# Notes on counting finite sets

Murray Eisenberg

February 26, 2009

## Contents

# 0 Introduction

The purposes of these notes are:

- to construct mathematical objects—"models"—for various kinds of things that may be counted systematically;
- to formulate basic counting methods as precise mathematical theorems about such mathematical objects; and
- to prove those mathematical theorems.

In short, the aim here is to construct the mathematical infrastructure for the subject of **combinatorics**.

The fundamental objects considered are sets and functions between sets. See the *Mathematica* notebook `SetsAndFunctions.nb` for information about sets, subsets, unions, intersections, etc., and about injective (one-to-one) functions, surjective ("onto") functions, and bijective functions (one-to-one correspondences).

Only a few motivating applications are included in this draft of these notes. Consult your textbook for many more examples.

For numerical calculations, you may want to use my notebook `Combinatorics.nb`, which references some of the functions from the *Mathematica* Add-On package `Combinatorica`\`.

# 1 What is a finite set?

The empty set—the set $\{\,\}$ having no elements whatsoever—is said to be finite. The idea that a nonempty set $A$ be finite is that it has exactly $n$ elements for some positive integer $n$. And this means that, for some positive integer $n$, the set $A$ can be expressed in the form

$$A = \{a_1, a_2, \ldots, a_n\} \tag{*}$$

subject to the restriction that distinct subscripts label distinct elements of $A$, that is,

$$a_i \neq a_j \text{ whenever } i \neq j. \tag{**}$$

Thus the elements of the "standard" finite set $\{1, 2, \ldots, n\}$ with $n$ elements can be used to **count** the elements $a_1, a_2, \ldots, a_n$ of $A$.

Subscripting elements of $A$ with the integers $1, 2, \ldots, n$ amounts to having a function

$$f\colon \{1, 2, \ldots, n\} \to A,$$

with

$$f(k) = a_k \qquad (k = 1, 2, \ldots, n).$$

Condition (*) means that $f\colon \{1, 2, \ldots, n\} \to A$ is surjective (that is, "onto"); condition (**) means that the function $f\colon \{1, 2, \ldots, n\} \to A$ is injective (that is, one-to-one). Thus the idea that a set be finite may be defined as follows.

**Definition 1.** A set $A$ is said to be **finite** when either $A$ is empty or else there is some positive integer $n$ and some bijection $f\colon \{1, 2, \ldots, n\} \to A$.

In short, a set $A$ is finite if and only if it is empty or else can be put into a one-to-one correspondence with $\{1, 2, \ldots, n\}$ for some positive integer $n$. We want to call such $n$ the "number of elements" of $A$, but before doing that we must know that there is only one such $n$. This is a consequence of the following proposition.

**Proposition 2.** *If $n$ and $m$ are positive integers with $n \neq m$, then there does* not *exist a bijection $h\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, m\}$.*

The preceding proposition can be proved, but the proof is somewhat complicated and rests upon some very fundamental properties of the natural numbers. We shall simply accept the truth of the proposition. From it we can deduce the result we want:

**Corollary 3.** *Let $A$ be a nonempty finite set. Suppose $n$ and $m$ are positive integers and suppose $f\colon \{1,2,\dots,n\} \to A$ and $g\colon \{1,2,\dots,m\} \to A$ are bijections. Then $m = n$.*

*Proof.* Let $n$ and $m$, $f$ and $g$ be in the statement. Then the inverse function $g^{-1}\colon A \to \{1,2,\dots,m\}$ is also bijective. Hence the composite function

$$g^{-1} \circ f\colon \{1,2,\dots,n\} \to \{1,2,\dots,m\}$$

is bijective. But this contradicts the preceding proposition unless $n = m$. □

In view of the preceding corollary, the following definition now makes sense.

**Definition 4.** Let $A$ be a finite set. If $A \neq \emptyset$, then there is a *unique* positive integer $n$ for which $A$ can be put into one-to-one correspondence with $\{1,2,\dots,n\}$; we call $n$ the **number of elements** of $A$ and write $n = \#A$.

If $A = \emptyset$, we say that 0 is the number of elements of $A$ and write $\# A = 0$.

The number of elements of a finite set $A$is also called its **cardinality**, denoted by $\mathrm{card}(A)$.

**Examples 5.** (1) For a positive integer $n$, the set $A = \{1,2,\dots,n\}$ is itself finite, and $\# A = n$. In fact, the identity function $i\colon \{1,2,\dots,n\} \to A$ is a bijection.

