1 The Inclusion-Exclusion Principle

Our next step in developing the twelvefold way will deal with the surjective functions. We'll build these through the use of *inclusion-exclusion.*

In its most basic form, inclusion-exclusion is a way of counting the membership of a union of sets. For two sets, it is easy to cinvince yourself that $|A \cup B| = |A| + |B| - |A \cap B|$. With a little bit more doing, we can show that $|A \cup B \cup C| = |A| + |B| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$. We add and subtract certain slices of sets until all overcounts and undercounts are eliminated. We can use these simple, small versions of the inclusion-exclusion principle in a simple example:

Question 1: *How many members of* $\{1, 2, 3, \ldots, 105\}$ *have nontrivial factors in common with 105?*

Answer 1: $105 = 3 \cdot 5 \cdot 7$*, so a nubmer shares factors with 105 if and only if it is divisible by 3, 5, or 7. Let A, B, and C be the members of* $\{1, 2, 3, \ldots, 105\}$ *divisible by 3, 5, and 7 respectively. Clearly* $|A| = 35$*,* $|B| = 21$*, and* $|C| = 15$*. Furthermore,* $A \cap B$ *consists of those numbers divisible by both and 5, i.e., divisible by 15. Likewise,* $A \cap C$ *and* $B \cap C$ *contain multiples of 21 and 35 respectively, so* $|A \cap B| = 7$*,* $|A \cap C| = 5$*, and* $|B \cap C| = 3$*. Finally,* $A \cap B \cap C$ *consists only of the nubmer* 105*, so it has 1 member total. Thus,*

$$|A \cup B \cup C| = 35 + 21 + 15 - 7 - 5 - 3 + 1 = 57$$

There are 3 simple presentations of the inclusion-exclusion principle, which you should be able to convince yourself are equivalent:

$$\begin{aligned}|A_1 \cup A_2 \cup \cdots \cup A_n| = {} & |A_1| + |A_2| + \cdots + |A_n| \\ & - |A_1 \cap A_2| - |A_1 \cap A_3| - \cdots - |A_1 \cap A_n| - |A_2 \cap A_3| - \cdots - |A_{n-1} \cap A_n| \\ & + |A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n| \\ & - \cdots \pm |A_1 \cap A_2 \cap \cdots \cap A_n|\end{aligned}$$

All those ellipses are a bit unhappy, so one of the following forms is a bit more explicit if a bit less readable:

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{i=1}^{n} (-1)^{i-1} \sum_{|S|=i; S \subseteq \{1,2,\ldots,n\}} \left| \bigcap_{j \in S} A_j \right|$$

which can be simplified a bit further, by choosing a set first and then worrying about its size:

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{S \subseteq \{1,2,\ldots,n\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right|$$

This form is also the easiest in which to prove the inclusion-exclusion principle.

1.1 Proof of Inclusion-Exclusion

Proposition 1. *For finite sets* $A_1, A_2, \ldots, A_n$,

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{S \subseteq \{1,2,\ldots,n\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right|$$

Proof. We prove this by induction on n. For $n = 1$, it is trivial:

$$|A_1| = \sum_{\emptyset \neq S \subseteq \{1\}} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right|$$

For our inductive step, we will take it as given that:

$$|A_1 \cup A_2 \cup \cdots \cup A_{n-1}| = \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right|$$

and thereby find $|A_1 \cup A_2 \cup \cdots \cup A_n|$:

$$\begin{aligned}
|A_1 \cup \cdots \cup A_n| &= |(A_1 \cup A_2 \cup \cdots \cup A_{n-1}) \cup A_n| \\
&= |A_1 \cup A_2 \cup \cdots \cup A_{n-1}| + |A_n| - |(A_1 \cup A_2 \cup \cdots \cup A_{n-1}) \cap A_n| \\
&= |A_1 \cup A_2 \cup \cdots \cup A_{n-1}| + |A_n| - |(A_1 \cap A_n) \cup (A_2 \cap A_n) \cup \cdots \cup (A_{n-1} \cap A_n)| \\
&= \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| + |A_n| - \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \cap A_n \right| \\
&= \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| + |A_n| - \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \left(\bigcap_{j \in S} A_j \right) \cap A_n \right| \\
&= \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| + |A_n| - \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in (S \cup \{n\})} A_j \right| \\
&= \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| + \sum_{S \subseteq \{1,2,\ldots,n-1\}} (-1)^{|S|} \left| \bigcap_{j \in (S \cup \{n\})} A_j \right| \\
&= \sum_{S \subseteq \{1,2,\ldots,n-1\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| + \sum_{S \subseteq \{1,2,\ldots,n\}; n \in S} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right| \\
&= \sum_{S \subseteq \{1,2,\ldots,n\}; S \neq \emptyset} (-1)^{|S|-1} \left| \bigcap_{j \in S} A_j \right|
\end{aligned}$$

