ALBANIAN JOURNAL OF MATHEMATICS
Volume 15, Number 2, Pages 73–83
ISSN: 1930-1235; (**2021**)

FROM REPEATED TOSSES OF A FAIR DIE TO THE RENEWAL THEOREM

JOSEPH BAK

*Department of Mathematics,
The City College of New York,
160 Convent Avenue New York, NY 10031.*

DANIELLA SHNAPS

*Jerusalem College of Technology
Havaad Haleumi 21, Givat Mordechai 91160, Jerusalem*

1. Introduction

Our interest in this topic began with a very basic question. If one tosses a fair die repeatedly, what is the probability that, at some point, the sum of the outcomes will equal exactly 1000? The answer to that question is a nondescript rational number with a denominator (in reduced form) of 6^{1000} . But it is also, interestingly and not coincidentally, equal to $\frac{2}{7}$, correct to well over one hundred decimal places. These assertions will emerge as we examine the bigger picture, investigating the probability of getting any positive integer as a partial sum of repeated trials of a given positive integer-valued random variable.

We begin with a "generalized die": a random variable X which assumes the values $1, 2, \ldots, k$ with equal probabilities $\frac{1}{k}$, and we let u_n denote the probability that n will belong to the sequence of partial sums $\{x_1, x_1 + x_2, \ldots\}$, obtained from repeated trials of X. Two recursive formulas and some basic laws of probability will lead to a simple closed form for u_n; $n = 1, 2, \ldots, k$. We then consider a more general random variable and obtain a complex (in both senses of the word) form for u_n which is valid for *all* $n \geq 1$. We also establish a very simple and highly intuitive value for $\lim_{n\to\infty} u_n$ and consider its relation to a famous problem of Frobenius.

The random variables we consider, along with their partial sums, fall under the broader category of renewal theory. "Renewal theory was developed in the first place for studying system reliability; namely, for solving problems related to the failure and replacement of components"; see [1, pg. 17]. A typical assumption in these cases is that a single component is in use at any time and a new identical replacement component is introduced as soon as the current one fails. Mathematically, then, renewal theory deals with a sequence of positive-valued *i.i.d.* random

E-mail addresses: jbak@ccny.cuny.edu, shnaps@jct.ac.il.

variables and their partial sums. Each of the identically distributed random variables X_i represents the "waiting time" for the ith arrival of a particular event (such as the failure of a component), while the partial sum

$$S_N = X_1 + X_2 + \cdots + X_N$$

represents the $N - th$ "arrival time".

Given the common distribution for the waiting times, a typical item of interest is the probability that there will be j arrivals before time n; i.e. $prob(S_j \leq n)$. The $i.i.d.$ waiting times X_i are usually continuously distributed, but they may be discrete, in which case we have a "discrete-time renewal process". In that case, we can assume that the values assumed are integers (that is, integral multiples of some fixed unit of time), and consider the nontrivial question of determining $prob(S_N = n)$, for any particular positive integers n and N. Finally, we can also consider the probability that n will be among the values assumed by the full sequence of arrival times $\{S_k\}$. This probability, then, is exactly what we have denoted above as u_n. We offer an explicit form for these probabilities using only elementary notions from linear algebra and combinatoric theory. Our formula for $\lim_{n\to\infty} u_n$ then offers an elementary proof of the renewal theorem (often referred to as the Erdös-Feller-Pollard Theorem) for discrete-time renewal processes with finitely many arrival times; see [1, p.29] and [4, pg.286].

2. The Generalized Die

Let X be a random variable which assumes the values $1, 2, \ldots, k$ with equal probabilities and let u_n be the probability that n will be among the partial sums generated by repeated trials of X. That possibility can be partitioned into the union of events E_i, each of which denotes that n appears as a partial sum with a final "toss", or summand, of $i; i = 1, 2, \ldots, k$. So, for $n \geq k + 1$, we have the recursive formula

$$u_n = \sum_{i=1}^{k} \Pr(E_i) = \sum_{i=1}^{k} \frac{1}{k} u_{n-i}. \tag{1}$$

In fact, Eq. (1) is valid for all $n \geq 1$ if we set $u_0 = 1; u_{-1} = u_{-2} = \cdots = u_{-k+1} = 0$. According to Eq. (1)

$$u_{n+1} = \frac{1}{k}[u_n + u_{n-1} + \cdots + u_{n-k+1}],$$

and

$$0 = u_n - \frac{1}{k}[u_{n-1} + u_{n-2} + \cdots + u_{n-k}],$$

so that adding the two equations yields the simpler recursive formula:

$$u_{n+1} = (\frac{k+1}{k})u_n - (\frac{1}{k})u_{n-k}, \text{ for } n \geq 1. \tag{2}$$

