PART II

COMBINATORICS

6

Stirling Numbers

Author: Thomas A. Dowling, Department of Mathematics, Ohio State University.

Prerequisites: The prerequisites for this chapter are basic counting techniques and equivalence relations. See Sections 5.3 and 8.5 of *Discrete Mathematics and Its Applications*.

Introduction

A set of objects can be classified according to many different criteria, depending on the nature of the objects. For example, we might classify

— accidents according to the day of the week on which they occurred

— people according to their profession, age, sex, or nationality

— college students according to their class or major

— positions in a random sequence of digits according to the digit appearing there

— misprints according to the page on which they occur

— printing jobs according to the printer on which they were done.

In each of these examples, we would have a function from the set of objects to the set of possible levels of the criterion by which the objects are classified. Suppose we want to know the number of different classifications that are possible, perhaps subject to additional restrictions. This would be necessary, for example, in estimating probabilities that a classification satisfied certain conditions. Determining the number of functions between finite sets, based on the sizes of the sets and certain restrictions on the functions, is an enumeration problem considered in Section 7.6 of *Discrete Mathematics and Its Applications*.

It may be the case, however, that the classification need not distinguish the various levels of the criterion, if such a distinction is even possible. There could be many levels, but the actual level on which an object falls is of no concern. Of interest might be how the set of objects is partitioned into disjoint subsets, with two objects being in the same subset whenever they fall at the same level of the criterion. A classification then represents a partition of the set of objects, where two objects belong to the same class of the partition whenever they fall at the same level. In this case, rather than counting functions, partitions of the set of objects would be counted. It is then that the Stirling numbers* become appropriate.

Along with the binomial coefficients $C(n,k)$, the Stirling numbers are of fundamental importance in enumeration theory. While $C(n,k)$ is the number of k-element *subsets* of an n-element set N, the Stirling number $S(n,k)$ represents the number of *partitions* of N into k nonempty subsets. Partitions of N are associated with sets of functions defined on N, and they correspond to equivalence relations on N.

Because of their combinatorial interpretation, the numbers $S(n,k)$ will be met first, despite their name as the Stirling numbers of the *second* kind. After we establish their role in a classification of functions defined on N, we consider their part in relating two sequences of polynomials (as they were originally defined by Stirling). We will then be led to the Stirling numbers $s(n,k)$ of the first kind. We will see why they are referred to as being of the first kind, despite the fact that some of these integers are negative, and thus cannot represent the size of a set.

Occupancy Problems

Many enumeration problems can be formulated as counting the number of distributions of *balls* into *cells* or *urns*. The balls here correspond to the objects and the cells to the levels. These counting problems are generally referred to as **occupancy** problems. Suppose n balls are to be placed in k cells, and assume that any cell could hold all of the balls. The question "In how many ways can this be done?" asks for the number of distributions of n balls to k cells.

As posed, the question lacks sufficient information to determine what is meant by a distribution. It therefore cannot be correctly answered. The num-

* They are named in honor of James Stirling (1692–1770), an eighteenth century associate of Sir Isaac Newton in England.

ber of distributions depends on whether the n balls and/or the k cells are *distinguishable* or *identical*. And if the n balls are distinguishable, whether the order in which the balls are placed in each cell matters. Let us assume that the order of placement does not distinguish two distributions.

Suppose first that both the balls and the cells are distinguishable. Then the collection of balls can be interpreted as an n-element set N and the collection of cells as a k-element set K. A distribution then corresponds to a function $f : N \to K$. The number of distributions is then k^n.

If, however, the balls are identical, but the cells are distinguishable, then a distribution corresponds to an n-combination with repetition from the set K of k cells (see Section 5.5 of *Discrete Mathematics and Its Applications*). The number of balls placed in cell j represents the number of occurrences of element j in the n-combination. The number of distributions in this case is $C(k + n - 1, n)$.

Occupancy problems can have additional restrictions on the number of balls that can be placed in each cell. Suppose, for example, that at most one ball can be placed in each cell. If both the balls and the cells are distinguishable, we would then be seeking the number of *one-to-one* functions $f : N \to K$. Such a function corresponds to an n-permutation from a k-set, so the number is $P(k, n) = k(k - 1) \cdots (k - n + 1)$. If instead there must be at least one ball in each cell, we want the number of *onto* functions $f : N \to K$. Using inclusion-exclusion, that number is equal to

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

(See Section 7.6 of *Discrete Mathematics and Its Applications*.)

