Western Washington University
Western CEDAR

Mathematics

College of Science and Engineering

2-2007

An Unexpected Limit of Expected Values

Branko Ćurgus
Western Washington University, branko.curgus@wwu.edu

Robert I. Jewett
Western Washington University

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Expo. Math. ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

EXPOSITIONES MATHEMATICAE

www.elsevier.de/exmath

An unexpected limit of expected values

Branko Ćurgus*, Robert I. Jewett

Department of Mathematics, Western Washington University, 516 High Street, Bellingham, WA 98225, USA

Received 4 October 2005

Abstract

Let $t \geqslant 0$. Select numbers randomly from the interval $[0, 1]$ until the sum is greater than t. Let $\alpha(t)$ be the expected number of selections. We prove that $\alpha(t) = e^t$ for $0 \leqslant t \leqslant 1$. Moreover, $\lim_{t\to+\infty} (\alpha(t) - 2t) = \frac{2}{3}$. This limit is a special case of our asymptotic results for solutions of the delay differential equation $f'(t) = f(t) - f(t-1)$ for $t > 1$. We also consider four other solutions of this equation that are related to the above selection process.

MSC 2000: Primary 34K06; 60G50; 34K60

Keywords: Linear delay differential equations; Asymptotic behavior; Sums of independent random variables; Random walks

1. Introduction

Consider the following question: Select numbers randomly from the interval $\mathbb{I} = [0, 1]$ until the sum is greater than 1. What is the expected number of selections?

The unexpected expected value is e.

This question appeared as a Putnam examination problem in 1958 [1, Part I, Problem 3]. It has attracted considerable attention, so that solutions appear in [1–5]. The question is also mentioned by John Derbyshire [6, p. 366]. We first learned about this question in a talk given by Erol Peköz [7].

* Corresponding author. Tel.: +1 360 650 3484; fax: +1 360 650 7788.
E-mail address: curgus@cc.wwu.edu (B. Ćurgus).

0723-0869/$ - see front matter
doi:10.1016/j.exmath.2006.01.004

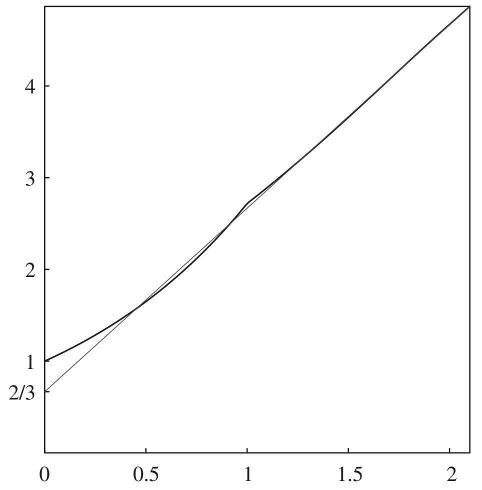

Fig. 1. $\alpha(t)$ and $2t + \frac{2}{3}$.

In this article we consider a more general question: Let $t \geqslant 0$. Select numbers randomly from the interval $\mathbb{I}$ until the sum is greater than t. What is the expected number $\alpha(t)$ of selections?

That more general question can be answered using a formula derived by James Uspensky [8, Example 3, p. 278]. Also, see [5]. Uspensky attributes that formula to Laplace. We use a different method. It turns out that $\alpha(t) = e^t$ for $0 \leqslant t \leqslant 1$. The answer for $t > 1$ is more complicated. In this article we investigate the function α.

It turns out that α satisfies a delay differential equation

$$f'(t) = f(t) - f(t-1), \quad t > 1.$$

This leads to an explicit formula for α. Surprisingly, the formula for α does not reveal its asymptotic behavior. Since the average value of a selection from $\mathbb{I}$ is $\frac{1}{2}$, one might reasonably guess that $\alpha(t)$ is approximately $2t$ for large values of t. But is this guess correct? The graph of α in Fig. 1 and its values in Table 1 strongly suggest that, in fact,

$$\lim_{t \to +\infty} (\alpha(t) - 2t) = \frac{2}{3}. \tag{1.1}$$

In addition to α we study the function β, where $\beta(t)$ is the expected value of the last selected number. We thought that $\beta(t) \to \frac{1}{2}$ as $t \to +\infty$. Then we discovered the following relation between α and β:

$$\beta(t) = t + 1 - \frac{\alpha(t)}{2} \quad \text{for } t \geqslant 0. \tag{1.2}$$

Table 1
Exact and approximate values of α

n	$\alpha(n)$	Approximate values
1	e	2.71828183
2	$e^2 - e$	4.67077427
3	$\frac{1}{2}\left(2e^3 - 4e^2 + e\right)$	6.66656564
4	$\frac{1}{6}\left(6e^4 - 18e^3 + 12e^2 - e\right)$	8.66660449
5	$\frac{1}{24}\left(24e^5 - 96e^4 + 108e^3 - 32e^2 + e\right)$	10.66666207
6	$\frac{1}{120}\left(120\,e^6 - 600\,e^5 + 960\,e^4 - 540\,e^3 + 80\,e^2 - e\right)$	12.66666714

From (1.1) and (1.2),

$$\lim_{t \to +\infty} \beta(t) = \frac{2}{3},$$

which is the unexpected limit in the title of this article. Notice that the value $\beta(1) = 2 - e/2 \approx 0.641$ is surprisingly close to $\frac{2}{3}$.

Our proofs of asymptotic properties of α and β are based on the fact that they belong to a class of functions which we call *delay functions*; see Definition 4.1. Sections 4–7 discuss the existence, uniqueness, asymptotic behavior, and oscillations of delay functions. In Sections 8 and 9 we prove that α and β are delay functions. As a consequence, we obtain the results about α and β stated above. In Section 10, we briefly discuss three more delay functions related to the process under consideration. We conclude the article with a heuristic argument for the asymptotic behavior of all five delay functions.

Apparently, asymptotic properties of α have not been considered in the literature, and β has not been studied at all.

An even more general question is as follows: let $a > 0$. Select numbers randomly from the interval $[0, a]$ until the sum is greater than t. What is the expected number $\alpha_a(t)$ of selections? This question can easily be reduced to the second question above: $\alpha_a(t) = \alpha(t/a)$.

