* $p_n(j) = P(D_n = 2^{-j}) = P(Z_n = j)$, $1 \leq j \leq n$, $n \geq 1$, *satisfy the recursive formulas*

$$p_1(1) = 1, \qquad p_{n+1}(j) = p_n(j) \frac{n}{n+1} + p_n(j-1) \frac{1}{n+1}. \tag{2}$$

The probability distribution $Z_{15} = Z_{1,15}$ is presented in Table 2.

3.4. *Proportional choice, splitting in half.* The next two facts give the recursive formulas for the probability distributions of D_n and Z_n and the relations between the moments.

PROPOSITION 5. *Under Assumptions* 3.4 *the random variables* D_n, Z_n *and probability distributions* $p_n(j) = P(D_n = 2^{-j}) = P(Z_n = j)$, $1 \le j \le n$, $n \ge 1$, *satisfy the following recursive formulas*:

$$D_1 = \frac{1}{2}, \qquad D_{n+1} = \begin{cases} D_n & \text{if } R_n > D_n, \\ \frac{1}{2} D_n & \text{if } R_n \le D_n, \end{cases}$$

$$(3) \qquad Z_1 = 1, \qquad Z_{n+1} = \begin{cases} Z_n & \text{if } Z_n < E_n, \\ 1 + Z_n & \text{if } Z_n \ge E_n, \end{cases}$$

$$p_1(1) = 1, \qquad p_{n+1}(j) = p_n(j)\left(1 - \frac{1}{2^j}\right) + p_n(j-1)\,\frac{1}{2^{j-1}}.$$

PROPOSITION 6. *Under Assumptions* 3.4 *the moments of* D_n, Z_n, T_n, $n \ge 1$, *satisfy the following recursive formulas*:

$$(4) \quad \mathrm{E}(D_1^r) = 2^{-r}, \quad \mathrm{E}(D_{n+1}^r) = \mathrm{E}(D_n^r) + (2^{-r} - 1)\mathrm{E}(D_n^{r+1}),$$

$$\mathrm{E}(Z_1) = 1, \qquad \mathrm{E}(Z_{n+1}) = \mathrm{E}(Z_n) + \mathrm{E}(D_n),$$

$$\mathrm{E}(T_1^r) = 2^r, \qquad \mathrm{E}(T_{n+1}^r) = \mathrm{E}(T_n^r) + (2^r - 1)\mathrm{E}(T_n^{r-1}), \qquad r \ge 1.$$

COROLLARY 1. *Under Assumptions* 3.4 *the moments* $E(D_n^r)$, $n \ge 1$, $r \ge 1$, *may be calculated from* (4), *and the following relations are satisfied*:

$$\mathrm{E}(D_1) = \tfrac{1}{2}, \qquad \mathrm{E}(D_{n+1}) = \mathrm{E}(D_n) - \tfrac{1}{2}\mathrm{E}(D_n^2),$$

$$\mathrm{E}(D_{n+1}) = \frac{1}{2}\Big(1 - \sum_{j=1}^{n} \mathrm{E}(D_j^2)\Big),$$

$$\mathrm{E}(Z_n) = 1 + \sum_{i=1}^{n-1} \mathrm{E}(D_i),$$

$$\mathrm{E}(T_n) = n + 1, \qquad \mathrm{E}(T_n^2) = \tfrac{3}{2}n(n+1) + 1, \qquad \mathrm{Var}(T_n) = \tfrac{1}{2}n(n-1),$$

$$\mathrm{E}(T_{n+1}^3) = \tfrac{7}{2}n(n+1)(n+2) + 7(n+1) + 1,$$

$$\mathrm{E}(T_{n+1}^r) = 1 + (2^r - 1)\sum_{i=0}^{n} \mathrm{E}(T_i^{r-1}), \quad \textit{where} \quad \mathrm{E}(T_0^r) = 1,$$

$$\mathrm{E}(T_n^r) = m_r n^r + o(n^r),\ n \to \infty, \quad \textit{where} \quad m_r = \prod_{j=1}^{r} (2^j - 1) j^{-1},\ r \ge 1.$$

