Aequat. Math.

https://doi.org/10.1007/s00010-024-01073-1

Aequationes Mathematicae

Probabilistic Stirling numbers and applications

José A. Adell and Beáta Bényi

Abstract. We introduce probabilistic Stirling numbers of the first kind $s_Y(n,k)$ associated with a complex-valued random variable Y satisfying appropriate integrability conditions, thus completing the notion of probabilistic Stirling numbers of the second kind $S_Y(n,k)$ previously considered by the first author. Combinatorial interpretations, recursion formulas, and connections between $s_Y(n,k)$ and $S_Y(n,k)$ are given. We show that such numbers describe a large subset of potential polynomials, on the one hand, and the moments of sums of i. i. d. random variables, on the other, establishing their precise asymptotic behavior without appealing to the central limit theorem. We explicitly compute these numbers when Y has a certain familiar distribution, providing at the same time their combinatorial meaning.

Mathematics Subject Classification. 05A18, 05A19, 60E05.

Keywords. Probabilistic Stirling numbers, Set partitions, Potential polynomials, Centered subordinators, Normal distribution.

1. Introduction

Stirling numbers of the first and second kinds, respectively denoted by $s(n,k)$ and $S(n,k)$, are among the most fascinating integer arrays in mathematics, having numerous applications in combinatorics, number theory, probability theory, and other fields. From a combinatorial viewpoint, $S(n,k)$ counts the number of partitions of an n element set into k distinct blocks, whereas $|s(n,k)|$ gives the number of permutations of $[n]=\{1,2,\ldots,n\}$ into k distinct cycles. Analytically, such numbers can be defined in various equivalent ways (cf. Abramowitz and Stegun [1, p. 824] and Comtet [8, Chapter 5]). For instance,

$$(x)_n=\sum_{k=0}^{n} s(n,k)x^k \quad \text{and} \quad x^n=\sum_{k=0}^{n} S(n,k)(x)_k, \tag{1}$$

Published online: 08 May 2024
Birkhäuser

where $(x)_n = x(x-1)\cdots(x-n+1)$, $n \geq 1$ ($(x)_0 = 1$), or via their generating function as

$$\frac{\log^k(1+z)}{k!} = \sum_{n=k}^{\infty} s(n,k)\frac{z^n}{n!}, \quad |z|<1 \quad \text{and}$$
$$\frac{(e^z-1)^k}{k!} = \sum_{n=k}^{\infty} S(n,k)\frac{z^n}{n!}, \quad z \in \mathbb{C}. \tag{2}$$

Stirling numbers have been generalized in many different ways (see, for instance, Hsu and Shiue [11], Luo and Srivastava [14], Cakić et al. [7], El-Desouky et al. [9], Kim et al. [12], and Bényi et al. [6], among many others).

Denote by $\mathcal{G}_0$ the set of complex-valued random variables Y such that

$$\mathbb{E}e^{|zY|} < \infty, \quad |z| \leq r, \tag{3}$$

for some $r > 0$, where $\mathbb{E}$ stands for mathematical expectation. Probabilistic Stirling numbers of the second kind $S_Y(n,k)$ associated with the random variable $Y \in \mathcal{G}_0$ were defined in [4] by

$$\frac{\left(\mathbb{E}e^{zY}-1\right)^k}{k!} = \sum_{n=k}^{\infty} S_Y(n,k)\frac{z^n}{n!}. \tag{4}$$

These numbers have found applications in analytic number theory, such as extensions of the classical formula for sums of powers on arithmetic progressions (cf. [4]) and explicit expressions for a large class of Appell polynomials (see [5]). Applications in probability theory, especially to computing moments of sums of independent identically distributed random variables and Edgeworth expansions are given in [2].

The aim of this paper is to complete definition (4) by introducing probabilistic Stirling numbers of the first kind $s_Y(n,k)$, showing at the same time various applications of them. To this end, consider the cumulant generating function of $Y \in \mathcal{G}_0$, that is,

$$\mathcal{K}_Y(z) = \log \mathbb{E}e^{zY} = \sum_{n=k}^{\infty} \kappa_n(Y)\frac{z^n}{n!}. \tag{5}$$

We define *probabilistic Stirling numbers of the first kind*, $s_Y(n,k)$, *associated with* $Y \in \mathcal{G}_0$ via their generating function as

$$\frac{(\mathcal{K}_Y(z))^k}{k!} = \sum_{n=k}^{\infty} (-1)^{n-k} s_Y(n,k)\frac{z^n}{n!}. \tag{6}$$

If $Y = 1$, we see from (4) and (6) that

$$s_Y(n,k) = \delta_{n,k} \quad \text{and} \quad S_Y(n,k) = S(n,k). \tag{7}$$

Suppose that Y has the exponential density $\rho(\theta) = e^{-\theta}$, $\theta \geq 0$, and denote the rising factorial by $\langle x \rangle_n = x(x+1)\cdots(x+n-1)$, $n \geq 1$ ($\langle x \rangle_0 = 1$). It will be shown in formula (45) below that in this case we have

$$s_Y(n,k) = s(n,k) \quad \text{and} \quad S_Y(n,k) = \binom{n}{k} \langle k \rangle_{n-k} = L(n,k), \tag{8}$$

where $L(n,k)$ are the Lah numbers (cf. [13]), sometimes also called the Stirling numbers of the third kind. Thus, probabilistic Stirling numbers generalize the classical ones, although associated with different random variables.

The paper is organized as follows. In the next section, we give a combinatorial meaning of probabilistic Stirling numbers in connection with binomial convolution of sequences. In Sect. 3, we show various properties of such numbers, namely, connections between $S_Y(n,k)$ and $s_Y(n,k)$, recursion formulas, computations of moments and cumulants in terms of such numbers, and their behavior with respect to sums of independent random variables. Two applications are considered in Sect. 4. On the one hand, we show that a large subset of potential polynomials can be explicitly written in terms of probabilistic Stirling numbers. On the other hand, we obtain explicit formulas for the moments of standardized sums of independent identically distributed random variables. This allows us to provide for such moments explicit rates of convergence with exact leading coefficients, without appealing to the central limit theorem. Finally, Sect. 5 is devoted to obtaining closed formulas for $S_Y(n,k)$ and $s_Y(n,k)$, together with their combinatorial interpretations, for some familiar random variables Y. Specifically, we consider centered subordinators, in particular, Poisson and gamma processes, and normally distributed random variables.

