The generating functions of Stirling numbers of the second kind derived probabilistically

George Kesidis * Takis Konstantopoulos † Michael Zazanis ‡

25 June 2018

Abstract

The Stirling numbers $S(n, r)$ of the second kind count the number of partitions of a finite set of size n into r disjoint nonempty subsets. Although the 2-variable integer function $S(n, r)$ cannot be expressed in terms of elementary functions, its ordinary and exponential generating functions with respect to n are both expressible in terms of simple functions. The goal of this short article is to shed a bit more light into these generating functions by deriving them probabilistically. We do this by linking them to Markov chains related to the classical coupon collector problem; coupons are collected in discrete time (ordinary generating function) or in continuous time (exponential generating function). In the process of doing so, we also review the shortest possible combinatorial derivations of these generating functions.

1 Introduction

The Stirling number $S(n, r)$ of the second kind is defined as the number of ways to place n labeled balls in r unlabeled boxes so that no box is empty. [1] The exponential and ordinary generating functions of the sequence $S(n, r)$, $n = 0, 1, \ldots$, are given by

$$\sum_{n=0}^{\infty} S(n,r)\frac{t^n}{n!} = \frac{(e^t-1)^r}{r!}, \tag{1}$$

$$\sum_{n=0}^{\infty} S(n,r)t^n = \frac{1}{(\frac{1}{t}-1)(\frac{1}{t}-2)\cdots(\frac{1}{t}-r)}, \tag{2}$$

respectively. See [3, Ch. 1, (24b), (24c)]. For more information on Stirling numbers, in relation to probability, see [2]. Mechanical demonstrations of these formulas can be based on the recursion

$$S(n, r) = S(n-1, r-1) + rS(n-1, r),$$

*CSE and EE Departments, The Pennsylvania State University; gik2@psu.edu

†Department of Mathematics, University of Liverpool; takiskonst@gmail.com

‡Department of Statistics, Athens University of Economics and Business; zazanis@aueb.gr

Keywords and phrases. Stirling numbers of the second kind; generating function; Markov chain; coupon collector problem

AMS 2010 *subject classification.* Primary 05A18,05-01; secondary 60J10,60J27

[1] Clearly, $S(0, r) = 0$ for all $r > 0$. Note that $S(n, 0) = 0$ if $n > 0$ because if you have 1 or more balls you need at least one box to put them in. But $S(0, 0) = 1$ because it is certainly true that you need no boxes if you have nothing to place inside. Hence $S(n, r)$ is defined for all integer $n, r \geq 0$ and, certainly, $S(n, r) = 0$ if $r > n$.

that is easily derived by a standard argument: for any placement of n labeled balls in r unlabeled boxes either there is a box containing ball n only or not. But mechanical proofs may not be too satisfactory. In this article we provide new probabilistic proofs of the two identities. For the sake of completeness, we also review combinatorial proofs. A proof is called combinatorial if it can be derived as a result of equality of two sets. A proof is called probabilistic if it is a consequence of equality between two random variables. In real life, there is not always a clear distinction between mechanical, combinatorial and probabilistic, but when we have a proof we do have a hunch of what type it is, mostly.

A combinatorial derivation of (1) is not very hard; see Section 2. However, a combinatorial derivation of (2), given in Section 6, is harder; but the probabilistic proof given in Section 5 is simpler and more intuitive.

A few words about notation: if A is a finite set then $|A|$ denotes its cardinality; $\mathscr{P}(A)$ denotes the set of all subsets of A; $\mathscr{P}_r(A)$ is the set of all $B \in \mathscr{P}(A)$ such that $|B| = r$; for a set (or logical clause) π, the indicator symbol $\mathbf{1}_\pi$ equals 1 on the set π (when the clause π is true) and 0 otherwise; if $f : A \to B$ is a function from the set A into the set B then $f^{-1}(b)$ is the set of all $a \in A$ such that $f(a) = b$; for a positive integer n we let $[n] := \{1, 2, \ldots, n\}$; we also let $(n)_r := n(n-1)(n-2)\cdots(n-r+1)$ and $\binom{n}{r} = (n)_r/r!$. If A_1, A_2 are disjoint sets and f_1, f_2 functions with domains A_1, A_2, respectively, we let $f_1 \times f_2$ be the function with domain $A_1 \cup A_2$ which is equal to f_1 on A_1 and f_2 on A_2. The set of surjective (onto) functions $f : A \to B$ is denoted by $\mathrm{sur}(A, B)$.