This paragraph concerns the book [9] by Diekmann et al. It gives a general treatment of delay differential equations, using methods of functional analysis. In particular see Exercise 3.15 and Example 5.13, both in Chapter IV. The formula at the top of page 109 can be used to derive the linear function $2t + \frac{2}{3}$ which appears in (1.1).

Sections 2–7 in this article use undergraduate mathematics only. In Sections 8 and 9 we integrate on the Hilbert cube.

We use $\mathbb{R}$ for the set of all real numbers and $\mathbb{N}$ for the set of all positive integers. As mentioned above, $\mathbb{I}$ stands for the closed interval $[0, 1]$. In this article we deal only with real numbers.

2. A partial solution

We first prove that $\alpha(1) = e$. Let $n \in \mathbb{N}$. Let $\mathbb{R}^n$ be n-dimensional Euclidean space. Consider two subsets of $\mathbb{R}^n$:

$$C_n = \left\{(x_1, \ldots, x_n) \in \mathbb{I}^n : x_1 + \cdots + x_n \leqslant 1\right\},$$
$$D_n = \left\{(y_1, \ldots, y_n) \in \mathbb{I}^n : y_1 \leqslant \cdots \leqslant y_n\right\}.$$

Let

$$p_0 := 1 \quad \text{and} \quad p_n := \operatorname{Vol}(C_n), \quad n \in \mathbb{N}.$$

That is, p_n is the n-dimensional volume of C_n. Notice that p_n is the probability that the sum of n randomly selected numbers from $\mathbb{I}$ does not exceed 1.

To evaluate p_n, define $L_n : \mathbb{R}^n \to \mathbb{R}^n$ to be the linear transformation given by

$$L_n (x_1, \ldots, x_n) = (x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_n).$$

The determinant of its matrix is 1 and $L_n(C_n) = D_n$. Therefore the volumes of C_n and D_n are equal. From any given set of n distinct numbers in $\mathbb{I}$, we can form $n!$ distinct n-tuples, only one of which is in increasing order. Therefore, the volume of D_n is $1/n!$. Thus, $p_n = 1/n!$ for all $n \in \mathbb{N}$.

Since $\{p_n\}$ is a nonincreasing sequence that converges to 0, the definition of expected value gives

$$\alpha(1) = \sum_{n=1}^{+\infty} n\,(p_{n-1} - p_n). \tag{2.1}$$

Notice that if $\{v_n\}$ is a sequence of numbers and $k \in \mathbb{N}$, then

$$\sum_{n=1}^{k} n\,(v_{n-1} - v_n) = \left(\sum_{n=0}^{k-1} v_n\right) - k v_k. \tag{2.2}$$

By (2.1), (2.2) and the fact that $kp_k \to 0$ as $k \to +\infty$,

$$\alpha(1) = \sum_{n=0}^{+\infty} \frac{1}{n!} = e.$$

To determine $\alpha(t)$ in general, for $n \in \mathbb{N}$ and $t \geqslant 0$, we define the sets

$$C_{n,t} = \left\{(x_1, \ldots, x_n) \in \mathbb{I}^n : x_1 + \cdots + x_n \leqslant t\right\} \tag{2.3}$$

and let

$$p_{n,t} := \operatorname{Vol}(C_{n,t}) \quad \text{for } n \in \mathbb{N} \tag{2.4}$$

and

$$p_{0,t} := 1.$$

Notice that $p_{n,t}$ is the probability that the sum of n randomly selected numbers from $\mathbb{I}$ does not exceed t. Clearly $p_{n,t}$ is nonincreasing in n and nondecreasing in t.

Let $n \in \mathbb{N}$, $t > 0$, and let $M_{n,t} : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation of scalar multiplication by t. Then

$$M_{n,t}(C_n) = \left\{ (x_1, \ldots, x_n) \in [0, t]^n : \sum_{j=1}^{n} x_j \leqslant t \right\}.$$

Since the determinant of $M_{n,t}$ is t^n, we conclude that

$$\operatorname{Vol}\left(M_{n,t}(C_n)\right) = t^n p_n = t^n/n!.$$

If $0 < t \leqslant 1$, then $M_{n,t}(C_n) \subset \mathbb{I}^n$, and therefore $M_{n,t}\,(C_n) = C_{n,t}$. Consequently,

$$p_{n,t} = \frac{t^n}{n!} \quad \text{for } 0 < t \leqslant 1, \quad n \in \mathbb{N}. \tag{2.5}$$

For $t > 1$ we have $C_{n,t} \subset M_{n,t}\,(C_n)$ and therefore

$$p_{n,t} < \frac{t^n}{n!} \quad \text{for } 1 < t, \quad n \in \mathbb{N}. \tag{2.6}$$

Since $\sum_{n=0}^{+\infty} t^n/n! = e^t$, we have

$$\sum_{n=0}^{+\infty} p_{n,t} = e^t \quad \text{for } 0 < t \leqslant 1,$$
$$\sum_{n=0}^{+\infty} p_{n,t} < e^t \quad \text{for } t > 1.$$

The value of $\alpha(t)$ for $t > 0$ is now calculated using (2.2) in the same way as before,

$$\alpha(t) = \sum_{n=1}^{+\infty} n \left(p_{n-1,t} - p_{n,t}\right) = \sum_{n=0}^{+\infty} p_{n,t}.$$

Thus,

$$\alpha(t) = e^t \quad \text{for } t \in \mathbb{I},$$
$$\alpha(t) < e^t \quad \text{for } t > 1. \tag{2.7}$$

It is clear from Definition (2.4) that the functions $t \mapsto p_{n,t}$ are continuous for each $n \in \mathbb{N}$. Inequalities (2.5), (2.6), and the Weierstrass M-test, imply that α is continuous on $[0, t]$ for each $t > 0$. Thus α is continuous on $[0, +\infty)$.