2. Combinatorial interpretation

Let $\mathbb{N}_0$ be the set of non-negative integers. Throughout this paper, we assume that $k, n \in \mathbb{N}_0$, $x \in \mathbb{R}$, and $z \in \mathbb{C}$ with $|z| \leq r$, for some $r > 0$, where r may change from line to line. The following definitions and properties can be found in [3]. Denote by $\mathbb{H}$ the set of complex sequences $\mathbf{u} = (u_n)_{n \geq 0}$ for which the generating function

$$G(\mathbf{u}, z) := \sum_{n=0}^{\infty} u_n \frac{z^n}{n!}$$

is well defined. If $\mathbf{u}, \mathbf{v} \in \mathbb{H}$, we define its binomial convolution $\mathbf{u} \times \mathbf{v} = ((u \times v)_n)_{n \geq 0}$ as

$$(u \times v)_n = \sum_{j=0}^{n} \binom{n}{j} u_j v_{n-j}. \tag{9}$$

It turns out that the sequence $\mathbf{u} \times \mathbf{v} \in \mathbb{H}$ and is characterized by

$$G(\mathbf{u} \times \mathbf{v}, z) = G(\mathbf{u}, z) G(\mathbf{v}, z). \tag{10}$$

In addition, $(\mathbb{H}, \times)$ is an abelian group with identity element $\mathbf{e} = (\delta_{n0})_{n \geq 0}$. Given $\mathbf{u}^{(m)} = (u_n^{(m)})_{n \geq 0} \in \mathbb{H}$, $m = 1, \ldots, k$, we have $\mathbf{u}^{(1)} \times \cdots \times \mathbf{u}^{(k)} \in \mathbb{H}$ and

$$(u^{(1)} \times \cdots \times u^{(k)})_n = \sum_{j_1 + \cdots + j_k = n} \frac{n!}{j_1! \cdots j_k!} u_{j_1}^{(1)} \cdots u_{j_k}^{(k)}. \tag{11}$$

Such a sequence is characterized by

$$G(\mathbf{u}^{(1)} \times \cdots \times \mathbf{u}^{(k)}, z) = G(\mathbf{u}^{(1)}, z) \cdots G(\mathbf{u}^{(k)}, z). \tag{12}$$

We define a partition of a set $[n] := \{1, 2, \ldots, n\}$ as a collection of pairwise disjoint subsets, called *blocks* whose union is $[n]$. Let Π_n denote the set of all set partitions of $[n]$. A set partition is given by the list of its blocks. We write the blocks in increasing order according to their least element, for example $\sigma = \{1, 7, 8\}\{2, 3, 9, 10\}\{4\}\{5, 6\} \in \Pi_{10}$. Let $\#\sigma$ denote the number of blocks in σ and $|B|$ the number of elements in the block $B \in \sigma$. If the order of the blocks is not arbitrary, we obtain an *ordered set partition.*

The right hand side of equality (11) counts the ordered set partitions with k distinct blocks such that the ith block has size j_i $|B_i| = j_i$ and is weighted by $u_{j_i}^{(i)}$. The weight of the set partition is defined as the product of the weights of its blocks.

To see that, let $j_1 + \cdots + j_k = n$ be a composition of n with k parts, that denote the sizes of the blocks. We choose first j_1 elements from n to create the first block in $\binom{n}{j_1}$ ways, then we choose j_2 elements for the second block in $\binom{n-j_1}{j_2}$ ways and so on, which gives $\binom{n}{j_1, j_2, \ldots, j_k}$ possibilities. Every such partition has the weight $u_{j_1}^{(1)} \cdots u_{j_k}^{(k)}$. Summing over all compositions of n, we get formula (11).

By ignoring the order of the blocks, from (11) and (12) the following combinatorial expression follows

$$\left[\frac{z^n}{n!}\right] \frac{G^k(\mathbf{u}, z)}{k!} = \sum_{\substack{\sigma \in \Pi_n \\ \#\sigma = k}} \prod_{B \in \sigma} u_{|B|}. \tag{13}$$

Remark 1. In terms of the symbolic method, this identity is nothing else but the so called $\mathtt{SET}_k$ construction. See for details the book of Flajolet and Sedgewick [10].

Considering the definition of probabilistic Stirling numbers of the second kind by the generating function (4) and setting $G(\mathbf{u}, z) = \mathbb{E}e^{zY} - 1$ ($u_n = \mathbb{E}Y^n$ in this setting) in (13), we obtain the following combinatorial definition of probabilistic Stirling numbers

$$S_Y(n,k) = \sum_{\substack{\sigma\in\Pi_n \\ \#\sigma=k}} \prod_{B\in\sigma} \mathbb{E}Y^{|B|}. \tag{14}$$

Note that the definition ensures that the composition makes sense, since the constant term in the power series $\mathbb{E}e^{zY} - 1$ is 0.

Similarly, from (6) and setting $G(\mathbf{u}, z) = \mathcal{K}_Y(z)$ ($u_n = \kappa_n(Y)$ in this setting) in (13), we obtain

$$s_Y(n,k) = (-1)^{n-k} \sum_{\substack{\sigma\in\Pi_n \\ \#\sigma=k}} \prod_{B\in\sigma} \kappa_{|B|}(Y). \tag{15}$$

In other words, $S_Y(n,k)$ is the weighted sum of all set partitions of n into k blocks, such that the weight of a block of size j is defined as the jth moment of the random variable Y, while $|s_Y(n,k)|$ is the weighted sum of all set partitions of n into k blocks, such that the weight of a block of size j is the jth cumulant of the random variable Y.

Example 2. Observe that

$$S_Y(4,2) = 4\mathbb{E}Y^3\mathbb{E}Y + 3(\mathbb{E}Y^2)^2,$$

since we have 4 partitions of 4 with blocks of size $3-1$:

$$1-234, \quad 2-134, \quad 3-124, \quad 4-123$$

and 3 partitions with blocks of sizes $2-2$:

$$12-34, \quad 13-24, \quad 14-23.$$

Moment sequences and the sequences of cumulants are sometimes known as enumerating sequences (or weighted enumerating sequences) of combinatorial objects. The above definitions lead to combinatorial interpretations of probabilistic Stirling numbers associated with a specific random variable Y.

