BETTER-THAN-AVERAGE UNIFORM RANDOM VARIABLES AND EULERIAN NUMBERS, OR: HOW MANY CANDIDATES SHOULD A VOTER APPROVE?

SVANTE JANSON AND WARREN D. SMITH

ABSTRACT. Consider n independent random numbers with a uniform distribution on $[0,1]$. The number of them that exceed their mean is shown to have an Eulerian distribution, i.e., it is described by the Eulerian numbers. This is related to, but distinct from, the well known fact that the integer part of the sum of independent random numbers uniform on $[0,1]$ has an Eulerian distribution. One motivation for this problem comes from voting theory.

1. INTRODUCTION

Consider $n \geqslant 2$ independent random numbers $U_1, \ldots, U_n$, each with a uniform distribution on $[0,1]$. Let $S_n := \sum_{i=1}^{n} U_i$ be their sum, and $\overline{U} := S_n/n$ their mean. Finally, let

$$W_n := \#\{i \leqslant n : U_i > \overline{U}\}, \tag{1.1}$$

i.e, the number of the numbers U_i that exceed their mean. Since we assume $n \geqslant 2$, a.s. (almost surely, i.e, with probability 1), the numbers $U_1, \ldots, U_n$ are distinct; thus $\overline{U}$ is strictly between the smallest and the largest U_i, and consequently $1 \leqslant W_n \leqslant n-1$. What is the distribution of the random variable W_n?

Our original motivation for this question comes from voting theory, and is briefly described in Section 4.

Our main purpose is to prove the following theorem, which shows that the distribution of W_n is described by the *Eulerian numbers* $\left\langle {m \atop k} \right\rangle$ (see Section 2 for definitions and references); we thus may say that W_n has an *Eulerian distribution*. The proof is in Section 3.

Theorem 1.1. *Define W_n by (1.1), for some $n \geqslant 2$. Then, for every $k \geqslant 0$,*

$$\mathbb{P}(W_n = k) = \frac{1}{(n-1)!} \left\langle {n-1 \atop k-1} \right\rangle. \tag{1.2}$$

The Eulerian number in (1.2) is non-zero if and only if $1 \leqslant k \leqslant n-1$, matching the observation above that $1 \leqslant W_n \leqslant n-1$ a.s.

Remark 1.2. Since we assume $n \geqslant 2$, we have a.s. $U_i \neq \overline{U}$ for every i; thus it does not make any difference if we replace ">" in the definition (1.1) by "$\geqslant$". Furthermore, by symmetry, we obtain the same distribution if we replace ">" by "<" or "$\leqslant$". Also, of course instead of "uniform on $[0,1]$" for (1.1) we could have made the summands be uniform on any particular real interval $[a,b]$ with $a < b$. □

Theorem 1.1 seems to be new, but we recall the following well known result:

Theorem 1.3. *The integer part, i.e. floor, of the sum $S_n = \sum_{i=1}^n U_i$ defined above has the Eulerian distribution*

$$\mathbb{P}(\lfloor S_n \rfloor = k) = \frac{1}{n!} \left\langle {n \atop k} \right\rangle, \qquad k \geqslant 0. \tag{1.3}$$

Note that $0 \leqslant \lfloor S_n \rfloor \leqslant n-1$ a.s. The formula (1.3) can also be interpreted as a formula for the volume of the slices $\{x \in [0,1]^n : k \leqslant x_1 + \cdots + x_n < k+1\}$ of the unit cube.

Remark 1.4. The distribution of the sum S_n, i.e, the probability $\mathbb{P}(S_n \leqslant x)$ for all real x, was found already by Laplace [12, pp. 257–260], see also e.g. Feller [6, XI.7.20] and [7, Theorem I.9.1]. The result (1.3), which connects this distribution at integer points x and the Eulerian numbers, was found much later; see e.g. [13] (implicitly), [18], [8, §2], [16], or [15, Chapter 7]. See also [11] for related results for non-integer x. □

Using Theorem 1.3, we see that Theorem 1.1 is equivalent to the identity in distribution

$$W_n - 1 \stackrel{\mathrm{d}}{=} \lfloor S_{n-1} \rfloor, \qquad n \geqslant 2. \tag{1.4}$$

(Note the shift from n to $n-1$.) In fact, our proof of Theorem 1.1 in Section 3 will prove (1.4) directly, then conclude (1.2) from Theorem 1.3.