3. The delay differential equation

For $t > 1$, calculating $p_{n,t}$ is complicated. Therefore, we derive a differential equation for $\alpha(t)$ to evaluate it. On average, it will take $\alpha(t)$ selections to reach $t > 0$. Given that the first selection is $u \in \mathbb{I}$, the expected number of additional selections to exceed $t > 1$ is $\alpha(t-u)$. It is plausible to surmise that $\alpha(t)$ equals 1 plus the average of $\alpha(t-u)$ as u varies over $\mathbb{I}$. That is,

$$\alpha(t) = 1 + \int_0^1 \alpha(t-u)\,\mathrm{d}u \quad \text{for } t > 1. \tag{3.1}$$

A rigorous derivation of this equation is given in Section 8. Eq. (3.1) can be rewritten as

$$\alpha(t) = 1 + \int_{t-1}^{t} \alpha(s)\,\mathrm{d}s \quad \text{for } t > 1. \tag{3.2}$$

Since α is continuous on $[0, +\infty)$, (3.2) implies that α is differentiable on $(1, +\infty)$. Therefore, to find α we have to solve the equation

$$\alpha'(t) = \alpha(t) - \alpha(t-1) \quad \text{for } t > 1 \tag{3.3}$$

with $\alpha(t) = e^t$ for $t \in \mathbb{I}$. Eq. (3.3) is an example of a *delay differential equation*. It can be used to find an explicit recursive formula for α. Multiplying by the integrating factor e^{-t}, we find that Eq. (3.3) is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-t}\alpha(t)\right) = -e^{-t}\alpha(t-1) \quad \text{for } t > 1. \tag{3.4}$$

Starting with

$$\alpha(t) = e^t, \quad t \in [0, 1],$$

Eq. (3.4) simplifies to $(\mathrm{d}/\mathrm{d}t)\left(e^{-t}\alpha(t)\right) = -1/e$ on $(1, 2]$. Hence,

$$\alpha(t) = e^t\left(1 + \left(\frac{1-t}{e}\right)\right), \quad t \in [1, 2]. \tag{3.5}$$

Proceeding successively on intervals $[2, 3]$ and $[3, 4]$ we find that the continuous solution of (3.4) is

$$\alpha(t) = e^t\left(1 + \left(\frac{1-t}{e}\right) + \frac{1}{2}\left(\frac{2-t}{e}\right)^2\right), \quad t \in [2, 3],$$

and

$$\alpha(t) = e^t\left(1 + \left(\frac{1-t}{e}\right) + \frac{1}{2}\left(\frac{2-t}{e}\right)^2 + \frac{1}{6}\left(\frac{3-t}{e}\right)^3\right), \quad t \in [3, 4].$$

The pattern is now clear; for $t \geqslant 0$ we define

$$A(t) := \sum_{k=0}^{\lfloor t \rfloor} \frac{1}{k!}\left(\frac{k-t}{e}\right)^k.$$

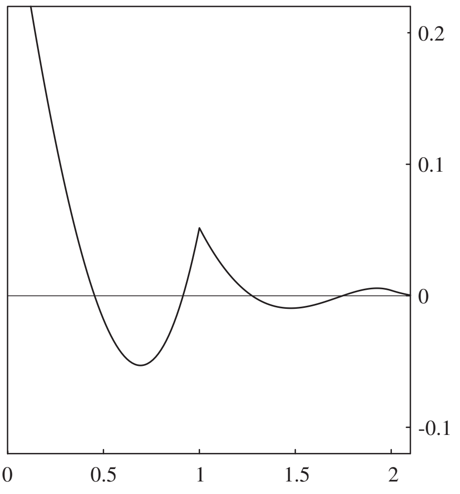

Fig. 2. $\alpha(t) - (2t + \frac{2}{3})$.

Here $\lfloor t \rfloor$ denotes the largest integer which does not exceed t. It is easy to show that A is continuous on $[0, +\infty)$ and $A(t) = 1$ for $t \in \mathbb{I}$. Differentiation of A for $t > 1$ and $t \notin \mathbb{N}$ gives

$$\begin{aligned} A'(t) &= -\frac{1}{e} \sum_{k=1}^{\lfloor t \rfloor} \frac{1}{(k-1)!} \left(\frac{k-t}{e} \right)^{k-1} \\ &= -\frac{1}{e} \sum_{j=0}^{\lfloor t-1 \rfloor} \frac{1}{j!} \left(\frac{j-(t-1)}{e} \right)^{j} \\ &= -\frac{1}{e} A(t-1). \end{aligned}$$

This formula and the continuity of A imply that A is differentiable on $(1, +\infty)$. Notice that A is not differentiable at 1. For $t > 1$ we now have

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(e^t A(t) \right) = e^t A(t) + e^t A'(t) = e^t A(t) - e^{t-1} A(t-1).$$

Thus, the function $t \mapsto e^t A(t)$, $t \geqslant 0$, satisfies (2.7) and (3.3). Assuming uniqueness, which is dealt with in Theorem 5.2,

$$\alpha(t) = e^t A(t).$$

However, this formula does not even suggest anything about the asymptotic behavior of α. But, the asymptotic behavior of α given in the Introduction is strongly suggested by the graph in Fig. 1 and the approximate values in Table 1. Furthermore, Figs. 1 and 2 and Table 1 suggest that the function $\alpha(t)$ oscillates about the line $2t + \frac{2}{3}$. We consider oscillations in Section 7. Table 1 also appears in [5].

The reasoning used to derive (3.3) can also be used on the interval $(0, 1)$ to obtain an alternative argument for (2.7). Although this reasoning is not rigorous, we find it useful for understanding α. To make it rigorous one would use the method of Sections 8 and 9.

Let $0 < t < 1$. The first selection exceeds t with probability $1 - t$ and the first selection does not exceed t with probability t. Given that the first selection u belongs to $[0, t]$, the expected number of additional selections to exceed t is $\alpha(t - u)$. Thus, the average number of selections for which the first selected number belongs to $[0, t]$ is

$$1 + \frac{1}{t}\int_0^t \alpha(t-u)\,\mathrm{d}u = 1 + \frac{1}{t}\int_0^t \alpha(s)\,\mathrm{d}s.$$

Therefore it is plausible that

$$\alpha(t) = (1-t)\,1 + t\left(1 + \frac{1}{t}\int_0^t \alpha(s)\,\mathrm{d}s\right) = 1 + \int_0^t \alpha(s)\,\mathrm{d}s, \quad 0 < t < 1.$$

Consequently, $\alpha'(t) = \alpha(t)$ for $0 < t < 1$. In Section 2 we proved that α is continuous. Since by definition $\alpha(0) = 1$, we obtain $\alpha(t) = e^t$ for $t \in [0, 1)$.