For instance, if $Y = 1$, the moment $\mathbb{E}Y^j = 1$ for all j and the cumulants κ_j are all zero except for $j = 1$, i.e., $\kappa_1(Y) = 1$ and $\kappa_j(Y) = 0$ for $j \geq 2$. Hence, (7) is immediate by the above definitions (14) and (15).

Similarly, if Y has the exponential density $\rho(\theta) = e^{-\theta}$, $\theta \geq 0$, it is known that $\mathbb{E}Y^j = j!$, the number of permutations of a j element set. Weighting a block by the number of permutations of its elements simply means that (14) counts in this special case in how many ways we can order the set $[n]$ into k linearly ordered subsets, which is exactly the combinatorial definition of the Lah numbers. On the other hand, the cumulants are given by $\kappa_j(Y) = (j-1)!$, and we see that in this case definition (15) gives the number of ways to partition $[n]$ into k blocks such that the elements in each block are ordered except one element, say the least. But this is the definition of a permutation written in cycle notation and having exactly k distinct cycles. This is the combinatorial definition of classical Stirling numbers of the first kind, so (8) follows. We will give some other examples in Sect. 5.

3. Main properties

Probabilistic Stirling numbers of the first and second kinds are closely connected. In fact, we can compute $S_Y(n,k)$ in terms of $s_Y(n,k)$ with the help of classical Stirling numbers of the second kind. A similar converse property is also true, as shown in the following result.

Theorem 3. *Let $Y \in \mathcal{G}_0$. Then,*

$$\sum_{k=0}^{n} S_Y(n,k)(x)_k = \sum_{k=0}^{n} (-1)^{n-k} s_Y(n,k) x^k. \tag{16}$$

As a consequence,

$$S_Y(n,k) = \sum_{j=k}^{n} (-1)^{n-j} s_Y(n,j) S(j,k) \tag{17}$$

and

$$s_Y(n,k) = (-1)^{n-k} \sum_{j=k}^{n} S_Y(n,j) s(j,k). \tag{18}$$

Proof. Let $z \in \mathbb{C}$ with $\left|\mathbb{E}e^{zY} - 1\right| < 1$. Using (4) and the binomial expansion, we have

$$(\mathbb{E}e^{zY})^x = (\mathbb{E}e^{zY} - 1 + 1)^x = \sum_{k=0}^{\infty} (x)_k \frac{(\mathbb{E}e^{zY} - 1)^k}{k!} \tag{19}$$

$$= \sum_{k=0}^{\infty} (x)_k \sum_{n=k}^{\infty} S_Y(n,k) \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{k=0}^{n} S_Y(n,k)(x)_k. \tag{20}$$

On the other hand, we get from (6)

$$(\mathbb{E}e^{zY})^x = e^{x\mathcal{K}_Y(z)} = \sum_{k=0}^{\infty} x^k \frac{(\mathcal{K}_Y(z))^k}{k!} = \sum_{k=0}^{\infty} x^k \sum_{n=k}^{\infty} (-1)^{n-k} s_Y(n,k) \frac{z^n}{n!} \tag{21}$$

$$= \sum_{n=0}^{\infty} \frac{z^n}{n!} \sum_{k=0}^{n} (-1)^{n-k} s_Y(n,k) x^k. \tag{22}$$

Thus, identity (16) follows by equating the coefficients in (19) and (21). Starting from (16), we have by the second equality in (1)

$$\begin{aligned}\sum_{k=0}^{n} S_Y(n,k)(x)_k &= \sum_{j=0}^{n-j}(-1)^{n-j}s_Y(n,j)x^j = \sum_{j=0}^{n}(-1)^{n-j}s_Y(n,j)\sum_{k=0}^{j}S(j,k)(x)_k \\ &= \sum_{k=0}^{n}(x)_k\sum_{j=k}^{n}(-1)^{n-j}s_Y(n,j)S(j,k).\end{aligned}$$

Therefore, (17) follows by equating the coefficients in the last expression. Formula (18) is shown in a similar way. □

It will be proved in Theorem 7 below that identity (16) is in fact the explicit expression of a certain potential polynomial of degree n. On the other hand, when applied to specific random variables, formulas (17) and (18) give rise to different identities. For instance, if $Y=1$ (resp. if Y is exponentially distributed), we obtain from (7) and (18) (resp. from (8) and (17)) the well known identities

$$\sum_{j=k}^{n} S(n,j)s(j,k) = \delta_{n,k} \quad \text{and} \quad \sum_{j=k}^{n}(-1)^{n-j}s(n,j)S(j,k) = L(n,k).$$

Also, if Y is exponentially distributed, we obtain from (1), (8), and (16)

$$\sum_{k=0}^{n} L(n,k)(x)_k = \sum_{k=0}^{n}(-1)^{n-k}s(n,k)x^k = (-1)^n(-x)_n = \langle x\rangle_n\,, \tag{23}$$

which is the known way to define Lah numbers as connecting coefficients between rising and falling factorials.

An easy consequence of Theorem 3 concerning moments and cumulants of random variables is given in the following result.

Corollary 4. *Let $Y \in \mathcal{G}_0$. Then,*

$$\mathbb{E}Y^n = \sum_{j=1}^{n}(-1)^{n-j}s_Y(n,j)$$

and

$$\kappa_n(Y) = \sum_{j=1}^{n}(-1)^{j-1}(j-1)!S_Y(n,j).$$

Proof. The first equality follows by setting $k=1$ in (17) and noting that

$$S_Y(n,1) = \mathbb{E}Y^n \quad \text{and} \quad S(j,1)=1, \quad j\geq 1. \tag{24}$$

The second one follows by choosing $k=1$ in (18) and observing that

$$(-1)^{n-1}s_Y(n,1) = \kappa_n(Y) \quad \text{and} \quad s(j,1) = (-1)^{j-1}(j-1)!, \quad j\geq 1. \tag{25}$$

The proof is complete. $\square$

Considering Corollary 4 from a combinatorial viewpoint, the first equality says that the nth moment can be calculated as the sum over *all* partitions such that a block of size j is weighted by the jth cumulant. More precisely,

$$\mathbb{E}Y^n = \sum_{\sigma\in\Pi_n} \prod_{B\in\sigma} \kappa_{|B|}.$$

This is in accordance with the combinatorial representation given by Speed [17]. Similarly, in the formula for the nth cumulant, *all* partitions are counted and the weight $\mathbb{E}Y^j$ is given to a block of size j, so we have

$$\kappa_n(Y) = \sum_{\sigma\in\Pi_n} (-1)^{\#\sigma-1}(\#\sigma-1)! \prod_{B\in\sigma} \mathbb{E}Y^{|B|}.$$

In Speed [17], this formula is derived by using the Möbius function of the partition lattice.