Since $u_1 = \frac{1}{k}$ and $u_{n-k} = 0$ for $n = 1, 2, \ldots, k-1$, Eq. (2) gives us the closed form:

$$u_n = (\frac{1}{k})(\frac{k+1}{k})^{n-1}, \text{ for } n = 1, 2, \ldots, k. \tag{3}$$

While recursive formulas will be a critical tool in almost all of our results, the closed form Eq. (3) can actually be derived directly. That is because u_n, for $1 \leq n \leq k$, can also be broken down, in the case of the generalized die as

$$u_n = \Pr(F_1) + \Pr(F_2) + \cdots + \Pr(F_n),$$

where F_i denotes the fact that n appears as a partial sum with exactly i summands. Note that any sequence of outcomes of length i will occur with probability $\frac{1}{k^i}$. Moreover, the number of such sequences of positive integers whose sum equals n is exactly $\binom{n-1}{i-1}$. So

$$u_n = \sum_{i=1}^{n} \binom{n-1}{i-1} \frac{1}{k^i} = \frac{1}{k} \sum_{j=0}^{n-1} \binom{n-1}{j} \frac{1}{k^j} = \frac{1}{k}\left(1 + \frac{1}{k}\right)^{n-1} = \frac{1}{k}\left(\frac{k+1}{k}\right)^{n-1}$$

Note that u_n is an increasing function of n for $1 \le n \le k$, and, for $n > k$, according to Eq. (1), u_n is the average of the previous k values of the sequence $\{u_j\}$. So, $Min_n u_n = u_1 = \frac{1}{k}$, and $Max_n u_n = u_k = (\frac{1}{k})(\frac{k+1}{k})^{k-1}$. (As a special case of a more general result in Section III, we will see that $\lim_{n\to\infty} u_n = \frac{2}{k+1}$)

Figure 1 shows the probabilities u_n when X is the generalized die with 10 "sides"; i.e., when X assumes the values $1, 2, \ldots, 10$ with equal probabilities. Note that u_n increases for $1 \le n \le 10$, and then levels off rather rapidly to $\lim_{n\to\infty} u_n$ which, according to the parenthetical remark, above, is $2/11$.

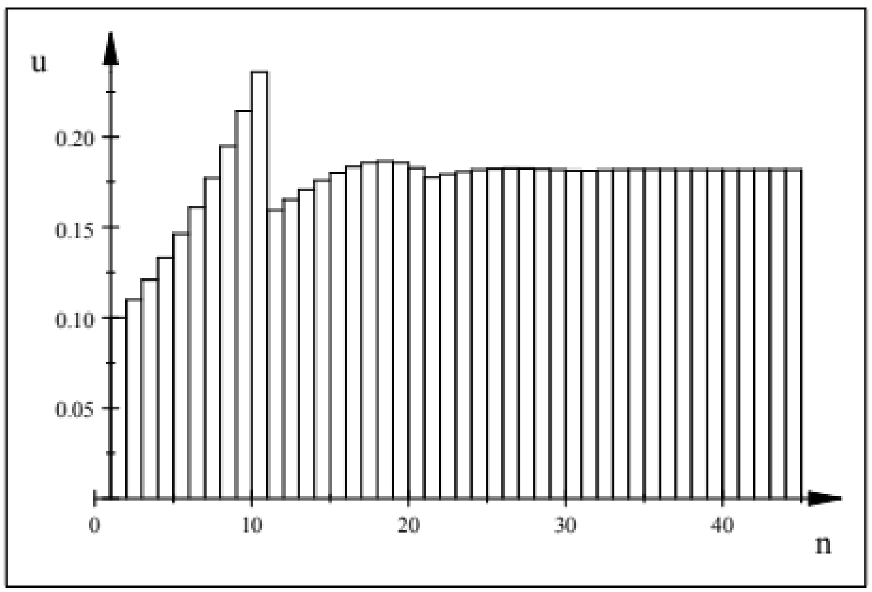

FIGURE 1. Probability of sums with a fair 10-sided die.

