3 Basic combinatorial techniques

In this chapter we bring together some systematic methods for solving counting problems.

3.1 Inclusion-exclusion

Let X be a set and A and B subsets of X. Our starting point is the obvious numerical identity,

$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{1}$$

In other words, to count the elements in $A \cup B$, it suffices to count the elements of A and B separately, and then subtract elements in the intersection (which have been counted twice).

If A, B, C are subsets of X then (1) generalizes to,

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|. \tag{2}$$

In this case, to get the count right, we subtract the elements counted by any two subsets then add on the elements in all three subsets (which have been added three times, but also subtracted three times). This process is called the inclusion-Exclusion principle, and it generalizes to n subsets $A_1, A_2, \ldots, A_n$ of the set X. That is

$$\begin{aligned} |A_1 \cup A_2 \cup \cdots \cup A_n| = &|A_1| + |A_2| + \cdots + |A_n| \\ &- |A_1 \cap A_2| - |A_1 \cap A_3| - \cdots |A_{n-1} \cap A_n| \\ &+ |A_1 \cap A_2 \cap A_3| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n| - \cdots \\ &(-1)^{n+1}|A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned} \tag{3}$$

Note. Identity (3) follows easily from (1) by induction.

Example 1. How many integers between 1 and 1000 are not divisible by any of 2,3,11 or 13?

This is easy once we realize that the number is

$$|X| - |A_2 \cup A_3 \cup A_{11} \cup A_{13}|,$$

where X is the set $\{1, 2, 3, \ldots, 1000\}$ and A_p is the subset of X containing those integers *divisible* by p. Now,

$$|A_2| = \left[\frac{1000}{2}\right] = 500, \quad |A_3| = \left[\frac{1000}{3}\right] = 333, \quad |A_{11}| = \left[\frac{1000}{11}\right] = 90,$$

$$|A_{13}| = \left[\frac{1000}{13}\right] = 76, \quad |A_2 \cap A_3| = \left[\frac{1000}{6}\right] = 166, \quad |A_2 \cap A_{11}| = 45,$$

$$|A_2 \cap A_{13}| = 38, \quad |A_3 \cap A_{11}| = 30, \quad |A_2 \cap A_{13}| = 25, \quad |A_{11} \cap A_{13}| = 6,$$

$$|A_2 \cap A_3 \cap A_{11}| = 15, \quad |A_2 \cap A_3 \cap A_{13}| = 12, \quad |A_2 \cap A_{11} \cap A_{13}| = 3,$$

$$|A_3 \cap A_{11} \cap A_{13}| = 2, \quad |A_2 \cap A_3 \cap A_{11} \cap A_{13}| = 1,$$

where $[x]$ is the integer part of $x \in \mathbb{R}$.

Hence, by the Inclusion-Exclusion principle,

$$\begin{aligned} |X| - |A_2 \cup A_3 \cup A_{11} \cup A_{13}| &= 1000 - \begin{pmatrix} 500 + 33 + 90 + 76 - 166 - 45 - 38 \\ -30 - 25 - 6 + 15 + 12 + 3 + 2 - 1 \end{pmatrix} \\ &= 1000 - 720 = 280. \end{aligned}$$

Example 2. (Euler's ϕ-function.) Two integers are *relatively prime* if they have no common divisors other than 1. For example, 15 and 26 are realtively prime, but 15 and 24 are not (3|15, 3|24).

The Euler ϕ-function is defined by letting $\phi(n)$ be the number of positive integers less than n that are relatively prime to n. [e.g. $\phi(15) = 8$ since 1, 2, 4, 7, 8, 11, 13, 14 are less than 15 and relatively prime to 15.]

We can use the Inclusion-Exclusion Principle to find a conveient formula for $\phi(n)$. Let $n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$ be the unique decomposition of n as a product of prime powers, where $a_1, a_2, \ldots, a_k$ are integers > 0.

Let $N = \{1, 2, 3, \ldots, n\}$ and let A_{p_i} be the subset of N containing those elements which are divisible by p. Then $|A_{p_i}| = \dfrac{n}{p_i}$, $|A_{p_i} \cap A_{p_j}| = \dfrac{n}{p_i p_i}$, ...

Hence, by the Inclusion-Exclusion Principle,

$$\begin{aligned}
\phi(n) &= n - \frac{n}{p_1} - \frac{n}{p_2} - \cdots - \frac{n}{p_k} + \frac{n}{p_1 p_2} + \cdots + \frac{n}{p_{k-1} p_k} \\
&\quad - \frac{n}{p_1 p_2 p_3} - \cdots - \frac{n}{p_{k-2} p_{k-1} p_k} + \cdots + (-1)^k \frac{n}{p_1 p_2 \cdots p_k} \\
&= n\Big(1 - \frac{1}{p_1}\Big)\Big(1 - \frac{1}{p_2}\Big) \cdots \Big(1 - \frac{1}{p_k}\Big). \qquad (4)
\end{aligned}$$

Check. $\phi(15) = 15(1 - \frac{1}{3})(1 - \frac{1}{5}) = 8$.

Example 3. (Derangements.) On a rainy day, n students leave their umbrellas (which are indistiguishable) outside their examination room. What is the probability that no student collects the correct umbrella when they finish the examination?