Probabilistic Stirling numbers can be recursively computed, as shown in the following result.

Theorem 5. *Let $Y \in \mathcal{G}_0$. For any $n \geq k \geq 1$, we have*

$$s_Y(n,k) = \frac{1}{k}\sum_{j=k-1}^{n-1} \binom{n}{j} s_Y(j,k-1)(-1)^{n-1-j}\kappa_{n-j}(Y) \tag{26}$$

and

$$S_Y(n,k) = \frac{1}{k}\sum_{j=k-1}^{n-1} \binom{n}{j} S_Y(j,k-1)\mathbb{E}Y^{n-j}. \tag{27}$$

In particular,

$$s_Y(k,k) = (\kappa_1(Y))^k \quad \text{and} \quad S_Y(k,k) = (\mathbb{E}Y)^k. \tag{28}$$

Proof. Starting with the identity

$$\frac{(\mathcal{K}_Y(z))^k}{k!} = \frac{(\mathcal{K}_Y(z))^{k-1}}{(k-1)!}\frac{\mathcal{K}_Y(z)}{k},$$

and applying (9) and (10), we see that

$$\begin{aligned}(-1)^{n-k}s_Y(n,k) &= \frac{1}{k}\sum_{j=0}^{n}\binom{n}{j}(-1)^{j-k-1}s_Y(j,k-1)(-1)^{n-j-1}s_Y(n-j,1)\\ &= \frac{1}{k}\sum_{j=0}^{n}\binom{n}{j}(-1)^{j-k-1}s_Y(j,k-1)\kappa_{n-j}(Y),\end{aligned}$$

where we have used (25) in the last equality. Hence, (26) follows from the fact that

$$S_Y(n,k) = s_Y(n,k) = 0, \quad n < k. \tag{29}$$

The proof of (27) follows a similar pattern, starting from the identity

$$\frac{\left(\mathbb{E}e^{zY} - 1\right)^k}{k!} = \frac{\left(\mathbb{E}e^{zY} - 1\right)^{k-1}}{(k-1)!} \frac{\mathbb{E}e^{zY} - 1}{k}$$

and using (24) and (29). Finally, we have from (26)

$$s_Y(k,k) = s_Y(k-1,k-1)\kappa_1(Y).$$

Thus, the first equality in (28) follows by induction on k. The second one is shown in a similar way. $\square$

We also provide a combinatorial proof for Theorem 5. We rewrite formula (26) in a slightly different form as

$$k s_Y(n,k) = \sum_{n-j=1}^{n-k+1} \left[\binom{n}{n-j} \kappa_{n-j}(Y) \right] \left[(-1)^{n-1-j} s_Y(j,k-1) \right].$$

The left hand side counts the weighted sum of partitions of $[n]$ into k blocks with a mark on one of the blocks. The right hand side does the same. First, we choose $n-j$ elements, form a block with them and associate with this block the weight $\kappa_{n-j}(Y)$. This block is the marked block. Then we create from the remaining j elements a partition into $k-1$ blocks and weight the blocks by the cumulants. Summing over all possible values of $n-j$ we get all the cases. The second formula can be proven in a similar manner.

Formulas (28) express the fact that there is only one way k elements can form a partition with k blocks, namely, that each block is a singleton.

Probabilistic Stirling numbers of the first kind have a simple behavior with respect to the sum of independent random variables. In this regard, if $(Y_k)_{k\geq 1}$ is a sequence of independent copies of $Y \in \mathcal{G}_0$, we denote

$$W_m = Y_1 + \cdots + Y_m, \quad m \in \mathbb{N}_0 \setminus \{0\}, \quad (W_0 = 0). \tag{30}$$

Theorem 6. *Let X and Y be two independent random variables in $\mathcal{G}_0$. Then,*

$$s_{X+Y}(n,k) = \sum_{m=0}^{k} \sum_{j=0}^{n} \binom{n}{j} s_X(j,m) s_Y(n-j,k-m). \tag{31}$$

In addition,

$$s_{W_m}(n,k) = m^k s_Y(n,k). \tag{32}$$

Proof. Since X and Y are independent, we have

$$\frac{\left(\mathcal{K}_{X+Y}(z)\right)^k}{k!} = \frac{\left(\mathcal{K}_X(z)+\mathcal{K}_Y(z)\right)^k}{k!} = \sum_{m=0}^{k} \frac{\left(\mathcal{K}_X(z)\right)^m}{m!} \frac{\left(\mathcal{K}_Y(z)\right)^{k-m}}{(k-m)!}.$$

Therefore, (31) follows from (9) and (10). Formula (32) readily follows from the fact that

$$\mathcal{K}_{W_m}(z) = m\mathcal{K}_Y(z).$$

The proof is complete. □

In [4, Th. 4.6], a very involved formula for $S_{X+Y}(n,k)$ was given. An easier alternative computing $S_{X+Y}(n,k)$ is to successively apply formulae (31) and (17).

4. Potential polynomials and moments

Let $B(z)$ be an analytic function at $z=0$ such that $B(0)=1$. The potential polynomials $\mathbb{P}(x) = (P_n(x))_{n\geq 0}$ associated with $B(z)$ are defined (see Comtet [8, Section 3.5] and Wang [21]) by means of the generating function

$$G(\mathbb{P}(x), z) = \sum_{n=0}^{\infty} P_n(x)\frac{z^n}{n!} = B(z)^x.$$

Note that $P_n(0) = 1$. In this paper, we consider the subset $\mathcal{P}$ of potential polynomials such that

$$\sum_{n=0}^{\infty} P_n(x)\frac{z^n}{n!} = \left(\mathbb{E}e^{zY}\right)^x, \tag{33}$$

for some $Y \in \mathcal{G}_0$. For such polynomials, we give the following result.