□

1.2 Applications of Inclusion-Exclusion

While finding members of a collection of sets union is a useful context in which to use inclusion-exclusion, it's more commonly used for finding sets of numbers which *lack* several properties. That is, if we have a set X, and we want to remove from it all elements of subsets A_1, A_2, etc., the subtractive principle says that our count will be

$$|X| - |A_1 \cup A_2 \cup \cdots \cup A_n| = |X| + \sum_{\emptyset \neq S \subseteq \{1,2,\ldots,n\}} (-1)^{|S|} \left| \bigcap_{j \in S} A_j \right|$$

This is particularly useful in cases where the sizes of A_i are all the same, and where the sizes of each $A_i \cap A_j$ are all the same, etc. If an intersection of j sets has size a_j, then we can simplify this to:

$$|X| + \sum_{i=1}^{n} (-1)^i \binom{n}{i} a_i$$

A classic problem along these lines is the so-called *derangement problem*:

Question 2: *A permutation of the set $\{1, 2, \ldots, n\}$ is known as a* derangement *if every element is in a position not equal to its value; that is, if for every value of i, the nubmer i does not appear in the ith position. How many derangements are there of $\{1, 2, \ldots, n\}$?*

Answer 2: *Let X be the set of all permutations of $\{1, 2, \ldots, n\}$; we know an enumeration statistic that counts this, and can say with confidence that $|X| = n!$. However, this is massively overcounting the derangements, since not every permutation is a derangement. We need to exclude all those permutations with 1 in the first position, 2 in the second position, etc. We may define sets of these undesirables: let A_i consist of all permutations of $\{1, 2, \ldots, n\}$ which place i in the ith position. Then the set of derangements will be $X - (A_1 \cup A_2 \cup \cdots \cup A_n)$, whose size we should be able to find by inclusion-exclusion.*

We can simplify the inclusion-exclusion, however, by making use of the fact that the sets A_i are, in a certain sense, symmetrical: there is nothing particularly special about a particular number i or ith position, so we would expect all the A_i to be the same size, and likewise all the $A_i \cap A_j$ to be the same size, and so forth. This will in fact be the case: A_i consists of those permutations with i in the ith position, and free selections for other positions, so elements of A_i are determined by permuting the remaining $n-1$ elements, and thus $|A_i| = (n-1)!$ regardless of what i is. Likewise, $A_i \cap A_j$ consists of permutations with the positions of i and j determined, so its elements are determined by permutations on $n-2$ elements; thus $|A_i \cap A_j| = (n-2)!$; by a similar argument,t he intersection of i different sets consists of permutations with i elements fixed in place, so they are determined by permutations among the remaining $n-i$ elements, of which there are $(n-i)!$. Thus, in the formulation of the inclusion-exclusing principle given above, $a_i = (n-i)!$ and we thus have:

$$|X| - |A_1 \cup A_2 \cup \cdots \cup A_n| = |X| + \sum_{i=1}^{n} (-1)^i \binom{n}{i} (n-i)! = n! + \sum_{i=1}^{n} (-1)^i \frac{n!}{(n-i)!i!} (n-i)! = \sum_{i=0}^{n} (-1)^i \frac{n!}{i!}$$

An interesting footnote to the derangement problem is the question of derangement probability: what is the probability that a random permutation of length n is a derangement? Taking the ratio of derangements to permutations, we have a probability of

$$\frac{\sum_{i=0}^{n} (-1)^i \frac{n!}{i!}}{n!} = \sum_{i=0}^{n} (-1)^i \frac{1}{i!}$$

which is equal to the nth partial sum of the extremely rapidly converging infinite series $\sum_{i=0}^{\infty} (-1)^i \frac{1}{i!}$, which converges to e^{-1}. Thus, for any moderately large n (even $n = 6$ is an excellent approximation), the probability that a permutation is a derangement is very close to $\frac{1}{e}$.