Partitions and Stirling Numbers of the Second Kind

An **ordered partition** of N with length k is a k-tuple $(A_1, A_2, \ldots, A_k)$ of disjoint subsets A_i of N with union N. By taking $K = \{1, 2, \cdots, k\}$, we can interpret any function $f : N \to K$ as an ordered partition of N with length k. We let A_i be the subset of N whose elements have image i. Then, since f assigns exactly one image in K to each element of N, the subsets A_i are disjoint and their union is N. Then A_i is nonempty if and only if i is the image of at least one element of N. Thus the function f is onto if and only if all k of the sets A_i are nonempty.

Recall that a **partition** of a set N is a set $P = \{A_i | i \in I\}$ of disjoint, *nonempty* subsets A_i of N that have union N. We shall call these subsets **classes**. An ordered partition of N of a given length may have empty subsets A_i, but the empty set is not allowed as a class in a partition. If it were, then even though the classes are disjoint, they would not necessarily be different sets, since the empty set might be repeated. Given an equivalence relation R on N, the distinct equivalence classes A_i are disjoint with union N. The reflexive property then implies that they are nonempty, so the equivalence classes form a partition of N. Conversely, given a partition $P = \{A_i | i \in I\}$, an equivalence relation R is defined by aRb if and only if there is some class A_i of P containing a and b. Thus, there is a one-to-one correspondence between the set of partitions of N and the set of equivalence relations on N.

A *binary relation* on N is defined as a subset of the set $N \times N$. Since $N \times N$ has n^2 elements, the number of binary relations on N is 2^{n^2}. The number of these that are equivalence relations, which is equal to the number of partitions of N, is called a *Bell number**, and denoted by B_n. For example, the number of ways that ten people can be separated into nonempty groups is the Bell number B_{10}. Two of the people satisfy the corresponding equivalence relation R if and only if they belong to the same group.

Suppose we want instead the number of ways the ten people can be separated into exactly three nonempty groups. That is, how many equivalence relations are there on a ten-element set that have exactly three equivalence classes? More generally, of the B_n equivalence relations, we might ask how many have exactly k equivalence classes? Since this is the number of partitions of N into k classes, we are led to our definition of a Stirling number $S(n,k)$ of the second kind. (Note that the Bell number B_n is then the sum over k of these numbers, that is, $B_n = \sum_{k=1}^{n} S(n,k)$.)

Definition 1 The *Stirling number of the second kind*, $S(n,k)$, is the number of partitions of an n-element set into k classes. □

Example 1 Find $S(4,2)$ and $S(4,3)$.

Solution: Suppose $N = \{a,b,c,d\}$. The 2-class partitions of the 4-element set N are

$$(\{a,b,c\},\{d\}),\quad (\{a,b,d\},\{c\}),\quad (\{a,c,d\},\{b\}),\quad (\{b,c,d\},\{a\})$$

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

It follows that $S(4,2) = 7$.

The 3-class partitions of N are

$$(\{a,b\},\{c\},\{d\}),\quad (\{a,c\},\{b\},\{d\}),\quad (\{a,d\},\{b\},\{c\})$$

---

* We won't pursue the Bell numbers here, but see any of the references and also Section 8.5, Exercise 68, of *Discrete Mathematics and Its Applications* for more about them.

$$(\{b,c\},\{a\},\{d\}),\quad (\{b,d\},\{a\},\{c\}),\quad (\{c,d\},\{a\},\{b\})$$

We therefore have $S(4,3)=6$. □

In occupancy problems, a partition corresponds to a distribution in which the balls are distinguishable, the cells are identical, and there is at least one ball in each cell. The number of distributions of n balls into k cells is therefore $S(n,k)$. If we remove the restriction that there must be at least one ball in each cell, then a distribution corresponds to a partition of the set N with at most k classes. By the sum rule the number of such distributions equals

$$S(n,1)+S(n,2)+\cdots+S(n,k).$$

Since the classes A_i of a partition $P=\{A_i|i\in I\}$ must be disjoint and nonempty, we have $S(n,k)=0$ for $k>n$. Further, since $(\{S\})$ itself is the only partition of N with one class, $S(n,1)=1$ for $n\geq 1$. Given these initial conditions, together with $S(0,0)=1$, $S(n,0)=0$ for $n\geq 1$, the Stirling numbers $S(n,k)$ can be recursively computed using the following theorem.