2. Preliminaries

The Eulerian numbers $\left\langle {n \atop k} \right\rangle$ were introduced by Euler [3; 4; 5], in a method to sum infinite series, see also [10]. They later were found to have many combinatorial interpretations. One common definition is to define $\left\langle {n \atop k} \right\rangle$ as the number of permutations of length n that have exactly k *descents*, see e.g. [2, pp. 267–269], [17, pp. 21–23], [19, §26.14], [15, §1.3]. They can also be defined by the generating function

$$\sum_{n,k=0}^{\infty} \left\langle {n \atop k} \right\rangle x^k \frac{t^n}{n!} = \frac{1-x}{e^{(x-1)t} - x}. \tag{2.1}$$

or the recursion

$$\left\langle {n \atop k} \right\rangle = (k+1) \left\langle {n-1 \atop k} \right\rangle + (n-k) \left\langle {n-1 \atop k-1} \right\rangle, \qquad n \geqslant 1, \ k \in \mathbb{Z}, \tag{2.2}$$

with the initial values $\left\langle {0 \atop k} \right\rangle = \mathbf{1}_{\{k=0\}}$. We use "$\mathbf{1}_{\{c\}}$ to mean "1 if c is true, otherwise 0." For these and other properties of the Eulerian numbers, see e.g. the monograph [15] and the further references there.

Remark 2.1. Other notations are also used, in particular $A_{n,k}$. Moreover, some authors shift the index k above by 1, so that $A_{n,k} = \left\langle {n \atop k-1} \right\rangle$. □

3. Proofs

We first consider a more general situation.

Let $U_1, \ldots, U_n$ be any n real numbers (random or not). As above, let $S_n := \sum_{i=1}^{n} U_i$ be their sum, $\overline{U} := S_n/n$ their mean, and now define

$$\widehat{W}_n := \#\{i \leqslant n : U_i < \overline{U}\}. \tag{3.1}$$

Furthermore, let $U_{(1)} \leqslant \ldots \leqslant U_{(n)}$ be $U_1, \ldots, U_n$ sorted into increasing order, then let

$$\Delta_i := U_{(i+1)} - U_{(i)}, \qquad 1 \leqslant i \leqslant n-1, \tag{3.2}$$

be the gaps between them.

Note that

$$U_{(i)} = U_{(1)} + \sum_{j=1}^{i-1} \Delta_j, \qquad 1 \leqslant i \leqslant n, \tag{3.3}$$

and thus

$$S_n = \sum_{i=1}^{n} U_{(i)} = nU_{(1)} + \sum_{i=1}^{n}\sum_{j=1}^{i-1} \Delta_j = nU_{(1)} + \sum_{j=1}^{n-1} (n-j)\Delta_j. \tag{3.4}$$

Remark 3.1. We use in this section $\widehat{W}_n$ rather than W_n in (1.1) for notational convenience below; as said in Remark 1.2, this makes no difference for our original problem. In general, we can interchange W_n and $\widehat{W}_n$ by changing the sign on each U_i, which reverses both the order and signs of $U_{(1)}, \ldots, U_{(n)}$, and thus changes Δ_i to Δ_{n-i}, so the increments are the same but in opposite order. $\square$

Lemma 3.2. *For any real numbers $U_1, \ldots, U_n$, we have, for $k = 0, \ldots, n-1$,*

$$\widehat{W}_n \leqslant k \iff \sum_{j=1}^{k} j\Delta_j - \sum_{j=k+1}^{n-1} (n-j)\Delta_j \geqslant 0. \tag{3.5}$$

Proof. By (3.1) and (3.3)–(3.4),

$$\widehat{W}_n \leqslant k \iff U_{(k+1)} \geqslant \overline{U} \iff nU_{(k+1)} \geqslant S_n \iff n\sum_{j=1}^{k} \Delta_j \geqslant \sum_{j=1}^{n-1} (n-j)\Delta_j$$

$$\iff \sum_{j=1}^{k} n\Delta_j - \sum_{j=1}^{n-1} (n-j)\Delta_j \geqslant 0. \tag{3.6}$$

Furthermore,

$$\sum_{j=1}^{k} n\Delta_j - \sum_{j=1}^{n-1} (n-j)\Delta_j = \sum_{j=1}^{k} j\Delta_j - \sum_{j=k+1}^{n-1} (n-j)\Delta_j, \tag{3.7}$$

and thus (3.5) follows. $\square$

Remark 3.3. In other words,

$$\widehat{W}_n = \min\left\{k \in \{0, \ldots, n-1\} : \sum_{j=1}^{k} j\Delta_j \geqslant \sum_{j=k+1}^{n-1} (n-j)\Delta_j\right\}. \tag{3.8}$$

$\square$

Proof of Theorem 1.1. We use the notation above. Now $U_1, \ldots, U_n$ are independent uniform variables in $[0, 1]$, and thus $U_{(1)}, \ldots, U_{(n)}$ are the corresponding order statistics. We define also $U_{(0)} = 0$ and $\Delta_0 := U_{(1)}$. (Thus (3.2) is now also for $i = 0$.)