Since $y = x^{-r}$, $r \ge 1$, is convex for $x \ge 0$, from the Jensen inequality we obtain the following result

PROPOSITION 7. *Under Assumptions* 3.4 *the moments of* D_n *and* T_n, $n \ge 1$, *satisfy*

$$(5) \qquad \mathrm{E}(D_n^r) \ge (\mathrm{E}(T_n^r))^{-1}, \qquad r \ge 1.$$

Proposition 8. *Under Assumptions* 3.4 *the moments of* Z_n *satisfy*

$$\text{(6)}\qquad \begin{aligned} &\mathrm{E}(Z_n^2) \le \log_2^2(n+2),\\ &\mathrm{E}((c_r+Z_n)^r) \le (c_r+\log_2(n+1))^r, \qquad r\ge 1,\end{aligned}$$

where $c_r = (r-1)\log_2 e - 1$.

Proof. Note that $y = \log_2^2 x$ is concave for $x \ge e$. Let $T_n^* = \max(e, T_n)$ and $Z_n^* = \max(\log_2 e, Z_n)$. Hence, $\mathrm{E}(T_n^*) = \mathrm{E}(T_n) + (e-2)p_n(1)$, where $p_n(1) = P(T_n = 2)$, $\mathrm{E}(Z_n^*) = \mathrm{E}(Z_n) + (\log_2 e - 1)p_n(1)$. Also $Z_n^2 < (Z_n^*)^2$. From the Jensen inequality we obtain

$$\mathrm{E}(Z_n^2) < \mathrm{E}((Z_n^*)^2) = \mathrm{E}(\log_2^2 T_n^*) \le \log_2^2(\mathrm{E}(T_n^*)) \le \log_2^2(n+2).$$

Note that $y = \log_2^r x$ is concave for $x \ge e^{r-1}$. Let $d_r = \frac{1}{2}e^{r-1}$ and $c_r = \log_2 d_r$. Because $T_n \ge 2$, it follows from the Jensen inequality that

$$\mathrm{E}((c_r+Z_n)^r) = \mathrm{E}(\log_2^r(d_r T_n)) \le \log_2^r(d_r \mathrm{E}(T_n)) = (c_r + \log_2(n+1))^r.$$

Corollary 2. *Under Assumptions* 3.4 *the moments of* Z_n, $n\ge 1$, *satisfy*

$$\begin{aligned} &\mathrm{E}(Z_n) \le \log_2(n+1),\\ &\mathrm{E}(Z_n^2) \le 2(\log_2 e - 1)(\log_2(n+1) - \mathrm{E}(Z_n)) + \log_2^2(n+1).\end{aligned}$$

We have $m_1 = 1$, $m_2 = 1.5$, $m_3 = 3.5$, $m_4 = 13.125$, $m_5 = 81.4375$, and so on. The well known (see [1, p. 174]) sufficient condition for expressing a probability distribution function by its moments is not satisfied, since the series $\sum_{r=1}^{\infty} m_r u^r (r!)^{-1}$ is not convergent for any $u > 0$.

Let $U_n = (n+1)^{-1}T_n$. Then $\mathrm{E}(U_n^r) \sim m_r$ as $n\to\infty$, for all $r\ge 1$. From Corollary 2 we obtain the following result:

Theorem 9. *Under Assumptions* 3.4 *the support of the random variable* Z_n *is* $\{1,2,\ldots,n\}$ *and* Z_n *has the following representation*:

$$Z_n - \log_2(n+1) \overset{\mathrm{d}}{=} \log_2 U_n.$$

In particular $\mathrm{E}(U_n) = 1$, $\mathrm{Var}(U_n) \sim \frac{1}{2}$ *as* $n\to\infty$.