2 Purely combinatorial derivation of the exponential generating function

Recall the definition of $S(n,r)$ as the number of ways to place n labeled balls into n unlabeled boxes so that no box is empty. Therefore, the number of ways to place n labeled balls in r *labeled* boxes so that no box is empty is

$$W_{n,r} = r!S(n,r) \tag{3}$$

because, to each placement in unlabeled boxes, there correspond r placements in labeled boxes. If we think of boxes as colors, then placing balls into boxes is the same as assigning colors to balls. With r different colors available, suppose that s_j of the balls have color j, $j = 1, \ldots, r$. Then the number of configurations is the multinomial coefficient

$$\binom{n}{s_1, \ldots, s_r} = \frac{n!}{s_1! \cdots s_r!}.$$

Hence

$$r!S(n,r) = \sum_{\substack{s_1,\ldots,s_r \geq 1 \\ s_1+\cdot+s_r=n}} \binom{n}{s_1, \ldots, s_r}. \tag{4}$$

The sum is over all integers $s_1, \ldots, s_r$ that sum up to n and are all positive because each color must be used at least once (no box can be left empty). This last identity is what is behind (1). The rest is mechanical. Multiply both sides by $t^n/n!$ and sum over all n:

$$\begin{aligned}\sum_{n\ge 0} r!S(n,r)\frac{t^n}{n!} &= \sum_{n\ge 0}\sum_{\substack{s_1,\dots,s_r\ge 1\\ s_1+\cdot+s_r=n}} \frac{t^n}{s_1!\cdots s_r!}\\ &= \sum_{s_1,\dots,s_r\ge 1}\frac{t^{s_1+\cdots+s_r}}{s_1!\cdots s_r!}\sum_{n\ge 0}\mathbf{1}_{s_1+\cdots+s_r=n}.\end{aligned}$$

But the last sum is clearly equal to 1 because $s_1+\cdots+s_r$ is a nonnegative integer. We are left with a sum over $s_1,\dots,s_r$ that, since "variables separate", splits into a product:

$$\sum_{s_1,\dots,s_r\ge 1}\frac{t^{s_1+\cdots+s_r}}{s_1!\cdots s_r!} = \sum_{s_1\ge 1}\frac{t^{s_1}}{s_1!}\cdots\sum_{s_r\ge 1}\frac{t^{s_r}}{s_r!} = (e^t-1)^r.$$

The last equality is due to the formula $e^t = 1+\sum_{s\ge 0} t^s/s!$.

3 Probabilistic derivation of the exponential generating function

Consider the classical coupon collecting problem. There are M distinct coupons. Select n of them at random with replacement and let Y_n be the number of distinct coupons selected. Then the distribution of Y_n is

$$\mathbb{P}(Y_n=r) = \frac{1}{M^n}\binom{M}{r} r!S(n,r). \tag{5}$$

The reason is that our sample space is the set of all ordered n-tuples of coupons, a set of cardinality M^n. On the other hand, the subset of the sample space defined by the event $\{Y_n=r\}$ is the set of all ordered n-tuples that use exactly r coupons. Given a subset of the set coupons of cardinality r, there are $r!S(n,r)$ ordered n-tuples that use these r coupons. Since there are $\binom{M}{r}$ ways to select the subset of coupons of size r, the formula follows.

Next observe that the sequence $Y_0 := 0, Y_1, Y_2, \dots$ is a Markov chain. For background on Markov chains see, e.g., [1]. We can think of this as the coupon selection process in discrete time. Let $N(t)$, $t\ge 0$, be an independent Poisson process with rate M and consider the stochastic process $Y_{N(t)}$, $t\ge 0$, a continuous time coupon selection process that is also a Markov chain. We can think of it in a different manner. Split the Poisson process N into M independent copies. Thus, let $N_1,\dots,N_M$ be independent rate-1 Poisson processes. Then $N(t)=\sum_{j=1}^M N_j(t)$, $t\ge 0$, is a rate-M Poisson process. Instead of having a single person collecting coupons at the ticks of the Poisson process N, we have M persons collecting coupons: person j collects coupons of type j at the points of N_j. Let τ_j be the first point of the Poisson process N_j. Then

$$\begin{aligned}\mathbb{P}(Y_{N(t)}=r) &= \mathbb{P}(r \text{ of the } \tau_j\text{'s are} \le t \text{ and the rest are} > t)\\ &= \binom{M}{r}\mathbb{P}(\tau_1,\dots,\tau_r\le t)\mathbb{P}(\tau_{r+1},\dots,\tau_M>t)\\ &= \binom{M}{r}(1-e^{-t})^r(e^{-t})^{M-r}.\end{aligned}$$