The vector $(U_{(1)}, \ldots, U_{(n)})$ is uniformly distributed in the simplex $\{0 \leqslant u_1 \leqslant \ldots \leqslant u_n \leqslant 1\}$. It follows that $U_{(1)}/U_{(n)}, \ldots, U_{(n-1)}/U_{(n)}$ have the same joint distribution as $n-1$ independent uniform random $U(0, 1)$ variables, arranged in increasing order. In particular, using (3.3) and (3.4) (with $n - 1$ instead of n),

$$S_{n-1} \overset{\mathrm{d}}{=} \sum_{i=1}^{n-1} \frac{U_{(i)}}{U_{(n)}} = \frac{\sum_{i=1}^{n-1} U_{(i)}}{U_{(n)}} = \frac{\sum_{j=0}^{n-2}(n-1-j)\Delta_j}{\sum_{j=0}^{n-1}\Delta_j} \tag{3.9}$$

and thus, for $k = 0 \ldots, n - 1$,

$$\lceil S_{n-1}\rceil \leqslant k \iff S_{n-1} \leqslant k \iff \sum_{j=0}^{n-1}(n-1-j)\Delta_j \leqslant k\sum_{j=0}^{n-1}\Delta_j$$

$$\iff \sum_{j=0}^{n-1}(j+k+1-n)\Delta_j \geqslant 0.$$

$$\iff \sum_{i=1}^{k} i\Delta_{i+n-1-k} + \sum_{i=k+1}^{n-1}(i-n)\Delta_{i-1-k} \geqslant 0. \tag{3.10}$$

Comparing this and (3.5), we see that the conditions are the same up to a (cyclic) permutation of the variables $\Delta_0, \ldots, \Delta_{n-1}$. It is well known, and easy to see, that for i.i.d. uniform random variables $U_1, \ldots, U_n$, the increments $\Delta_0, \ldots, \Delta_{n-1}$ have an *exchangeable* distribution; thus we obtain from (3.5) and (3.10), for $k = 0, \ldots, n-1$,

$$\mathbb{P}\big(\widehat{W}_n \leqslant k\big) = \mathbb{P}\left(\sum_{j=1}^{k} j\Delta_j - \sum_{j=k+1}^{n-1}(n-j)\Delta_j \geqslant 0\right)$$

$$= \mathbb{P}\left(\sum_{i=1}^{k} i\Delta_{i+n-1-k} + \sum_{i=k+1}^{n-1}(i-n)\Delta_{i-1-k} \geqslant 0\right)$$

$$= \mathbb{P}\big(\lceil S_{n-1}\rceil \leqslant k\big). \tag{3.11}$$

Consequently, $\widehat{W}_n \overset{\mathrm{d}}{=} \lceil S_{n-1}\rceil$. Finally, $W_n \overset{\mathrm{d}}{=} \widehat{W}_n$ by Remark 1.2, and $\lceil S_{n-1}\rceil = \lfloor S_{n-1}\rfloor + 1$ almost surely. This proves the identity in distribution $W_n \overset{\mathrm{d}}{=} \lfloor S_{n-1}\rfloor + 1$, which as remarked in the introduction is equivalent to Theorem 1.1 by the well known Theorem 1.3. $\qquad\square$

We note an alternative proof which calculates the probability of the event in (3.5) directly, using properties of exponential random variables. While we regard this proof as more complicated than the proof above, it might be of independent interest.

Alternative proof of Theorem 1.1. As in the proof above, let $U_1, \ldots, U_n$ be i.i.d. uniform random variables in $[0, 1]$, and let the order statistics $U_{(i)}$ and increments Δ_i be as above.

Let $\xi_1, \xi_2, \ldots$ be i.i.d. random variables with a standard exponential density e^{-x} for $x > 0$. Let $\zeta_1, \zeta_2, \ldots$ be the partial sums $\zeta_k := \sum_{i=1}^{k} \xi_i$. (These are the points of a rate 1 Poisson point process on $[0, \infty)$.) It is well known that conditioned on ζ_{n+1}, the points $\zeta_1, \ldots, \zeta_n$ are distributed as n independent uniformly distributed points on

$[0, \zeta_{n+1}]$ that have been ordered, and thus $(\zeta_1/\zeta_{n+1}, \ldots, \zeta_n/\zeta_{n+1}) \stackrel{\mathrm{d}}{=} (U_{(1)}, \ldots, U_{(n)})$. Hence $(\Delta_0, \ldots, \Delta_{n-1}) \stackrel{\mathrm{d}}{=} (\xi_1/\zeta_{n+1}, \ldots, \xi_n/\zeta_{n+1})$, and it follows by Lemma 3.2 and Remark 1.2 that, for $k = 0, \ldots, n-1$,

$$\mathbb{P}(W_n \leqslant k) = \mathbb{P}\left(\sum_{j=1}^{k} j\Delta_j - \sum_{j=k+1}^{n-1} (n-j)\Delta_j \geqslant 0 \right)$$

$$= \mathbb{P}\left(\sum_{j=1}^{k} j\frac{\xi_{j+1}}{\zeta_{n+1}} - \sum_{j=k+1}^{n-1} (n-j)\frac{\xi_{n-j+1}}{\zeta_{n+1}} \geqslant 0 \right)$$

$$= \mathbb{P}\left(\sum_{j=1}^{k} j\xi_{j+1} - \sum_{j=k+1}^{n-1} (n-j)\xi_{n-j+1} \geqslant 0 \right). \qquad (3.12)$$