(2) Let $A$ be the set of all *even* positive integers that are less than 15; that is,

$$A = \{2,4,6,8,10,12,14\}.$$

Then $A$ is finite and $\# A = 7$ because the function $f\colon \{1,2,3,4,5,6,7\} \to A$ defined by

$$f(j) = 2j \qquad (1 \leq j \leq 7)$$

is bijective.

(3) The set

$$A = \{0,1,2,3\}$$

is finite because the function $f\colon \{1,2,3,4\} \to A$ defined by

$$f(j) = j - 1 \qquad (1 \leq j \leq 4)$$

is bijective.

(4) The set

$$\mathbb{N}^* = \{1,2,3,\dots\}$$

of all natural numbers is *not* finite. In fact, just suppose $\mathbb{N}^*$ is finite and let $n = \# \mathbb{N}^*$. Since $\mathbb{N}^*$ is not empty, then $n > 0$. Since $n = \# \mathbb{N}^*$, there is some bijection

$$f\colon \{1,2,\dots,n\} \to \mathbb{N}^*.$$

Construct a new bijection

$$g\colon \{1,2,\dots,n,n+1\} \to \mathbb{N}^*$$

by the formula

$$g(j) = \begin{cases} 1 + f(j) & \text{if } 1 \leq j \leq n \\ 1 & \text{if } j = n+1. \end{cases}$$

Then the function

$$g^{-1} \circ f\colon \{1,2,\dots,n\} \to \{1,2,\dots,n,n+1\}$$

is also a bijection. But this is impossible according to Proposition 2.

**Definition 6.** A set is said to be **infinite** if it is not finite.

According the preceding example, the set $\mathbb{N}^*$ of all positive integers is infinite. You may show, similarly, that the set $\mathbb{N}$ of all natural numbers—0 together with all positive integers—is also infinite.

Since the composition of two bijections is itself a bijection, any set $A'$ that can be put into one-to-one correspondence with a given finite set $A$ is also finite and has the same number of elements as $A$:

**Proposition 7.** *Let $A$ and $A'$ be sets, and suppose there is some bijection $g\colon A \to A'$. Then $A$ is finite if and only if $A'$ is finite, and in this case*

$$\#(A') = \#(A).$$

We shall accept the following result without proving it.

**Proposition 8.** *Let $A$ and $B$ be sets $A \subset B$. If $B$ is finite, then $A$ is also finite; moreover, in this case $\# A \leq \# B$.*

**Corollary 9.** *Let $A$ and $B$ be sets $A \subset B$. If $A$ is infinite, then $B$ is also infinite.*

According to this corollary, each of the sets $\mathbb{Z}$ (the set of all integers), $\mathbb{Q}$ (the set of all rational numbers), $\mathbb{R}$ (the set of all real numbers), and $\mathbb{C}$ (the set of all complex numbers) is infinite. In fact, each has the infinite set $\mathbb{N}$ as a subset.

Eventually, you will learn how to count the $k$-element sets of a finite set. Here is a start.

**Example 10.** Let $S$ be a finite set, with $\#(S) = n$.

(a) How many 0-element subsets does $S$ have? Answer: Just 1, namely, the empty subset of $S$—the subset $\{\,\}$ that has no elements whatsoever. (The empty set is often denoted by $\emptyset$, but the notation $\{\,\}$ is very suggestive.)

(b) How many 1-element subsets does $S$ have? Answer: $n$. Explanation: Let $\mathcal{P}_1(S)$ be the set of 1-element subsets of $S$. Then the function

$$f\colon S \to \mathcal{P}_1(S),$$

defined by

$$f(x) = \{x\} \qquad\qquad (x \in S)$$

is bijective.

For example, suppose $S = \{a, b, c\}$ is a 3-element set. Then $\mathcal{P}_1(S) = \big\{\{a\}, \{b\}, \{c\}\big\}$ has 3 elements, as does $S$. However, the elements of $\mathcal{P}_1(S)$ are *not* the same as the elements of $S$ itself: no matter what $x$ is,

$$x \neq \{x\}.$$

In fact, $\{x\}$ is a set having exactly 1 element, namely, $x$, whereas $x$ is just the single element of that set.

(c) How many $n-1$-element subsets does $S$ have. Answer: $n$. Can you supply the reason?

## 2 Counting unions and cartesian products

This section concerns the finiteness of various sets—unions, cartesian products, etc.—formed from given finite sets, and the number of elements in the sets that result.

