*On leave from the Centre National de la Recherche Scientifique.

UNBIASED DIE ROLLING WITH A BIASED DIE

by

P. Camion*

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 920
April, 1974

UNBIASED DIE ROLLING WITH A BIASED DIE

by P. Camion

C. N. R. S. & University of North Carolina

1. *Introduction:* We will here use some algebraic methods lectured by M. P. Schützenberger at Toulouse University in 1965, some of these were introduced in [6].

Let X^* be the monoïd freely generated by a finite alphabet X. $p : X \to \mathbb{R}_+$ for which $\sum_{x \in X} p(x) = 1$ is defined and may be interpreted as a probability distribution. A word in X^* may be viewed as a finite sequence of trials corresponding to the tossing of a die that finite number of times. p is extended to the set X^n of words of length n by $p(u) = \prod_{j \in [1,n]} p(\alpha_{i_j})$ where $u = \alpha_{i_1} \dots \alpha_{i_n}$. The first aim is to build up a set $E \subset X^*$ which is partitioned into $E = E^{(1)} + \dots + E^{(k)}$, with, for k the cardinality of X :

$$(1) \qquad \sum_{u \in E_n^{(i)}} p(u) = \sum_{u \in E_n^{(j)}} p(u) \quad , \forall \ i,j \in [1,k] \ , \forall \ n \in \mathbb{N}$$

where $F_n = F \cap X^n$, for any $F \subset X^*$. We must have, for whatever p

$$\lim_{n\to\infty} \sum_{u\in X^n \setminus EX^*} p(u) = 0 . \tag{2}$$

Such an E has been built up by von Neumann for $X = \{0,1\}$ [8]. That particular construction is extended here. The generating series of probability for the values under consideration in (2) is given, for any X and any p. (2) is proved and the generating series for Card $(X^n \setminus EX^*)$ is found explicitly. A computable formula for the mean delay is given as well.

The advantage of this procedure is that only a small memory capacity and a few computations are required. Actually the number of comparisons required is less than the value of the mean delay. But the efficiency is very poor. In the next paragraph we give a procedure with high efficiency and a reasonable amount of computation.

2. *Von Neumann sequences in the set* $\{0,1\}$

A classical solution for $X = \{0,1\}$ is to take E to be the set of words with even length having a right factor in $\{01,10\}$ and without a proper left factor with the same property (i.e. the property of having even length and having a right factor in $\{01,10\}$). Here we write

$$E = E^{(1)} + E^{(2)} \tag{4}$$

where $E^{(1)}$ is the set of words in E ending with 0 and $E^{(2)}$ is the set of words in E ending with 1. Clearly, the mapping $\sigma : E^{(1)} \to E^{(2)}$ defined by

$$\sigma(u01) = u10 ,$$

is one to one and since $p(u10) = p(u01)$, for every p , (1) is verified. E is the set of von Neumann sequences. Clearly (2) is satisfied since

$$(4) \qquad \sum_{u \in X^n \setminus EX^*} p(u) = (1 - 2p(0)p(1))^n \ .$$

Moreover (4) may be easily computed for all n , whatever p be. However the set E here described is not the best possible, as proved by W. Hoeffding and G. Simons [5] who define several other sets with better values of (4).

3. *An extension of von Neumann sequences in the case where X has more than two symbols*

3.1 *Construction of the set E of sequences producing the output symbols*

Let $\text{Card } X = k$.

We define in X^* a prefix code C , that is a subset of X^* for which

$$(5) \qquad C \cap CXX^* = \{\phi\} \ .$$

We write the partition

$$(6) \qquad C = C_1 \cup C_2 \cup \ldots \cup C_i \cup \ldots \cup C_k \ ,$$

where C_i is the set of words of length n in C . $C_1 = \phi$. C_2 is the set of words $u = \alpha\alpha$, that is the words with two equal symbols. Now, recursively C_i is the set of all words in C of length i having at least two equal symbols and no left factor in any C_j , $j < i$. (Thus, C_i has exactly two equal symbols, all others are distincts.) Then, by definition, the requirement (5) for a prefix code is met.