Theorem 1 Let n and k be positive integers. Then

$$S(n+1,k)=S(n,k-1)+k\,S(n,k). \tag{1}$$

Proof: We give a combinatorial proof. Let S be an $(n+1)$-element set. Fix $a\in S$, and let $S'=S-\{a\}$ be the n-element set obtained by removing a from S. The $S(n+1,k)$ partitions of S into k classes can each be uniquely obtained from either

(i) a partition P' of S' with $k-1$ classes by adding a singleton class $\{a\}$, or

(ii) a partition P' of S' with k classes by first selecting one of the k classes of P' and then adding a to that class.

Since the cases are exclusive, we obtain $S(n,k)$ by the sum rule by adding the number $S(n,k-1)$ of partitions of (i) to the number $k\,S(n,k)$ of partitions in (ii). ■

Example 2 Let P be one of the partitions of $N=\{a,b,c,d\}$ in Example 1. Then P is obtained from the $S(3,1)=1$ one-class partition $(\{a,b,c\})$ of $N'=N-\{d\}=\{a,b,c\}$ by adding $\{d\}$ as a second class, or from one of the $S(3,2)=3$ two-class partitions

$$(\{a,b\},\{c\}),\quad (\{a,c\},\{b\}),\quad (\{b,c\},\{a\})$$

of N' by choosing one of these partitions, then choosing one of its two classes, and finally by adding d to the class chosen. □

Example 3 Use Theorem 1 to find $S(5,3)$.

Solution: We have $S(4,2) = 7$ and $S(4,3) = 6$ from Example 1. Then we find by Theorem 1 that $S(5,3) = S(4,2) + 3\ S(4,3) = 7 + 3 \cdot 6 = 25$. □

The Pascal recursion $C(n+1,k) = C(n,k-1) + C(n,k)$ used to compute the binomial coefficients has a similar form, but the coefficient in the second term of the sum is 1 for $C(n,k)$ and k for $S(n,k)$.

Using the recurrence relation in Theorem 1, we can obtain each of the Stirling numbers $S(n,k)$ from the two Stirling numbers $S(n-1,k-1)$, $S(n-1,k)$ previously found. These are given in Table 1 for $1 \le k \le n \le 6$.

n \ k	**1**	**2**	**3**	**4**	**5**	**6**
1	1					
2	1	1				
3	1	3	1			
4	1	7	6	1		
5	1	15	25	10	1	
6	1	31	90	65	15	1

Table 1. Stirling numbers of the second kind, S(n,k).

Inverse Partitions of Functions

Recall that a function $f : N \to K$ is *onto* if every element $b \in K$ is an image of some element $a \in N$, so $b = f(a)$. If we denote the set of images by $f(N) = \{f(a) | a \in N\}$, then f is onto if and only if $f(N) = K$.

Example 4 Suppose $N = \{a, b, c, d\}$, $K = \{1, 2, 3\}$, and the function f is defined by $f(a) = f(b) = f(d) = 2$, $f(c) = 1$. Then $f(N) = \{1, 2\} \ne K$, so f is not onto. □

Every function $f : N \to K$ determines an equivalence relation on N under which two elements of N are related if and only if they have the same image under f. Since the equivalence classes form a partition of N, we make the following definition.

Definition 2 The *inverse partition* defined by a function $f : N \to K$ is the partition $P(f)$ of N into equivalence classes with respect to the the equivalence relation R_f defined by $aR_f b$ if and only if $f(a) = f(b)$. □

Given the function f, every element $b \in f(N)$ determines an equivalence class of R_f. Suppose we denote by $f^{-1}(b)$ the set of elements in N that have b as their image*. Then $f^{-1}(b)$ is an equivalence class, called the **inverse image** of b. It follows that

$$a \in f^{-1}(b) \text{ if and only if } b = f(a).$$

Thus, the equivalence classes $f^{-1}(b)$ for $b \in f(N)$ form the inverse partition $P(f)$ of N, and the number of equivalence classes is the size of the set $f(N)$. Therefore $f : N \to K$ is onto if and only if $P(f)$ has k classes.

Example 5 Find the inverse partition of N defined by the function f of Example 4.

Solution: For the function f in Example 4, the set of elements of N with image 1 is $f^{-1}(1) = \{a, b, d\}$, while the set of elements with image 2 is $f^{-1}(2) = \{c\}$. Since no elements have 3 as an image, the set $f^{-1}(3)$ is empty, so it is not an equivalence class. Thus the inverse partition of N defined by f is the 2-class partition $P(f) = (\{a, b, d\}, \{c\})$. □

An Identity for Stirling Numbers

We can classify the k^n distributions of n distinguishable balls into k distinguishable cells by the number j of nonempty cells. Such a distribution is determined by a j-class partition P of the set of balls together with an injective function φ from the j classes of P to the k cells. The classification obtained in this way leads to an identity involving the Stirling numbers of the second kind. Denote by $F(N, K)$ the set of functions $f : N \to K$. Recall that the size of $F(N, K)$ is k^n. Let us classify the functions $f \in F(N, K)$ according to the size of their image sets $f(N)$, or equivalently, according to the number of classes in their inverse partitions $P(f)$.