Number the students and their umbrellas, $1, 2, 3, \ldots, n$. Each permutation of $\{1, 2, \ldots, n\}$ corresponds to a possible assignment of umbrellas to students. The permutations which fix no numbers correspond to the cases in which no student collects the correct umbrella. These cases are called *derangements*.

To determine the number of derangements, D_n, we let X be the set of all permutations of $\{1, 2, \ldots, n\}$, and A_i the subset of X of those permutations which fix the number i.

The total number of derangements is then,

$$D_n = |X| - |A_1 \cup A_2 \cup \cdots \cup A_n|,$$

which we can calculate using the Inclusion-Exclusion Principle.

$$\begin{aligned}
&|X| = n!, \qquad |A_i| = (n-1)!, \\
&|A_{i_1} \cap A_{i_2} \cap \cdots A_{i_k}| = (n-k)!, \quad \text{for } k = 1, 2, \ldots, n,
\end{aligned}$$

Hence,

$$\begin{aligned}
D_n &= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \cdots + (-1)^n \binom{n}{n}(n-n)! \\
&= n!\Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}\Big). \qquad (5)
\end{aligned}$$

The probability that no student collects the correct umbrella is thus

$$\frac{D_n}{n!} = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}.$$

As n tends to ∞, the probability tends to $\frac{1}{e} \approx 0.368$ (or 36.8%). Even for small n, the probability is close to this, since the error is determined by $\frac{1}{n!}$, the last term of the series.

Example 4. **(Stirling numbers.)** Let us count the onto functions $f\colon N \to M$, where $|N| = n$ and $M = [m] = \{1, 2, \ldots, m\}$. On the one hand, we know from Chapter 1 (12), that the number is $m!S(n, m)$. We can also count the onto functions using the Inclusion-Exclusion principle.

Let X be the set of all function $f\colon N \to M$ and for each i, A_i the subset of X of those functions whose image does not include i. Then,

$$|A_i| = (m-1)^n, \qquad \text{and} \qquad |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = (m-p)^n.$$

Hence,

$$\begin{aligned} |X| - |A_1 \cup A_2 \cup \cdots \cup A_m| &= m^n - \binom{m}{1}(m-1)^n + \binom{m}{2}(m-2)^n - \cdots \\ &\qquad + (-1)^{m-2}\binom{m}{m-2}2^n + (-1)^{m-1}\binom{m}{m-1} \\ &= \sum_{k=0}^{m} (-1)^{m-+k}\binom{m}{k}(m-k)^n. \end{aligned}$$

Hence, letting $j = m - k$, we obtain,

$$m!S(n, m) = \sum_{j=0}^{m} (-1)^j \binom{m}{j} j^m.$$

3.1.1 Exercises with Answers

1. Find the number of integers between 1 and 1000 which are divisible by none of the numbers 2,3,4 or 5. [266]

2. Ten people enter a lift on floor 1. The lift stops at each of the floors 2,3,4 and 5 to dicharge passengers (at least one), until it empties on floor 5. In how any ways can this happen? [818,528]

3. In how many ways can n pairs of socks be hung on a line so that adjacent socks are from different pairs, if socks within a pair are indistingiushable and each pair is different. [$\frac{1}{2^n}\sum_{k=0}^{n}(-2)^k\binom{n}{k}(2n-k)!$]

4. If D_n is the number of derangements of an n-set, show that

(i) $D_n = nD_{n-1} + (-1)^n$;

(ii) $D_{n+2} = (n+1)D_{n+1} + D_n$.

3.2 Generating functions

In this section we introduce one of the most successful devices for studying a sequence of numbers, by treating them as coefficients in a *formal power series* If $a_0, a_1, \ldots, = \{a_n\}_{n=0}^{\infty}$ is a sequence, we call,

$$G(t) = a_1 + a_1 t + a_2 t^2 + \cdots,$$

the *generating function* for the sequence. [The indeterminate t may be replaced by any other symbol.]

Generating functions may be multiplied by scalars (usually real numbers), added together and multiplied just like infinite series. [In short, the set of all generating functions of sequences of real numbers form an algebra over the real numbers.] These operations are often sufficient for us to deduce

See the paper by Ivan Niven, *Formal Power Series*, American Math. Monthly (1969), 871-889, for a systematic and self contained-account of the theory of generating functions.

the essential properties of sequences of numbers which arise in counting. On the other hand, most students will be familiar with the basic properites of infinite series and convergences of these, so it is a good idea to think of generating functions as if they were functions (almost all of these we consider will have a non-zero radius of convergence, and so are functions on a non-zero interval). This will allow us to freely use the operations that hold for infinite series.

Example 5. (Binomial coefficients.) The binomial coefficients, $\binom{n}{0}$, $\binom{n}{1}$, $\binom{n}{2}$, …, $\binom{n}{n}$ have the generating function,

$$\begin{aligned} G(t) &= \binom{n}{0} + \binom{n}{1}t + \binom{n}{2}t^2 + \cdots + \binom{n}{n}t^n \\ &= (1+t)^n. \end{aligned} \tag{7}$$

The identitites arising from the binomial theorem of §Chapter 1 can be thought of as arising from the combinatorics of generating functions of binomial coefficients.