Theorem 7. *Let $\mathbb{P}(x) \in \mathcal{P}$. Then,*

$$P_n(x) = \sum_{k=0}^{n} S_Y(n,k)(x)_k = \sum_{k=0}^{n}(-1)^{n-k}s_Y(n,k)x^k. \tag{34}$$

In addition, we have for any $m \in \mathbb{N}_0$

$$\mathbb{E}W_m^n = P_n(m) = \sum_{k=0}^{n} S_Y(n,k)(m)_k = \sum_{k=0}^{n}(-1)^{n-k}s_Y(n,k)m^k. \tag{35}$$

Proof. Formula (34) readily follows from (19), (21), and (33). Choosing $x=m$ in (33), we get from (30)

$$\sum_{n=0}^{\infty} P_n(m)\frac{z^n}{n!} = \left(\mathbb{E}e^{zY}\right)^m = \mathbb{E}e^{zW_m} = \sum_{n=0}^{\infty}\mathbb{E}W_m^n\frac{z^n}{n!}.$$

This shows (35) and completes the proof. $\square$

Theorem 7 has both an analytic and a probabilistic reading. On the one hand, formula (34) gives us explicit expressions for the potential polynomials in $\mathcal{P}$ in terms of probabilistic Stirling numbers. A similar expression to the first equality in (34) in terms of Bell polynomials was given by Comtet [8, p.141] and [21]. On the other hand, formula (35) allows us to compute moments of sums of independent identically distributed random variables by means of probabilistic Stirling numbers.

In this last respect, formulas become simpler for centered random variables. In fact, denote by $\mathcal{G}_1$ the set of random variables $Y \in \mathcal{G}_0$ such that $\mathbb{E}Y = 0$. On the other hand, if Z is a random variable having the standard normal density, it is well known that

$$\mathbb{E}Z^{2n} = \frac{(2n)!}{n!2^n}. \tag{36}$$

Lemma 8. *Assume that $Y \in \mathcal{G}_1$ and denote by $\sigma^2 = \mathbb{E}Y^2$. Then,*

$$s_Y(n,k) = S_Y(n,k) = 0, \quad n < 2k. \tag{37}$$

As a consequence,

$$(-1)^n s_Y(2n,n) = S_Y(2n,n) = \mathbb{E}(\sigma Z)^{2n}, \tag{38}$$

as well as

$$(-1)^{n+1} s_Y(2n+1,n) = S_Y(2n+1,n) = n(2n+1)\frac{\mathbb{E}Y^3}{3\sigma^2}\mathbb{E}(\sigma Z)^{2n}. \tag{39}$$

Proof. In [2], it was shown that $S_Y(n,k) = 0$ for $n < 2k$. Thus, (37) readily follows from (18). On the other hand, the second equalities in (38) and (39) were shown in [2]. Therefore, formulae (38) and (39) follow from (18) and (37), after observing that $S(k,k) = 1$ for $k \in \mathbb{N}_0$. $\square$

Suppose that $Y \in \mathcal{G}_1$ is real valued. It is well known (cf. [20]) that the standardized sums $W_m/(\sigma\sqrt{m})$ fulfill the asymptotic result

$$\lim_{m\to\infty} \mathbb{E}\left(\frac{W_m}{\sigma\sqrt{m}}\right)^n = \mathbb{E}Z^n, \quad n \in \mathbb{N}_0. \tag{40}$$

The usual way to prove (40) is based on the central limit theorem satisfied by $W_m/(\sigma\sqrt{m})$. Note that if n is odd, the limit in (40) is zero and therefore no information about the speed of convergence for odd moments can be inferred from (40). Here, we provide the following non asymptotic result without appealing to the central limit theorem.

Corollary 9. *Assume that $Y \in \mathcal{G}_1$ and let $\sigma^2 = \mathbb{E}Y^2$. Then,*

$$\mathbb{E}\left(\frac{W_m}{\sigma\sqrt{m}}\right)^{2n} = \frac{1}{\sigma^{2n}}\sum_{k=0}^{n} S_Y(2n,k)\frac{(m)_k}{m^n} = \frac{1}{\sigma^{2n}}\sum_{k=0}^{n}(-1)^k s_Y(2n,k)\frac{1}{m^{n-k}}$$
$$= \mathbb{E}Z^{2n} + O\left(\frac{1}{m}\right),$$

and

$$\mathbb{E}\frac{W_m^{2n+1}}{(\sigma\sqrt{m})^{2n}} = \frac{1}{\sigma^{2n}}\sum_{k=0}^{n} S_Y(2n+1,k)\frac{(m)_k}{m^n} = \frac{1}{\sigma^{2n}}\sum_{k=0}^{n}(-1)^k s_Y(2n+1,k)\frac{1}{m^{n-k}}$$
$$= n(2n+1)\frac{\mathbb{E}Y^3}{3\sigma^2}\mathbb{E}Z^{2n} + O\left(\frac{1}{m}\right).$$

Proof. The result follows from (35), Lemma 8, and some simple computations. □

Whenever $\mathbb{E}Y^3 \neq 0$, Corollary 9 tells us that the order of magnitude of $\mathbb{E}W_m^{2n}$ and $\mathbb{E}W_m^{2n+1}$ is m^n, as $m \to \infty$, providing at the same time the exact leading constants.

Another interesting consequence for the potential polynomials is the following.

Corollary 10. *Let $\mathbb{P}(x) \in \mathcal{P}$ with associated random variable $Y \in \mathcal{G}_1$. Then,*

$$P_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} S_Y(n,k)(x)_k = \sum_{k=0}^{\lfloor n/2 \rfloor} (-1)^{n-k} s_Y(n,k)x^k,$$

where $\lfloor x \rfloor$ stands for the integer part of x.

Proof. Apply (37) to formula (34). □

According to Corollary 10 and (38), $P_{2n}(x)$ has exact degree n, whereas $P_{2n+1}(x)$ has exact degree n, whenever $\mathbb{E}Y^3 \neq 0$, as follows from (39). If $\mathbb{E}Y^3 = 0$, then $P_{2n+1}(x)$ has degree at most $n-1$.