3. THE GENERAL POSITIVE INTEGER-VALUED RANDOM VARIABLE

While the closed form Eq. (3) for $\{u_n\}$, $1 \le n \le k$, is especially simple, we can obtain a more complex closed form for u_n which is valid for *all* n, and which applies much more generally. Let X assume any positive integer values $m_1, m_2, \ldots, m_k$ with probabilities $p_1, p_2, \ldots, p_k$, respectively (and $\sum p_i = 1$). Once again we let u_n denote the probability that n will belong to the sequence of partial sums $\{x_1, x_1 + x_2, \ldots\}$ where $x_1, x_2, \ldots$are the outcomes of repeated trials of X.

Suppose that the set of integer values, $\{m_1, m_2, \ldots, m_k\}$, is not relatively prime and that $g > 1$ equals $\gcd(m_1, m_2, \ldots, m_k)$. Then the partial sums of repeated trials of X cannot assume any values other than integer multiples of g. Moreover, u_{gn} is exactly equal to u_n for the "reduced" random variable $Y = X/g$ which assumes the relatively prime set of integer values $m_1/g, m_2/g, \ldots, m_k/g$. So it suffices to assume that $\{m_1, m_2, \ldots, m_k\}$ is relatively prime.

As we noted for the generalized die, the fact that n is achieved as a partial sum of a sequence of values of the *i.i.d.* random variables X_i can be partitioned into the union of events E_i; $i = 1, 2, \ldots, k$, where E_i denotes that n was achieved as a partial sum with a final summand of m_i. So

$$u_n = \sum_{i=1}^{k} \Pr(E_i) = \sum_{i=1}^{k} p_i u_{n-m_i},\ n > m_k \tag{4}$$

and formula Eq. (4) is equally valid for all $n \geq 1$ if we set $u_0 = 1$, $u_{-1} = u_{-2} = \cdots = u_{-m_k+1} = 0$. Let $M = m_k$. To solve the recursive formula Eq. (4); (*i.e.*, to obtain a closed form for u_n), we recall the well-known formula for the solution of an $M - th$ order linear homogeneous difference equation, namely:

Theorem 1. *With every Mth order linear homogeneous difference equation for the sequence $\{y_n\}$:*

$$a_M y_{n+M} + a_{M-1} y_{n+M-1} + \cdots + a_0 y_n = 0 \tag{5}$$

we associate the polynomial: $P(z) = a_M z^M + a_{M-1} z^{M-1} + \cdots + a_0$

If $P(z)$ has M distinct zeroes: $z_1, z_2, \ldots, z_M$, the general solution of Eq. (5) has the closed form

$$y_n = c_1 (z_1)^n + c_2 (z_2)^n + \cdots + c_M (z_M)^n \tag{6}$$

see [6, pg. 518].

If any of the M zeroes of P are of higher order; e.g., if $z_2 = z_3 = z_4$, the repeated terms in the solution are replaced by the derivatives of z^n for that value of z. So, in the given example, $c_2(z_2)^n + c_3(z_3)^n + c_4(z_4)^n$ would be replaced by $c_2(z_2)^n + c_3 n (z_2)^{n-1} + c_4 n(n-1)(z_2)^{n-2}$; see [6, pg. 519].

In either case, any *particular* solution $\{y_n\}$ is given by Eq. (6), or by its modified form when P has multiple zeroes, and with the coefficients $c_1, c_2, \ldots, c_M$ chosen so that Eq. (6) is valid for the initial M values of n.

□

So to solve the difference equation Eq. (4), we let $y_n = u_n$, and obtain the standard form

$$u_{m_k+n} - p_1 u_{m_k-m_1+n} - p_2 u_{m_k-m_2+n} - \cdots - p_k u_n = 0 \tag{7}$$

with the associated polynomial

$$P(z) = z^{m_k} - p_1 z^{m_k-m_1} - p_2 z^{m_k-m_2} - \cdots - p_k$$

Note that P has no zeroes z with $|z| > 1$ since, in that case, $|z^{m_k}| > |z^{m_k-m_i}|$ and hence

$$|z^{m_k}| > |p_1 z^{m_k-m_1} + p_2 z^{m_k-m_2} + \cdots + p_k|.$$

P has a simple zero at $z = 1$ since $P(1) = 1 - \sum p_i = 0$ and

$$P'(1) = m_k - \sum_{i=1}^{k} p_i (m_k - m_{i)} = \sum_{i=1}^{k} p_i m_i = E(X) > 0.$$

As we will see in Lem. 2, below, $z = 1$ is the only zero of $P(z)$ with absolute value 1.