And finally, we can bring inclusion-exclusion into play in addressing twelvefold-path enumeration statistics:

Question 3: *How many 6-element strings of 1s, 2s, 3s, and 4s contain at least one 1, at least one 2, at least one 3, and at least one 4?*

Answer 3: *There are 4^6 6-element strings in total. This will be our set X in the above template. We wish to exclude any string which lacks 1s; we can call the set of such strings A_1; likewise, let us identify A_2, A_3, and A_4 as containing those strings without 2s, 3s, and 4s respectively. Since A_j consists of those strings lacking js, each element of the string msut be a number other than j, of which there are 3 possibilities, regardless of what j is. Thus $|A_j| = 3^6$ for all j. Likewise, $A_i \cap A_j$ consists of strings lacking both i and j, so there would be only 2 possibilites for each element, and $|A_i \cap A_j| = 2^6$ regardless of what distinct values i and j are. Likewise, $|A_i \cap A|j \cap A_k| = 1^6$ and $|A_1 \cap A_2 \cap A_3 \cap A_4| = 0$ (which is unsurprising, since there are no strings of 1, 2, 3, and 4 which have none of 1, 2, 3, or 4 in them).*

Using the formulation above, we will find that $a_1 = 3^6$, $a_2 = 2^6$, $a_3 = 1^6$, and $a_4 = 0^6$. Thus,

$$|X - (A_1 \cup A_2 \cup A_3 \cup A_4)| = 4^6 - \binom{4}{1}3^6 + \binom{4}{2}2^6 - \binom{4}{3}1^6 + \binom{4}{4}0^6 = 1560$$

A string of numbers, in which each number has to appear at least once, corresponds to one of our twelvefold-way boxes! Strings have order, and we are considering distinct elements of the string as distinguishable, so we are considering the first row; the requirement to have every element appear at least once is the third row, so we are filling in cell (3) here. Generalizing the above argument, we can ask the following question formally:

Question 4: *How many ways are there to put k labeled balls in n boxes such that each box contains at least one ball?*

Answer 4: *Let X consist of the set of all ways of putting k balls in n boxes (so that $|X| = n^k$, by known enumeration statistics); let A_i consist of those ways which leave box i empty. Then the quantity described above is $|X - (A_1 \cup A_2 \cup \cdots \cup A_n)|$.*

Let a_i be defined as $\| \cap_{j \in S} A_j \|$ for $|S| = i$. We shall see that the actual membership of S is irrelevant. If S has i members, then a_i consists of the intersection of i different A_j, so this number counts those distributions which leave i specific boxes empty. Thus, a_i counts distributions of k balls among the $n - i$ boxes which are not predestined to be empty; there are $(n - i)^k$ such distributions, so $a_i = (n - i)^k$.

Using the inclusion-exclusion principle, it thus follows that

$$|X - (A_1 \cup A_2 \cup \cdots \cup A_n)| = |X| + \sum_{i=1}^{n} (-1)^i \binom{n}{i} a_i = n^k + \sum_{i=1}^{n} (-1)^i \binom{n}{i} (n-i)^k = \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n-i)^k$$

So, rejoice! We have a new entry in our table.

Choosing k el'ts from a set of size n	Free choice	No repeats	Each el't at least once
Ordered choice, distinguished el'ts	n^k	$k!\binom{n}{k}$	$\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$
Unordered choice, distinguished el'ts	$\binom{n+k-1}{n-1}$	$\binom{n}{k}$	$\binom{k-1}{n-1}$
Ordered choice, undistinguished el'ts	?	$\begin{cases} 1 \text{ if } n \geq k \\ 0 \text{ if } n < k \end{cases}$	?
Unordered choice, undistinguished el'ts	?	$\begin{cases} 1 \text{ if } n \geq k \\ 0 \text{ if } n < k \end{cases}$	?