The last probability is calculated by Lemma A.1, with $\ell = n - k - 1$, $a_i = i$ and $b_j = j$ (after reordering), and thus, by (A.1),

$$\mathbb{P}(W_n \leqslant k) = \sum_{m=1}^{k} m^{n-2} \frac{(-1)^{k-m} m!}{(m-1)!\,(k-m)!\,(m+n-k-1)!}$$

$$= \sum_{m=1}^{k} m^{n-1} \frac{(-1)^{k-m}}{(k-m)!\,(n-1+m-k)!}$$

$$= \sum_{i=0}^{k-1} (k-i)^{n-1} \frac{(-1)^i}{i!\,(n-1-i)!}$$

$$= \frac{1}{(n-1)!} \sum_{i=0}^{k-1} (-1)^i \binom{n-1}{i} (k-i)^{n-1}. \qquad (3.13)$$

Hence, by a simple calculation,

$$\mathbb{P}(W_n = k) = \mathbb{P}(W_n \leqslant k) - \mathbb{P}(W_n \leqslant k-1)$$

$$= \frac{1}{(n-1)!} \sum_{i=0}^{k-1} (-1)^i \binom{n-1}{i} (k-i)^{n-1} - \frac{1}{(n-1)!} \sum_{i=0}^{k-2} (-1)^i \binom{n-1}{i} (k-1-i)^{n-1}$$

$$= \frac{1}{(n-1)!} \sum_{i=0}^{k-1} (-1)^i \binom{n}{i} (k-i)^{n-1} = \frac{1}{(n-1)!} \left\langle {n-1 \atop k-1} \right\rangle, \qquad (3.14)$$

where the final inequality is a standard formula for Eulerian numbers, see e.g. [2, (6.38)], [15, Corollary 1.3], or [19, 26.14.6].

In fact, the right-hand side of (3.13) equals the probability $\mathbb{P}(\lceil S_{n-1} \rceil \leqslant k) = \mathbb{P}(S_{n-1} \leqslant k)$ as calculated in e.g. [12] and [7], and (3.14) just repeats a calculation in e.g. [18] or [8]. (We may also use (3.10) and the method above to obtain a possibly new proof of this formula for $\mathbb{P}(\lceil S_{n-1} \rceil \leqslant k)$ and thus of Theorem 1.3.) $\qquad \square$

4. APPLICATION IN (OR MOTIVATION FROM) VOTING THEORY

One motivation for the study of the random variable W_n comes from voting theory.

Consider an election (of 1 person) by *Approval Voting* [1]. This means that each voter *approves* (i.e, votes for) an arbitrary set of candidates, and the candidate approved by the most voters wins. (Ties are broken randomly.)

We are interested in the best strategy of a single voter. We assume each voter has evaluated all n candidates and therefore assigned a personal real "utility" U_i to each candidate $i = 1, \ldots, n$. The voter does not know the preferences (or strategies) of the other voters, so from the perspective of this voter, the other votes can be regarded as random, with some distribution. We assume also that the voter tries to maximize the expected utility of the winner.

Let us further assume that the voter is completely ignorant of the choices of other voters, so the assumed probability distribution of the aggregated votes of all others is symmetric under the group of $n!$ permutations of the n candidates. Since ties are broken randomly we may simplify the analysis by assuming that this is done by giving each candidate a (secret) random number of extra votes $X_i \in (0, 1)$ before the election, with $X_1, \ldots, X_n$ i.i.d. and uniformly distributed. We include these extra votes in our probability model of the other voters, thus (a.s.) eliminating ties. Let V_i be the number of votes from all other voters for candidate i, including the X_i extra ones. (Thus V_i is a real number.) The assumed probability distribution of $V_1, \ldots, V_n$ is still symmetric. Hence there is a (presumably small) probability p such that for each pair (i, j), V_i is the largest number of votes, *but* $V_i - V_j < 1$ so that our voter could swing the outcome. Let us also assume that the probability that there is more than one such "swing pair" (i, j) is neglectibly small – which is reasonable if the number of voters is large. Assume our voter approves a subset A of the candidates. Then let I_i denote the indicator-function $\mathbf{1}_{\{i \in A\}}$ that candidate i is approved. Then the expected change that this vote causes to the utility of the winner is, using the notations S_n and $\overline{U}$ as in Section 3,

$$\sum_{i \neq j} p I_j (1 - I_i)(U_j - U_i) = p \sum_{i,j=1}^{n} I_j (1 - I_i)(U_j - U_i) = p \sum_{i,j=1}^{n} I_j (U_j - U_i)$$

$$= pn \sum_{j=1}^{n} I_j U_j - p \sum_{j=1}^{n} I_j S_n = pn \sum_{j=1}^{n} I_j (U_j - \overline{U}). \qquad (4.1)$$

Since the objective is to maximize the expected utility of the winner, or equivalently to maximize (4.1), the optimal strategy is to choose $I_j := \mathbf{1}_{\{U_j > \overline{U}\}}$; in other words:

The voter should approve each candidate with a utility exceeding the average utility (all utilities as reckoned by that voter).

