$$(1+x_1)(1+x_2)\dots(1+x_n) = \sum_{I\subseteq\{1,2,\dots,n\}} \Bigl(\prod_{i\in I} x_i\Bigr). \tag{2.17}$$

Contemplate what this formula says (write it out for $n = 1, 2, 3$, say) and why it holds.

In order to prove the inclusion–exclusion principle, let us denote $A = A_1 \cup A_2 \cup \dots \cup A_n$, and let $f_i\colon A \to \{0,1\}$ be the characteristic function of the set A_i, which means that $f_i(a) = 1$ for $a \in A_i$ and $f_i(a) = 0$ otherwise. For every $a \in A$, we have $\prod_{i=1}^n (1 - f_i(a)) = 0$ (don't we?), and using (2.17) with $x_i = -f_i(a)$ we get

$$\sum_{I\subseteq\{1,2,\dots,n\}} (-1)^{|I|} \prod_{i\in I} f_i(a) = 0.$$

By adding all these equalities together for all $a \in A$, and then by interchanging the summation order, we arrive at

$$\begin{aligned} 0 &= \sum_{a\in A}\Bigl(\sum_{I\subseteq\{1,2,\dots,n\}} (-1)^{|I|}\prod_{i\in I} f_i(a)\Bigr)\\ &= \sum_{I\subseteq\{1,2,\dots,n\}} (-1)^{|I|}\Bigl(\sum_{a\in A}\prod_{i\in I} f_i(a)\Bigr). \end{aligned} \tag{2.18}$$

Now it suffices to note that the $\prod_{i\in I} f_i(a)$ is the characteristic function of the set $\bigcap_{i\in I} A_i$, and therefore $\sum_{a\in A}\prod_{i\in I} f_i(a) = \left|\bigcap_{i\in I} A_i\right|$. In particular, for $I = \emptyset$, $\prod_{i\in\emptyset} f_i(a)$ is the empty product, with value 1 by definition, and so $\sum_{a\in A}\prod_{i\in\emptyset} f_i(a) = \sum_{a\in A} 1 = |A|$. Hence (2.18) means

$$|A| + \sum_{\emptyset\neq I\subseteq\{1,2,\dots,n\}} (-1)^{|I|}\Bigl|\bigcap_{i\in I} A_i\Bigr| = 0,$$

and this is exactly the inclusion–exclusion principle. An expert in algebra can thus regard the inclusion–exclusion principle with mild contempt: a triviality, she might say. □

Bonferroni inequalities. Sometimes we can have the situation where we know the sizes of all the intersections up to m-fold ones, but we do not know the sizes of intersections of more sets than m. Then we cannot calculate the size of the union of all sets exactly. The so-called *Bonferroni inequalities* tell us that if we leave out all terms with $k > m$ on the right-hand side of the inclusion–exclusion principle (2.15) then the *error* that we make in this way in the calculation of the size of the union *has the same sign as the first omitted term*. Written as a formula, for every *even* q we have

$$\sum_{k=1}^{q} (-1)^{k-1} \sum_{I\in\binom{\{1,2,\dots,n\}}{k}} \Bigl|\bigcap_{i\in I} A_i\Bigr| \le \Bigl|\bigcup_{i=1}^{n} A_i\Bigr| \tag{2.19}$$

and for every *odd* q we have

$$\sum_{k=1}^{q} (-1)^{k-1} \sum_{I\in\binom{\{1,2,\dots,n\}}{k}} \Bigl|\bigcap_{i\in I} A_i\Bigr| \ge \Bigl|\bigcup_{i=1}^{n} A_i\Bigr|. \tag{2.20}$$

This means, for instance, that if we didn't know how many diligent persons are simultaneously in all the three clubs in Example 2.7.1, we could still estimate that the total number of members in all the clubs is at least 32. We do not prove the Bonferroni inequalities here.

Exercises

1. Explain why the formulas (2.15) and (2.16) express the same equality.
2. *Prove the Bonferroni inequalities. If you cannot handle the general case try at least the cases $q = 1$ and $q = 2$.
3. (Sieve of Eratosthenes) How many numbers are left in the set $\{1, 2, \dots, 1000\}$ after all multiples of 2, 3, 5, and 7 are crossed out?
4. How many numbers $n < 100$ are not divisible by a square of any integer greater than 1?
5. *How many orderings of the letters A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P are there such that we cannot obtain any of the words BAD, DEAF, APE by crossing out some letters? What if we also forbid LEADING?
6. How many ways are there to arrange 4 Americans, 3 Russians, and 5 Chinese into a queue, in such a way that no nationality forms a single consecutive block?

2.8 The hatcheck lady & co.

2.8.1 Problem (Hatcheck lady problem). Honorable gentlemen, n in number, arrive at an assembly, all of them wearing hats, and they deposit their hats in a cloak-room. Upon their departure, the hatcheck lady, maybe quite absent-minded that day, maybe even almost blind after many years of service in the poorly lit cloak-room, issues one hat to each gentleman at random. What is the probability than none of the gentlemen receives his own hat?