5. Centered subordinators and other examples

Let $\mathbb{X} = (X(t))_{t\geq 0}$ be a centered subordinator, that is, a stochastic process starting at the origin, having independent stationary increments, and such that $\mathbb{E}X(t) = t$, $t \geq 0$. The characteristic function of this process is given by (cf. Steutel and van Harn [18, p. 107] and [2])

$$\mathbb{E}e^{i\xi X(t)} = \exp\left(t\mathbb{E}\left(\frac{e^{i\xi T-1}}{T}\right)\right) = \exp\left(i\xi t\mathbb{E}e^{i\xi UT}\right), \quad \xi \in \mathbb{R}, \tag{41}$$

where U is a random variable uniformly distributed on $[0,1]$ and independent of the non-negative random variable T, which determines $\mathbb{X}$. Here, we assume that $T \in \mathcal{G}_0$.

Let $(U_j)_{j\geq 1}$ and $(T_j)_{j\geq 1}$ be two sequences of independent copies of U and T, respectively. Assume that both sequences are mutually independent and for $k \in \mathbb{N}_0 \setminus \{0\}$ denote

$$V_k = U_1T_1 + \cdots + U_kT_k \quad (V_0 = 0). \tag{42}$$

With the preceding notations and assumptions, we give the following result.

Proposition 11. *For any $t \geq 0$, we have*

$$s_{X(t)}(n,k) = (-1)^{n-k} t^k \binom{n}{k} \mathbb{E}V_k^{n-k}.$$

Proof. By (41), we have

$$\mathcal{K}_{X(t)}(z) = tz\mathbb{E}e^{zUT},$$

thus implying, by virtue of (42), that

$$\frac{\left(\mathcal{K}_{X(t)}(z)\right)^k}{k!} = \frac{t^kz^k}{k!}\mathbb{E}e^{zV_k} = \frac{t^kz^k}{k!}\sum_{m=0}^{\infty}\mathbb{E}V_k^m\frac{z^m}{m!} = t^k\sum_{n=k}^{\infty}\binom{n}{k}\mathbb{E}V_k^{n-k}\frac{z^n}{n!}.$$

Therefore, the conclusion follows from (6). □

We provide a combinatorial interpretation of Proposition 11.

Proposition 12. *For any $t \geq 0$, $(-1)^{n-k}s_{X(t)}(n,k)$ is the weighted sum of partitions of $[n]$ into k distinct blocks such that block B_i has the weight $t\mathbb{E}T_i^{|B_i|}$.*

Proof. We have by the multinomial theorem and by the independence of the random variables involved

$$\begin{aligned}\mathbb{E}V_k^{n-k} &= \sum_{j_1+\cdots+j_k=n-k}\binom{n-k}{j_1,\ldots,j_2}\mathbb{E}\left(U_1T_1\right)^{j_1}\cdots\mathbb{E}(U_kT_k)^{j_k}\\ &= \sum_{j_1+\cdots+j_k=n-k}\binom{n-k}{j_1,\ldots,j_2}\mathbb{E}U_1^{j_1}\cdots\mathbb{E}U_k^{j_k}\mathbb{E}T_1^{j_1}\cdots\mathbb{E}T_k^{j_k}.\end{aligned}$$

Since $\mathbb{E}U^j = 1/(j+1)$ for a random variable U uniformly distributed on $[0,1]$, we have

$$t^k\binom{n}{k}\mathbb{E}V_k^{n-k} = \sum_{j_1+\cdots+j_k=n-k}\binom{n}{k}\binom{n-k}{j_1,\ldots,j_2}\frac{1}{j_1+1}\cdots\frac{1}{j_k+1}t^k\mathbb{E}T_1^{j_1}\cdots\mathbb{E}T_k^{j_k}. \tag{43}$$

We count the partitions of $[n]$ into k non-empty blocks by enumerating the possible cases with given block sizes. More precisely, given a composition of

n into k parts $j_1' + \cdots + j_k' = n$, we count how many partitions there are with $|B_i| = j_i'$, $i \in [k]$ (note that $j_i' > 0$). First, we choose k elements from n that we "put" into different blocks (ordering the so created—at the moment one-element blocks—in increasing order). Let us call these elements in a block the *indicator element* of the block. Now we choose from the remaining $n-k$ elements $j_1' - 1$ elements and put them into the first block, then we choose $j_2' - 1$ elements from the remaining $n - k - j_1' + 1$ elements, and so on. This can be done in

$$\binom{n}{k}\binom{n-k}{j_1'-1,\ldots,j_2'-1}$$

ways. However, in a block any of the elements could be the indicator element. Therefore, in the enumeration above each partition is counted more times, namely $j_1'j_2'\cdots j_k'$ times. We introduce the notation $j_i = j_i' - 1$ (note that $j_i \geq 0$) and obtain for given $j_1, \ldots, j_k$ with $j_1 + \cdots + j_k = n - k$ that the number of partitions is

$$\binom{n}{k}\binom{n-k}{j_1,\ldots,j_2}\frac{1}{j_1+1}\cdots\frac{1}{j_k+1}.$$

Associating with each block the weight $t\mathbb{E}T_i^{|B_i|}$ and summing over all possible compositions, we obtain (43) and the proof is complete. □

Probabilistic Stirling numbers of the second kind $S_{X(t)}(n,k)$ can be computed using (17) and Proposition 11.

When considering particular examples of centered subordinators, interesting formulas can be obtained. As an illustration, we give the following two examples.

5.1. The standard Poisson process $(N(t))_{t\geq 0}$

This process satisfies (41) with $T = 1$. Explicit expressions for its associated Stirling numbers are given in the following result.

Proposition 13. *For any $t \geq 0$, we have*

$$s_{N(t)}(n,k) = (-1)^{n-k}S(n,k)t^k \quad \textit{and} \quad S_{N(t)}(n,k) = \sum_{j=k}^{n} S(n,j)S(j,k)t^j.$$

Proof. Choosing $T = 1$ in (41), we get

$$\mathcal{K}_{N(t)}(z) = t(e^z - 1).$$

Thus, the first equality follows from (2) and (6), whereas the second one follows from (17). The proof is complete. □

Comparing Proposition 11 with the first equality in Proposition 13, we obtain the following probabilistic representation for classical Stirling numbers of the second kind

$$S(n,k) = \binom{n}{k} \mathbb{E}(U_1 + \cdots + U_k)^{n-k}.$$

Such a representation was already shown by Sun [19]. On the other hand, we have from Corollary 4 and the first identity in Proposition 13

$$\mathbb{E}N(t)^n = \sum_{j=1}^{n} S(n,j) t^j, \quad t \geq 0. \tag{44}$$

This is known in the literature as Dobiński's formula (cf. Pinsky [15]).