Now let R be the ring $\mathbb{Z}\langle X\rangle$, that is the ring over the rational integers of the monoid X^*. To every subset F of X^* corresponds a polynomial $\sum_{u\in E} u \in \mathbb{Z}\langle X\rangle$. For notational simplicity, we just write F for such a polynomial. We now consider $R[t]$ and as well the ring of formal series $R[[t]]$. $1 - \sum_{1\le i\le k} C_i t^i$ belongs to $R[t]$. Its inverse is

$$1 + \sum_{n\ge 1} \Big(\sum_{1\le i\le k} C_i t^i \Big)^n = \sum_{i\ge 0} A_i t^i \tag{7}$$

and the monomials that we find in the A_i are all possible products of words in C. Those monomials all have coefficient one since every word factorizable in C has a unique factorization in C, as is well known, a consequence of (5). Now, the mapping $\phi : X \to \mathbb{Z}$ defined by $\phi(x) = 1$, $\forall\ x \in X$ extends into a morphism of $\mathbb{Z}\langle X\rangle$ into $\mathbb{Z}$ and further into a morphism of $R[[t]]$ into $\mathbb{Z}[[t]]$. Then

$$\phi(1 - \textstyle\sum C_i t^i)^{-1} = \phi \sum_{i\ge 0} A_i t^i ,$$

or

$$(1 - \textstyle\sum c_i t^i)^{-1} = \sum_{i\ge 0} a_i t^i ,$$

where c_i denotes $\phi C_i = \mathrm{Card}\, C_i$ and a_i denotes $\phi A_i = \mathrm{Card}\, A_i$, by our previous remark. Thus a_i is the number of words of length i in $C^* = \bigcup_{i\ge 0} A_i$. We will also consider the natural extension of $\phi_p : \mathbb{Z}\langle X\rangle \to \mathbb{R}$, defined by $p : X \to \mathbb{R}_+$, into a morphism of $R[[t]]$ into $\mathbb{R}[[t]]$. Denote $\phi_p C_i$ by c_{ip}, and for $A(t) = \sum_{i\in\mathbb{N}} A_i t^i$, denote $\phi_p A_i$ by a_{ip}.

One has, for example,

$$(1 - \sum_{\ell\le i\le k} c_{ip} t^i)^{-1} = \sum_{i\ge 0} a_{ip} t^i = a_p(t) \tag{8}$$

and since $a_{ip} = \phi_p A_i = \sum_{u \in A_i} p(u)$, the formal serie (8) gives, for every i , the probability for a word of length i to be in C^* .

Now let P be the $k!$ words of X^* which is the set of all sequences of the k distincts letters in X . Let $E(t) = A(t)Pt^k$. Denote $E(1)$ by E .

$$E = \sum_{x \in X} E^{(x)} \tag{9}$$

where $E^{(x)}$ is the subset of E ending with the letter x . (9) is a partition of X into k subsets. For any distribution of probability

$$p : X \to \mathbb{R}_+ , \sum_{x \in X} p(x) = 1 ,$$

we see that if $E_n^{(x)}$ denotes the subset of words in $E^{(x)}$ with length n , $\sum_{u \in E_n^{(x)}} p(u)$ does not depend on x .

3.2 *The generating series*

Property 1 One has

$$c_1 = 0 , c_i = k(k-1) \dots (k-i+2)(i-1) , 2 \le i \le k . \tag{10}$$

This is a straightforward consequence of the definition of C . We observe that $c_k = k(k-1) \dots 2(k-1) = (k-1)k!$ Following M. P. Schützenberger, we say that a finite prefix code F is complete when, if n is the largest length of a word in F , every word in X^n has a left factor in F . Then complete prefix codes satisfy the polynomial equality

$$\sum_{1 \le i \le n} F_i X^{n-i} = X^n . \tag{11}$$

Property 2: $C \cup P$ is a complete prefix code.

Proof: If a word of X^k has no two equal symbols, it belongs to P. If it has two equal symbols, it has a left factor in C.

Also $C \cup P$ verifies $(C \cup P)X^* \cap (C \cup P)XX^* = \phi$. We write in place of (11)

$$\sum_{1 \le i \le k} C_i X^{k-i} = X^k - P \tag{12}$$

Property 3: For every probability distribution p all of the roots of the polynomial $f(t) = 1 - \sum_{1 \le i \le k} c_{i,p} t^i$ have a modulus larger than one.

The reciprocal polynomial of $f(t)$ is a Frobenius polynomial, i.e. its companion matrix is non-negative. Thus its largest positive root has the largest absolute value among all its roots.