## 2.1 Sum rules

Suppose you are going to order a single beverage—either a cup of coffee or else a bottle of soda, but not both. If there are 8 kinds of coffee and 5 flavors of soda from which you can choose, then you have a total of $48 + 5 = 13$ possible choices of beverage. Why? This situation may be modelled by the union $A \cup B$, where set $A$ represents the 8 kinds of coffee and set $B$ represents the 5 flavors of soda. Then the sets $A$ and $B$ are disjoint—they have no element in common (at least if the none of sodas is coffee-flavored). Then the total number of choices is what is the sum indicated in the following proposition.

**Proposition 11** (Basic Sum Rule)**.** *Let $A$ and $B$ be* disjoint *finite sets. Then the union $A \cup B$ of $A$ and $B$ is also finite, and*

$$\#(A \cup B) = \#(A) + \#(B).$$

*Proof.* The result is obvious if $A$ is empty or $B$ is empty. So suppose that neither $A$ nor $B$ is empty. Let $m = \#(A)$ and $n = \#(B)$. By definition, there exist bijections $f\colon \{1, 2, \ldots, m\} \to A$ and $g\colon \{1, 2, \ldots, n\} \to B$. Now use $f$ and $g$ to construct a bijection $h\colon \{1, 2, \ldots, m + n\} \to A \cup B$. [*Hint:* Write $a_i = f(i)$ for each $i = 1, 2, \ldots, m$ and $b_j = g(j)$ for each $j = 1, 2, \ldots, n$, so that

$$A \cup B = \{a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n\}. \tag{*}$$

Now define $h$ so that, for each $k = 1, 2, \ldots, m + n$, the value $h(k)$ will be the $k$th element in the list (*). Then verify that the $h$ so defined is actually a bijection.] □

When $A$ and $B$ are not disjoint—when they have one or more elements in common—it is still true that $A \cup B$ is finite (see Proposition 14).

Given a list $A_1, A_2, \ldots, A_n$ of sets, their **union** is the set denoted by $\bigcup_{i=1}^{n} A_i$ or just $A_1 \cup A_2 \cup \cdots \cup A_n$ and defined to be the set

$$\{\, x : x \in A_i \text{ for at least one } i \text{ among } 1, 2, \ldots, n \,\}$$

consisting of those elements belonging to one or more of the individual sets $A_1, A_2, \ldots, A_n$.

A list $A_1, A_2, \ldots, A_n$ of sets is said to be **pairwise disjoint** when no two of them have an element in common, that is, when each two of them are disjoint.

**Theorem 12** (Sum Rule)**.** *Let $n$ be a positive integer and let $A_1, A_2, \ldots, A_n$ be finite sets that are pairwise disjoint. Then their union $\bigcup i = 1^n A_i$ is also finite, and*

$$\#\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \#(A_i).$$

*Proof.* There is nothing to prove when $n = 1$. Now use induction on the number $n$ of sets in the list. The Base Step is just the Basic Sum Rule (Proposition 11). □

If $A$ and $B$ are sets, then their **set difference**, denoted by $A \setminus B$, is defined to be the set of all elements of $A$ that are not elements of $B$. The set difference $A \setminus B$ is also called the **complement of** $A$ **in** $B$, especially in the case that $B$ is a subset of $A$.

**Proposition 13** (Difference Rule)**.** *Let $A$ be a finite set and let $B$ be a* subset *of $A$. Then the complement $A \setminus B$ of $B$ in $A$ is also finite, and*

$$\#(A \setminus B) = \#(A) - \#(B).$$

*Proof.* Both $B$ and the complement $A \setminus B$ are subsets of $A$, so according to the Difference Rule 8 these two sets are finite. Further, these two sets $B$ and $A \setminus B$ are disjoint, and

$$A = (A \setminus B) \cup B.$$

By the Basic Sum Rule (Proposition 11),

$$\#(A) = \#(A \setminus B) + \#(B).$$

The stated result follows at once. $\square$

For example, if a class of 34 students includes exactly 6 math majors, then it includes exactly $34 - 6 = 28$ students who are not math majors.

**Proposition 14** (Union Rule)**.** *Let $A$ and $B$ be finite sets (not necessarily disjoint). Then the union $A \cup B$ of $A$ and $B$ is also finite, and*

$$\#(A \cup B) = \#(A) + \#(B) - \#(A \cap B).$$

*Proof.* Apply the Sum Rule to the three pairwise disjoint subsets $A \setminus B$, $B \setminus A$, and $A \cap B$. Then apply the Difference Rule (Proposition 13). $\square$

For example, suppose a class includes 6 students who are majoring in math and 10 students who are majoring in computer science; among those are 4 students with a double major in both math and computer science. How many students in the class are majoring in math or computer science? Answer: $6 + 10 - 4 = 12$ students. Explanation: Apply the Union Rule (Proposition 14).