* What we have denoted by f^{-1} is actually a function from K to the power set $P(N)$ of N, and not the inverse function of f. The function $f : N \to K$ has an inverse function from K to N if and only if it is a one-to-one correspondence.

Given a j-class partition $P = \{A_1, , A_2, \ldots, A_j\}$ of N, consider the set of functions $f \in F(N, K)$ for which $P(f) = P$. Two elements in N can have the same image if and only if they belong to the same A_i. Hence a function f has $P(f) = P$ if and only if there is a one-to-one function $\varphi : \{A_1, A_2, \ldots, A_j\} \to K$ such that $f(a) = \varphi(A_i)$ for all i and all $a \in A_i$. The number of one-to-one functions from a j-set to a k-set is the **falling factorial**

$$(k)_j = P(k, j) = k(k-1)\cdots(k-j+1).$$

Hence, for any j-class partition P, the number of functions $f \in F(N, K)$ such that $P(f) = P$ is $(k)_j$. Since there are $S(n, j)$ j-class partitions of N, by the product rule there are $S(n, j)(k)_j$ functions $f \in F(N, K)$ such that $f(N)$ has size j. Summing over j gives us

$$k^n = \sum_{j=1}^{k} S(n, j)(k)_j. \tag{2}$$

The jth term in the sum is the number of functions f for which $|f(N)| = j$.

We noted earlier that the onto functions correspond to distributions in the occupancy problem where the balls and the cells are distinguishable and there must be at least one ball in each cell. If we remove the restriction that there must be at least one ball in each cell, then a distribution corresponds to a function $f : N \to K$. Classification of the k^n distributions according to the number of nonempty cells gives (2). If we consider the term where $j = k$ in (2), and note that $(k)_k = k!$, we have the following theorem.

Theorem 2 The number of functions from a set with n elements onto a set with k elements is $S(n, k)k!$. ■

The inclusion-exclusion principle can be used to show that the number of onto functions from a set with n elements to a set with k elements is

$$\sum_{i=0}^{k} (-1)^i C(k, i)(k-i)^n. \tag{3}$$

By Theorem 2, we may equate (3) to $S(n, k)k!$, and we obtain

$$S(n, k) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i C(k, i)(k-i)^n. \tag{4}$$

Example 6 Let $n = 5$, $k = 3$. Then if the 3^5 functions $f : N \to K$ are classified according to the number of classes in their inverse partitions, we obtain

$$3^5 = \sum_{j=1}^{3} S(5,j)(3)_j = 1 \cdot 3 + 15 \cdot 6 + 25 \cdot 6 = 243,$$

which is of course easily verified. Theorem 2 shows that the number of these functions that are onto is $S(5,3)3! = 25 \cdot 6 = 150$. Alternatively, by the inclusion-exclusion expression (3), we have

$$\begin{aligned}\sum_{i=0}^{3}(-1)^k C(3,i)(3-i)^5 &= 3^5 - 3 \cdot 2^5 + 3 \cdot 1^5 \\ &= 243 - 96 + 3 = 150.\end{aligned}$$

□

Before turning to polynomials in a real variable x, we need to write the summation in (2) in a different form. Since $S(n,j) = 0$ for $j > n$ and $(k)_j = 0$ for an integer $j > k$, the term $S(n,j)(k)_j$ in (2) can be nonzero only when $j \leq n$ and $j \leq k$. Thus we can replace the upper limit k in (2) by n, which gives

$$k^n = \sum_{j=1}^{n} S(n,j)(k)_j. \qquad (2')$$

Stirling Numbers and Polynomials

Recall that a *polynomial of degree* n in a real variable x is a function $p(x)$ of the form