4. Delay functions

The delay differential equation (3.3) is the motivation for the following definition.

Definition 4.1. A function f defined on $[0, +\infty)$ will be called a *delay function* if:

(i) f is continuous on $[0, +\infty)$,
(ii) f is differentiable on $(1, +\infty)$,
(iii) $f'(t) = f(t) - f(t-1)$ for all $t > 1$.

The following proposition is easy to prove.

Proposition 4.2. *A linear combination of delay functions is a delay function. If f is a delay function and $c > 0$, then the function $g(t) = f(t + c)$, $t \geqslant 0$, is a delay function.*

Theorem 4.3. *Let f be a continuous function on $[0, +\infty)$. Set*

$$a = f(1) - \int_0^1 f(s)\,\mathrm{d}s \quad \textit{and} \quad b = \int_0^1 sf(s)\,\mathrm{d}s.$$

The following four statements are equivalent:

(a) $at + (b - a) = \int_0^1 (1-u) f(t-u)\,\mathrm{d}u$ *for* $t \geqslant 1$.
(b) $f(t) = a + \int_0^1 f(t-u)\,\mathrm{d}u$ *for* $t \geqslant 1$.
(c) $f(t) = at + b + \int_0^1 uf(t-u)\,\mathrm{d}u$ *for* $t \geqslant 1$.
(d) *f is a delay function.*

Proof. First, we set the notation for the functions that appear in the first three statements. For $t \geqslant 1$ we put

$$\begin{aligned}
\eta(t) &= \int_0^1 (1-u) f(t-u)\,\mathrm{d}u,\\
\varphi(t) &= f(t) - \int_0^1 f(t-u)\,\mathrm{d}u,\\
\psi(t) &= f(t) - \int_0^1 u f(t-u)\,\mathrm{d}u.
\end{aligned} \tag{4.1}$$

A straightforward calculation gives

$$\eta(1) = b, \quad \varphi(1) = a, \quad \psi(1) = a + b.$$

The change of variable $t - u = s$ in (4.1) yields

$$\eta(t) = \int_{t-1}^{t} (1 + s - t) f(s)\,\mathrm{d}s \quad \text{for } t \geqslant 1.$$

Then, for $t > 1$,

$$\begin{aligned}
\eta'(t) &= (1 + t - t) f(t) - (1 + t - 1 - t) f(t-1) + \int_{t-1}^{t} (-1) f(s)\,\mathrm{d}s\\
&= f(t) - \int_{t-1}^{t} f(s)\,\mathrm{d}s\\
&= f(t) - \int_0^1 f(t-u)\,\mathrm{d}u.
\end{aligned}$$

Thus,

$$\eta' = \varphi \quad \text{and} \quad \psi - \eta = \varphi,$$

the second equation being a straightforward consequence of the definitions.

We proceed by proving the following equivalences:

$$\text{(a)} \Leftrightarrow \text{(b)}, \quad \text{(b)} \Leftrightarrow \text{(c)}, \quad \text{and} \quad \text{(b)} \Leftrightarrow \text{(d)}.$$

Assume (a). Then $a = \eta' = \varphi$. Hence (b) holds.

Assume (b). Then (a) follows from $\eta' = \varphi$ and $\eta(1) = b$.

Now assume (c): $\psi(t) = at + b$. Then $\varphi(t) = at + b - \eta(t)$. Differentiating, we get $\varphi' = a - \eta' = a - \varphi$. Also, $\varphi(1) = a$. Therefore $\varphi = a$. This proves (b).

Assume (b) again. This implies (a). Therefore $\psi(t) = \eta(t) + \varphi(t) = at + (b-a) + a = at + b$. This proves (c).

Now assume (d). The equation in Definition 4.1(iii) implies that f' is continuous on $(1, +\infty)$ and the one-sided limit at 1 exists. Therefore for $t > 1$ we calculate:

$$\begin{aligned} f(t) &= f(1) + \int_1^t f'(s)\,\mathrm{d}s \\ &= f(1) + \int_1^t f(s)\,\mathrm{d}s - \int_1^t f(s-1)\,\mathrm{d}s \\ &= f(1) + \int_1^t f(s)\,\mathrm{d}s - \int_0^{t-1} f(s)\,\mathrm{d}s \\ &= f(1) - \int_0^1 f(s)\,\mathrm{d}s + \int_{t-1}^t f(s)\,\mathrm{d}s \\ &= a + \int_0^1 f(t-u)\,\mathrm{d}u. \end{aligned}$$

This proves (b).

To finish the proof assume (b) again. A change of variable in (b) leads to

$$f(t) = a + \int_{t-1}^t f(s)\,\mathrm{d}s \quad \text{for } t \geqslant 1.$$

Since f is continuous, it follows that f is differentiable on $(1, +\infty)$. The differentiation yields (iii) in Definition 4.1. Thus f is a delay function. $\square$

Definition 4.4. By $\mathscr{F}(a, b)$ we denote the set of all delay functions f which satisfy the equivalent conditions of Theorem 4.3.

A straightforward calculation proves the following lemma.

Lemma 4.5. *Any linear function $\ell(t) = mt + k$ is in $\mathscr{F}(a, b)$, where $a = m/2$ and $b = k/2 + m/3$.*

The most importation example of a delay function in this article is the function α. We proved that α is continuous on $[0, +\infty)$ in Section 2. The proof that α is a delay function will be completed in Section 8 with the rigorous proof of (3.3). Then, using (2.7), one easily verifies that $\alpha \in \mathscr{F}(1, 1)$.

We close this section with a lemma which combines the three equations of Theorem 4.3 in a single statement about linear functions. For convenience, we change the interval of integration from $[0, 1]$ to $[t-1, t]$.