Fig. 1 represents the graph of $\widehat{p}_n = \max_{1\le j\le n} p_n(j)$ for $n \le 2000$. It suggests that Z_n does not converge in distribution as $n\to\infty$: the fluctuations do not disappear. Selected distributions are presented in Table 1. If $n+1 = 2^m$, $m\ge 1$, then the support of $Z_n - m$ is a set of integers, and this subsequence is convergent.

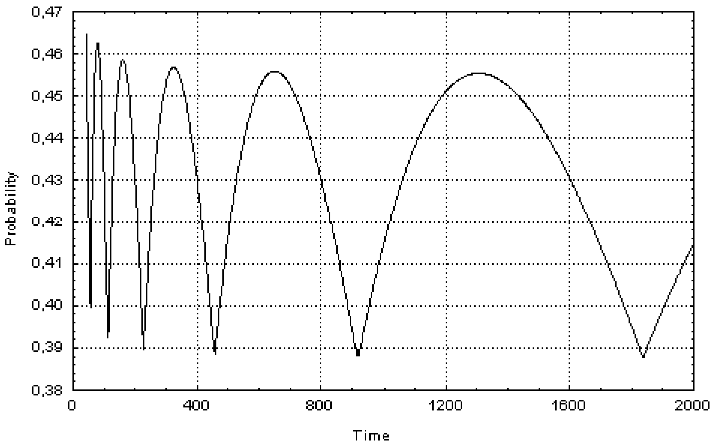

Fig. 1. The probabilities $\widehat{p}_n = \max_{1 \le j \le n} p_n(j)$ in the distribution of D_n

Table 1. Selected distributions

Probabilities $P(Z_n = j)$						
j	Z_{326}	Z_{458}	Z_{1024}	Z_{1303}	Z_{1855}	Z_{2048}
6	0.020	0.003				
7	0.226	0.090	0.001			
8	0.457	0.389	0.060	0.021	0.003	0.001
9	0.249	0.388	0.353	0.228	0.090	0.060
10	0.045	0.118	0.415	0.455	0.388	0.343
11	0.003	0.013	0.145	0.248	0.387	0.421
12		0.001	0.018	0.045	0.118	0.153
13			0.001	0.003	0.013	0.019
14					0.001	0.001
$\mathrm{E}(Z_n)$	8.081	8.570	9.728	10.076	10.17	10.727
$\mathrm{D}(Z_n)$	0.870	0.871	0.872	0.873	0.873	0.873

3.5. *The distribution of a piece chosen from the collection.* Consider the sequence $\mathbf{D}_n = (D_{1,n}, \ldots, D_{n+1,n})$, $n \ge 1$, under Assumptions 3.3 or 3.4. In the case of splitting in half, the components of $\mathbf{D}_n$ are not identically distributed. For example, in both cases considered $D_{1,2}$ takes the values $1/2$ and $1/4$ with probabilities $1/2$ and $D_{2,2}$ takes the value $1/4$ with probability 1. Therefore, the length of a given piece in the collection, $D_{i,n}$, depends on the index i. The number of elements in the support of $\mathbf{D}_n$ increases rapidly. The lengths of the pieces after the nth split belong to the set $\{2^{-j}; 1 \le j \le n\}$. Therefore, it is more convenient to analyse the chain $\mathbf{M}_n = (M_{1,n}, \ldots, M_{n,n})$, $n \ge 1$, defined by counting the number of pieces of the same length:

$$M_{j,n} = \sum_{i=1}^{n+1} \mathbf{1}(D_{i,n} = 1/2^j), \qquad 1 \le j \le n. \tag{7}$$

Because (7) is a partition of $[0,1]$, it follows that $M_{j,n} \ge 0$ for $1 \le j \le n$, and

$$\sum_{j=1}^{n} M_{j,n} = n+1, \qquad \sum_{j=1}^{n} \frac{1}{2^j} M_{j,n} = 1. \tag{8}$$