Lemma 1. *If $\{t_1, t_2, \ldots, t_k\}$ is a relatively prime set of positive integers, there are no integers $\{s_1, s_2, \ldots, s_k\}$ such that $0 < s_n < t_n$ for all n, and*

$$\frac{s_i}{t_i} = \frac{s_j}{t_j}; \quad 1 \le i < j \le k.$$

Proof. Assume that p^e divides t_1, for some prime p and positive integer e. Then, if $\{t_1, t_2, \ldots, t_k\}$ is a relatively prime set of integers, at least one of the integers t_j is not divisible by p. Since

$$\frac{s_1}{t_1} = \frac{s_j}{t_j},$$

$s_1 t_j = t_1 s_j$ and it follows that p^e divides s_1. Since this is true for all prime factors of t_1, s_1 must be a multiple of t_1, proving the lemma.

□

Lemma 2. *With $\{m_1, m_2, \ldots, m_k\}$ and $\{p_1, p_2, \ldots, p_k\}$ as above, $z = 1$ is the only zero of*

$$P(z) = z^{m_k} - p_1 z^{m_k - m_1} - p_2 z^{m_k - m_2} - \cdots - p_k$$

with absolute value 1.

Proof. Suppose $|z| = 1$, and z is a zero of P. Then $|z^{m_k}| = |p_1 z^{m_k - m_1} + p_2 z^{m_k - m_2} + \cdots + p_k| = 1$. But

$$|\ p_1 z^{m_k - m_1} + p_2 z^{m_k - m_2} + \cdots + p_k| \ \le |\ p_1 + p_2 + \cdots + p_k| \ = 1.$$

and the "triangle inequality" above becomes an equality if and only if all the terms on the left side have the same argument. Since all the probabilities are positive real numbers, this means that all of the powers $z^{m_k - m_i}$; $i = 1, 2, \ldots, k-1$ would have to equal 1. But then the same is true for z^{m_k} and (by taking quotients), $z^{m_i} = 1$ for all i. Suppose then that $z = e^{i\theta}$, $0 \le \theta < 2\pi$, is a zero of P, and that $z^{m_i} = 1$. Then

$$z^{m_i} = e^{i\theta m_i} = 1$$

and $\theta m_i = 2\pi n_i$, for some integer n_i, with $n_i < m_i$, and $\frac{\theta}{2\pi} = \frac{n_i}{m_i}$ for $i = 1, 2, \ldots, k$ so that

$$\frac{n_i}{m_i} = \frac{n_j}{m_j}; \quad 1 \le i < j \le k.$$

Since $\{m_1, m_2, \ldots, m_k\}$ is a relatively prime set of integers, according to **??**, that is only possible if $\theta = 0$; *i.e.* if $z = 1$.

□

Combining the above lemma with our previous remarks gives:

Theorem 2. *The solution of Eq.* (7),

$$u_{m_k} - p_1 u_{m_k - m_1} - p_2 u_{m_k - m_2} - \cdots - p_k u_0 = 0, \ n \ge 1,$$

is

$$u_n = c_1 (z_1)^n + c_2 (z_2)^n + \cdots + c_M (z_M)^n, \tag{8}$$

or the indicated modification in the case of multiple roots, where $z_1, z_2, \ldots, z_M$ are the $M = m_k$ zeroes of P, $z_1 = 1$, $|z_i| < 1$, for $i = 2, \ldots, M$, and

$$c_1 = \frac{1}{P'(1)} = \frac{1}{E(X)}$$

Proof. As we saw in Thm. 1, Eq. (6) is valid for all positive integers n as long as the coefficients $c_1, c_2, \ldots, c_M$ are chosen so that Eq. (6) is valid for the first M values of n.

If we set $u_0 = 1$, and $u_n = 0$ for $n = -1, -2, \ldots, 1-M$, then the recursive formula Eq. (7) will give the correct values of u_n for all $n \geq 1$. So the values of the coefficients c_i can be found by solving the equations which guarantee that Eq. (6) is valid for $n = 0, -1, -2, \ldots, -M+1$. To simplify our notation, however, we will determine the coefficients $d_1, d_2, \ldots, d_M$ so that

$$w_n = d_1(z_1)^n + d_2(z_2)^n + \cdots + d_M(z_M)^n \tag{9}$$

satisfies $w_0 = w_1 = \cdots = w_{M-2} = 0$; $w_{M-1} = 1$. The resulting sequence $\{w_n\} = \{u_{n-M+1}\}$ so that $u_n = w_{n+M-1}$. It follows that $c_i = d_i(z_i)^{M-1}$. In particular, since $z_1 = 1$, $c_1 = d_1$.