Note that this argument does not depend on any assumptions (beyond its symmetry) about the shape of the probability distribution for the other voters, Hence this strategy is optimal if we are ignorant about the other voters; and we do not have to construct any specific probability distribution describing their voting.

Consider now a number of voters such that each of them follows this optimal strategy, where the voters make their own assignments of utilities. A simple model is that each voter's utilities are n i.i.d. random numbers uniform in $[0, 1]$. Then the number of candidates that the voter approves is given by W_n in (1.1), whereupon Theorem 1.1 shows that

The number of approved candidates has an Eulerian distribution.

A more realistic model than "total ignorance about the other voters" might be to assume that there is a certain "good" subset of the candidates who have a reasonable chance to win, and whom almost all voters regard as better than the complement

"bad" candidate-subset. If so, then we'd predict that the number of approved candidates has distribution approximately Eulerian among the good-subset only.

Also, much more generally: a voter *not* wholly ignorant of all the other voters, typically[1] should approve all candidates whose utility (to her) exceeds the expected utility (to her) of the winner. To determine the latter, the voter needs to estimate the win-probabilities for all the candidates; and the more knowledge she has about how the other voters are likely to behave, the better such estimates she can make. We have discussed above only the special case where the voter has zero knowledge about the other voters; but many other special cases also can be treated, including with the aid of Lemma A.1, e.g. see [22].

It is interesting to study the distribution of the random variable W_n also for other distributions of $(U_1, \ldots, U_n)$. We give here a quite different case where the distribution can be found explicitly.

Example 4.1. We may start with any positive random variables $\Delta_1, \ldots, \Delta_{n-1}$. We then define $U_{(1)}, \ldots, U_{(n)}$ by (3.3), with, say, $U_{(1)} = 0$, and finally let $U_1, \ldots, U_n$ be a random permutation of $U_{(1)}, \ldots, U_{(n)}$. This yields random utilities with the increments $\Delta_1, \ldots, \Delta_{n-1}$.

Consider the case when the increments Δ_i are i.i.d. random variables with a positive (strictly) stable distribution of index $\alpha \in (0, 1)$; thus their Laplace transform is $\mathbb{E}\, e^{-t\Delta_i} = e^{-ct^\alpha}$ for $t \geqslant 0$ and a constant $c > 0$ (which is unimportant by rescaling). We find the distribution of W_n by using Lemma 3.2, noting that $W_n \stackrel{\mathrm{d}}{=} \widehat{W}_n$ by the sign-change argument in Remark 3.1. In this case

$$\sum_{j=1}^{k} j\Delta_j \stackrel{\mathrm{d}}{=} s_k^{1/\alpha}\Delta_1, \tag{4.2}$$

where

$$s_k := \sum_{j=1}^{k} j^\alpha. \tag{4.3}$$

Similarly,

$$\sum_{j=k+1}^{n-1} (n-j)\Delta_j = \sum_{i=1}^{n-k-1} i\Delta_{n-i} \stackrel{\mathrm{d}}{=} s_{n-k-1}^{1/\alpha}\Delta_1. \tag{4.4}$$

Furthermore, the two sums in (3.5) are independent, and it follows that, for $0 \leqslant k \leqslant n-1$,

$$\mathbb{P}\big(W_n \leqslant k\big) = \mathbb{P}\big(s_k^{1/\alpha}\Delta_1 \geqslant s_{n-k-1}^{1/\alpha}\Delta_2\big) = \mathbb{P}\left(\frac{\Delta_2}{\Delta_1} \leqslant \left(\frac{s_k}{s_{n-k-1}}\right)^{1/\alpha}\right). \tag{4.5}$$

In the special case $\alpha = 1/2$ (when the distribution of Δ_i is known as the Lévy distribution), we can take $\Delta_i = Z_i^{-2}$, where $Z_1, Z_2, \ldots$ are i.i.d. standard normal

---

[1]Actually this strategy is not always optimal. A counterexample: Suppose there are three candidates A, B, C. My utilities are 10, 6, 0. The other $2m$ voters are very polarized, with exactly half approving A but not B, and half approving B but not A. But they all toss fair coins to decide whether to approve C. Then A, B, C's win-chances are (in the $m \to \infty$ limit) 1/4, 1/4, 1/2, so my expected utility is 4. Hence the recommended strategy approves A, B, yielding expected utility still 4. But here a better strategy is approving only A, yielding win probabilities 1/2, 0, 1/2 so that my expected utility is 5. More of this sort of thing: [20],[21].

random variables. We thus obtain

$$\mathbb{P}(W_n \leqslant k) = \mathbb{P}\left(\left(\frac{Z_2}{Z_1}\right)^{-2} \leqslant \left(\frac{s_k}{s_{n-k-1}}\right)^2\right) = \mathbb{P}\left(\frac{|Z_1|}{|Z_2|} \leqslant \frac{s_k}{s_{n-k-1}}\right)$$

$$= \frac{2}{\pi}\arctan\frac{s_k}{s_{n-k-1}} = \frac{2}{\pi}\arctan\frac{\sum_{j=1}^{k}\sqrt{j}}{\sum_{j=1}^{n-k-1}\sqrt{j}}, \qquad (4.6)$$

since the argument of the complex number $|Z_2| + \mathrm{i}|Z_1|$ is uniformly distributed on $[0, \frac{\pi}{2}]$.