Let $p_{\mathbf{m}_n,\mathbf{m}_{n+1}}$ denote the transition probabilities for the vector $\mathbf{M}_n$ indexed in an appropriate way. Note that

$$M_{j,n+1}|\mathbf{M}_n = \begin{cases} M_{j,n}+2 & \text{with probability } P_{n,j}M_{j-1,n}, \\ M_{j,n}-1 & \text{with probability } P_{n,j}M_{j,n}, \\ M_{j,n} & \text{otherwise}, \end{cases} \tag{9}$$

where

$$P_{n,j} = \begin{cases} (n+1)^{-1} & \text{under Assumptions 3.3}, \\ 2^{-j} & \text{under Assumptions 3.4}. \end{cases}$$

Then

$$p_{\mathbf{m}_n,\mathbf{m}_{n+1}} = \sum_{j=1}^{n+1} p_{\mathbf{m}_n,\mathbf{m}_{n+1}}(j),$$

$$p_{\mathbf{m}_n,\mathbf{m}_{n+1}}(j) = P(\mathbf{M}_{n+1} = (m_1,\dots,m_{j-1},m_j-1,m_{j+1}+2, m_{j+2},\dots,m_{n+1}) \mid \mathbf{M}_n = (m_1,\dots,m_n)), \qquad 1 \le j \le n,$$

where $m_j \ge 1$, $n \ge 1$.

PROPOSITION 10. *The chain* $\mathbf{M}_n$ *is Markovian*; $P(M_{11} = 2) = 1$,

$$p_{\mathbf{m}_n,\mathbf{m}_{n+1}}(j) = \begin{cases} m_{j,n}(n+1)^{-1} & \textit{under Assumptions } 3.3, \\ m_{j,n}2^{-j} & \textit{under Assumptions } 3.4. \end{cases}$$

Define the random variables $\overline{D}_n$ and $\widetilde{D}_n$, $n \ge 1$, by

$$P\left(\overline{D}_n = \frac{1}{2^j}\right) = \frac{1}{n+1} \sum_{i=1}^{n+1} P\left(D_{i,n} = \frac{1}{2^j}\right),$$

$$P\left(\widetilde{D}_n = \frac{1}{2^j}\right) = \frac{1}{2^j} \sum_{i=1}^{n+1} P\left(D_{i,n} = \frac{1}{2^j}\right), \qquad 1 \le j \le n.$$

Using (7), we obtain

$$\begin{aligned} P\left(\overline{D}_n = \frac{1}{2^j}\right) &= \frac{1}{n+1}\,\mathrm{E}(M_{j,n}), \\ P\left(\widetilde{D}_n = \frac{1}{2^j}\right) &= \frac{1}{2^j}\,\mathrm{E}(M_{j,n}), \qquad 1 \le j \le n. \end{aligned} \tag{10}$$

We interpret the random variable $\overline{D}_n$ as a randomly chosen component of $\mathbf{D}_n$ chosen uniformly from the set of $n+1$ components. We interpret the random variable $\widetilde{D}_n$ as a randomly chosen component of $\mathbf{D}_n$, with each component being chosen with a probability proportional to its value. Using (10), we formulate the following proposition:

PROPOSITION 11. *Under Assumptions* 3.3 *or* 3.4 *we have*

$$\mathrm{E}(\overline{D}_n) = \frac{1}{n+1}, \qquad \mathrm{E}(\overline{D}_n^2) = \frac{1}{n+1}\sum_{i=1}^{n+1} \mathrm{E}(D_{i,n}^2), \qquad \mathrm{E}(\widetilde{D}_n) = (n+1)\mathrm{E}(\overline{D}_n^2).$$

THEOREM 12. *Under Assumptions* 3.3 *we have* $\widetilde{D}_n \stackrel{\mathrm{d}}{=} D_{1,n}$.