On the other hand, using (5) and the fact that N is independent of (Y_n),

$$\begin{aligned}\mathbb{P}(Y_{N(t)} = r) &= \sum_{n=0}^{\infty} \frac{(Mt)^n}{n!} e^{-Mt} \mathbb{P}(Y_n = r) \\ &= \sum_{n=0}^{\infty} \frac{(Mt)^n}{n!} e^{-Mt} \frac{1}{M^n} \binom{M}{r} r! S(n,r).\end{aligned}$$

Equating the last two displays we obtain (1).

4 Coda

Let A, B be two nonempty finite sets. The set B^A of *all functions* from A to B has cardinality $|B|^{|A|}$.

The set of all *injective* (one-to-one) functions from A to B has cardinality

$$|B|(|B|-1)\cdots(|B|-|A|+1),$$

a number that is, obviously, zero if $|A| > |B|$.

The set of all *surjective* (onto) functions from A to B has cardinality

$$|B|!\, S(|A|, |B|)$$

(a number obviously 0 if $|B| > |A|$), because we can think of a surjective function $f : A \to B$ as a placement of labeled balls (the elements of A) into labeled boxes (the elements of B) so that no box is left empty. See (3).

Each function $f : A \to B$ is obviously a surjective function onto some nonempty set, namely the set $f(A)$. Hence

$$\begin{aligned}B^A &= \bigcup_{C \subset B, C \neq \varnothing} \{f \in B^A : f(A) = C\} \\ &= \bigcup_{r=1}^{|B|} \bigcup_{C \subset B, |C| = r} \{f \in B^A : f(A) = C\},\end{aligned}$$

a splitting of B^A according to the image of each of its elements. Both unions are disjoint in the last display. Hence

$$|B|^{|A|} = \sum_{r=1}^{|B|} \sum_{C \subset B, |C| = r} |C|!\, S(|A|, |C|) = \sum_{r=1}^{|B|} r! S(|A|, r) \sum_{C \subset B, |C| = r} 1 = \sum_{r=1}^{|B|} r! S(|A|, r) \binom{|B|}{r}$$

and this verifies that (5) is a probability function.

5 Probabilistic derivation of the ordinary generating function

It suffices to prove (2) for $t = 1/M$, $M = 1, 2, \ldots$. Consider the probability distribution defined by (5) and sum over n:

$$\sum_{n=0}^{\infty} \mathbb{P}(Y_n = r) = \binom{M}{r} r! \sum_{n=0}^{\infty} S(n,r) M^{-n} = (M)_r \sum_{n=0}^{\infty} S(n,r) M^{-n}, \tag{6}$$

where $(M)_r := M(M-1)\cdots(M-r+1)$. As mentioned earlier, the random sequence Y_n, $n = 0, 1, 2, \ldots$, is a Markov chain. We refer to the index n as 'time'. We can easily compute the transition probabilities:

$$\mathbb{P}(Y_{n+1} = r+1 | Y_n = r) = \frac{M-r}{M}, \quad r = 0, 1, \ldots$$

Let $\sigma(r)$ be the total time that the chain remains in state r. Then $\sigma(r)$ is a geometric random variable with parameter $\frac{M-r}{M}$:

$$\mathbb{P}(\sigma(r) = k) = \left(\frac{r}{M}\right)^{k-1} \frac{M-r}{M}, \quad k = 1, 2, \ldots$$

Hence

$$\mathbb{E}\sigma(r) = M/(M-r).$$

But $\sigma(r) = \sum_{n=0}^{\infty} \mathbf{1}_{Y_n=r}$ and so

$$\mathbb{E}\sigma(r) = \sum_{n=0}^{\infty} \mathbb{P}(Y_n = r).$$

Hence

$$\sum_{n=0}^{\infty} S(n,r) M^{-n} = \frac{1}{(M)_r} \frac{M}{M-r} = \frac{1}{(M-1)(M-2)\ldots(M-r)},$$

and this yields (2) for $t = 1/M$.