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0,$$

where $a_n, a_{n-1}, \ldots, a_0$ are constants with $a_n \neq 0$. Suppose we extend the definition of the falling factorial $(k)_n$ from the set $\mathcal{N} \times \mathcal{N}$ to the set $\mathcal{R} \times \mathcal{N}$ by defining the function $(x)_n$ for any *real number* x and any nonnegative integer n to be $(x)_n = 1$ when $n = 0$ and $(x)_n = x(x-1)\cdots(x-n+1)$ when $n \geq 1$. Equivalently, define $(x)_n$ recursively by $(x)_0 = 1$ and $(x)_n = (x-n+1)(x)_{n-1}$. As the product of n linear factors, $(x)_n$ is a polynomial of degree n. We can then replace k by x in $(2')$ to obtain

$$x^n = \sum_{j=1}^{n} S(n,j)(x)_j. \qquad (5)$$

Subtracting the right-hand side from each side gives

$$x^n - \sum_{j=1}^{n} S(n,j)(x)_j = 0. \qquad (6)$$

But $(x)_j = x(x-1)\cdots(x-j+1)$ is a polynomial of degree j, for $1 \le j \le n$. Hence the left-hand side of (6) is a polynomial $p(x)$ of degree at most n. But since $p(x) = 0$ for every positive integer x and since a polynomial of degree n has at most n real roots, $p(x)$ must be identically zero. Then we obtain for every real number x

$$x^n = \sum_{j=1}^{n} S(n,j)(x)_j. \tag{7}$$

Equation (7) represents (2′) extended from positive integers k to arbitrary real numbers x. Since k does not appear in (7), we may replace the index variable j by k, rewriting (7) as

$$x^n = \sum_{k=1}^{n} S(n,k)(x)_k. \tag{7′}$$

Equation (7′) represents (7) in a form convenient for comparison with (10) in the next section.

Example 7 Let $n = 3$. Then $(x)_1 = x$, $(x)_2 = x(x-1) = x^2 - x$, and $(x)_3 = x(x-1)(x-2) = x^3 - 3x^2 + 2x$. Substituting these into the right-hand side of (5) gives

$$\begin{aligned} S(3,1)(x)_1 + {} & S(3,2)(x)_2 + S(3,3)(x)_3 \\ &= 1\cdot x + 3\cdot(x^2 - x) + 1\cdot(x^3 - 3x^2 + 2x) \\ &= x^3. \end{aligned}$$

□

Stirling Numbers of the First Kind

There is an interpretation of (7) that will be understood more readily by readers who are familiar with linear algebra. The set of polynomials $p(x)$ with real coefficients (or with coefficients in any field) forms a *vector space**. A *fundamental basis* of this vector space is a sequence of polynomials $\{p_k(x) | k = 0, 1, \ldots\}$ such that $p_k(x)$ has degree k. Every polynomial can be uniquely expressed as a linear combination of the polynomials in a fundamental basis. Two important fundamental bases are the *standard basis* $\{x^k | k = 0, 1, \ldots\}$ and the *falling factorial basis* $\{(x)_k | k = 0, 1, \ldots\}$. The standard basis is in fact the one used in the definition of a polynomial in x, since a polynomial is defined as a linear combination of the powers of x. Then Equation (7) expresses the basis $\{x^k\}$ in terms of the basis $\{(x)_k\}$.

* This vector space has infinite (but countable) dimension.

In fact, (7) was Stirling's original definition of the numbers $S(n,k)$ as the coefficients in changing from the falling factorial basis $\{(x)_k\}$ to the standard basis $\{x^k\}$. These numbers were said to be of the *second kind* since they change a basis *to* the standard basis. More commonly, we would be changing *from* the standard basis to another basis, and the coefficients would be of the *first kind*. When the other basis is the sequence $\{(x)_k\}$ of falling factorial polynomials, we encounter the Stirling numbers of the first kind.

To illustrate these ideas, let us first suppose we want to express the basis $\{(x-1)^k\}$ in terms of the standard basis $\{x^k\}$. Using the Binomial Theorem, we can express the polynomial $(x-1)^n$ as

$$(x-1)^n = \sum_{k=0}^{n} (-1)^{n-k} C(n,k) x^k. \tag{8}$$

The coefficients $(-1)^{n-k}C(n,k)$ in (8), that change the standard basis to the other basis (and which are the coefficients when the product $(x-1)^n$ of n identical factors is expanded), by analogy would be called binomial coefficients of the first kind. On the other hand, if we express x as $(x-1)+1$ and apply the Binomial Theorem, we obtain

$$x^n = ((x-1)+1)^n = \sum_{k=0}^{n} C(n,k)(x-1)^k. \tag{9}$$

Thus, the binomial coefficients $C(n,k)$ are used in (9) to change from the other basis to the standard basis, and would therefore be referred to as the binomial coefficients of the second kind.