Next, we also prove Proposition 13 based on the combinatorial definitions of probabilistic Stirling numbers. In the case of the standard Poisson process, the cumulants are $\kappa_n = t$ for all $n \geq 1$. Substituting this into the combinatorial definition of Stirling numbers of the first kind, we get

$$s_Y(n,k) = (-1)^{n-k} \sum_{\substack{\sigma \in \Pi_n \\ \#\sigma = k}} \prod_{B \in \sigma} t.$$

This expression can be written in a simpler way. We associate with each block in a partition a weight t, then we sum over all partitions with exactly k distinct blocks. Since we fixed the number of blocks in the partition, each partition that we take into account has weight t^k. Hence,

$$s_Y(n,k) = (-1)^{n-k} \sum_{\sigma \in \Pi_n} t^k,$$

which is the first formula in Proposition 13.

Similarly, as given in (44), the moments of $N(t)$ are the values of the Touchard polynomials and our combinatorial definition gives

$$S_Y(n,k) = \sum_{\substack{\sigma \in \Pi_n \\ \#\sigma = k}} \prod_{B \in \sigma} \sum_{j=1}^{|B|} S(|B|, j) t^j,$$

which has the following combinatorial interpretation. Given n objects, color them with exactly k colors (each color should occur), i.e., partition them into k distinct blocks. Create now a partition from the colored elements into j monochrome blocks. Associate with each block the weight t. This coincides with the second formula in Proposition 13, since we can also partition the n elements into j blocks, with weights t^j, and then color the blocks with k colors.

5.2. The gamma process $(Y(t))_{t\geq 0}$

For each $t>0$, $Y(t)$ has the gamma density

$$\rho_t(\theta)=\frac{\theta^{t-1}}{\Gamma(t)}e^{-\theta},\quad \theta>0.$$

It turns out (cf. [2]) that $(Y(t))_{t\geq 0}$ is a centered subordinator satisfying (41), where T has the exponential density $\rho_1(\theta)$.

Proposition 14. *For any $t\geq 0$, we have*

$$s_{Y(t)}(n,k)=s(n,k)t^k \quad \textit{and} \quad S_{Y(t)}(n,k)=\sum_{j=k}^{n}(-1)^{n-j}s(n,j)S(j,k)t^j.$$

Proof. Since

$$\mathcal{K}_{Y(t)}(z)=-t\log(1-z),\quad |z|<1,$$

we have from (2)

$$\frac{\left(\mathcal{K}_{Y(t)}(z)\right)^k}{k!}=t^k\sum_{n=k}^{\infty}(-1)^{n-k}s(n,k)\frac{z^n}{n!}.$$

This, together with (6), shows the first identity in Proposition 14. The second one follows from (17).

For the random variable $Y(1)$ having the exponential density, it was proved in [4] that

$$S_{Y(1)}(n,k)=\binom{n}{k}\langle k\rangle_{n-k}=L(n,k). \tag{45}$$

By the first identity in Proposition 14, we also have $s_{Y(1)}(n,k)=s(n,k)$.

We provide an argument for the first equality in Proposition 14 based on the combinatorial interpretation in Proposition 12. If T has the exponential density, then $\mathbb{E}T^j=j!$, $j\geq 0$. According to the definition of the weights of the blocks, this gives a factor $t(|B_i|-1)!$, for $i\in[k]$, which means that we actually count the possible arrangements of the elements in each block after the indicator element. In other words, we create now cycles, and $(-1)^{n-k}s_{X(t)}(n,k)$ is the total number of permutations with k cycles, which is $(-1)^{n-k}s(n,k)t^k$, as stated. Stirling numbers of the second kind, $S_{Y(t)}(n,k)$ counts partitions of $[n]$ into k linearly ordered blocks. The elements in block B_i are in one-to-one correspondence with a permutation of $\{1,2,\ldots,|B_i|\}$. The associated weight of the block B_i is now the number of cycles in this permutation. One can check that both the original (combinatorial) definition of $S_{Y(t)}(n,k)$ and the second expression in Proposition 14 indeed count these objects.

To conclude this section, we consider a random variable Z having the normal density

$$\rho(\theta) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(\theta-\mu)^2/(2\sigma^2)}, \quad \theta \in \mathbb{R},\ \mu \in \mathbb{R},\ \sigma > 0.$$

Proposition 15. *We have*

$$s_Z(n,k) = (-1)^{n-k} \binom{k}{n-k} \frac{n!}{k!} \left(\frac{\sigma^2}{2}\right)^{n-k} \mu^{2k-n}, \quad n = k, \dots, 2k, \tag{46}$$

whereas $s_Z(n,k) = 0$, *otherwise.*

In particular, if $\mu = 0$, *then*

$$s_Z(2k,k) = (-1)^k \mathbb{E}(\sigma Z)^{2k}, \quad s_Z(n,k) = 0,\ n \neq 2k, \tag{47}$$

and

$$S_Z(2n,k) = S(n,k)\mathbb{E}(\sigma Z)^{2n}. \tag{48}$$

whenever $k \leq n$. *Otherwise,* $S_Z(n,k) = 0$.

Proof. Formula (46) follows from (6), the fact that

$$\mathcal{K}_Z(z) = \mu z + \frac{\sigma^2}{2} z^2, \tag{49}$$

and some simple computations. On the other hand, suppose that $\mu = 0$. From (36) and (49), we see that

$$\frac{(\mathcal{K}_Z(z))^k}{k!} = \frac{\sigma^{2k}}{2^k} \frac{(2k)!}{k!} \frac{z^{2k}}{(2k)!} = \mathbb{E}(\sigma Z)^{2k} \frac{z^{2k}}{(2k)!},$$

which shows (47). Finally, we have from (17) and (47)

$$S_Z(2n,k) = \sum_{j=k}^{2n} (-1)^j s_Z(2n,j) S(j,k) = (-1)^n s_Z(2n,n) S(n,k) = S(n,k)\mathbb{E}(\sigma Z)^{2n},$$

whenever $k \leq n$. This shows (48) and completes the proof. $\square$

Combinatorially, (46) can be seen as follows. Since in the case of the normal distribution only two cumulants are nonzero, $\kappa_1(Z) = \mu$ and $\kappa_2(Z) = \sigma^2$, only blocks with 1 or 2 elements will have a weight other than zero. Actually, such sets are in one-to-one correspondence with matchings of the complete graph K_n on the vertices $\{1, 2, \dots, n\}$. Such a partition with k blocks has $n-k$ pairs and k singletons. Choose $n-k$ elements in $\binom{n}{n-k}$ ways and then choose for each chosen element a pair from the remaining k elements in $k(k-1)\cdots(k-(n-k)+1)$ ways. However, the order of the elements in the block does not matter, so we have to divide by 2^k. Associating the weight μ with each singleton and σ^2 with each pair, we obtain the factor $(\sigma^2)^{n-k}\mu^{2k-n}$ and we have shown (46).