Lemma 4.6. *Let f be a delay function in $\mathscr{F}(a, b)$ and let ℓ be a linear function, with $\ell(t) = mt + k$. If $t \geqslant 1$, then*

$$\int_{t-1}^t \ell(s) f(s)\,\mathrm{d}s = \ell(t-1) f(t) + (mb - ka)\,.$$

Proof. Let $t \geqslant 1$ and $f \in \mathscr{F}(a, b)$. By Theorem 4.3,

$$\begin{aligned}
\int_{t-1}^{t} \ell(s) f(s)\,\mathrm{d}s &= \int_0^1 \ell(t-u) f(t-u)\,\mathrm{d}u \\
&= \int_0^1 (mt - mu + k)\, f(t-u)\,\mathrm{d}u \\
&= (mt+k)\int_0^1 f(t-u)\,\mathrm{d}u - m\int_0^1 u f(t-u)\,\mathrm{d}u \\
&= (mt+k)\,(f(t)-a) - m\,(f(t) - (at+b)) \\
&= (m(t-1)+k)\, f(t) + (mb - ka) \\
&= \ell(t-1) f(t) + (mb-ka)\,. \qquad \square
\end{aligned}$$

5. Existence and uniqueness of delay functions

Lemma 5.1. *Let f be a delay function and let $c \geqslant 1$. If $f = 0$ on $[c-1, c]$, then $f = 0$ on $[0, +\infty)$.*

Proof. Assume that $c \geqslant 1$ and $f = 0$ on $[c-1, c]$. By Definition 4.1(iii)

$$f' = f \quad \text{on } (c, c+1].$$

Since $f(c) = 0$, this implies that $f(t) = 0$ for all $c < t \leqslant c+1$. By induction, $f(t) = 0$ for all $t \geqslant c - 1$.

Let $d \geqslant c$ be an integer. Since $f(t) = f'(t) = 0$ for $t > d - 1$, (iii) in Definition 4.1 implies

$$0 = 0 - f(t-1) \quad \text{for } d - 1 < t < d.$$

Thus, $f(t) = 0$ for $d-2 < t < d-1$. By induction we conclude that $f(t) = 0$ for $0 < t < c-1$. Since f is continuous, this proves the lemma. $\square$

Theorem 5.2. *If ϕ is a continuous function on $\mathbb{I}$, then there exists a unique delay function f such that $f = \phi$ on $\mathbb{I}$.*

Proof. Existence can be proved using an inductive argument starting with ϕ and proceeding as we did for α using (3.4).

For uniqueness, let f and g be delay functions such that $f(t) = g(t) = \phi(t)$ for all $t \in \mathbb{I}$. Then $h = f - g$ is a delay function such that $h(t) = 0$ for all $t \in \mathbb{I}$. By Lemma 5.1, $h(t) = 0$ for all $t \geqslant 0$. Thus $f = g$. $\square$

6. Asymptotic properties of delay functions

Lemma 6.1. *Let w be a nonnegative continuous function on $\mathbb{I}$ such that $\int_0^1 w(u)\,\mathrm{d}u = 1$. Let g be a continuous function on $[0, +\infty)$ such that*

$$g(t) = \int_0^1 w(u) g(t-u)\,\mathrm{d}u \quad \textit{for } t \geqslant 1. \tag{6.1}$$

Let

$$m = \min_{s\in\mathbb{I}} g(s) \quad \textit{and} \quad M = \max_{s\in\mathbb{I}} g(s).$$

Then $m \leqslant g(t) \leqslant M$ for $t \geqslant 0$.

Proof. We prove $g(t) \leqslant M$ for $t \geqslant 0$ by contradiction. Suppose $g(t_1) > M$ for some $t_1 > 1$. Set

$$t_0 := \min\{t \in [0, +\infty) : g(t) = g(t_1)\}.$$

Clearly $t_0 > 1$ and

$$g(t) < g(t_0) \quad \text{for } 0 \leqslant t < t_0.$$

By (6.1) we have

$$g(t_0) = \int_0^1 w(u) g(t_0 - u)\,\mathrm{d}u < \int_0^1 w(u) g(t_0)\,\mathrm{d}u = g(t_0) \int_0^1 w(u)\,\mathrm{d}u = g(t_0),$$

which is a contradiction. The proof of $g(t) \geqslant m$ for $t \geqslant 0$ is analogous. □

Lemma 6.2. *Suppose that $f \in \mathscr{F}(0,0)$. Set*

$$m = \min\left\{e^s f(s) : s \in \mathbb{I}\right\} \quad \textit{and} \quad M = \max\left\{e^s f(s) : s \in \mathbb{I}\right\}.$$

Then

$$m e^{-t} \leqslant f(t) \leqslant M e^{-t} \quad \textit{for } t \geqslant 0.$$

Proof. By Theorem 4.3(c) for $t \geqslant 1$ we have

$$f(t) = \int_0^1 u f(t-u)\,\mathrm{d}u.$$

Define $g(t) = e^t f(t),\ t \geqslant 0$. Then

$$g(t) = e^t \int_0^1 u e^{-(t-u)} g(t-u)\,\mathrm{d}u = \int_0^1 u e^u g(t-u)\,\mathrm{d}u \quad \text{for } t \geqslant 1.$$

Since $\int_0^1 u e^u\,\mathrm{d}u = 1$, Lemma 6.1 implies that

$$m \leqslant g(t) \leqslant M \quad \text{for } t \geqslant 0.$$

This proves the lemma. □

Theorem 6.3. *If $f \in \mathscr{F}(a,b)$, then*

$$\lim_{t\to+\infty} (f(t) - 2at) = 2b - \frac{4}{3}a.$$

Proof. Set $g(t) = 2at - 2b + 4a/3$. By Lemma 4.5, $g \in \mathscr{F}(a,b)$. Set $h = f - g$. The linearity property of integration yields that $h \in \mathscr{F}(0,0)$. The theorem now follows since $h(t) \to 0$ as $t \to +\infty$ by Lemma 6.2. □

Proposition 6.4. *Let $f \in \mathscr{F}(a,b)$. Then f is bounded if and only if $a = 0$. A bounded delay function reaches its minimum and its maximum on $\mathbb{I}$.*

Proof. The first claim is an immediate consequence of Theorem 6.3. The second claim follows from Lemma 6.1 since, by Theorem 4.3(b),

$$f(t) = \int_0^1 f(t-u)\,\mathrm{d}u \quad \text{for } t \geqslant 1. \qquad \square$$

7. Oscillations of delay functions

Definition 7.1. Let f be a real function defined on an open interval J. We say that f *changes sign at least once* in J if it takes on both positive and negative values. We say that f *changes sign at least twice* in J if there exist $t_1, t_2, t_3 \in J$ such that

$$t_1 < t_2 < t_3$$

and the numbers

$$f(t_1), \quad f(t_2), \quad f(t_3)$$

are nonzero and alternate in sign.