As stated, this is a toy problem, but mathematically it is quite remarkable, and a few hundred years back, it occupied some of the best mathematical minds of their times. First we reformulate the problem using permutations. If we number the gentlemen (our apologies)

$1, 2, \ldots, n$, and their hats too, then the activity of the hatcheck lady results in a random permutation π of the set $\{1, 2, \ldots, n\}$, where $\pi(i)$ is the number of the hat returned to the ith gentleman. The question is, what is the probability of $\pi(i) \neq i$ holding for all $i \in \{1, 2, \ldots, n\}$? Call an index i with $\pi(i) = i$ a *fixed point* of the permutation π. So we ask: what is the probability that a randomly chosen permutation has no fixed point? Each of the $n!$ possible permutations is, according to the description of the hatcheck lady's method of working, equally probable, and so if we denote by $D(n)$ the number of permutations with no fixed point[6] on an n-element set, the required probability equals $D(n)/n!$.

Using the inclusion–exclusion principle, we derive a formula for $D(n)$. We will actually count the "bad" permutations, i.e. those with at least one fixed point. Let S_n denote the set of all permutations of $\{1, 2, \ldots, n\}$, and for $i = 1, 2, \ldots, n$, we define $A_i = \{\pi \in S_n \colon \pi(i) = i\}$. The bad permutations are exactly those in the union of all the A_i.

Here we suggest that the reader contemplate the definition of the sets A_i carefully—it is a frequent source of misunderstandings (their elements are permutations, not numbers).

In order to apply the inclusion–exclusion principle, we have to express the size of the k-fold intersections of the sets A_i. It is easy to see that $|A_i| = (n-1)!$, because if $\pi(i) = i$ is fixed, we can choose an arbitrary permutation of the remaining $n - 1$ numbers. Which permutations lie in $A_1 \cap A_2$? Just those with both 1 and 2 as fixed points (and the remaining numbers can be permuted arbitrarily), and so $|A_1 \cap A_2| = (n-2)!$. More generally, for arbitrary $i_1 < i_2 < \cdots < i_k$ we have $|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = (n-k)!$, and substituting this into the inclusion–exclusion formula yields

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} (n-k)! = \sum_{k=1}^{n} (-1)^{k-1} \frac{n!}{k!}.$$

We recall that we have computed the number of bad permutations (with at least one fixed point), and so

$$D(n) = n! - |A_1 \cup \cdots \cup A_n| = n! - \frac{n!}{1!} + \frac{n!}{2!} - \cdots + (-1)^n \frac{n!}{n!},$$

which can still be rewritten as

[6] Such permutations are sometimes called *derangements*.

$$D(n) = n! \left(1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + (-1)^n \frac{1}{n!}\right). \tag{2.21}$$

As is taught in calculus, the series in parentheses converges to e^{-1} for $n \to \infty$ (where e is the Euler number), and it does so very fast. So we have the approximate relation $D(n) \approx n!/e$, and the probability in the hatcheck lady problem converges to the constant $e^{-1} = 0.36787\ldots$. This is what also makes the problem remarkable: the answer almost doesn't depend on the number of gentlemen!

The Euler function φ. A function denoted usually by φ and named after Leonhard Euler plays an important role in number theory. For a natural number n, the value of $\varphi(n)$ is defined as the number of natural numbers $m \leq n$ that are relatively prime to n; formally

$$\varphi(n) = |\{m \in \{1, 2, \ldots, n\} \colon \gcd(n, m) = 1\}|.$$

Here $\gcd(n, m)$ denotes the greatest common divisor of n and m; that is, the largest natural number that divides both n and m. As an example of application of the inclusion–exclusion principle, we find a formula which allows us to calculate $\varphi(n)$ quickly provided that we know the factorization of n into prime factors.

The simplest case is when $n = p$ is a prime. Then every $m < p$ is relatively prime to p, and so $\varphi(p) = p - 1$.

The next step towards the general solution is the case when $n = p^\alpha$ ($\alpha \in \mathbf{N}$) is a prime power. Then the numbers not relatively prime to p^α are multiples of p, i.e. $p, 2p, 3p, \ldots, p^{\alpha-1}p$, and there are $p^{\alpha-1}$ such multiples not exceeding p^α (in general, if d is an any divisor of some number n, then the number of multiples of d not exceeding n is n/d). Hence, there are $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = p^\alpha(1 - 1/p)$ remaining numbers that are relatively prime to p^α.