We may note that it follows from (4.6) that as $n \to \infty$, W_n/n converges in distribution to a random variable W^* with

$$\mathbb{P}(W^* \leqslant x) = \frac{2}{\pi}\arctan\left(\frac{x}{1-x}\right)^{3/2}, \qquad 0 \leqslant x \leqslant 1, \qquad (4.7)$$

and thus density function

$$f_{W*}(x) = \frac{3}{\pi} \cdot \frac{\sqrt{x(1-x)}}{x^3 + (1-x)^3}, \qquad 0 < x < 1. \qquad (4.8)$$

$\square$

Example 4.2. Another interesting case is when the numbers $U_1, \ldots, U_n$ are i.i.d. standard normal random variables. Let $Y_i := U_i - \overline{U}$; then $(Y_1, \ldots, Y_n)$ has a normal distribution with mean 0 and covariances $\mathrm{Var}(Y_i) = \frac{n-1}{n}$, $\mathrm{Cov}(Y_i, Y_j) = -\frac{1}{n}$, $i \neq j$, and thus correlations $\mathrm{Corr}(Y_i, Y_j) = -\frac{1}{n-1}$, $i \neq j$. Note that the vector $(Y_1, \ldots, Y_n)$ lies in the $(n-1)$-dimensional subspace $\mathcal{V} := \{(x_i)_1^n : \sum_1^n x_i = 0\}$ of $\mathbb{R}^n$.

Let again W_n be given by (1.1). By symmetry, for $1 \leqslant k \leqslant n-1$,

$$\mathbb{P}(W_n = k) = \binom{n}{k}\mathbb{P}(Y_1, \ldots, Y_k > 0 > Y_{k+1}, \ldots, Y_n). \qquad (4.9)$$

For each $i = 1, \ldots n$, the intersection of the hyperplane $x_i = 0$ with $\mathcal{V}$ defines a hyperplane H_i in $\mathcal{V}$. Similarly, we define a half-space A_i in $\mathcal{V}$ as $\mathcal{V}$ intersected with $\{x_i > 0\}$ if $1 \leqslant i \leqslant k$, and $\{x_i < 0\}$ if $k+1 \leqslant i \leqslant n$. The intersection $C_k := \bigcap_{i=1}^n A_i$ of these half-spaces is a polyhedral cone, bounded by (parts of) the hyperplanes H_i. Color these hyperplanes *red* for $1 \leqslant i \leqslant k$ and *blue* for $k+1 \leqslant i \leqslant n$. The interior angle between H_i and H_j ($i \neq j$) is $\mathrm{arcsec}(n-1)$ if i and j have the same color and $\mathrm{arcsec}(1-n)$ otherwise.

By intersecting the cone C_k and the $(n-2)$-dimensional unit sphere in $\mathcal{V}$, we obtain a spherical polyhedron T_k, and if $f(n, k)$ is its volume, then (4.9) yields

$$P(W_n = k) = \binom{n}{k}\frac{f(n,k)}{\Omega(n-1)} = \binom{n}{k}\frac{f(n,k)}{2\pi^{(n-1)/2}/\Gamma((n-1)/2)}. \qquad (4.10)$$

where $\Omega(n-1)$ is the surface area of the unit sphere in R^{n-1}.

For $n = 4$ and $k = 1$, T_k is a spherical triangle with all three angles $\mathrm{arcsec}(3)$, and thus area $f(4, 1) = 3\,\mathrm{arcsec}(3) - \pi$, yielding

$$\mathbb{P}(W_4 = 1) = 4\frac{3\,\mathrm{arcsec}(3) - \pi}{4\pi} = \frac{3\,\mathrm{arcsec}(3)}{\pi} - 1 = \frac{\arccos(23/27)}{\pi} \doteq 0.175479656$$

$$(4.11)$$

By symmetry, $\mathbb{P}(W_4 = 3) = \mathbb{P}(W_4 = 1) \doteq 0.175479656$, and thus $\mathbb{P}(W_4 = 2) \doteq 1 - 2\,\mathbb{P}(W_4 = 1) = 3 - 6\,\mathrm{arcsec}(3)/\pi \doteq 0.649040689$. These results agree excellently

with Monte Carlo experiments. This distribution, not surprisingly, differs from the Eulerian distribution found in Theorem 1.1 for *uniform* random variables.