To find the value of d_1, note that in order to get the desired starting values for w_n, $0 \leq n \leq M-1$, the coefficients d_i must satisfy the system of equations:

$$\begin{aligned}
&d_1 + d_2 + \cdots + d_M = 0\\
&d_1 z_1 + d_2 z_2 + \cdots + d_M z_M = 0\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
&d_1(z_1)^{M-2} + d_2(z_2)^{M-2} + \cdots + d_M(z_M)^{M-2} = 0\\
&d_1(z_1)^{M-1} + d_2(z_2)^{M-1} + \cdots + d_M(z_M)^{M-1} = 1
\end{aligned}$$

(or the modified form in the case of repeated roots).

According to Cramer's Rule, then,

$$d_1 = \frac{D_1}{D}, \text{ with } D = \begin{vmatrix} 1 & 1 & 1 & \cdot & 1 \\ z_1 & z_2 & z_3 & \cdot & z_M \\ z_1^2 & z_2^2 & z_3^2 & \cdot & z_M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ z_1^{M-1} & z_2^{M-1} & z_3^{M-1} & \cdot & z_M^{M-1} \end{vmatrix}$$

and

$$D_1 = \begin{vmatrix} 0 & 1 & 1 & \cdot & 1 \\ 0 & z_2 & z_3 & \cdot & z_M \\ 0 & z_2^2 & z_3^2 & \cdot & z_M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & z_2^{M-1} & z_3^{M-1} & \cdot & z_M^{M-1} \end{vmatrix}$$

(or the modified forms in the case of multiple roots).

To evaluate the above determinants, note that expansion by cofactors of the first column shows that

$$D(z) = \begin{vmatrix} 1 & 1 & 1 & \cdot & 1 \\ z & z_2 & z_3 & \cdot & z_M \\ z^2 & z_2^2 & z_3^2 & \cdot & z_M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ z^{M-1} & z_2^{M-1} & z_3^{M-1} & \cdot & z_M^{M-1} \end{vmatrix}$$

is a polynomial of degree $M-1$. It has zeroes at all the zeroes of $P(z)$, except at $z_1 = 1$, which are precisely the same zeroes as the $(M-1)$-st degree polynomial $R(z) = \frac{P(z)}{z-1}$.

[It is not hard to see that $D(z)$ has the same zeroes as $R(z) = \frac{P(z)}{z-1}$ even in the case of multiple zeroes. For example,

$$D^*(z) = \begin{vmatrix} 1 & 1 & 1 & \cdot & 1 \\ z & z_2 & 1 & \cdot & z_M \\ z^2 & z_2^2 & 2z_2 & \cdot & z_M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ z^{M-1} & z_2^{M-1} & (M-1)z_2^{M-2} & \cdot & z_M^{M-1} \end{vmatrix}$$

has a double zero at z_2. This follows from the fact that both $D^*(z)$ and its derivative have two identical columns, and therefore equal zero, when $z = z_2$.]

Since P and R both have a leading coefficient of 1,

$$D(z) = AR(z)$$

where A is the leading coefficient of $D(z)$; namely, the cofactor of z^{M-1}. Note then that $A = D_1$ and $D = D(z_1)$. So, by the continuity of R, (and the fact that $z_1 = 1$)

$$c_1 = d_1 = \frac{D_1}{D} = \frac{1}{R(z_1)} = \lim_{z \to 1} \frac{z-1}{P(z)} = \frac{1}{P'(1)} = \frac{1}{E(x)}.$$

□

Corollary 1. *Suppose a random variable X assumes the relatively prime set of positive integer values: $m_1, m_2, \ldots, m_k$, with probabilities $p_1, p_2, \ldots, p_k$, respectively. Let u_n denote the probability that n will belong to the sequence of partial sums $\{x_1, x_1 + x_2, \ldots\}$ generated by repeated trials of X. Then*

$$\lim_{n \to \infty} u_n = \frac{1}{E(X)}.$$

Proof. The proof follows immediately from the fact that ${z_i}^n$ (and $n{z_i}^{n-1}, n(n-1){z_i}^{n-2}, \ldots$) all approach 0 as $n \to \infty$, as long as $|z_i| < 1$. Hence

$$\lim_{n \to \infty} u_n = \lim_{n \to \infty} [c_1(z_1)^n + c_2(z_2)^n + \cdots + c_M(z_M)^n] = c_1 = \frac{1}{E(X)}$$

□

Note that the above result is highly intuitive since

$$E(X_1 + X_2 + \cdots + X_N) = N * E(x).$$

So, "on average", of the first $N * E(X)$ positive integers, only N will appear as sums of the outcomes of repeated trials of X. Thus, *assuming that u_n approaches some limit as $n \to \infty$*, it stands to reason that the limit should be $1/E(X)$.