From (12) by using the morphism ϕ_p, one obtains

$$\sum_{2 \le i \le k} c_{ip} = 1 - k! \prod_{x \in X} p(x) \tag{13}$$

which proves that $t^k - \sum_{1 \le i \le k} c_{ip} t^{k-i}$ has a real root ζ in $]0,1[$, since it takes a negative value for $t = 0$. On the other hand, it does not have any other positive root since its quotient by $t - \zeta$ is a polynomial with non-negative coefficients.

In (9) we have defined E as C^*P. Let D be the set $X^* \setminus EX^*$, that is the set of words with no left factor in E. We denote by d_{np} the probability for a word of length n to be in D_n. We have

THEOREM 1

$$\sum_{n \in \mathbb{N}} d_{np} t^n = \frac{1 - \sum_{2 \le i \le k} c_{ip} t^i - k! \prod_{x \in X} p(x) t^k}{(1 - \sum_{2 \le i \le k} c_{ip} t^i)(1 - t)}, \tag{14}$$

and $\sum_{n \in \mathbb{N}} d_n$ *converges.*

Let $D(t)$ be the formal serie corresponding to the set D. One has

$$(15) \qquad (1 - E(t))^{-1}D(t) = \sum_{i\in\mathbb{N}} X^i t^i .$$

Applying the morphism ϕ, we get, with $a_p(t)$ defined in (8),

$$(16) \qquad (1 - k! \prod_{x\in X} p(x)t^k a_p(t))^{-1}d(t) = (1 - t)^{-1}$$

then by (8), we obtain (14).

Then by (13), the numerator of (14) has 1 as a root so that the denominator of (14) reduces to

$$(17) \qquad f(t) = 1 - \sum_{2\le i\le k} c_{ip}t^i .$$

Hence property 3 completes the proof.

Remark: For $p(x) = p(y)$, $\forall\ x,y \in X$, $k^n d_n$ is the cardinal of D_n and since we then have $k^j c_{ip} = c_i$, c_i being given by (10), one has

$$(18) \qquad \sum_{n\in\mathbb{N}} \text{Card } D_n t^n = \frac{1 - \sum_{i<k} k(k-1)\dots(k-i+2)(i-1)t^i - kk!t^k}{(1 - \sum_{i\le k} k(k-1)\dots(k-i+2)(i-1)t^i)(1-kt)^{-1}} .$$

If we write as well $E(t) = \sum_{n\in\mathbb{N}} E_n t^n$ and $\ell_p(t) = \sum_{n\in\mathbb{N}} \ell_{n,p} t^n$, where $\ell_{n,p} = \phi_p E_n$. By theorem 1 we have $\sum_{n\in\mathbb{N}} \ell_{n,p} = 1$ since $\sum_{0\le n\le m} E_n X^{m-n} = X^m - D_m$, and, from applying ϕ_p :

$$\sum_{0\le n\le m} \ell_{n,p} = 1 - d_m$$

ℓ_{np} may be interpreted as the probability that the sequence of symbols obtained at time n be a word in E. If $|v|$ denotes the number of symbols in the word v we see that $\ell_{|v|,p}$ is the distribution of the $|v|$ for $v \in E$. By the same argument as that one used by M. P. Schützenberger in [6] III.7 page 1209 it can be proved easily that this distribution is dominated by an exponential distribution and then has moments of every order. This argument relies on the fact that there exists a word $u \in X^*$ such that $X^*u \subset EX^*$. This actually occurs, since, if v is any word in P, $u = v^2$ is easily seen to have this property. However, we give a proof which is valid for any prefix code C, $C \cup P$ being a complete prefix code. (*)

THEOREM 2: The distribution $\ell_{|v|,p}$ of the $|v|$ for $v \in E$ has moments of every order.

(*) Gordon Simons shows that such a word u always exists in the situation under consideration. Here is his statement and proof.

THEOREM: Let P be a complete finite prefix code for a finite alphabet X and let A be a nonempty subset of P. There exists a word $u \in X^*$ for which $X^*uX^* \subset P^*AX^*$.