Lemma 7.2. *Let f be a nonzero function in $\mathscr{F}(0,0)$. For each $c \geqslant 1$, f changes sign at least twice in $(c-1,c)$.*

Proof. Let $c \geqslant 1$. Set

$$A = \{x \in [c-1,c] : f(x) < 0\} \quad \text{and} \quad B = \{x \in [c-1,c] : f(x) > 0\}.$$

If ℓ is linear, then by Lemma 4.6,

$$\int_{c-1}^{c} \ell(s) f(s)\,\mathrm{d}s = \ell(c-1) f(c).$$

First, let $\ell(t) = t - (c-1)$. Then $\ell > 0$ on $(c-1,c)$, and

$$\int_{c-1}^{c} \ell(s) f(s)\,\mathrm{d}s = 0.$$

Hence, f changes sign at least once in $(c-1, c)$. That is, A and B are nonempty (and clearly disjoint).

Second, assume that f does not change sign at least twice in $(c-1, c)$. Then there exists $d \in (c-1, c)$ such that d separates A and B. We assume that A is to the left of d. If not, consider $-f$. Let $\ell(t) = t - d$. Then $\ell f \geqslant 0$ on $[c-1, c]$ and

$$\int_{c-1}^{c} \ell(s) f(s)\,\mathrm{d}s = \ell(c-1) f(c) = (c-1-d) f(c) \leqslant 0.$$

Thus $\ell f = 0$ on $[c-1, c]$ and that is a contradiction. □

Theorem 7.3. *Let f be a nonlinear function in $\mathscr{F}(a, b)$. For each $c \geqslant 1$, the nonzero function*

$$t \mapsto f(t) - \left(2at + 2b - \frac{4}{3}a\right) \tag{7.1}$$

changes sign at least twice in $(c-1, c)$.

Proof. The linearity of integration and Lemma 4.5 imply that the function in (7.1) is a nonlinear function in $\mathscr{F}(0, 0)$. Now the theorem follows from Lemma 7.2. □

8. The function α

The function α is defined in the Introduction. For $t \geqslant 0$, we considered a process of selecting numbers randomly from the interval $\mathbb{I}$ until the sum is greater than t and defined $\alpha(t)$ to be the expected number of selections. Instead, in the rest of this article we assume that the selection process goes on forever. The associated sequence of partial sums is nondecreasing. The mathematical framework for this is provided by the probability space $\mathbb{I}^{\mathbb{N}}$, which consists of all sequences in $\mathbb{I}$. This is the *Hilbert cube*.

The following definition is analogous to (2.3). For $n \in \mathbb{N}$ and $t \geqslant 0$ we put

$$B_{n,t} = \left\{ x \in \mathbb{I}^{\mathbb{N}} : x_1 + \cdots + x_n \leqslant t \right\} = C_{n,t} \times \mathbb{I}^{\mathbb{N}}$$

and

$$B_{0,t} = \mathbb{I}^{\mathbb{N}}.$$

Clearly $B_{n,t}$ is closed and its measure in $\mathbb{I}^{\mathbb{N}}$ is equal to $\mathrm{Vol}(C_{n,t})$. Let $B \subset \mathbb{I}^{\mathbb{N}}$ be the set of all sequences in $\mathbb{I}$ for which the corresponding series converges. That is,

$$B = \bigcup_{k=1}^{+\infty} \bigcap_{n=1}^{+\infty} B_{n,k}.$$

By (2.5) and (2.6), the measure of B in $\mathbb{I}^{\mathbb{N}}$ is 0.

Theorem 8.1. *The function α is a delay function in $\mathscr{F}(1, 1)$ and*

$$\lim_{t \to +\infty} (\alpha(t) - 2t) = \frac{2}{3}. \tag{8.1}$$

Proof. Let $t \geqslant 0$. Define the random variable $F_t : \mathbb{I}^{\mathbb{N}} \to \mathbb{N} \cup \{+\infty\}$ by

$$F_t(x) := \min\{n \in \mathbb{N} : x_1 + \cdots + x_n > t\}.$$

Here we adopt the convention that $\min \emptyset = +\infty$. Since B has measure 0 and

$$F_t^{-1}(\{+\infty\}) = \bigcap_{n=1}^{+\infty} B_{n,t} \subset B,$$

F_t is finite almost everywhere on $\mathbb{I}^{\mathbb{N}}$. For $n \in \mathbb{N}$ we have

$$F_t^{-1}(\{n\}) = B_{n-1,t} \backslash B_{n,t}.$$

Therefore, $F_t : \mathbb{I}^{\mathbb{N}} \to \mathbb{N} \cup \{+\infty\}$ is a Borel function. By definition, α is the expected value of F_t:

$$\alpha(t) = \int_{\mathbb{I}^{\mathbb{N}}} F_t(x)\, \mathrm{d}x.$$

The following property of F_t, which is the key to the proof that α is a delay function, is an immediate consequence of its definition. If $t \geqslant 1$ and

$$x = (x_1, x_2, x_3, \ldots) = (u, v_1, v_2, \ldots) = (u; v) \in \mathbb{I}^{\mathbb{N}}, \tag{8.2}$$

then $2 \leqslant F_t(x) \leqslant +\infty$ and

$$F_t(x) = F_t(u; v) = 1 + F_{t-u}(v). \tag{8.3}$$

Using Fubini's Theorem [10, Theorem 4.19] first, and than (8.3), we get

$$\begin{aligned}
\alpha(t) &= \int_{\mathbb{I}^{\mathbb{N}}} F_t(x)\, \mathrm{d}x \\
&= \int_0^1 \left(\int_{\mathbb{I}^{\mathbb{N}}} F_t(u; v)\, \mathrm{d}v \right) \mathrm{d}u \\
&= \int_0^1 \left(1 + \int_{\mathbb{I}^{\mathbb{N}}} F_{t-u}(v)\, \mathrm{d}v \right) \mathrm{d}u \\
&= 1 + \int_0^1 \alpha(t-u)\, \mathrm{d}u.
\end{aligned}$$

Hence, (3.1) is proved. Since in Section 2 we proved that α is continuous, by Theorem 4.3, α is a delay function. Moreover, $\alpha \in \mathscr{F}(1,1)$, as is easily verified using the fact that $\alpha(t) = e^t$ on $\mathbb{I}$, which is proved in Section 2. Finally, equality (8.1) is a consequence of Theorem 6.3. □

The asymptotic behavior of α suggested in Table 1 and Fig. 1 is now proved. To explain the oscillations, let $g(t) = 2t + \frac{2}{3}$ and recall Theorem 7.3. Then for $c \geqslant 1$ the function $f - g$ changes sign at least twice in the interval $(c-1, c)$, as suggested by Table 1 and Figs. 1 and 2.