The renewal theorem for continuous renewal processes offers a similarly intuitive result. Let $N(T)$ denote the expected number of sums (or arrival times) S_k in the interval $(0, T)$. The renewal theorem for continuous renewal processes, sometimes referred to as Blackwell's Theorem, asserts that, for any $h > 0$,

$$\lim_{T \to \infty} [N(T+h) - N(T)] = h/E(X)$$

see [2, pg. 145].

Cor. 1, in greater generality, is often referred to as the Erdös-Feller-Pollard Theorem, based on their article [3]. Interestingly, the only indication that [3] has anything to do with renewal theory (or probability in general) is its first line which says: "The following theorem is suggested by a theorem in probability" along with a footnote which adds: "To be published elsewhere". The theorem itself, as the title of the article suggests, is simply about power series. With very slight modifications (to avoid confusion with similar notation used differently in this article, we have changed the symbols p_k and P in [3] to q_k and Q, respectively, and to clarify that it represents a complex variable, the symbol x in [3] has been changed to z), it states:

Theorem 3 (Erdös-Feller-Pollard)**.** *Let q_k be a sequence of non-negative numbers for which $\sum_0^\infty q_k = 1$, and let $m = \sum_1^\infty kq_k \leq \infty$. Suppose further that*

$$Q(z) = \sum_0^\infty q_k z^k$$

is not a power series in z^t, for any integer $t > 1$. Then $1 - Q(z)$ has no zeros in the circle $|z| < 1$, and the series

$$U(z) = \frac{1}{1 - Q(z)} = \sum_0^\infty u_k z^k$$

has the property:

$$\lim_{n\to\infty} u_n = \frac{1}{m}.$$

If $m = \infty$, we define $1/m$ to be zero; see [3].

The connection between the theorem, above, and the renewal theorem is the notion of the "generating function" for a sequence; i.e., $\sum_0^\infty u_k z^k$ is the "generating function" for $\{u_k\}$. Suppose then that $\{u_k\}$ has the same meaning in the theorem as it does throughout this article, and that the *i.i.d.* random variables $X_i = X$ assume nonnegative integral values with $prob(X_i = k) = q_k$. Then, the recursive formula for $\{u_k\}$ makes it fairly easy to verify that $[1 - Q(z)]U(z) = 1$, so that the full hypothesis of the theorem is satisfied and we conclude that, indeed,

$$\lim_{n\to\infty} u_n = \frac{1}{m} = \frac{1}{E(X)}.$$

So, the renewal theorem is proven for all discrete-time renewal processes even if there are infinitely many arrival times; i.e., if $q_i > 0$ for infinitely many values of i, and even if $E(X) = \sum_1^\infty kq_k = \infty$. There are two proofs of the theorem in [3], both based on analytic function theory. But neither proof gives a closed form for $\{u_k\}$ so additional information, such as the rate at which $u_n \to \frac{1}{m}$, is not included.

On the other hand, if X assumes only finitely many distinct values with positive probabilty, the equation

$$\frac{1}{1 - Q(z)} = \sum\nolimits_0^\infty u_k z^k$$

can be solved for $\{u_k\}$ in terms of the poles of $\frac{1}{1-Q(z)}$; see for example [5], where formula Eq. (8) is derived by using a (complex) partial fraction decomposition of the rational function $\frac{1}{1-Q(z)}$.[Note that in this case, the zeroes of $1 - Q(z)$ are the reciprocals of the zeroes of the polynomial which we labeled $P(z)$ in our proof of Thm. 2. Assume, as we did in Thm. 2, that X takes only the finitely many positive integers $m_1, m_2, \ldots, m_k$ with positive probabilities $q_1, q_2, \ldots, q_k$, respectively. Then

$$H(z) = 1 - Q(z) = 1 - q_1 z^{m_1} - q_2 z^{m_2} - \cdots - q_k z^{m_k}$$

while

$$P(z) = z^{m_k} - q_1 z^{m_k - m_1} - q_2 z^{m_k - m_2} - \cdots - q_k = z^{m_k} H(\frac{1}{z}).]$$