Proof: Since P is complete, there exists a finite set $S = \{s_1, \ldots, s_n\} \subset X^*$ for which $X^* = P^*S$ and $P^* \cap S = \{e\}$, where e is the empty word. There exists a $u_1 \in X^*$ for which $P^*s_1u_1 \subset P^*AX^*$, and, in turn, a $u_2 \in X^*$ for which $P^*s_2u_1u_2 \subset P^*AX^*$. Proceeding recursively, one can obtain $u_1, u_2, \ldots u_j$ for which $P^*s_ju_1u_2 \ldots u_j \subset P^*AX^*$, $j = 1, \ldots, n$. Let $u = u_1u_2 \ldots u_n$. Then $P^*s_ju \subset P^*AX^*$ for each j, and $X^*uX^* = P^*SuX^* \subset P^*AX^*$.

The denominator of the rational fraction giving $\ell_p(t)$ verifies Property 3 as well as its powers which will appear in the derivatives. Then all series under consideration converge.

Example: $k = 3$

(19) $f(t) = 1 - 3t^2 - 12t^3$, $(f(t) - 3!t^3)(1 - 3t)^{-1} = 6t^2 + 3t + 1$.

Then, knowing the first three Card D_i , $1 \le i \le 3$, we may compute the following by the recurrence relation

(20) $$\text{Card } D_i = 3 \text{ Card } D_{i-2} + 12 \text{ Card } D_{i-3} .$$

We finally obtain the array. (See page 9a.)

Since $f^*(t)$ has a real root $a^{-1} = 0{,}367392$ and two complex conjugated roots b^{-1} and c^{-1} , we know from (18) that the first member of (18) has the form

(21) $$x_1(1-a^{-1}t)^{-1} + x_2(1-b^{-1}t)^{-1} + x_3(1-c^{-1}t)^{-1}$$

and since $|b^{-1}| \le a^{-1}$, $|c^{-1}| \le a^{-1}$, one has

(22) $$\text{Card } D_n \le (|x_1| + 2|x_2|)a^n .$$

We have a Vandermond system of linear equations:

$$x_1 + x_2 + x_3 = 1$$
$$ax_1 + bx_2 + cx_3 = 3$$
$$a^2x_1 + b^2x_2 + c^2x_3 = 9$$

from which

i	3^i	Card D_i	Card $D_i/3^i$	$(x_1+2\|x_2\|)a^i/3^i$	$(1-3!/3^3)^{i/3}$
1	3	3	1	1,1834	
2	9	9	1	1,0737	
3	27	21	0,777	0,9741	0,777
4	81	63	0,777	0,8838	
5	243	171	0,703	0,8020	
6	729	441	0,605	0,7276	0,605
7	2.187	1269	0,580	0,6601	
8	6.561	3.375	0,514	0,5990	
9	19.683	9.099	0,462	0,5434	0,470
10	59.049	25.353	0,429	0,4930	
11	177.147	67.797	0,382	0,4473	
12	531.441	185.247	0,348	0,4058	0,366
13	1.594.323	507.327	0,318	0,3682	
14	4.782.969	1.368.405	0,286	0,3340	
15	14.348.907	3.742.245	0,260	0,3031	0,285
16	43.046.721	10.185.039	0,236	0,2750	
17	129.139.163	27.647.595	0,214	0,2495	
18	387.417.48×	75462057	0,194	0,2264	0,221
19	$1162252,4.10^2$	20516324×	0,176	0,2054	
20	$3486757,2.10^2$	5.581.5.731×	0,160	0,1863	
21	$1.0.46.0271,10^2$	$15.210.344.10^2$	0,145	0,1690	0,172
22	31380813.10^2	$41.364.307.10^2$	0.132	0.1534	
23	94142439.10^2	$11.260.990.10^3$	0,120	0,1392	
24	28242731.10^3	$30.661.704.10^3$	0,108	0,1262	0,134
25	8472819193.10^3	$83.420,1310^6$	0,098	0,1146	

$$x_1 = \frac{(3-b)(3-c)}{(a-b)(a-c)} \ , \ x_2 = \frac{(a-3)(3-c)}{(a-b)(b-c)} \ , \ x_3 = \frac{(a-3)(b-3)}{(a-c)(b-c)} \ ,$$

from which, after computing,

$$a = 2{,}721892 \qquad a^{-1} = 0{,}367392.$$

$$x_1 = 1{,}189301 \ , \ |x_2| = |x_3| = 0{,}05703.$$

$$|x_1| + 2|x_2| = 1{,}3043071 \ .$$

We also compute $(1 - k!/k^k)^{i/k}$, $i \in \mathbb{N}$, which is the ratio $(\text{Card } D_i)/k^i$ wher $D = X^* \backslash EX^*$, $E = \bigcup_{n \in \mathbb{N}} X^{3n}P$, which gives a natural extension of von Neumann sequences but is inferior to the one described here.