Example 6. (Fibonacci numbers.) The Fibonacci nmbers, $F_0, F_1, F_2, \ldots$, occur so often in mathematics that an entire journal, the *Fibonacci Quarterly*, is devoted to publishing new discoveries involving them. The simplest way of defining them is by the recurrence relation,

$$F_{n+2} = F_{n+1} + F_n,$$

with initial conditions, $F_0 = 0$, $F_1 = 1$.
Let $G(t) = F_0 + F_1 t + F_2 t^2 + F_3 t^3 + \cdots$ be the generating function for the Fibonacci numbers.

$$\begin{array}{lrr} \text{Then} & G(t) = & F_0 t + F_1 t^2 + F_2 t^3 + \cdots \\ \text{and} & t^2 G(t) = & F_0 t^2 + F_1 t^3 + \cdots, \end{array}$$

by multiplication of generating functions. Hence, by subtraction,

$$\begin{aligned} G(t) - tG(t) - t^2G(t) &= F_0 + (F_1 - F_0)t + (F_2 - 1F_1 - F_0)t^2 + \cdots \\ &\qquad + (F_{n+2} - F_{n+1} - F_n)t^{n+2} + \cdots \\ &= t. \end{aligned}$$

That is, $(1 - t - t^2)G(t) = t$. We can thus write $G(t)$ in the closed form,

$$G(t) = \frac{t}{1 - t - t^2}. \tag{8}$$

The right hand side of (8) decomposes into partial fractions,

$$\frac{1/\sqrt{5}}{1 - at} - \frac{1/\sqrt{5}}{1 - bt},$$

where $a = \frac{1+\sqrt{5}}{2}$ and $b = \frac{1+\sqrt{5}}{2}$.
It follows that,

$$\begin{aligned} G(t) &= \frac{1}{\sqrt{5}}\Big(\frac{1}{1 - at} - \frac{1}{1 - bt}\Big) \\ &= \frac{1}{\sqrt{5}}\big(1 + at + a^2t^2 + \cdots\big) - \frac{1}{\sqrt{5}}\big(1 + bt + b^2t^2 + \cdots\big) \\ &= \frac{(a-b)}{\sqrt{5}}t + \frac{(a^2 - b^2)}{\sqrt{5}}tt^2 + \frac{(a^3 - b^3)}{\sqrt{5}}t^3 + \cdots + \frac{(a^n - b^n)}{\sqrt{5}}t^n + \cdots. \end{aligned}$$

Hence we obtain the following formula for the nth Fibonacci number:

$$F_n = \frac{1}{\sqrt{5}}\Big(\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \Big(\frac{1-\sqrt{5}}{2}\Big)^n\Big). \tag{9}$$

Note. Since $\frac{1-\sqrt{5}}{2} \approx -0.618$, it follows that, for $n > 0$, F_n is the integer closest to $\frac{1}{\sqrt{5}}\Big(\frac{1+\sqrt{5}}{2}\Big)^n$.

Exxample 7. (Catalan Numbers.) The nth Catalan number, c_n, is the number of ways a product of $n+1$ numbers can be bracketed by n pairs of brackets, so that each subproduct is a product of exactly two factors.

For example, the products of four numbers, a, b, c, d, can be bracketed in the following five ways:

$$(a(b(cd))), \qquad ((ab)(cd)), \qquad (((ab)c)d), \qquad ((a(bc)d), \qquad (a((bc)d)).$$

So $c_3 = 5$. It's easy to see that $c_1 = 1$, $c_2 = 2$, $c_4 = 14$. It is convenient to define also $c_0 = 1$.

Notice that every bracketing of $n+1$ numbers contains an outside pair of brackets which multiplies together a product of $k+1$ numbers and a product of $n-k$ numbers, where k may be $0, 1, 2, 3, \ldots$ or n. So for each k, there are $c_k c_{n-k-1}$ bracketings. Hence we obtain the following recurrence relation for the Catalan numbers

$$c_n = c_0 c_{n-1} + c_1 c_{n-2} + c_2 c_{n-2} + \cdots + c_{n-1} c_0. \tag{10}$$

Now if,

$$G(t) = c_0 + c_1 t + c_2 t^2 + \cdots,$$

is the generating function for the Catalan numbers, then

$$\begin{aligned} G(t)^2 &= c_0^2 + (c_0c_1 + c_1c_0)t + (c_0c_2 + c_1c_2 + c_2c_0)t^2 + \cdots \\ &\qquad + (c_0c_{n-1} + c_1c_{n-2} + \cdots + c_{n-1}c_0)t^n + \cdots \\ &= c_1 + c_2 t + c_3 t^2 + \cdots. \end{aligned}$$

Thus,

$$\begin{aligned} tG(t)^2 &= c_1 t + c_2 t^2 + c_3 t^3 + \cdots \\ &= G(t) - 1. \end{aligned}$$

Hence,

$$tG(t)^2 - G(t) + 1 - 0. \tag{11}$$

If we solve (11) using the quadratic formula, we find

$$G(t) = \frac{1+\sqrt{1-4t}}{2t} \quad \text{or} \quad \frac{1-\sqrt{1-4t}}{2t}. \tag{12}$$