As a bonus, we now have a statement which is combinatorially trivial, but arithmetically quite non-obvious:

$$\text{For } k < n,\ \sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k = 0$$

1.3 Two freebies: cells (7) and (9); the Stirling and Bell numbers

It turns out that the division principle will give us cell (9) of the twelvefold way based on cell (3). Cell (3), recall, enumerates the number of ways to put numbered balls in numbered boxes so that no box is empty. Now, since every box has at least one uniquely identified ball within, any permutation of the boxes results in a different configuration (note that this is not true when boxes can be empty, since swapping two empty boxes will result in a different permutation). Thus, the situations enumerated in cell (3) can be organized into groups of $n!$ up to permutation-equivalence; each of these groups corresponds to a placement of numbered balls into *unlabeled* boxes, which is what cell (9) counts. Thus, by the division principle, the enumeration statistic for cell (9) is $\frac{1}{n!}\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$.

Note that this gives us another combinatorially straightforward proof of an arithmetically difficult result! It is far from obvious, from an artihmentic point of view, that $\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$ should be divisible by $n!$, but since we have found a set enumerated by their quotient, it must be an integer.

The above formula for the contents of cell (9) is regarded as sufficiently fundamental that it is given a name: These are known as *Stirling numbers of the second kind*, and denoted as such:

$$S(k,n) = \frac{1}{n!}\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$$

Thus, for instance, cell (3) is canonically written as $n!S(k,n)$.

Let us consider the contents of cell (7) now. This counts the number of ways of putting k labeled balls into n unlabeled boxes, permitting some of the boxes to be empty. This is easily determined by breaking down into cases. Any number of boxes between 0 and $n-1$ can be empty, so that any number of boxes between 1 and n can be used. If we consider the case where exactly i boxes are used, then we are distributing k balls between i boxes, which can be done in $S(k,i)$ ways: Adding together the cases corresponding to different values of i, we get that cell (7) consists of $S(k,1)+S(k,2)+\cdots+S(k,n) = \sum_{i=1}^{n} S(k,i)$.

We may now assert a nearly-complete list of twelvefold-way enumerations:

Choosing k el'ts from a set of size n	Free choice	No repeats	Each el't at least once
Ordered choice, distinguished el'ts	n^k	$k!\binom{n}{k}$	$n!S(k,n) = \sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$
Unordered choice, distinguished el'ts	$\binom{n+k-1}{n-1}$	$\binom{n}{k}$	$\binom{k-1}{n-1}$
Ordered choice, undistinguished el'ts	$\sum_{i=1}^{n} S(k,i)$	$\begin{cases} 1 \text{ if } n \geq k \\ 0 \text{ if } n < k \end{cases}$	$S(k,n) = \frac{1}{n!}\sum_{i=0}^{n}(-1)^i\binom{n}{i}(n-i)^k$
Unordered choice, undistinguished el'ts	?	$\begin{cases} 1 \text{ if } n \geq k \\ 0 \text{ if } n < k \end{cases}$	?

Note that the Stirling numbers can also be thought of as enumerating the ways that a set of size k can be partitioned into n nonempty subsets. For instance, $S(4,2) = 7$ corresponds to the fact that $\{1,2,3,4\}$ can be partitioned into two sets as $\{\{1\},\{2,3,4\}\}$, $\{\{2\},\{1,3,4\}\}$, $\{\{3\},\{1,2,4\}\}$, $\{\{4\},\{1,2,3\}\}$, $\{\{1,2\},\{3,4\}\}$, $\{\{1,3\},\{2,4\}\}$, or $\{\{1,4\},\{2,3\}\}$. A question sometimes worth asking is how many ways a set of size n can be partitioned, regardless of how many parts the partition is into. This can easily be answered as $\sum_{i=1}^{n} S(n,i)$. This quantity is known as the *Bell number*, denoted B_n. For instance, $B_3 = S(3,1) + S(3,2) + S(3,3) = 1 + 3 + 1 = 5$, corresponding to the five ways to partition $\{1,2,3\}$: $\{\{1,2,3\}\}$, $\{\{1,2\},\{3\}\}$, $\{\{1,3\},\{3\}\}$, $\{\{1\},\{2,3\}\}$, and $\{\{1\},\{2\},\{3\}\}$.