An arbitrary n can be written in the form

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_r^{\alpha_r},$$

where $p_1, p_2, \ldots, p_r$ are distinct primes and $\alpha_i \in \mathbf{N}$. The "bad" $m \leq n$, i.e. those not contributing to $\varphi(n)$, are all multiples of some of the primes p_i. Let us denote by $A_i = \{m \in \{1, 2, \ldots, n\} \colon p_i | m\}$ the set of all multiples of p_i. We have $\varphi(n) = n - |A_1 \cup A_2 \cup \cdots \cup A_r|$. The inclusion–exclusion principle commands that we find the sizes of the intersections of the sets A_i. For example, the intersection $A_1 \cap A_2$ contains the numbers divisible by both p_1 and p_2, which are exactly the multiples of $p_1 p_2$, and hence $|A_1 \cap A_2| = n/(p_1 p_2)$. The same argument gives

$$|A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}| = \frac{n}{p_{i_1} p_{i_2} \cdots p_{i_k}}.$$

Let us look at the particular cases $r = 2$ and $r = 3$ first. For $n = p_1^{\alpha_1} p_2^{\alpha_2}$ we have

$$\varphi(n) = n - |A_1 \cup A_2| = n - |A_1| - |A_2| + |A_1 \cap A_2|$$
$$= n - \frac{n}{p_1} - \frac{n}{p_2} + \frac{n}{p_1 p_2} = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right).$$

Similarly, for $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3}$ we get

$$\varphi(n) = n - \frac{n}{p_1} - \frac{n}{p_2} - \frac{n}{p_3} + \frac{n}{p_1 p_2} + \frac{n}{p_1 p_3} + \frac{n}{p_2 p_3} - \frac{n}{p_1 p_2 p_3}$$
$$= n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\left(1 - \frac{1}{p_3}\right).$$

This may raise a suspicion concerning the general formula.

2.8.2 Theorem. *For $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$, we have*

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\dots\left(1 - \frac{1}{p_r}\right). \qquad (2.22)$$

Proof. For an arbitrary r, the inclusion–exclusion principle (we use, to our advantage, the short formula (2.16)) gives

$$\varphi(n) = n - \sum_{\emptyset \neq I \subseteq \{1,2,\dots,r\}} (-1)^{|I|-1} \frac{n}{\prod_{i \in I} p_i} = n \cdot \sum_{I \subseteq \{1,2,\dots,r\}} \frac{(-1)^{|I|}}{\prod_{i \in I} p_i}.$$

We claim that this frightening formula equals the right-hand side of Eq. (2.22). This follows from the formula (2.17) for expanding the product $(1+x_1)(1+x_2)(1+x_3)\dots$ by substituting $x_i = -1/p_i$, $i = 1, 2, \dots, r$. □

Exercises

1. There are n married couples attending a dance. How many ways are there to form n pairs for dancing if no wife should dance with her husband?
2. (a) Determine the number of permutations with exactly one fixed point.

   (b) Count the permutations with exactly k fixed points.
3. What is wrong with the following inductive "proof" that $D(n) = (n-1)!$ for all $n \geq 2$? Can you find a false step in it? For $n = 2$, the formula holds, so assume $n \geq 3$. Let π be a permutation of $\{1, 2, \dots, n-1\}$ with no fixed point. We want to extend it to a permutation π' of $\{1, 2, \dots, n\}$ with no fixed point. We choose a number $i \in \{1, 2, \dots, n-1\}$, and we define $\pi'(n) = \pi(i)$, $\pi'(i) = n$, and $\pi'(j) = \pi(j)$ for $j \neq i, n$. This defines a permutation of $\{1, 2, \dots, n\}$, and it is easy to check that it has no fixed point. For each of the $D(n-1) = (n-2)!$ possible choices of π, the index i can be chosen in $n-1$ ways. Therefore, $D(n) = (n-2)! \cdot (n-1) = (n-1)!$.
4. *Prove the equation

$$D(n) = n! - nD(n-1) - \binom{n}{2} D(n-2) - \dots - \binom{n}{n-1} D(1) - 1.$$

5. (a) *Prove the recurrent formula $D(n) = (n-1)[D(n-1) + D(n-2)]$. Prove the formula (2.21) for $D(n)$ by induction.

   (b) *Calculate the formula for $D(n)$ directly from the relation derived in (a). Use an auxiliary sequence given by $a_n = D(n) - n\,D(n-1)$.
6. How many permutations of the numbers $1, 2, \dots, 10$ exist that map no even number to itself?
7. (Number of mappings onto) Now is the time to calculate the number of mappings of an n-element set *onto* an m-element set (note that we have avoided it so far). Calculate them

   (a) for $m = 2$

   (b) for $m = 3$.

   (c) *Write a formula for a general m; check the result for $m = n = 10$ (what is the result for $n = m$?). Warning: The resulting formula is a sum, not a "nice" formula like a binomial coefficient.

   (d) *Show, preferably without using part (c), that the number of mappings onto an m-element set is divisible by $m!$.
8. (a) *How many ways are there to divide n people into k groups (or: how many equivalences with k classes are there on an n-element set)? Try solving this problem for $k = 2, 3$ and $k = n-1, n-2$ first. For a general k, the answer is a sum.

   (b) What is the total number of equivalences on an n-element set? (Here the result is a double sum.)

   (c) *If we denote the result of (b) by B_n (the *nth Bell number*), prove the following (surprising) formula:

$$B_n = \frac{1}{e} \sum_{i=0}^{\infty} \frac{i^n}{i!}.$$

9. *Prove the formula (2.22) for the Euler function in a different way. Suppose it holds for $n = p^\alpha$ (a prime power). Prove the following *auxiliary claim*: if m, n are relatively prime, then $\varphi(mn) = \varphi(m)\varphi(n)$.