The generating function will be the Maclaurin's series of one of the two functions in (12). To determine which, notice that $\sqrt{1-4t} = (1-4t)^{\frac{1}{2}}$ has the binomial series

$$\sqrt{1-4t} = 1 + \frac{\big(\frac{1}{2}\big)}{1!}(-4t) + \frac{\big(\frac{1}{2}\big)\big(-\frac{1}{2}\big)}{2!}(-4t)^2 + \frac{\big(\frac{1}{2}\big)\big(-\frac{1}{2}\big)\big(-\frac{3}{2}\big)}{3!}(-4t)^3 + \cdots,$$

where all terms, except the first, are negative. Hence the first candidate in (12) cannot be the generating function for the Catalan numbers. Hence,

$$\begin{aligned}
G(t) &= \frac{1}{\sqrt{1-4t}}2t \\
&= \frac{1}{2t}\Big(2t + \frac{1}{2^2 2!}4^2t^2 + \frac{1\cdot 3}{2^3 3!}4^3t^3 + \cdots \frac{1\cdot 3\cdot 5\cdots(2n-3)}{2^n n!}4^nt^n + \cdots\Big) \\
&= 1 + \frac{2t}{2!} + \frac{1\cdot 3}{3!}2^2t^2 + \frac{1\cdot 3\cdot 5}{4!}2^3t^3 + \cdots + \frac{1\cdot 3\cdot 5\cdots(2n-3)}{n!}2^{n-1}t^{n-1} + \cdots .
\end{aligned}$$

Hence,

$$\begin{aligned}
c_n &= \frac{1\cdot 3\cdot 5\cdots(2n-1)}{(n+1)!}2^n = \frac{1\cdot 2\cdot 3\cdot(2n-1)(2n)}{(n+1)!n!2^n}2^n \\
&= \frac{1}{n+1}\binom{2n}{n}. \qquad (13)
\end{aligned}$$

Check. $c_3 = \frac{1}{4}\binom{6}{3} = 5.$

3.2.1 Exercises with Answers

1. Show that F_{n+1}/F_n approaches $\frac{1+\sqrt{5}}{2} \approx 1..6.18$, the *gorlden ratio*, as n approaches ∞.

2. Prove the following identitites:

(i) $F_1 + F_3 + F_5 + \cdots + F_{2n-1} = F_{2n}$,

(ii) $F_1^2 + F_2^2 + F_3^2 + \cdots + F_n^2 = F_nF_{n+1}$.

3. The sequence of number $a_0, a_1, \ldots,$ is defined by the recurrence relations,

$$a_{n+2} = 5a_{n+1} - 6a_n,$$

with initial conditions $a_0 = 0$, and $a_1 = 1$.

(i) Determine the generating function for this sequence. $[G(t) = \frac{1}{1-5t+6t^2}]$

(ii) Find an explicit formula for a_n (in terms of n, not in terms of $a_0, a_1, a_2, \ldots, a_{n-1}$). $[3^n - 2^n]$

3.3 Binomial inversion

From Chapter 1, (15) we know that

$$m^n = \sum_{k=0}^{n}\binom{m}{k}k!S(n,k) = \sum_{k=0}^{m}\binom{m}{k}k!S(n,k). \qquad (14)$$

As we have seen in Chapter 1 Example 3, this identity can be used to find the $S(n,k)$'s, but not easily. Our aim in this section will be to show how such identiteis may be *inverted*, in order to, for example, express the Stirling numbers in terms of nth powers.

We first formulate the problem as one of inverting matrices. It will tunr out that these are easy to invert using only our knowledge of polynomials. To be more precise, we define a *polynomial sequence* to be an ordered set $\{p_n(x)\}_{n=0}^{\infty}$ of polynomials

$$p_0(x), \quad p_1(x), \quad p_2(x), \quad p_3(x), \quad \ldots$$

having real coefficients and satisfying the properties

(i) $p_0(x)$ is a nonzero constant,

(ii) $\deg p_n(x) = n$.

Note. If $\{q_n(x)\}$ is a polynomial sequence then

$$\{q_0(x), q_x(x), q_2(x), \ldots, q_n(x)\} \tag{15}$$

is a *basis* for the vector space of polynomials having degree at most n.

In particular, any polynomial of degree n can be written uniquely as a linear combination of the $q_i(x)$'s in (15). That is there are constants $a_{n,0}, a_{n,1}, \ldots, a_{n,n}$ such that

$$p_n(x) = \sum_{k=0}^{n} a_{n,k} q_k(x), \qquad \text{for } n = 0, 1, 2, \ldots. \tag{16}$$

[The $a_{n,k}$ are sometimes called *connection coefficients*.]
Similarly, if $\{p_n(x)\}$ is a polynomial sequence, then there are constants $b_{n,0}, b_{n,1}, \ldots, b_{n,n}$ such that

$$q_n(x) = \sum_{k=0}^{n} b_{n,k} p_k(x), \qquad \text{for } n = 0, 1, 2, \ldots. \tag{17}$$

Now we can write the first $m+1$ equations defined by (16) as the matrix equation

$$\begin{pmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_m(x) \end{pmatrix} = \begin{pmatrix} a_{0,0} & 0 & \cdots & 0 \\ a_{1,0} & a_{1,1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,0} & a_{m,1} & \cdots & a_{m,m} \end{pmatrix} \begin{pmatrix} q_0(x) \\ q_1(x) \\ \vdots \\ q_m(x) \end{pmatrix}$$

or more simply,

$$P = AQ, \tag{18}$$

where P and Q are the column vectors $(p_k(x))$ and $(q_k(x))$ respectively, and $A = (a_{i,j})$.
Similarly, the first $m+1$ equations defined by (17) can be written as the matrix equation

$$Q = BP, \tag{19}$$

where P and Q are as before but the matrix of coefficients is $B = (b_{i,j})$.
Note. Notice that A and B are inverse matrices, as we see by substituting the Q of (19) into (18) and the P of (18) into (19) to get

$$P = ABP \qquad \text{and} \qquad Q = BAQ.$$

The linear independence of the polynomoials of P and Q now forces $AB = I = BA$.