6 Combinatorial derivation of the ordinary generating function

The combinatorial derivation of the ordinary generating function (2) is harder. It is suggested as an exercise in Stanley's book: see E [3, Ch. 1, p. 46, Exercise 16]. By multiplying both sides of (2) by $r!$ and by expanding the right-hand side we see that proving correctness of (2) is equivalent to showing that

$$r! S(n,r) = \sum_{\substack{k_1,\ldots,k_r \geq 1 \\ k_1+\cdot+k_r=n}} 1^{k_1} 2^{k_2} \cdots r^{k_r}. \tag{7}$$

As noted in (4), the number $r!S(n,r)$ is the cardinality of the set $\mathrm{SUR}([n],[r])$ of surjections from a set of size n onto a set of size r. We will show that the right-hand side of (7) counts the same thing by exhibiting a bijection

$$\Phi : \mathrm{SUR}([n],[r]) \to \bigcup_{\substack{k_1,\ldots,k_r \geq 1 \\ k_1+\cdot+k_r=n}} \left([1]^{k_1} \times [2]^{k_2} \times \cdots \times [r]^{k_r}\right). \tag{8}$$

Let $f : [n] \to [r]$ be a surjective function. Consider the set $[n] = \{1, \ldots, n\}$ in its natural order and let $i_1 < i_2 < \cdots < i_r$ be the points at which the function takes a new value for the first time. That is, let

$$\begin{aligned} i_1 &:= 1 \\ i_{p+1} &:= \min\{i > i_m : f(i) \neq f(i_1), \ldots, f(i_p)\}, \quad p = 1, \ldots, r-1. \end{aligned}$$

Then the numbers $f(i_1), \ldots, f(i_r)$ are just the numbers $1, \ldots, r$ in a different order. Let $L : [r] \to [r]$ be then defined by

$$L(f(i_p)) := p, \quad p = 1, \ldots, r,$$

a relabeling of the values of f. Let

$$K_1 := [i_1, i_2),\ K_2 := [i_2, i_3), \ldots, K_r := [i_r, n]$$

and define $g_p \in [p]^{K_p}$ by

$$g_p(i) := L(f(i)), \quad i \in K_p.$$

Finally let

$$\Phi f := g_1 \times \cdots \times g_r.$$

(See the word on notation at the end of Section 1.) It is easy to see that Φ is a bijection from $\mathrm{SUR}([n], [r])$ onto $\bigcup_{K_1,\ldots,K_r} \left([1]^{K_1} \times [2]^{K_2} \times \cdots \times [r]^{K_r}\right)$, where the union is taken over disjoint intervals $K_1, \ldots, K_r$ such that $K_1 = [1, i_1),\ K_2 = [i_2, i_3), \ldots, K_r = [i_r, n]$ for some $1 = i_1 < i_2 < \cdots < i_r \leq n$, and this union is the right-hand side of (8). Since all sets in the union are pairwise disjoint, we have that the cardinality of it is the right-hand side of (7).

6.1 Discussion

Note that for the combinatorial derivation of the ordinary generating function we used the ordering of $[n]$, something that was not done for the exponential generating function. Combining (4) and (7) we have

$$\sum_{\substack{s_1,\ldots,s_r \geq 1 \\ s_1+\cdot+s_r=n}} \binom{n}{s_1, \ldots, s_r} = \sum_{\substack{s_1,\ldots,s_r \geq 1 \\ s_1+\cdot+s_r=n}} 1^{s_1} 2^{s_2} \cdots r^{s_r}$$

a rather curious identity whose direct proof (that is, showing that the two sides are equal without showing that they are both equal to a known quantity) is unknown (to the authors).

References

[1] Pierre Brémaud (2008). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer-Verlag.

[2] Jim Pitman(2002). *Combinatorial Stochastic Processes:* In Ecole d'Eté de Probabilités de Saint Flour XXXII, ed. by J. Picard. Lecture Notes in Math. Springer-Verlag

[3] Richard P. Stanley (2011). *Enumerative Combinatorics*, Vol. 1. Cambridge University Press.