To count the union of three or more pairwise disjoint finite sets is ore complicated than the formula of the Union Rule (Prop. 14). Then one needs the Inclusion-Exclusion Principle: see Section 2.4.

## 2.2 Product rules: ordered selection with independent choices

How many possible outcomes are there if you first roll a die and then toss a coin? Represent the outcomes of just rolling a die by the set $A = \{1, 2, 3, 4, 5, 6\}$ of the six number that can be rolled; represent the outcomes of just tossing a coin by the set $B = \{H, T\}$, where $H$ represents heads and $T$ represents tails. Then the set of possible outcomes from first rolling the die and then tossing the coin is:

$$\{(1, H), (2, H), (3, H), (4, H), (5, H), (6, H), (1, T), (2, T), (3, T), (4, T), (5, T), (6, T)\}$$

This set consists of all the possible ordered pairs that can be formed by first selecting an element of $A$ and then selecting an element of $B$.

In general, the **(cartesian) product** of sets $A$ and $B$, denoted by $A \times B$, is defined to be the set $\{\, (a, b) : a \in A, b \in B \,\}$ consisting of all *ordered* pairs whose first entry belongs to $A$ and whose second entry belongs to $B$.

**Proposition 15** (Basic Product Rule)**.** *Let $A$ and $B$ be finite sets. Then the cartesian product $A \times B$ of $A$ and $B$ is also finite, and*

$$\#(A \times B) = \#(A) \cdot \#(B).$$

*Proof.* In case $B$ is empty, then so is $A \times B$; in this case $A \times B$ is certainly finite, and $\#(A \times B) = 0 = \#(A) \cdot 0 = \#(A) \cdot \#(B)$.

Consider now the case that $B$ is not empty. Let $n = \#(B)$. Then we may write

$$B = \{b_1, b_2, \ldots, b_n\}$$

with $b_i \neq b_j$ whenever $i \neq j$.

(In case $n = 1$, there is an obvious bijection $f \colon A \to A \times B = A \times \{b_1\}$, but there is actually no need to consider this case separately.)

Now suitably apply the Sum Rule (Theorem 12) to the union of pairwise disjoint sets whose union is the desired cartesian product. For guidance, you may wish to examine the case $n = 2$ first. $\square$

The example of rolling a die and then tossing a coin was modelled by a cartesian product because the situation, as described, involved an **order**—first one thing and then another. But suppose you do not care whether you roll the die first and then toss the coin or, instead, toss the coin first and then roll the die. In this case, the outcomes of interest would no longer be ordered pairs but simply 2-element sets $\{a, b\}$ consisting of one of the 6 numbers on the die and one of the two sides, $H$ or $T$ of the coin. Of course there would still be exactly $6 \cdot 2 = 12$ outcomes, that is, 12 such 2-element sets. But the sets $\{5, H\}$ and $\{H, 5\}$ are equal and represent precisely the same outcome.

Suppose, however, you are going to toss two coins. It would be *wrong* to model this situation by the set $\big\{\{H, H\}, \{H, T\}, \{T, T\}\big\}$ of 2-element sets. Even though the sets $\{H, T\}$ and $\{T, H\}$ are equal, there are actually $2 \cdot 2 = 4$ possible outcomes: you must resort to a model with ordered pairs. To understand why, imagine painting one of the coins green and the other red before tossing them.

**Definition 16.** Given a list $B_1, B_2, \ldots, B_n$ of sets, their **(cartesian) product** is the set denoted by $B_1 \times B_2 \times \cdots \times B_n$ and defined to be the set

$$\{\, (x_1, x_2, \ldots, x_n) : x_i \in B_i \text{ for each } i = 1, 2, \ldots n, \}$$

consisting of ordered $n$-tuples $(x_1, x_2, \ldots, x_n)$ having the property that each entry $x_i$ belongs to the corresponding set $B_i$. Sometimes the notation $\prod_{i=1}^{n} B_i$ is used for such a cartesian product.

In the following proposition, the notation $\prod_{i=1}^{n} c_i$ is used to denote the product of numbers $c_1, c_2, \ldots, c_n$. (This notation is the analog for multiplication of the sigma notation $\sum_{i=1}^{n} c_i$ for the sum of numbers.)