3.3 *Computing the mean delay*

By the construction of (9), the probability of producing an output symbol at the arrival of the i^{th} input symbol is the coefficient of degree i of the series

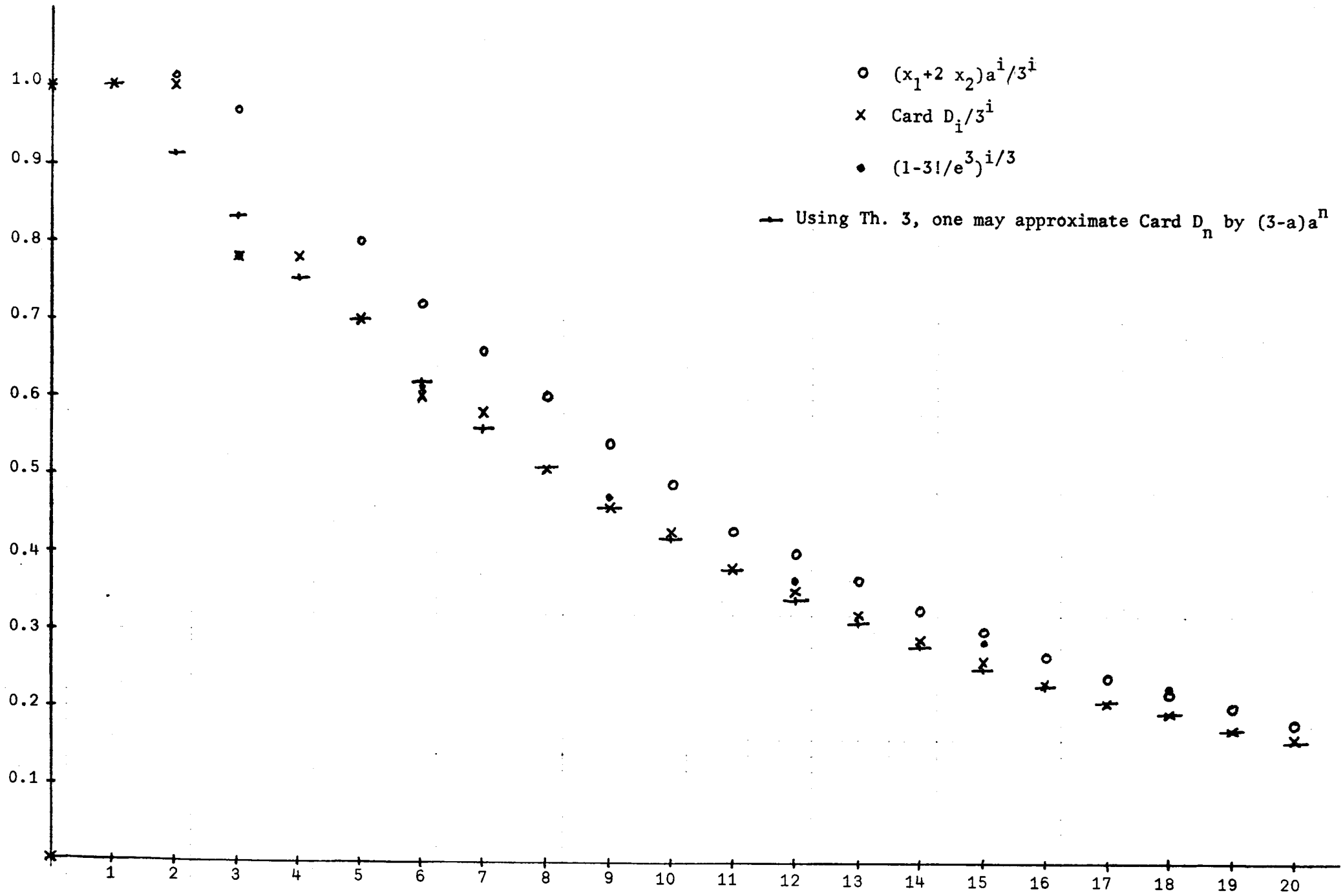
$$g(t) = k!\pi t^k \sum_{n \geq 0} a_{np} t^n = k!\pi t^k f_p^{-1}(t) \tag{23}$$

where $\pi = \prod_{x \in X} p(x)$, $f_p(t) = 1 - \sum_{2 \leq i \leq k} c_{ip} t^i$. This coefficient is $k!\pi a_{i-k,p}$ and the mean delay is:

$$g'(1) = k!\pi f_p^{-1}(1)(k - f_p^{-1}(1).f_p'(1))$$

$$g'(1) = k - \frac{f_p'(1)}{k!\pi} \ , \tag{24}$$

since $f_p(1) = k!\pi$, by (13).