4. An Application to the Frobenius Problem

An interesting problem of Frobenius seeks to find a formula for the greatest integer which is *not* expressible as a sum of nonnegative integral multiples of the relatively prime integers $m_1, m_2, \ldots, m_k$. A related question is how many positive integers in total are not expressible as nonnegative integral combinations of $m_1, m_2, \ldots, m_k$. (The answer is well-known if $k = 2$. The largest integer not expressible as a positive integral combination of a and b is $ab - (a + b)$, and the total number of non-expressible integers is $(a - 1)(b - 1)/2$. But there seem to be no simple closed formulas to answer either of the questions if $k > 2$.) Both questions, of course, assume that every "sufficiently large" integer is expressible as a positive integral combination of $m_1, m_2, \ldots, m_k$. This result is cited and used in the proof of the Erdös-Feller-Pollard Theorem [3, pg. 203]. Interestingly, as we will see below, it can be also derived as another corollary of Thm. 2.

Corollary 2. *For every set of relatively prime integers $m_1, m_2, \ldots, m_k$, there exists an integer $L = L(m_1, m_2, \ldots, m_k)$ such that any integer $n > L$ can be expre ssed as nonnegative integral combination of $m_1, m_2, \ldots, m_k$.*

Proof. Since $\lim_{n\to\infty} u_n = \frac{1}{E(X)} > 0$, there exists an integer L such that $n > L$ implies that $u_n > 0$. But then all such values of n are obviously expressible as positive linear combinations of $m_1, m_2, \ldots, m_k$.

□

Fig. 2 shows the values of u_n when the random variable X assumes the values 5 and 7 with equal probabilities. Note that the greatest of the twelve positive integers n with $u_n = 0$ is 23. In this case, $\lim_{n\to\infty} u_n = 1/6$.

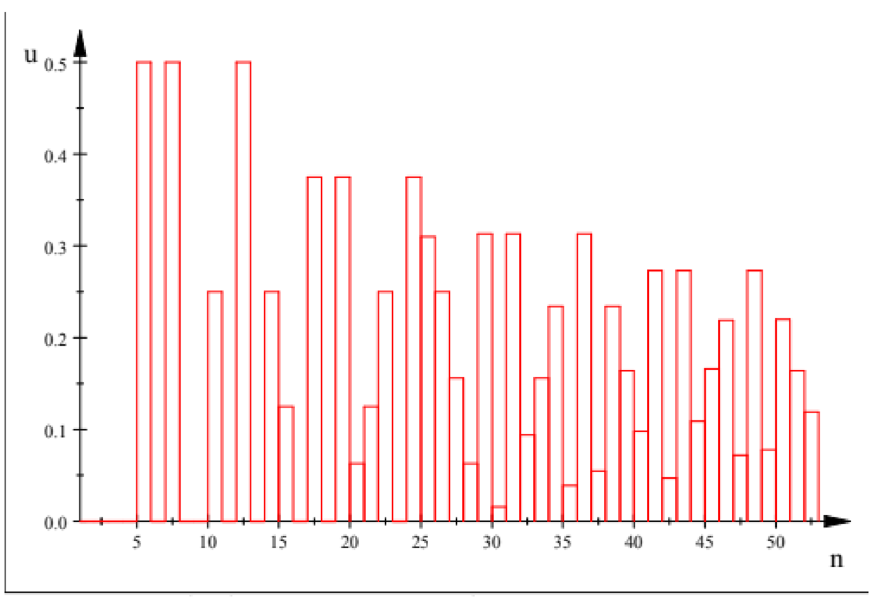

Figure 2. Values of u if X assumes the values of 5 and 7 with equal probability.

5. Some Additional Observations for the Generalized Die

If we return once again to the generalized die, we can find several special properties aside from the closed form Eq. (3). For one thing, it follows easily from the recursive formula Eq. (1) that u_n is a rational number with a denominator equal to k^n and a numerator which is congruent to $1 \mod k$. In addition, the polynomial associated with the difference equation for u_n:

$$P(z) = z^k - \frac{1}{k}(z^{k-1} + z^{k-2} + \cdots + 1) = \frac{1}{k}(kz^k - z^{k-1} - z^{k-2} - \cdots - 1)$$

has k distinct zeroes. This follows from the facts that

(i) if $S(z) = k(z-1)P(z) = kz^{k+1} - (k+1)z^k + 1$, the only zeroes of

$$S'(z) = k(k+1)z^{k-1}(z-1)$$

are $z = 0$ and $z = 1$;

(ii) $P'(1) = E(X) > 0$ so $z = 1$ is a simple zero of P, and

(iii) with the exception of $z = 1$, z is a multiple zero of P if and only if it is a multiple zero of S (whose only multiple zero is $z = 1$).