If $\mu = 0$ only pairs survive, so in this case all blocks in the partition have exactly two elements, which are in one-to-one correspondence with perfect matchings of the complete graph K_{2k}. It is well known that the number of perfect matchings is given by the double factorial ([16, A001147])

$$(2k-1)!! = 1 \cdot 3 \cdot 5 \cdots (2k-1) = \frac{(2j)!}{j!2^j}.$$

We associate with each pair the weight σ^2. We obtain (47) after substituting formula (36).

For seeing (48) combinatorially, note that the combinatorial definition of probabilistic Stirling number of the second kind implies that $S_Z(2n,k)$ counts partitions of $2n$ into k blocks such that each block includes an *even* number of elements (since only moments of even order are nonzero). Such partitions can be constructed by first creating pairs (in $(2n-1)!!$ ways), and then partitioning these n pairs into k blocks (in $S(n,k)$ ways). This is exactly formula (48), after considering the weights and writing again (36) instead of the double factorial.

Acknowledgements

The first author was supported by Research Project DGA E48_23R.
We thank the referees for their careful reading of the manuscript and comments that led to improvements in the presentation of the paper.

Author contributions All authors wrote the main manuscript and worked on the revision.

Funding Open access funding provided by National University of Public Service.

Data Availibility No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no Conflict of interest.

References

[1] Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York (1992)

[2] Adell, J.A.: Probabilistic Stirling numbers of the second kind and applications. J. Theor. Probab. **35**(1), 636–652 (2022). https://doi.org/10.1007/s10959-020-01050-9

[3] Adell, J.A., Leukona, A.: Binomial convolution and transfomrations of Appell polynomials. J. Math. Anal. Appl. **456**(1), 16–33 (2017). https://doi.org/10.1016/j.jmaa.2017.06.077

[4] Adell, J.A., Leukona, A.: A probabilistic generalization of the Stirling numbers of the second kind. J. Number Theory **194**, 335–355 (2019). https://doi.org/10.1016/j.jnt.2018.07.003

[5] Adell, J.A., Leukona, A.: Explicit expressions for a certain subset of Appell polynomials: a probabilistic perspective, Integers 21 Paper No. A60 (2021)

[6] Bényi, B., Mendez, M., Ramírez, J.L., Wakhare, T.: Restricted r-Stirling numbers and their combinatorial applications. Adv. Math. Comput. **348**, 186–205 (2019). https://doi.org/10.1016/j.amc.2018.11.047

[7] Cakić, N.P., Desouky, B.S., Milanović, G.V.: Explicit formulas and combinatorial identities for generalized Stirling numbers. Mediterr. J. Math. **10**(1), 57–72 (2013). https://doi.org/10.1007/s00009-011-0169-x

[8] Comtet, L.: Advanced Combinatorics, the Art of Finite and Infinite Expansions. Reidel, Dordrecht (1974). https://doi.org/10.1007/978-94-010-2196-8

[9] El-Desouky, B.S., Cakić, N.P., Shiha, F.A.: New families of Whitney numbers. Filomat **31**(2), 309–320 (2017)

[10] Flajolet, P., Sedgewick, R.: Analytic Combinatorics. Cambridge University Press, Cambridge (2009). https://doi.org/10.1017/CBO9780511801655

[11] Hsu, L.C., Shiue, P.J.S.: A unified approach to generalized Stirling numbers. Adv. Appl. Math. **20**(3), 366–384 (1998). https://doi.org/10.1006/aama.1998.0586

[12] Kim, D.S., Kim, T., Jang, G.W.: A note on degenerate Stirling numbers of the first kind. Proc. Jangjeon Math. Soc. **21**, 393–404 (2018)

[13] Lah, I.: A new kind of numbers and its applications in the actuarial mathematics. Boletim do Instituto dos Atuários Portugueses **9**, 7–15 (1954)

[14] Luo, Q.M., Srivastrava, H.M.: Some generalizations of the Apostol–Genocchi polynomials and the Stirling numbers of the second kind. Appl. Math. Comput. **217**(2), 5702–5728 (2011). https://doi.org/10.1016/j.amc.2010.12.048

[15] Pinsky, R.G.: Some connections between permutation cycles and Touchard polynomials and between permutations that fix a set and covers of multisets. Electron. Commun. Probab. **22**(17), 1–9 (2017). https://doi.org/10.1214/17-ECP49

[16] Sloane, N.J.A.: The online encyclopedia of integer sequences. https://oeis.org

[17] Speed, T.P.: Cumulants and partition lattices. Austr. J. Stat. **25**(2), 378–388 (1983)

[18] Steutel, F.W., van Harn, K.: Infinite Divisibility of Probability Distributions on the Real Line. CRC Press, Boca Raton (2003). https://doi.org/10.1201/9780203014127

[19] Sun, P.: Product of uniform distribution and Stirling numbers of the first kind. Acta Math. Sin. **21**(6), 1435–1442 (2005). https://doi.org/10.1007/s10114-005-0631-4

[20] von Bahr, B.: On the convergence of moments in the central limit theorem. Ann. Math. Stat. **36**(3), 808–818 (1965). https://doi.org/10.1214/aoms/1177700055

[21] Wang, W.: A determinantal approach to Sheffer sequences. Linear Algebra Appl. **463**, 228–254 (2014). https://doi.org/10.1016/j.laa.2014.09.009

José A. Adell
Departamento de Métodos Estadísticos, Facultad de Ciencias
Universidad de Zaragoza
Zaragoza
Spain
e-mail: adell@unizar.es

Beáta Bényi
Faculty of Water Sciences
University of Public Service
Baja
Hungary
e-mail: benyi.beata@uni-nke.hu

Received: September 21, 2023
Revised: April 12, 2024
Accepted: April 18, 2024