Example 8. (Application to inversion of combinatorial identitites.) Suppose that we have two sequences of numbers $u_0, u_1, u_2, \ldots$ and $v_0, v_1, v_2, \ldots,$ which are related by the equations:

$$u_n = \sum_{k=0}^{n} a_{n,k} v_k, \qquad n = 0, 1, 2, \ldots,$$

where $A = (a_{i,j})$ as in (18).

Then we can immediately write the v_k's in terms of the u_k's, using the inverse matrix B. That is, we have immediately

$$v_n = \sum_{k=0}^{n} b_{n,k} u_k.$$

Example 9. (Binomial inversion.) Consider the polynomial sequences $\{x^n\}$ and $\{(x-1)^n\}$. We can write polynomials of one sequence in terms of the other as follows:

$$x^n = (x-1+1)^n = \sum_{i=0}^{n} \binom{n}{i} (x-1)^i \tag{20}$$

and

$$(x-1)^n = \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} x^i. \tag{21}$$

These equations show that the matrices

$$\begin{pmatrix} \binom{0}{0} & 0 & \cdots & 0 \\ \binom{1}{0} & \binom{1}{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \binom{m}{0} & \binom{m}{1} & \cdots & \binom{m}{m} \end{pmatrix}$$

and

$$\begin{pmatrix} (-1)^0\binom{0}{0} & 0 & \cdots & 0 \\ (-1)^1\binom{1}{0} & (-1)^0\binom{1}{1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^m\binom{m}{0} & (-1)^{m-1}\binom{m}{1} & \cdots & (-1)^0\binom{m}{m} \end{pmatrix}$$

are inverse matrices for each nonnegative integer m.

Returning to our original problem, we now see that (14) can be inverted to get

$$m!S(n,m) = \sum_{k=0}^{m} (-1)^{m-k} \binom{m}{k} k^n. \tag{22}$$

Example 10. (Stirling Inversion.) By definition of the Stirling numbers of the first and second kind we have the identities:

$$x_{(n)} = \sum_{k=0}^{n} s(n,k) x^k \tag{23}$$

and

$$x^n = \sum_{k=0}^{n} S(n,k) x_{(k)}. \tag{24}$$

Let A and B be matrices whose (m,k)th elements are $s(m,k)$ and $S(m,k)$ respectively. Then A and B are inverse matrices and we have the identity

$$\sum_{k=0}^{m} s(l,k) S(k,j) = \delta_{l,j}. \tag{25}$$

3.4 Sieve formulas

In §1 we were able to use the Inclusion-Exclusion Principle to find the number of elements in a union of sets, or the complement of such a union. In this section we apply the inversion methods of §3 to obtain some important refinements. The results of this section and §1 are examples of "sieve methods" which work to count the number of elements of a set by subtracting unwanted elements of some larger set.[1]

Let X be a finite set and $A_1, A_2, \ldots, A_n$ subsets of X. We define two sets of numbers: $s_0, s_1, \ldots, s_n$ and $e_0, e_1, \ldots, e_n$, associated with these subsets.

The s_k's are defined by:

$$\begin{aligned} s_0 &= |X|, \\ s_1 &= |A_1| + |A_2| + \cdots + |A_n|, \\ s_2 &= |A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n|, \\ &\vdots \\ s_k &= \sum_{i_1 \le i_2 \le \cdots \le i_k} |A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| \qquad (26) \\ &\vdots \\ s_n &= |A_1 \cap A_2 \cap \cdots \cap A_n|. \end{aligned}$$

[Roughly the s_k count the elements in *at least* k subsets, possibly many times.]

On the other hand, e_k is the number of elements which are contained in *exactly* k of the subsets $A_1, A_2, \ldots, A_n$.

The Inclusion-Exclusion Principle is just the identity,

$$e_0 = s_0 - s_1 + s_2 - \cdots + (-1)^n s_n. \qquad (27)$$

The sieve formulas of this section express e_k in terms of $s_0, s_1, \ldots, s_n$ (which, in many applications, are easy to calculate).

The idea is to first exress the s_k's in terms of the e_k's and then invert.

Clearly,

$$s_0 = |X| = e_0 + e_1 + e_2 + \cdots + e_n,$$

and

$$s_1 = e_1 + 2e_2 + \cdots + ne_n$$

(elements in exactly k subsets are counted in s_1 by each of the subsets).

In general, an element in exactly m subsets will be counted by s_k a total of $\binom{m}{k}$ times (since there are $\binom{m}{k}$ terms in (26) with the A_{i_j} selected from the original m subsets).

Hence we have the following important identity,

$$s_k = \binom{k}{k} e_k + \binom{k+1}{k} e_{k+1} + \cdots + \binom{n}{k} e_n. \qquad (28)$$

[1] See Stanley, Richard P., *Enumerative Combinatorics Volume I*, Wadsworth & Brooks/Cole Mathematics Series 1986, Chapter 2, for a comprehensive treatment of sieve methods.