For $n = 5$, we similarly can obtain $\mathbb{P}(W_5 = 1) = \mathbb{P}(W_5 = 4)$ from the volume of a spherical tetrahedron with all face-pair dihedral angles $\operatorname{arcsec}(4)$. Computing this volume using Murakami's formula [14], we find $0.1931420068473869869689612878 3\ldots$ yielding $\mathbb{P}(W_5 = 1) = \mathbb{P}(W_5 = 4) = 0.048923441862084390572624516 67\ldots$ agreeing with 0.048923 ± 0.000003 from Monte Carlo. (Warning: some implementations of Murakami's formula trigger bugs in Mathematica$^{\mathrm{TM}}$; and Murakami's formula has severe numerical problems for regular tetrahedra with dihedral angle $\pi - \epsilon$ for small enough $\epsilon > 0$.) Surprisingly, we noticed that this volume apparently equals $\operatorname{arccos}(1 - 3/4^3)\pi/5$, which Murakami [personal communication] then confirmed to hundreds of decimal places accuracy – but it remains unproven. If so, then $\mathbb{P}(W_5 = 1) = \mathbb{P}(W_5 = 4) = \operatorname{arccos}(61/64)/(2\pi)$. $\qquad\square$

Appendix A. A probability for exponential variables

We give here a calculation for linear combinations of exponential variables used in the alternative proof in Section 3. We give it in a general form. The assumption that $a_1, \ldots, a_\ell$ are distinct is necessary for the formula as stated here to make sense; however, the proof extends to cases with repeated a_i, but then the residue calculations will be more complicated.

Lemma A.1. *Let $\xi_1, \xi_2, \ldots$ and $\eta_1, \eta_2, \ldots$ be i.i.d. standard-exponential distributed random variables, i.e. probability density $exp(-x)$ Let $a_1, \ldots, a_k$ and $b_1, \ldots, b_\ell$ be positive real numbers, with $k, \ell \geqslant 0$, $k + \ell \geqslant 1$, and $a_1, \ldots, a_k$ distinct. Then*

$$\mathbb{P}\left(\sum_{i=1}^{k} a_i \xi_i \geqslant \sum_{j=1}^{\ell} b_j \eta_j\right) = \sum_{m=1}^{k} \frac{a_m^{k+\ell-1}}{\prod_{i \neq m}(a_m - a_i) \prod_{j=1}^{\ell}(a_m + b_j)}. \qquad (\text{A.1})$$

Proof. The cases $(k, \ell) = (0, 1)$ or $(1, 0)$ are trivial, so we may assume $k + \ell \geqslant 2$.

Let $X := \sum_{i=1}^{k} a_i \xi_i - \sum_{j=1}^{\ell} b_j \eta_j$. Then X has the moment generating function

$$\mathbb{E}\, e^{zX} = \prod_{i=1}^{k} \frac{1}{1 - a_i z} \prod_{j=1}^{\ell} \frac{1}{1 + b_j z} =: \Psi(z), \qquad (\text{A.2})$$

for all complex z with $|\operatorname{Re} z|$ sufficiently small. In particular, the characteristic function of X is $\varphi(t) = \Psi(\mathrm{i}t)$, $-\infty < t < \infty$.

We let (A.2) define $\Psi(z)$ in the entire complex plane $\mathbb{C}$, and note that $\Psi(z)$ is a meromorphic function with poles at $1/a_i$ and $-1/b_j$. We have the estimates, for some constants C and A,

$$|\Psi(z)| \leqslant C|\operatorname{Im} z|^{-k-\ell}, \qquad z \in \mathbb{C}, \qquad (\text{A.3})$$

$$|\Psi(z)| \leqslant C|z|^{-k-\ell}, \qquad |z| \geqslant A. \qquad (\text{A.4})$$

Since $k + \ell \geqslant 2$, (A.3) implies that the characteristic function is integrable, and X has a continuous density function given by

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\mathrm{i}xt} \varphi(t)\, \mathrm{d}t = \frac{1}{2\pi\mathrm{i}} \int_{-\infty\mathrm{i}}^{\infty\mathrm{i}} e^{-xz} \Psi(z)\, \mathrm{d}z, \qquad x \in \mathbb{R}, \qquad (\text{A.5})$$

where the last integral is along the imaginary axis.

Let $x > 0$. We move the line of integration to $\operatorname{Re} z = r$, using the estimate (A.3) and Cauchy's residue theorem; we then let $r \to \infty$, which makes the integral tend to 0 by (A.3)–(A.4), and obtain

$$f(x) = -\sum_{m=1}^{k} \operatorname{Res}_{z=1/a_m} \left(e^{-xz} \Psi(z) \right) = \sum_{m=1}^{k} e^{-x/a_m} \frac{1}{a_m} \prod_{i \neq m} \frac{1}{1 - a_i/a_m} \prod_{j=1}^{\ell} \frac{1}{1 + b_j/a_m}. \tag{A.6}$$

(This is where we use the assumption that $a_1, \ldots, a_m$ are distinct, which means that the poles $1/a_m$ are simple.) Integrating (A.6) yields

$$\mathbb{P}(X \geqslant 0) = \int_0^\infty f(x)\,\mathrm{d}x = \sum_{m=1}^{k} \prod_{i \neq m} \frac{1}{1 - a_i/a_m} \prod_{j=1}^{\ell} \frac{1}{1 + b_j/a_m}, \tag{A.7}$$

which is equivalent to (A.1). $\qquad\qquad\square$

References

[1] Steven Brams and Peter C. Fishburn: *Approval Voting*, 2nd ed, Springer, New York, 2007.