Since the zeroes of P are distinct, we can find all of the coefficients d_i using Cramer's Rule in the same way that we found d_1 in the proof of Thm. 2; that is,

$$d_i = \frac{D_i}{D}$$

where

$$D = \begin{vmatrix} 1 & 1 & 1 & \cdot & 1 \\ z_1 & z_2 & z_3 & \cdot & z_k \\ z_1^2 & z_2^2 & z_3^2 & \cdot & z_k^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ z_1^{k-1} & z_2^{k-1} & z_3^{k-1} & \cdot & z_k^{k-1} \end{vmatrix}$$

is the familiar Vandermonde determinant which is equal to $\prod(z_j - z_l)$, with the product taken over all j, l with $1 \le l < j \le k$, and where

$$D_i = \begin{vmatrix} 1 & 1 & 0 & \cdot & 1 \\ z_1 & z_2 & 0 & \cdot & z_M \\ z_1^2 & z_2^2 & 0 & \cdot & z_M^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ z_1^{M-1} & z_2^{M-1} & 1 & \cdot & z_M^{M-1} \end{vmatrix}$$

has the ith column of D replaced by the column of starting values $0, 0, \ldots, 0, 1$.

Expansion by the ith column shows that $D_i = \pm\prod'(z_j - z_l)$, where the product is taken over all j, l with $1 \le l < j \le k$, and $l \neq i$, $j \neq i$, so that

$$d_i = \frac{D_i}{D} = \frac{\pm 1}{\prod(z_j - z_i)} \tag{10}$$

with the product taken over all $j \neq i$.

For the generalized die X, $E(X) = \frac{k+1}{2}$, so according to Eq. (6)

$$u_n = c_1(z_1)^n + c_2(z_2)^n + \cdots + c_k(z_k)^n = \frac{2}{k+1} + c_2(z_2)^n + \cdots + c_k(z_k)^n.$$

Hence

$$|u_n - \frac{2}{k+1}| = |c_2(z_2)^n + \cdots + c_k(z_k)^n|,$$

and we can obtain an upper bound for the latter expression by noting that $|c_i| < |d_i|$ for all i, and by obtaining estimates for $z_2, z_3, \ldots, z_k$. According to Eq. (9), this will lead to upper bounds for $|d_i|$ as well as for $|z_i|^n$.

In particular, if we let $k = 6$; *i.e.*, if we return to the original question regarding repeated tosses of an ordinary fair die, we find that the zeroes of P are $z_1 = 1$, and, correct to 4 decimal places, $z_2 = -0.6703$; $z_3, z_4 = 0.2942 \pm 0.6684\, i$; $z_5, z_6 = -0.3757 \pm 0.5702\, i$. It follows that $|z_j - z_i\,| > \frac{1}{2}$ for all $j \neq i$. According to Eq. (10), then, $\mid d_i| < 32$ and $Max \mid z_i| = \mid z_3| < 0.731$. Finally, since $0.731^{1000} < 10^{-136}$, $|u_{1000} - \frac{2}{7}| < 160 * 10^{-136}$, and the probability of getting 1000 is equal to $\frac{2}{7}$, correct to well over 100 decimal places.

REFERENCES

[1] Vlad Stefan Barbu and Nikolaos Limnios, *Semi-Markov chains and hidden semi-Markov models toward applications*, Lecture Notes in Statistics, vol. 191, Springer, New York, 2008. Their use in reliability and DNA analysis. MR2452304

[2] David Blackwell, *A renewal theorem*, Duke Math. J. **15** (1948), 145–150. MR24093

[3] P. Erdös, W. Feller, and H. Pollard, *A property of power series with positive coefficients*, Bull. Amer. Math. Soc. **55** (1949), 201–204. MR27867

[4] William Feller, *An introduction to probability theory and its applications. Vol. I*, Third, John Wiley & Sons, Inc., New York-London-Sydney, 1968. MR0228020

[5] S. P. Lalley, *Renewal theory*, University of Chicago, 2016.

[6] K. Rosen, *Discrete mathematics and its applications*, McGraw-Hill, 2012.