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
$(x_1+2\ x_2)a^i/3^i$
Card $D_i/3^i$
$(1-3!/e^3)^{i/3}$
Using Th. 3, one may approximate Card D_n by $(3-a)a^n$

THEOREM 3: The mean delay is $\sum_{n\in\mathbb{N}} d_{np}$.

$\sum_{n\in\mathbb{N}} d_{np}$ converge since the denominator of the second member of (14) is actually a polynomial whose roots are outside the unit circle. If we write $f_p(t) - k!\pi f^k = (1 - t)h(t)$, we have by derivation $h(1) = kk!\pi - f_p'(1)$. Thus by (14), (13) and (24)

$$\sum_{n\in\mathbb{N}} d_{np} = h(1)/f_p(1) = h(1)/k!\pi = g'(1) .$$

4. *Toward an efficient easily computable procedure*

4.1 *The efficient construction*

P. Elias [2] published a procedure (independantly obtained, he says, by J. A. Lechner and J. Gill) which is proved to approach the best possible efficie (= the expected number of output digits per input digit) which is $H_k(p) = - \sum_{x\in X} p(x)\mathrm{Log}_k p(x)$ where k is the cardinality of X .

We describe the procedure in the binary case. The sequence of input is factorized into words of length n . Now a mapping $\delta : X^n \to X^*$ is defined and the images of the factors of the input sequence are concatenated. The $\binom{n}{i}$ words of X^n with i symbols 1 and $n - i$ symbols o form a set in which any two words are equiprobable. If the binary writing of $\binom{n}{i}$ is $2^{j_1} + 2^{j_2} + \ldots + 2^{j_s}$, then $\mathrm{Card}\,(X^{j_1} \cup X^{j_2} \cup \ldots \cup X^{j_s}) = \binom{n}{i}$ and the one-to-one image under δ of those words is $X^{j_1} \cup \ldots \cup X^{j_s}$. If $j_s = 0$, $X^{j_s} = \{\ell\}$ the empty word, and any word of this set may be mapped on the empty word. We compute the best possible efficiency ρ for $n = 2,4,8$. We have respectively $\rho = 25\%;\ 40,\ 62\%;\ 55,\ 18\%$.

It seems that for high efficiency n has to be large and the decoding by the mapping δ will need some computations.

4.2 *Decoding with permutation groups:*

We first remind the reader that if k is a prime and n a power of k, then $\binom{n}{i} \equiv 0$ modulo k for $i \neq 0, n$. Also for $i_1 + \ldots + i_k = n$

$$\frac{n!}{i_1! \ldots i_k!} \equiv 0 \bmod k, \tag{1}$$

if at least one of the i_j is not 0 or n.

So the set $S_{i_1, \ldots, i_k}$ of words of length n with i_ℓ repetition of the ℓ^{th} symbol of X, $\ell \in [1,k]$ may be partitioned into subsets with respective cardinalities $k^{j_1}, \ldots, k^{j_s}$, with $j_1 \geq \ldots \geq j_s > 0$. Actually the best possible partition for defining δ is given by the procedure given above, however any set of integers $\{j_1, \ldots, j_s\}$ with $\text{Card } S_{i_1,\ldots,i_k} = k^{j_1} + \ldots + k^{j_s}$ will allow the definition of a suitable δ. For example, any group G of permutations on Ω, Card $\Omega = n$, with Card $G = k^s$ (any power of k) will define such a partition. Let us denote again by G, with some notational abuse, the group of permutations on X^n induced by G. X^n will be partitioned into orbits under G and since

$$\text{Card } G_u \text{ Card } u^G = \text{Card } G$$

where G_u is the subgroup of G in which the word u is fixed and u^G is the orbit of u under G, we see that Card u^G divides k^s. This partition then allows a suitable δ. Practically, the problem reduce to the following. Given a $u \in X^n$, find u^G, then determine the location of u for the lexicographic ordering of u^G and define δu as the corresponding word in X^{j_u},

where $j_u = \mathrm{Log}_k \mathrm{Card}\ u^G$. The group G that we will use is the subgroup of order k^{2t-1} of the affine group of $\mathbb{Z}_n$, $n = k^t$, k a prime. Denote the permutation