[2] Ronald L. Graham, Donald E. Knuth and Oren Patashnik. *Concrete Mathematics*. 2nd ed, Addison-Wesley, Reading, MA, 1994.

[3] Leonhard Euler. Methodus universalis series summandi ulterius promota. *Commentarii academiae scientiarum imperialis Petropolitanae* **8** (1736), St. Petersburg, 1741, pp. 147–158. `http://www.math.dartmouth.edu/~euler/pages/E055.html`

[4] Leonhard Euler. *Institutiones calculi differentialis cum eius usu in analysi finitorum ac doctrina serierum. Vol I*. St. Petersburg, 1755. `http://www.math.dartmouth.edu/~euler/pages/E212.html`

[5] L. Euler. Remarques sur un beau rapport entre les séries des puissances tant directes que réciproques, *Memoires de l'Académie Royale des Sciences et des Belles-Lettres* **17**, in *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin 1761*, Berlin, Haude et Spener, 1768, pp. 83–106. (Paper read to the academy in 1749.) `http://www.math.dartmouth.edu/~euler/pages/E352.html`

[6] William Feller. *An Introduction to Probability Theory and its Applications, Volume I*. 2nd ed, Wiley, New York, 1957.

[7] William Feller. *An Introduction to Probability Theory and its Applications, Volume II*. 2nd ed, Wiley, New York, 1971.

[8] Dominique Foata. Distributions eulériennes et mahoniennes sur le groupe des permutations. *Higher Combinatorics (Proc. NATO Advanced Study Institute, Berlin, 1976)*, Reidel, Dordrecht, 1977, pp. 27–49.

[9] Dominique Foata. Eulerian polynomials: from Euler's time to the present. *The Legacy of Alladi Ramakrishnan in the Mathematical Sciences*, 253–273, Springer, New York, 2010.

[10] Dominique Foata. Les polynômes eulériens, d'Euler à Carlitz. In *Leonhard Euler. Mathématicien, physicien et théoricien de la musique*, eds. Xavier Hascher and Athanase Papadopoulos, 413–432, CNRS Éditions, Paris, 2015.

[11] Svante Janson. Euler–Frobenius numbers and rounding. *Online J. Anal. Comb.* No. 8 (2013), Paper 5, 34 pp.

[12] Pierre-Simon de Laplace. *Théorie analytique des probabilités.* 3rd ed, Courcier, Paris 1820; *Œuvres completes de Laplace, vol. 7*, Gauthier-Villars, Paris, 1886. `http://gallica.bnf.fr/ark:/12148/bpt6k775950`

[13] Werner Meyer and R. von Randow. Ein Würfelschnittproblem und Bernoullische Zahlen. *Math. Ann.* **193** (1971), 315–321.

[14] Jun Murakami. Volume formulas for a spherical tetrahedron. *Proc. Amer. Math. Soc.* **140** (2012), no. 9, 3289–3295.

[15] T. Kyle Petersen. *Eulerian Numbers.* Birkhäuser/Springer, New York, 2015.

[16] Richard P. Stanley. Eulerian partitions of a unit hypercube. *Higher Combinatorics (Proc. NATO Advanced Study Institute, Berlin, 1976)*, Reidel, Dordrecht, 1977, p. 49.

[17] Richard P. Stanley. *Enumerative Combinatorics, Volume I.* Cambridge Univ. Press, Cambridge, 1997.

[18] S. Tanny. A probabilistic interpretation of Eulerian numbers. *Duke Math. J.* **40** (1973), 717–722. Corrigenda, *ibid* **41** (1974), 689.

[19] *NIST Handbook of Mathematical Functions.* Edited by Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert and Charles W. Clark. Cambridge Univ. Press, 2010.
Also available as *NIST Digital Library of Mathematical Functions*, `http://dlmf.nist.gov/`

[20] Warren D. Smith. Examples in which best Range Voting strategy is "dishonest" approval-style. `http://rangevoting.org/RVstrat2.html`.

[21] Warren D. Smith. Completion of Gibbard-Satterthwaite impossibility theorem; range voting and voter honesty, 2006. `http://rangevoting.org/WarrenSmithPages/homepage/newgibbsat.pdf`

[22] Warren D. Smith. How many candidates should a voter approve? 2022. `http://www.rangevoting.org/HowManyApprove.html`

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Email address: `svante.janson@math.uu.se`
URL: `http://www2.math.uu.se/~svante/`

Email address: `warren.wds@gmail.com`
URL: `http://RangeVoting.org`