Or, using matrices,

$$\begin{pmatrix} s_0 \\ s_1 \\ \vdots \\ s_n \end{pmatrix} = \begin{pmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \binom{n}{0} \\ 0 & \binom{1}{1} & \binom{2}{1} & \cdots & \binom{n}{1} \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \binom{n}{n} \end{pmatrix} \begin{pmatrix} e_0 \\ e_1 \\ \vdots \\ e_n \end{pmatrix},$$

summarized by the matrix equations,

$$\mathbf{s} = A\mathbf{e}, \tag{29}$$

where A is the matrix of coefficients, $\mathbf{s}$, and $\mathbf{e}$ the vectors of s_k's and e_k's respectively.

Equation (29) expresses the s_k's in terms of the e_k's. Now invert.

$$\mathbf{e} = A^{-1}\mathbf{s}. \tag{30}$$

But A is the transpose of the matrix of binomial coefficients in Example 8. Hence we find that the inverse of A is,

$$\begin{pmatrix} (-1)^0\binom{0}{0} & (-1)^1\binom{1}{0} & (-1)^2\binom{2}{0} & \cdots & (-1)^n\binom{n}{0} \\ 0 & (-1)^0\binom{1}{1} & (-1)^1\binom{2}{1} & \cdots & (-1)^{n-1}\binom{n}{1} \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & (-1)^0\binom{n}{n} \end{pmatrix}.$$

Hence we obtain the following expression for e_k in terms of the s_i's:

$$e_k = s_k - \binom{k+1}{k} s_{k+1} + \binom{k+2}{k} s_{k+2} - \cdots + (-1)^{n-k}\binom{n}{k} s_n. \tag{31}$$

Generating functions

Let $G(t) = e_0 + e_1 t + e_2 t^2 + \cdots + e_n t^n$, be the generating function for the e_k's. By (31), we can write this in terms of s_k's.

$$\begin{aligned} G(t) &= s_0 - s_1 + s_2 - \cdots + (-1)^k s_k + \cdots + (-1)^n s_n \\ &\quad + \left(s_1 - \binom{2}{1} s_2 + \cdots + (-1)^{k-1}\binom{k}{1} s_k \pm \cdots + (-1)^{n-1}\binom{n}{1} s_n\right) t \\ &\quad + \left(s_2 - \cdots + (-1)^{k-2}\binom{k}{2} s_k \pm \cdots + (-1)^{n-2}\binom{n}{2} s_n\right) t^2 \\ &\quad \vdots \\ &\quad + \left((-1)^{k-k}\binom{k}{k} s_k \pm \cdots + (-1)^{n-k}\binom{n}{k} s_n\right) t^k \\ &\quad \vdots \\ &\quad + \left((-1)^{n-n}\binom{n}{n} s_n\right) t^n \\ &= s_0 + s_1(t-1) + s_2(t-1)^2 + \cdots + s_k(t-1)^k + \cdots + s_n(t-1)^n. \end{aligned} \tag{32}$$

Elements in an even or odd number of subsets

We can use the above generating function to calculate the number of elements that occur in an even number (respectively, odd) number of the subsets $A_1, S_2, \ldots, A_n$ of X. First observe from (**??**),

$$G(1) = s_0 \qquad (33)$$
$$G(-1) = s_0 - 2s_1 + 2^2 s_2 + \cdots + (-1)^n 2^n s_n.$$

But also

$$G(1) = e_0 + e_1 + e_2 + \cdots + e_n, \qquad (34)$$
$$G(-1) = e_0 - e_1 + e_2 + \cdots + (-1)^n e_n.$$

Equations (33) and (34) now yield the identities,

$$\begin{aligned} e_0 + e_2 + e_4 + \cdots &= \tfrac{1}{2}(G(1) + G(-1)) \\ &= \tfrac{1}{2}\Big(s_0 + \sum_{k=0}^{n} (-2)^k s_k\Big), \qquad (35) \\ e_1 + e_3 + e_5 + \cdots &= \tfrac{1}{2}(G(1) - G(-1)) \\ &= \tfrac{1}{2}\Big(s_0 - \sum_{k=0}^{n} (-2)^k s_k\Big), \qquad (35) \end{aligned}$$

Example 11. RNA chains consist of molecules U, A, C, G (uracil, adenine, cytosine and quanine). How many such chains of length n, have an even number of A's?

Let X be the set of all sequences of length n that can be formed from the four kinds of molecules. Let A_i be the subset of chains having an A in the ith position. Then

$$|X| = 4^n, \qquad |A_i| = 4^{n-1}, \quad |A_i \cap A_j| = 4^{n-2}, \qquad \text{etc.}$$

So,

$$s_0 = 4^n, \qquad s_1 = n4^{n-1}, \qquad s_k = \binom{n}{k} 4^{n-k}.$$

Hence, by (35), the number of chains with an even number of A's is,

$$e_0 + e_2 + e_4 + \cdots = \tfrac{1}{2}\Big(4^n + \sum_{k=0}^{n} \binom{n}{k} 4^{n-k}(-2)^k\Big) = \tfrac{1}{2}(4^n + 2^n).$$

3.4.1 Exercises with Answers

1. How many positive integers ≤ 250 are not divisible by any of $2, 5$ or 7? By exactly on of 2, 5 or 7? By exactly two of 2, 5, or 7? By all three of $2, 5$ and 7?