2 Using the Twelvefold Way

2.1 Partition numbers: what about cells (10) and (12)?

Even though we lack the means at present to come up with a good formula for them, it's worth discussing what cells (10) and (12) measure. The numbers on the third row are often known as *set partitions*, since, as seen above, the Stirling numbers can be considered as the number of ways to partition a set. Here, we are not looking at distinguishable elements, so while we are still performing partitions, instead of partitioning a set into smaller sets, we are partitioning an integer into smaller integers.

Let us denote by $p_n(k)$ the number of ways of expressing k as an unordered sum of n nonzero numbers. This will be our cell (12) statistic. For example, $p_3(8) = 5$, because there are 5 such partitions: $8 = 6+1+1 = 5+2+1 = 4+3+1 = 4+2+2 = 3+3+2$.

Using a similar argument to that giving the relationship between cells (7) and (9), we can come up with a formula for cell (10) as well: this cell counts the nubmer of partitions of k into n parts, some of which may be zero. There are $p_i(k)$ ways to have i nonzero terms in the partition, so ranging over all possible values of i, we see that the statistic in cell (10) will be $p_1(k) + p_2(k) + \cdots + p_n(k)$.

In addition, we have an analogue of the Bell numbers, to answer the question: how many ways can n be broken down into a sum of any number of nonzero integers. This is denoted $p(n)$, and is clearly equal to $p_1(n) + p_2(n) + \cdots + p_n(n)$. For instance, it is easy to see that $p(5) = 7$ by brute force listing of every partition:

$$5 = 4+1 = 3+2 = 3+1+1 = 2+2+1 = 2+1+1+1 = 1+1+1+1+1$$

2.2 Slight variations on known distributions

We are now equipped to answer a variety of problems, using the tools we developed here. These are slight variations on established results:

Question 5: *How many ways are there to distribute 25 identical coins among a captain, first mate, and 6 pirates if each person receives at least 2 coins, the first mate gets 3, and the captain 4?*

Answer 5: *We can start with some pre-emptive distribution: we must hand out 4 coins to the captain, 3 to the mate, and 2 to each of the sailors, so 19 of our coins are spoken for and we might as well give them out immediately. We are left with 6 coins to be freely distributed among 8 individuals. This statistic we know to be* $\binom{8+6-1}{8-1} = \binom{13}{7}$.

Note that this problem would have been a lot harder if we'd asked about distinguishable coins! There are still many problems that are difficult.

Notice that we could phrase the question above another way, to wit: **Question 6:** *How many non-negative integer solutions are there to the linear equation* $x_1+x_2+x_3+x_4+x_5+x_6+x_7+x_8 = 25$*, if each* $x_i \geq 2$*, and furthermore* $x_1 \geq 4$ *and* $x_2 \geq 3$*?*

Answer 6: *This is the same question as above, merely abstracted.*

Another likely question might be probabalistic:

Question 7: *15 balls are randomly distributed among 6 boxes. What is the probability that no box is empty?*

Answer 7: *Note that we said nothing about box or ball distinguishability, but we have a good idea of the process in question: drop a ball in a random box, drop another ball in a random box, etc. The standard probability calculation of dividing the desired set size by the size of the total set of possibilities has an implicit assumpion: namely that every member of the event space has identical probability. We might run into trouble using, for instance, unlabeled balls or boxes for this: for instance, putting two balls randomly into four boxes, the two possibilities 1+1+0+0 and 2+0+0+0 are not equally likely; 1+1+0+0 occurs with probability* $\frac{3}{4}$ *while 2+0+0+0 occurs with probability* $\frac{1}{4}$*. Compare to the labeled-balls-and-boxes paradigm, where each of the 16 possible configurations is equally likely. Thus, in dealing with a probabalistic question, we usually will want a distinct-balls-into-distinct-boxes statistic. Here we are looking at the probability that a random selection from among* 6^{15} *possible configurations is one of the* $S(15, 6)$ *configurations in which every box has at least one ball, so the probability is* $\frac{6!S(15,6)}{6^15} = \frac{302899156560}{470184984576} \approx 0.644$.