(2) $$i \rightsquigarrow i + 1 \bmod n$$

by σ and the permutation

(3) $$i \rightsquigarrow a\, i \bmod n$$

by a , a being a unit of $\mathbb{Z}_n$ with order a power of k . Then every element in G has the form $a\sigma^j$, since we know that the cyclic group (σ) is an invariant subgroup of G . But we may also write $\sigma^j a$ for the generic element of G . This means that every permutation in G may be obtained by first applying a permutation of type (3) and afterward a cyclic shift. Denote by H the cyclic subgroup of order n of G and by K the subgroup of permutations with type (3). $G = HK = KH$ $u^G = u^{KH} = (u^K)^H$, and K has exactly one word in each orbit of the set u^G under H .

Then the decoding algorithm will be the following.

1. For every word v of u^K , find the $v' \in v^H$ with the highest lexicographic order. If u' is the word with the highest order in u^H , determine the order of u' among all v' . This is an integer between 0 and $\mathrm{Card}\ u^K$, which is a divisor, say, k^i of k^{t-1} . This integer written in basis k , has i digits. It is the left factor of δu .

2. Determine the order of u in u^H . Since $\mathrm{Card}\ u^H$, say, p^j divides p^t , we find an integer which will be written with j digits and concatenated with the known left factor of δu to form δu .

Here is an example. $k = 2$ $n = 8$. We have determined all leading words of the orbits v^H , for v having not more than four ones. The set is partitioned into orbits under G (which are found by letting operate the elements of (k) . We find a class of four leaders, with cyclic order 8, which means that a corresponding u will be decoded into a word of five digits.

10000000 | 8	10101000 | 8
11000000 | 8 10010000 |	11110000 | 8 11010010 |
10100000 | 8	11101000 | 11100010 | 8 11010100 |
10001000 | 4	11001010 |
11100000 | 8 10100100 |	11100100 | 8
	11011000 | 8
11010000 | 8 11000010 |	
	11001100 | 4
11001000 | 8 11000100 |	10101010 | 2

The figure gives the cyclic order of the elements in the corresponding class. The best possible efficiency of the procedure is

$$\rho = (5.2^5 + 9.4.2^4 + 8.3.2^3 + 3.2.2^2 + 2) \backslash 8.2^8 = 46{,}58\% ,$$

which is 86% more than the efficiency of the procedure of von Neumann.

When k is an odd prime, the group K is known to be cyclic, as a subgroup of a cyclic group [7]. This makes the computation easier. However suppose for example that K has two generators. Every element in K has the form $a^i b^j$. We compute u^a , u^{a^2} ... u^{a^s} ; where s is the smallest integer with $u^{a^s} \in u^H$. s has to be a power of 2 (maybe 2^0). This means $(a^s) \subset K_{u^H}$

We find similarly s' and then $(a^s, b^{s'}) = K_{u^H}$. (This is true because H is an invariant subgroup of $HK = G$). Elements of u^K are computed simultaniously and as soon as s' is determined, they are all known.

Remark: As a final remark we observe that for m a prime and whatever k an easily computable procedure with poor efficiency consists in the following. First factorize the given sequence into words of length m. If u is not a repeated symbol, the m cyclic shifts of each word u are distinct. If X is the set of input symbols, which may be linearly ordered, u may be assigned an integer in $[1,m]$ corresponding to its lexicographic rank among its cyclic shifts. This integer is the decoding symbol of u. This procedure is also an extension of the procedure of von Neumann.

Acknowledgement. I am grateful to Gordon Simons who introduced me to this problem and helped me find the references.

Appendix.

See N. Bourbaki "Polynonus et fraction rationnelles" for more details.

Some algebraic justifications.