[86; 121; 40; 3]

2. How many 10-molecule RNA chains are there which have no U's and an even number of G's?

[29,525]

3. In how many ways can eight letters be taken from their envelopes, read and then replaced at random, so that none of the letters will be in its correct envelope? At least one will be in its correct envelope? At least two?

[14,833; 25,487; 10,655]

3.5 Finite integration

Combinatorial identities often involve finite sums, expressed in closed form. In this section we will find a relatively simple method which will allow us, in principle, to sum many finite series. The method is based on the fact that, just as integration and differentiation are inverse processes involving the derivative operator D, summing finite series and taking differences are inverse processes involving the *forward difference operator*, Δ. This process is called *Finite Integration*.

Definition. The *forward difference operator*, Δ, maps functions to functions, just as D does. It is defined by,

$$\Delta f\colon x \mapsto f(x+1) - f(x), \tag{37}$$

for any function f whose domain contains $x+1$ whenever it contains x. [We will consider mainly functions with domain the positive integers.]

The Fundamental Theorem

Suppose we wnat to find the following sum:

$$f(a) + f(a+1) + f(a+2) + \cdots + f(n).$$

If we can find $F(x)$ such that

$$\Delta F(x) = f(x),$$

then

$$\begin{aligned} f(a) &= F(a+1) - F(a), \\ f(a+1) &= F(a+2) - F(a+1), \\ f(a+2) &= F(a+3) - F(a+2), \\ &\vdots \\ f(n-1) &= F(n) - F(n-1), \\ f(n) &= F(n+1) - F(n). \end{aligned} \tag{38}$$

By adding both sides of (38) (observing that the right-hand sum collapses), we obtain the following:

The Fundamental Theorem for Finite Integration

$$\begin{aligned} f(a) + f(a+1) + \cdots + f(n) = F(n+1) - F(a) = F(x)\Big|_a^{n+1} \\ = \Delta^{-1} f(x)\Big|_a^{n+1} \end{aligned} \tag{39}$$

Example 12. Consider the sum:

$$1 + 3 + 5 + 7 + \cdots + (2n+1) = \sum_{k=0}^{n} (2k+1),$$

where n is any natural number.

Noice that $\Delta x^2 = (x+1)^2 - x^2 = 2x+1$. Hence by the Fundamental Theorem:

$$\begin{aligned} 1+3+5+7+\cdots+(2n+1) &= \sum_{k=0}^{n}(2k+1) = \Delta^{-1}(2x+1)\Big|_0^{n+1} \\ &= x^2\Big|_0^{n+1} = (n+1)^2 - 0^2 \\ &= (n+1)^2. \end{aligned}$$

Example 13. Since

$$\Delta x^3 = (x+1)^3 - x^3 = 3x^2+3x+1,$$

we find from (39) that

$$\begin{aligned} 1+7{+}19+37+\cdots+(3n^2+3n+1) \\ &= \sum_{k=0}^{n}(3k^2+3k+1) = x^3\Big|_0^{n+1} \\ &= (n+1)^2. \end{aligned}$$

Note. The role of the standard polynomials, x^n, in ordinary differential and integral calculus, is replaced by the *falling factorial polynomials*, $x_{(n)} = x(x-1)(x-2)\cdots(x-n+1)$. For example,

$$\begin{aligned} \Delta x_{(n)} &= (x+1)_{(n)} - x_{(n)} \\ &= (x+1)x(x-1)\cdots(x+1-n+1) - x(x-1)\cdots(x-n+1) \\ &= \big((x+1)-(x-n+1)\big)x(x-1)\cdots(x-n+2) \\ &= nx_{(n-1)}. \end{aligned}$$

Hence $\Delta\frac{x_{(n)}}{n} = x_{(n-1)}$ and we have the simple finite integration formula

$$\boxed{\Delta^{-1}x_{(n)} = \frac{x_{(n+1)}}{n+1}.} \tag{41}$$

Example 14. Find $\displaystyle\sum_{k=2}^{n}\binom{k}{2}$.

First, using (39), we have,

$$\begin{aligned} 2\cdot 1+3\cdot 2+\cdots+n(n-1) &= \sum_{k=2}^{n}k_{(2)} = \Delta^{-1}x_{(2)}\Big|_2^{n+1} = \frac{1}{3}x_{(2)}\Big|_2^{n+1} \\ &= \tfrac{1}{3}(n+1)n(n-1). \end{aligned}$$

But,

$$\begin{aligned} \sum_{k=2}^{n}\binom{k}{2} &= \binom{2}{2}+\binom{3}{2}+\cdots\binom{n}{2} \\ &= \tfrac{1}{2}(2\cdot 1+3\cdot 2+\cdots n(n-1)) = \tfrac{1}{6}(n+1)n(n-1) \\ &= \binom{n}{3}. \end{aligned}$$

Example 15. To find an expression for the sum of any series of the form

$$1^m + 2^m + 3^m + \cdots + n^m$$

we need only find $\Delta^{-1}x^m$. This involves expressing x^m in terms of falling factorials. In other words, we need to find coefficients $S(m,k)$ such that

$$x^m = S(m,0)x_{(0)} + S(m,1)x_{(1)} + \cdots + S(m,m)x_{(m)}, \tag{42}$$

for all x.

The coefficients $S(m,k)$ in (42) are the Stirling numbers of the second kind which we studied in Chapter 1 §4.