9. The function β

The function β is defined in the Introduction. Here, $\beta(t)$ is the expected value of the first selection for which the partial sum exceeds t.

Theorem 9.1. *The function β is a delay function in $\mathscr{F}\left(0, \frac{1}{3}\right)$. Moreover,*

$$\beta(t) = t + 1 - \frac{\alpha(t)}{2} \quad \textit{for } t \geqslant 0 \tag{9.1}$$

and

$$\lim_{t \to +\infty} \beta(t) = \frac{2}{3}. \tag{9.2}$$

Proof. Let $t \geqslant 0$. Define the random variable $G_t : \mathbb{I}^{\mathbb{N}} \backslash B \to \mathbb{I}$ by

$$G_t(x) := x_{F_t(x)}, \quad x \in \mathbb{I}^{\mathbb{N}} \backslash B.$$

Since B is of measure 0, G_t is defined almost everywhere on $\mathbb{I}^{\mathbb{N}}$. In other words, $G_t(x)$ is the first selected number for which the corresponding partial sum exceeds t. Equivalently,

$$G_t(x) = x_n \quad \text{if and only if} \quad x \in B_{n-1,t} \backslash \left(B_{n,t} \cup B\right).$$

Hence, the restriction of G_t to each Borel set $B_{n-1,t} \backslash \left(B_{n,t} \cup B\right)$, $n \in \mathbb{N}$, is a continuous function. Since the union of these Borel sets coincides with the domain of G_t, it follows that G_t is a Borel function. By definition, $\beta(t)$ is the expected value of G_t:

$$\beta(t) = \int_{\mathbb{I}^{\mathbb{N}}} G_t(x)\,\mathrm{d}x.$$

Let $x \in \mathbb{I}^{\mathbb{N}} \backslash B$. Then $F_t(x) \in \mathbb{N}$ and $G_t(x) \in \mathbb{I}$. Let, in addition, $0 \leqslant s < t$ and compare $|G_t(x) - G_s(x)|$ with $F_t(x) - F_s(x)$. Clearly, $|G_t(x) - G_s(x)| \leqslant 1$ and $F_t(x) - F_s(x) \in \mathbb{N} \cup \{0\}$. Since by definition $F_t(x) - F_s(x) = 0$ implies $G_t(x) = G_s(x)$, we conclude that

$$|G_t(x) - G_s(x)| \leqslant F_t(x) - F_s(x).$$

Therefore,

$$\begin{aligned}
|\beta(t) - \beta(s)| &= \left| \int_{\mathbb{I}^{\mathbb{N}}} G_t(x)\,\mathrm{d}x - \int_{\mathbb{I}^{\mathbb{N}}} G_s(x)\,\mathrm{d}x \right| \\
&\leqslant \int_{\mathbb{I}^{\mathbb{N}}} |G_t(x) - G_s(x)|\ \mathrm{d}x \\
&\leqslant \int_{\mathbb{I}^{\mathbb{N}}} (F_t(x) - F_s(x))\ \mathrm{d}x \\
&= \alpha(t) - \alpha(s).
\end{aligned}$$

Since α is continuous, the last inequality implies that β is continuous as well.

The following property of G_t, which is the key to the rest of the proof that β is a delay function, is a consequence of the definitions of F_t and G_t. For $t \geqslant 0$ and x given by (8.2) we have

$$G_t(x) = G_t(u; v) = \begin{cases} u & \text{for } 0 \leqslant t < u, \\ G_{t-u}(v) & \text{for } u \leqslant t. \end{cases} \tag{9.3}$$

We first consider β on $\mathbb{I}$. For $0 < t < 1$, Fubini's theorem and (9.3) imply that

$$\begin{aligned} \beta(t) &= \int_0^1 \left(\int_{\mathbb{I}^{\mathbb{N}}} G_t(u; v)\, \mathrm{d}v \right) \mathrm{d}u \\ &= \int_0^t \left(\int_{\mathbb{I}^{\mathbb{N}}} G_{t-u}(v)\, \mathrm{d}v \right) \mathrm{d}u + \int_t^1 u\, \mathrm{d}u \\ &= \int_0^t \beta(t-u)\, \mathrm{d}u + \frac{1}{2} - \frac{t^2}{2} \\ &= \int_0^t \beta(s)\, \mathrm{d}s + \frac{1}{2} - \frac{t^2}{2}. \end{aligned}$$

Since β is continuous, differentiating the last equality leads to

$$\beta'(t) = \beta(t) - t \quad \text{for } 0 < t < 1.$$

In addition, by definition of β we have $\beta(0) = \frac{1}{2}$. Therefore

$$\beta(t) = t + 1 - e^t/2 \quad \text{for } t \in \mathbb{I}.$$

A straightforward calculation yields

$$\beta(1) - \int_0^1 \beta(s)\, \mathrm{d}s = 0 \quad \text{and} \quad \int_0^1 s\beta(s)\, \mathrm{d}s = \frac{1}{3}. \tag{9.4}$$

For $t > 1$, applying again Fubini's theorem and (9.3) we find

$$\beta(t) = \int_0^1 \left(\int_{\mathbb{I}^{\mathbb{N}}} G_{t-u}(v)\, \mathrm{d}v \right) \mathrm{d}u = \int_0^1 \beta(t-u)\, \mathrm{d}u.$$

Thus, by Theorem 4.3, β is a delay function and $\beta \in \mathscr{F}\left(0, \frac{1}{3}\right)$ by (9.4).