For R any ring with unity, we denote by $R[[t]]$ the ring of all formal series of the form $\sum_{i\in\mathbb{N}} a_i t^i$, where $a_i \in R$, $\forall\, i \in \mathbb{N}$. We have

$$\sum_{i\in\mathbb{N}} a_i t^i + \sum_{i\in\mathbb{N}} b_i t^i = \sum_{i\in\mathbb{N}} (a_i + b_i) t^i$$

and

$$\sum_{i\in\mathbb{N}} a_i t^i \sum_{i\in\mathbb{N}} b_i t^i = \sum_{k\in\mathbb{N}} \Big(\sum_{0\le i\le k} a_{k-i} b_i \Big) t^k .$$

The formal derivative D is defined by

$$D \sum_{i\ge 0} a_i t^i = \sum_{i\ge 1} i a_i t^{i-1} ,$$

and the property

$$\forall\, u, v \in R[[t]] , \quad D(u.v) = uDv + (Du)v ,$$

is easily verified. From now on, R is assumed to be commutative.

Now suppose v,w are polynomials (i.e. formal series with almost all zero coefficients). It is known that v is invertible in $R[[t]]$ iff the coefficie of t^0 is a unit of R. For simplicity, let this coefficient be 1. $w = vu$, $Dw = vDu + uDv$ and, consequently $Du = v^{-2}(vDw - wDv)$.

Suppose $v^2 = 1 - c_1 t - c_2 t^2 \ldots - c_k t^k$. Then

$$v^{-2} = 1 + c_1 t + \ldots + c_k t^k + (c_1 t + \ldots + c_k t^k)^2 + \ldots .$$

D_u , which may be directly computed from the definition of D is also obtained by multiplying this series on the right by $vDw - wDv$. Suppose now that R is the ring of real numbers and that we have the situation

$$q = gh$$

where h is a polynomial and where the series of coefficients in g converges. Then, if in general, $v = \sum_{i \in \mathbb{N}} V_i t^i$,

$$\sum_{0 \le i \le n} q_i = \sum_{0 \le k \le n} \sum_{0 \le i \le k} g_{k-i} h_i = \sum_{0 \le i \le k} h_i \sum_{i \le k \le n} g_{k-}$$

$$= \sum_{0 \le i \le k} h_i \sum_{0 \le j \le n-i} g_j \qquad \lim_{n \to \infty} \sum_{0 \le i \le n} q_i = \left(\sum_{0 \le i \le k} h_i \right) \lim_{n \to \infty} \sum_{0 \le j \le n} g_j \ .$$

Hence $\sum q_i$ converges. If g is the series h^{-1} and if $\sum g_i$ converges we see that $\sum_{0 \le i \le \infty} g_i = 1/h(1)$. Applying this to the case where $q = Du$, $g = v^{-2}$ $h = vDw - wDv$, we see that if the series $\sum g_i$ converges we are allowed to write, with the usual notation for derivatives

$$\sum_{0 \le i \le \infty} q_i = (v(1)w'(1) - w(1)v'(1))/v^2(1)$$

if only we know that the series of coefficients of v^{-2} converges. This will be the case in our paper since no roots of v are in the closed unit circle.

Bibliography

[1] Dwass, Meyer "Unbiased coin tossing with discrete random variables", *Ann. Math. Stat.*, 1972, 860 - 864.

[2] Elias, Peter "The efficient construction of an unbiased random sequence", *Ann. Math. Stat.*,1972, Vol. 43, 865 - 870.

[3] Lechner, James "Efficient techniques for unbiasing a Bernoulli generator" (abstract), *Ann. Math. Stat.*, 1971, page 2171.

[4] Bernard, Jacques and Letac, Gerard "Construction d'evenements equiprobables et coefficients multinomiaux modulo p^n", *Illinois J. of Math.*, 1973, 317 - 332.

[5] Hoeffding, Wassily and Simons, Gordon "Unbiased coin tossing with a biased coin", *Ann. Math. Stat.*, 1970, Vol. 41, No. 2, 341 - 352.

[6] Schützenberger, M. P. "On a special class of recurrent events", *Ann. Math. Stat.*, 1961, Vol. 32, 1201 - 1213.

[7] Albert, A. "Fundamental concepts of Algebra", *Chic. Univ. Press.*

[8] von Neumann, John (1951), "Various techniques used in connection with random digits. Monte Carlo Method", *Applied Mathematics Series* No. 12, 36 - 38, U. S. National Bureau of Standards, Washington, D. C.