From the first five rows of Stirling numbers given in Chapter 1 §4 we can immediately write down the identities:

$$\begin{aligned} x^1 &= x_{(1)}, \\ x^2 &= x_{(1)} + x_{(2)}, \\ x^3 &= x_{(1)} + 3x_{(2)} + x_{(3)}, \\ x^4 &= x_{(1)} + 15x_{(2)} + 25x_{(3)} + 10x_{(4)} + x_{(5)}, \end{aligned}$$

These identities give the following summations

$$\begin{aligned} \sum_{k=1}^{n} k &= \tfrac{1}{2}x_{(2)}\Big|_1^{n+1} = \tfrac{1}{2}(n+1)n, \\ \sum_{k=1}^{n} k^2 &= \tfrac{1}{2}x_{(2)} + x_{(3)} + \tfrac{1}{4}x_{(4)}\Big|_1^{n+1} = \left(\tfrac{1}{2}(n+1)n\right)^2, \quad \text{(after some multiplication)} \\ \sum_{k=1}^{n} k^3 &= \tfrac{1}{2}x_{(2)} + 5x_{(3)} + \tfrac{25}{4}x_{(4)} + 2x_{(5)} + \tfrac{1}{6}x_{(6)}\Big|_1^{n+1} \\ &= \tfrac{1}{2}(n_1)_{(2)} + 5(n+1)_{(3)} + \tfrac{25}{4}(n+1)_{(4)} + 2(n+1)_{(5)} + \tfrac{1}{6}(n+1)_{(6)}. \end{aligned}$$

Similarly, we can find an expression for the sum

$$1^m + 2^m + \cdots + n^m$$

in terms of m if we extend the triangle of Stirling numbers down to the m-th row.

Summation by parts

We now take a brief look at sums like

$$1 \cdot 2 + 2 \cdot 2^2 + 3 \cdot 2^3 + \cdots + n \cdot 2^n. \tag{43}$$

The terms of this sum are products of terms of two simpler series which we can sum. We need a rule for handling such products.

Observe first that for functions $f(x)$ and $g(x)$,

$$\begin{aligned} \Delta(f(x)g(x)) &= f(x+1)g(x+1) - f(x)g(x) \\ &= f(x+1)g(x+1) - f(x+1)g(x) + f(x+1)g(x) - f(x)g(x) \\ &= f(x+1)\Delta g(x) + (\Delta f(x))g(x). \end{aligned}$$

Hence we have the important rule:

The rule for summation by parts

$$\Delta^{-1}[g(x)\Delta f(x)] = f(x)g(x) - \Delta^{-1}[f(x+1)\Delta g(x)] \tag{44}$$

Remark. This rule is analogous to the familiar integration by parts rule of the integral calculus.

Example 16. Let $f(x) = \frac{a^x}{a-1}$, $\quad g(x) = x, \quad a > 1$, then

$$\Delta f(x) = a^x \qquad \text{and} \qquad \Delta g(x) = 1.$$

Hence, (44) tells us that

$$\Delta^{-1}(xa^x) = \frac{xa^x}{a-1} - \Delta^{-1}\frac{a^{x+1}}{a-1}.$$

That is

$$\Delta^{-1}(xa^x) = \frac{xa^x}{a-1} - \frac{a^{x+1}}{(a-1)^2}.$$

Hence

$$\begin{aligned}1\cdot a^1 + 2\cdot a^2 + \cdots + n\cdot a^n &= \Delta^{-1}(xa^x)\Big|_1^{n+1} \\ &= \frac{(n+1)a^{n+1}}{a-1} - \frac{a^{n+2}}{(a-1)^2} - \frac{a}{a-1} + \frac{a^2}{(a-1)^2} \\ &= \frac{a^{n+1}(na-n-a)+a}{(a-1)^2}.\end{aligned}$$

In particular, setting $n = 2$ we find that the sum (43) is

$$\begin{aligned}1\cdot 2 + 2\cdot 2^2 + 3\cdot 2^3 + \cdots + n\cdot 2^n &= (n+1)2^{n+1} - 2^{n+1} + 2 \\ &= (n-1)2^{n+1} + 2.\end{aligned}$$

3.5.1 Exercises with answers

1. Show that $x(x(+2)(x+3) = (x+3)_{(3)} - (x+2)_{(2)}$ and hence sum the series,

$$1\cdot 3\cdot 4 + 2\cdot 4\cdot 5 + 3\cdot 5\cdot 6 + \cdots + n(n+2)(n+3).$$

$[\frac{1}{12}(n+4)(n+3)(n+2)(3n-1)+2]$

2. Sum the following series

$$1^2\cdot 2 + 2^2\cdot 2^2 + 3^2\cdot 2^2 + 4^2\cdot 2^3 + 4^2\cdot 2^4 + \cdots + n^2\cdot 2^n.$$

$[(n^2-2n+3)2^{n+1}-6]$

3. Use finite integration to sum the following series:

(i) $\binom{m}{m} + \binom{m+1}{m} + \cdots + \binom{n}{m}$ $\qquad [\binom{n+1}{m+1}]$

(ii) $m\binom{m}{m} + (m+1)\binom{m+1}{m} + \cdots + n\binom{n}{m}$ $\qquad [(n+1)\binom{n+1}{m+1} - \binom{n+2}{m+2}]$