The coefficients needed to change the falling factorial basis $\{(x)_k\}$ to the standard basis $\{x^k\}$ can be obtained by simply expanding the product $(x)_n = x(x-1)\ldots(x-n+1)$.

Example 8 Write $(x)_3$ in terms of the standard basis $\{x^k\}$.

Solution: Let $n=3$. Then expanding the product $(x)_3$ and reversing the order of the terms gives

$$\begin{aligned}(x)_3 &= x(x-1)(x-2)\\ &= x^3 - 3x^2 + 2x\\ &= 2x - 3x^2 + x^3.\end{aligned}$$

□

Let us now define the Stirling numbers of the first kind.

Definition 3 The *Stirling numbers of the first kind* $s(n,k)$ are the numbers satisfying

$$(x)_n = \sum_{k=0}^{n} s(n,k)x^k. \tag{10}$$

That (10) uniquely determines the coefficients $s(n,k)$ follows either from the definition of multiplication of polynomials or from the fact that every polynomial can be uniquely expressed as a linear combination of the polynomials in a basis.

Example 9 Find the Stirling numbers which express $(x)_3$ in terms of the polynomials $1, x, x^2$.

Solution: Using the definition of $s(n,k)$ and Example 8, we have

$$\begin{aligned}(x)_3 &= s(3,0)\cdot 1 + s(3,1)x + s(3,2)x^2 + s(3,3)x^3\\ &= 2x - 3x^2 + x^3.\end{aligned}$$

Hence, the Stirling numbers of the first kind which express $(x)_3$ in terms of the polynomials $1, x, x^2$ and x^3 in the standard basis are $s(3,0) = 0$, $s(3,1) = 2$, $s(3,2) = -3$, $s(3,3) = 1$. □

Note that in Example 9 some of the Stirling numbers $s(n,k)$ of the first kind can be negative. Hence, unlike the Stirling numbers of the second kind, those of the first kind cannot represent the sizes of sets. (But their absolute values do. See Exercise 16.)

Consider for a fixed n the sequence of numbers $s(n,k)$. These are zero when $n > k$. We can determine their sum over $1 \leq k \leq n$ immediately from (10). The polynomial $(x)_n$ vanishes whenever x is a positive integer and $n > x$. Thus, on setting $x = 1$ in (10), we get

$$\sum_{k=1}^{n} s(n,k) = 0 \quad \text{for} \quad n \geq 2. \tag{11}$$

There is a recurrence relation for $s(n,k)$ analogous to the recurrence relation in Theorem 1 for $S(n,k)$, but we will obtain this one by a method different from the combinatorial argument given there. We start by finding the boundary values where $k = 0$. Using the convention that $(x)_0 = 1$ if $x = 0$ and 0 otherwise, we have $s(0,0) = 1$. And since x is a factor of $(x)_n$ for $n \geq 1$, the constant term is 0, so $s(n,0)$ for $n \geq 1$.

The analogue of Theorem 1 is then given by the following theorem.

Theorem 3 Let n and k be positive integers. Then

$$s(n+1,k) = s(n,k-1) - n\,s(n,k). \tag{12}$$

Proof: We make use of the fact that $(x)_{n+1} = x(x-1)\cdots(x-n+1)(x-n) = (x-n)(x)_n$. Then by (10),

$$(x)_{n+1} = \sum_{k=0}^{n+1} s(n+1,k)x^k. \tag{13}$$

But

$$\begin{aligned}(x)_{n+1} &= (x-n)(x)_n \\ &= \sum_{j=0}^{n} s(n,j)x^{j+1} - n\sum_{j=0}^{n} s(n,j)x^j.\end{aligned} \tag{14}$$

When we equate the coefficients of x^k in (13) and (14), we find that $s(n+1,k) = s(n,k-1) - n\,s(n,k)$. ■

Example 10 Use recurrence relation (12) to find $s(5,3)$.

Solution: We see from Example 9 that $s(3,1) = 2$, $s(3,2) = -3$, and $s(3,3) = 1$. Thus, by (12), we find $s(4,2) = 2-3\cdot(-3) = 11$ and $s(4,3) = -3-3\cdot 1 = -6$. We then obtain $s(5,3) = 11 - 4\cdot(-6) = 35$. □

The values of $s(n,k)$ for $1 \le k \le n \le 6$ shown in Table 2 were computed using (12).

n \ k	**1**	**2**	**3**	**4**	**5**	**6**
1	1					
2	-1	1				
3	2	-3	1			
4	-6	11	-6	1		
5	24	-50	35	-10	1	
6	-120	274	-225	85	-15	1

Table 2. Stirling numbers of the first kind, s(n,k).