As a linear combination of delay functions, the function

$$g(t) = t + 1 - \alpha(t)/2, \quad t \geqslant 0,$$

is a delay function by Proposition 4.2 and Lemma 4.5. Since β and g coincide on $\mathbb{I}$, Theorem 5.2 implies (9.1). Equality (9.2) is a consequence of Theorem 8.1. □

Theorem 7.3 implies that for each $c \geqslant 1$ the function $\beta - \frac{2}{3}$ changes sign at least twice in the interval $(c-1, c)$.

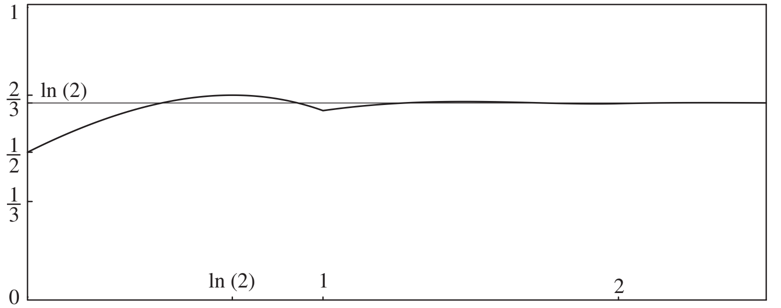

Fig. 3. The function β.

By Proposition 6.4 the function β reaches its maximum and its minimum on ▯; see Fig. 3. It is easy to check that

$$\min \beta = \beta(0) = \tfrac{1}{2},$$
$$\max \beta = \beta(\ln 2) = \ln 2.$$

10. Three other delay functions

1. For $t \geqslant 0$, let $\gamma(t)$ be the expected value of the first partial sum that exceeds t. The method used in Sections 8 and 9 could be used here, but we shall give a different treatment.

Recall that $p_{n,t}$ is the probability that the nth partial sum does not exceed t. Hence, $p_{n,t}$ is the contribution to $\alpha(t)$ of the nth selection. Therefore,

$$\alpha(t) = \sum_{n=0}^{+\infty} p_{n,t}.$$

This formula was derived in Section 2 in a different way. Since the expected value of the nth selection is $\frac{1}{2}$,

$$\gamma(t) = \sum_{n=0}^{+\infty} \frac{1}{2} \cdot p_{n,t} = \frac{1}{2}\,\alpha(t).$$

Thus, γ is a delay function, $\gamma \in \mathscr{F}\left(\frac{1}{2}, \frac{1}{2}\right)$ and

$$\lim_{t \to +\infty} (\gamma(t) - t) = \frac{1}{3}.$$

2. For $t \geqslant 0$, let $\delta(t)$ be the expected value of the last partial sum that does not exceed t. We set $\delta(0) = 0$. Clearly

$$\delta + \beta = \gamma.$$

Thus δ is a delay function and

$$\delta(t) = \alpha(t) - (t+1), \quad \delta \in \mathscr{F}\left(\tfrac{1}{2}, \tfrac{1}{6}\right), \quad \lim_{t\to+\infty} (\delta(t) - t) = -\frac{1}{3}.$$

3. For $t \geqslant 0$, let $\varepsilon(t)$ be the expected number of partial sums in the interval $(t, t+1]$. In particular $\varepsilon(0) = e - 1$. Clearly,

$$\varepsilon(t) = \alpha(t+1) - \alpha(t).$$

By Proposition 4.2, ε is a delay function. If $t \in \mathbb{I}$, then by (2.7) and (3.5) we have

$$\varepsilon(t) = e^{t+1} + e^t(-t) - e^t = (e - 1 - t)e^t.$$

A simple computation shows that

$$\varepsilon \in \mathscr{F}(0, 1) \quad \text{and} \quad \lim_{t\to+\infty} \varepsilon(t) = 2.$$

Notice that $\varepsilon(t) = \alpha'(t+1)$ if $t > 0$.

11. A heuristic argument

Here it is convenient to write $f(t) \approx g(t)$ to mean that

$$\lim_{t\to+\infty} (f(t) - g(t)) = 0.$$

After proving that

$$\alpha(t) \approx 2t + \tfrac{2}{3},$$

we looked for a heuristic argument for that result and this is what we found.

The expected value of any selection is $\frac{1}{2}$. It is plausible that

$$\varepsilon(t) \approx 2.$$

The following statement is well known and easy to prove: *For the random selection of two numbers from* $\mathbb{I}$, *the expected value of the minimum is* $\frac{1}{3}$ *and the expected value of the maximum is* $\frac{2}{3}$.

This statement suggests, but does not imply, that

$$\gamma(t) \approx t + \tfrac{1}{3} \quad \text{and} \quad \delta(t) \approx (t-1) + \tfrac{2}{3} = t - \tfrac{1}{3}.$$

Since $\beta = \gamma - \delta$, we have

$$\beta(t) \approx \tfrac{2}{3}.$$

It is also plausible that $\alpha(t) \approx 2\gamma(t)$. If so,

$$\alpha(t) \approx 2t + \tfrac{2}{3}.$$

References

[1] Bush LE. The William Putnam mathematical competition. Amer Math Monthly 1961;68:18–33.
[2] Shultz HS. An expected value problem. Two-Year College Math J 1979;10:179.
[3] MacKinnon N. Another surprising appearence of e. Math Gazete 1990;74:167–9.
[4] Schwartzman S. An unexpected expected value. Math Teacher 1993; 118–20.
[5] Weisstein EW. Uniform sum distribution. From MathWorld – A Wolfram Web Resource. ⟨http://mathworld.wolfram.com/UniformSumDistribution.html⟩.
[6] Derbyshire J. Prime obsession: Bernhard Riemann and the greatest unsolved problem in mathematics. Joseph Henry Press; 2003.
[7] Peköz EA. Probability, paradoxes, and games. Talk given at Department of Mathematics, Western Washington University, March 10, 1998.
[8] Uspensky JV. Introduction to Mathematical Probability. New York: McGraw-Hill; 1937.
[9] Diekmann O, van Gils SA, Verduyn Lunel SM, Walther H. Delay equations: functional-, complex-, and nonlinear analysis. Applied mathematical sciences, vol. 110. Berlin: Springer; 1995.
[10] Royden HL. Real analysis, 3rd ed., New York: MacMillan; 1988.