THEOREM 13. *Under Assumptions* 3.4 *we have* $\widetilde{D}_n \stackrel{\mathrm{d}}{=} D_{1,n}$.

Proof of Theorem 12. Let $\delta(p)$ be a binary 0-1 random variable: $P(\delta(p)=1) = E(\delta(p)) = p$. If Π is a random variable, $0 \le \Pi \le 1$, then $\delta(\Pi)$ is a mixed binary random variable and we have

$$P(\delta(\Pi)=1) = \mathrm{E}(\delta(\Pi)) = \mathrm{E}(\Pi).$$

From (9) assuming 3.3 we have

$$M_{j,n+1}|\mathbf{M}_n \stackrel{\mathrm{d}}{=} M_{j,n} + 2\delta_1\left(\frac{1}{n+1}\,M_{j-1,n}\right) - \delta_2\left(\frac{1}{n+1}\,M_{j,n}\right),$$

where δ_1, δ_2 are binary random variables for which

$$\delta_1\left(\frac{1}{n+1}\,M_{j-1,n}\right)\delta_2\left(\frac{1}{n+1}\,M_{j,n}\right) = 0.$$

We have

$$\mathrm{E}(M_{j,n+1}) = \mathrm{E}(M_{j,n}) + \frac{2}{n+1}\,\mathrm{E}(M_{j-1,n}) - \frac{1}{n+1}\,\mathrm{E}(M_{j,n}),$$

and from (10) we obtain

$$P\left(\widetilde{D}_{n+1} = \frac{1}{2^j}\right) = \frac{n}{n+1}\,P\left(\widetilde{D}_n = \frac{1}{2^j}\right) + \frac{1}{n+1}\,P\left(\widetilde{D}_n = \frac{1}{2^{j-1}}\right).$$

Comparing this equation with (2) completes the proof. ∎

The *proof of Theorem 13* is analogous.

Table 2. The probability distributions of $\overline{D}_{15}$ and $\widetilde{D}_{15}$ for Models 3.3 and 3.4

Probabilities $P(\overline{Z}_n = j)$ and $P(\widetilde{Z}_n = j)$				
	Model 3.3		Model 3.4	
j	$\overline{D}_{15}$	$\widetilde{D}_{15} \stackrel{\mathrm{d}}{=} D_{1,15}$	$\overline{D}_{15}$	$\widetilde{D}_{15} \stackrel{\mathrm{d}}{=} D_{1,15}$
1	0.008	0.067		
2	0.054	0.217	0.009	0.035
3	0.150	0.300	0.170	0.340
4	0.238	0.238	0.460	0.460
5	0.244	0.122	0.299	0.150
6	0.173	0.043	0.058	0.015
7	0.082	0.011	0.004	
8	0.033	0.002		
9	0.009			
10	0.002			
E	4.762	3.318	4.240	3.769
D	1.557	1.318	0.846	0.801

The increasing number of states causes some obstacle in the numerical analysis of $\mathbf{M}_n$. Define $\overline{Z}_n = -\log_2 \overline{D}_n$ and $\widetilde{Z} = -\log_2 \widetilde{D}_n$. Table 2 gives the distributions of $\overline{D}_{15}$ and $\widetilde{D}_{15} \stackrel{\mathrm{d}}{=} D_{1,15}$, their expected values E and dispersions D.

Acknowledgements. I am grateful to the referee for suggesting many improvements in the exposition.

References

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton, 1946.

[2] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York, 1978.

[3] H. Steinhaus, *One Hundred Problems in Elementary Mathematics*, Basic Book, New York, and PWN and Pergamon Press, Warszawa, 1963.

Mathematical Institute
Wrocław University
Pl. Grunwaldzki 2/4
50-384 Wrocław, Poland
E-mail: ibk@math.uni.wroc.pl

Received on 20.3.2003;
revised version on 27.10.2003 (1676)