Comments

When we observe the tables of values of the Stirling numbers for $1 \le k \le n \le 6$, we notice that the numbers in certain columns or diagonals come from wellknown sequences. We will examine some of these. Proofs will be left for the exercises.

First consider the numbers $S(n,2)$ in the second column of Table 1. For $2 \le n \le 6$ these are $1, 3, 7, 15, 31$. It appears that $S(n,2)$ has the form $2^{n-1} - 1$ (Exercise 14). The numbers $S(n, n-1)$ just below the main diagonal in Table 1 are $1, 3, 6, 10, 15$ for $2 \le n \le 6$, suggesting (Exercise 13) that $S(n, n-1)$ is the **triangular number**

$$C(n,2) = \frac{1}{2}n(n-1).$$

Next consider Table 2, which gives the values $s(n,k)$ of the Stirling numbers of the first kind. We can easily verify that the row sums are zero for $2 \le n \le 6$, as given for all $n \ge 2$ by (9). Observe that the numbers $s(n,1)$ in the first column of Table 2 satisfy $s(n,1) = (-1)^{n-1}(n-1)!$ for $1 \le n \le 6$. This in fact holds for all n (Exercise 14). The numbers $s(n, n-1)$ just below the main diagonal in Table 2 are the negatives of the Stirling numbers $S(n, n-1)$ of the second kind, which are apparently the triangular numbers $C(n,2)$ (Exercise 13).

Note also that the numbers $s(n,k)$ alternate in sign in both rows and columns. It is not difficult to prove (Exercise 16) that the sign of $s(n,k)$ is $(-1)^{n-k}$, so that the absolute value of $s(n,k)$ is

$$t(n,k) = (-1)^{n-k}s(n,k).$$

These integers $t(n,k)$ are the **signless Stirling numbers of the first kind.** As nonnegative integers, these numbers have a combinatorial interpretation (Exercise 18).

Suggested Readings

A basic treatment of the Stirling numbers is in reference [2]. A great deal of information about the generating functions of the Stirling numbers and further properties of these numbers can be found in the more advanced books [1], [3], [4], and [5].

1. M. Aigner, *Combinatorial Theory*, Springer, New York, 1997.

2. K. Bogart, *Introductory Combinatorics*, Brooks/Cole, Belmont, CA, 2000.

3. L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Springer, New York, 1974.

4. R. Stanley, *Enumerative Combinatorics*, Volume 1, Cambridge University Press, New York, 2000.

5. H. Wilf, *generatingfunctionology*, Third Edition, A K Peters, Wellesley, MA, 2005.

Exercises

1. For $n = 7$ and $1 \leq k \leq 7$, find the Stirling numbers $S(n,k)$ and $s(n,k)$.

2. Express $(x)_4$ as a polynomial by expanding the product $x(x-1)(x-2)(x-3)$.

3. Express the polynomial x^4 as a linear combination of the falling factorial polynomials $(x)_k$.

4. Find the number of onto functions from an eight-element set to a five-element set.

5. Let $n = 3$ and $x = 3$. Classify the $x^n = 27$ functions f from $N = \{1, 2, 3\}$ to $X = \{a, b, c\}$ according to the size of the image set $f(N)$ to verify that there are $S(3, j)(3)_j$ such functions f with $|f(N)| = j$, for $1 \leq j \leq 3$.

6. In how many ways can seven distinguishable balls be placed in four identical boxes so that
a) there is at least one ball in each box?
b) some box(es) may be empty?

7. Suppose each of a group of six people is assigned one of three tasks at random. Find the probability (to three decimal places) that
a) task 3 is not assigned.
b) all three tasks are assigned.
c) exactly two of the tasks are assigned.

8. In how many ways can a set of 12 people be divided into three (nonempty) subsets.

9. Suppose a k-sided die is rolled until each of the numbers $1, 2, \ldots, k$ have appeared at least once, at which time the rolls are stopped. Give an expression for the number of possible sequences of n rolls that satisfy the condition.

10. A computer is programmed to produce a random sequence of n digits.
a) How many possible sequences are there?
b) How many of these sequences have each of the 10 digits appearing?

11. A pool table has four corner pockets and two center pockets. There is one white ball (the cue ball) and 15 colored balls numbered $1, 2, \ldots, 15$. In a game each of the numbered balls is sunk into a pocket (and remains there) after being struck by the cue ball, which the player has propelled with a cue stick. Thus a game produces a distribution of the numbered balls in the set of pockets.
a) Assuming the pockets are distinguishable, how many distributions are there?

b) Suppose we assume the corner pockets are identical and the center pockets are identical, but that a corner pocket is distinguishable from a center pocket. Give an expression for the number of distributions in which all of the numbered balls are sunk in corner pockets or they are all sunk in center pockets.

12. Let n and k be positive integers with $n+1 \geq k$. Prove that $S(n+1,k) = \sum_{j=0}^{n} C(n,j)S(n-j,k-1)$.

13. Prove that
a) $S(n,n-1) = C(n,2)$.
b) $s(n,n-1) = -C(n,2)$.

14. Prove that
a) $S(n,2) = 2^{n-1} - 1$.
b) $s(n,1) = (-1)^{n-1}(n-1)!$.

15. Let m and n be nonnegative integers with $n \leq m$. The *Kronecker delta function* $\delta(m,n)$ is equal to 1 if $m = n$ and 0 otherwise. Prove that

$$\sum_{k=n}^{m} S(m,k)s(k,n) = \delta(m,n).$$

In Exercises 16–20, let n and k be positive integers with $k \leq n$.

16. Prove that the sign of the Stirling number of the first kind is $(-1)^{n-k}$. (Thus the signless Stirling number of the first kind $t(n,k) = |s(n,k)|$ is equal to $(-1)^{n-k}s(n,k)$.)

17. Find a recurrence relation for the signless Stirling numbers $t(n,k)$ of the first kind similar to the recurrence relation (12) satisfied by the Stirling numbers $s(n,k)$ of the first kind.

18. A **cyclic permutation** (or **cycle**) of a set A is a permutation that cyclically permutes the elements of A. For example, the permutation σ, with $\sigma(1) = 2$, $\sigma(2) = 4$, $\sigma(3) = 1$, $\sigma(4) = 5$, $\sigma(5) = 3$, is cyclic, and denoted by (12453) or any of its cyclic equivalents, such as (24531) or (31245). It can be shown that every permutation of a set N can be expressed as a product of cycles that cyclically permute disjoint subsets of N, unique except for the order of the cycles. Thus, for example, (153)(24) is the **cyclic decomposition** of the permutation σ given by $\sigma(1) = 5$, $\sigma(2) = 4$, $\sigma(3) = 1$, $\sigma(4) = 2$, $\sigma(5) = 3$. Prove that $t(n,k)$ is the number of permutations of an n-element set that have exactly k cycles in their cyclic decomposition. *Hint*: Use Exercise 17.

19. Suppose a computer is programmed to produce a random permutation of n different characters.

a) Give an expression for the probability that the permutation will have exactly k cycles.

b) Calculate the probabilities with $n = 8$ that the permutation will have two cycles and that it will have three cycles. Which is more likely?

$\star$**20.** Let $a(n, k)$ be the sum of the products $1^{c_1}2^{c_2} \cdots k^{c_k}$, taken over all k-tuples $(c_1, c_2, \ldots, c_k)$ such that c_i is a nonnegative integer for $i = 1, 2, \ldots, k$ and the sum of the c_is is $n-k$. Prove that $a(n, k)$ satisfies the initial conditions and the recurrence relation (1) for the Stirling numbers $S(n, k)$ of the second kind. (This shows that $S(n, k) = a(n, k)$.)

21. (Requires calculus.) Let k be a fixed positive integer. It can be shown that the ratios $S(n, k-1)/S(n, k)$ and $s(n, k-1)/s(n, k)$ both approach 0 as a limit as n approaches ∞. Use these facts to find

a) $\lim_{n\to\infty} S(n+1, k)/S(n, k)$.

b) $\lim_{n\to\infty} s(n+1, k)/s(n, k)$.

$\star$**22.** (Requires calculus.) The *exponential generating function* of a sequence (a_n) is defined to be the power series

$$A^*(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}.$$

Let k be a fixed positive integer and let $a_n = S(n, k)$. Prove that

$$A^*(x) = \frac{(e^x - 1)^k}{k!}.$$

$\star$**23.** Find the expansion of the polynomial $P_k(x) = (1-x)(1-2x)\ldots(1-kx)$.

Computer Projects

1. Write a computer program that uses the recurrence relation (1) to compute Stirling numbers of the second kind.

2. Write a computer program that uses the recurrence relation (12) to compute Stirling numbers of the first kind.