**Theorem 17** (Product Rule)**.** *Let $n$ be a positive integer and let $B_1, B_2, \ldots, B_n$ be finite sets. Then their cartesian product $B_1 \times B_2 \times \cdots \times B_n$ is also finite, and*

$$\#\,(B_1 \times B_2 \times \cdots \times B_n) = \prod_{i=1}^{n} \#(B_i).$$

*Proof.* There is nothing to prove when $n = 1$. Now use induction on the number $n$ of sets in the list. The Base Step is just the Basic Product Rule (Proposition 15). □

A cartesian product $B_1 \times B_2 \times \cdots \times B_n$ can be used to represent *ordered* samples formed by first selecting an element of $B_1$, then an element of $B_2$, etc., and finally an element of $B_n$—with the selection of each subsequent entry in the ordered sample being independent of the selection of every entry already selected.

For example, suppose I can choose from 6 different colors of shirts, 3 different styles of pants, 5 different patterns of ties, and 2 different jackets to wear. Then according to the Product Rule, the total number of outfits—shirt, pants, tie, and jacket—that I can choose is $6 \cdot 3 \cdot \cdot 5 \cdot 2 = 180$.

Suppose in particular that all the $B_i$ are the same set $B$: $B_1 = B_2 = \cdots = B_n = B$. Then the cartesian product $B_1 \times B_2 \times \cdots \times B_n$ consists of all ordered $n$-tuples $(b_1, b_2, \ldots, b_n)$ of elements of the set $B$. Such an ordered $n$-tuple can be regarded as a record of first selecting an element of $B$; after replacing that element back into $B$, selecting an element of $B$; after replacing that element back into $B$, selecting an element of $B$; etc. In other words, the cartesian product $B_1 \times B_2 \times \cdots \times B_n$ can be regarded as representing all possible **ordered samples of length $n$ from $A$ with replacement**.

For example, how many 5-letter "words" are there, where the letters are among the 26 lower-case letters of the alphabet (but a "word" does not actually have to be a real word having meaning to English speakers). Answer: $26 \cdot 26 \cdot 26 \cdot 26 \cdot 26 = 26^5 = 11,881,376$ words.

Another, superficially different, model for ordered sampling with replacement is formulated in Section 5. That model is expressed in terms of functions.

Often the Product Rule is used in conjunction with the Union Rule (Proposition 14) or other rules of counting.

**Example 18.** How many 8-bit bytes are there that begin with 1 or end with 00? To count these, let $A$ be the set of bytes that begin with 1; let $B$ be the set of bytes that end with 00. The number desired is $\#(A \cup B)$. By the Product Rule, $\#(A) = 1 \cdot 2^7 = 128$ and $\#(B) = 2^6 \cdot 1 \cdot 1 = 64$. Now $A \cap B$ is the set of bytes that both begin with 1 *and* end with 00; by the Product Rule, $\#(A \cap B) = 1 \cdot 2^5 \cdot 1 \cdot 1 = 32$. Finally, by the Union Rule, $\#(A \cup B) = \#(A) + \#(B) - \#(A \cap B) = 128 + 64 - 32 = 160$.

## 2.3 Generalized product rules: ordered selection with dependent choices

Suppose the selection of an entry in an ordered sample *does* depend on what was selected as a previous entry.

**Example 19.** How many 2-letter words are there in which no letter is repeated?

To answer that question, it is correct to say that, since the first letter can be any one of 26 and the second any one of the remaining $26 - 1 = 25$, then the total number of such words is $26 \cdot 25 = 650$. It is tempting—and *wrong!*—to explain the answer by means of the Product Rule (Theorem 17 for the case of two sets (that case is just the Basic Product Rule, Proposition 15). And the reason that explanation would be wrong is that the set being counted is *not* the cartesian product of two sets! When you represent a set of ordered selections by a cartesian product, you are tacitly assuming that the second choice does not depend upon the first choice. But in this problem about 2-letter words having different letters, the choice of second letter definitely *does* depend upon what the first letter is. So what is the model for the set of all these words with different letters?

Represent the alphabet by the set

$$A = \{a_1, a_2, \dots, a_{26}\}.$$

For $i = 1, 2, \dots, 26$, let

$$B_i = A \setminus \{a_i\},$$

the remaining set of 25 letters. Then the set of all 2-letter words having two different letters is represented by the union

$$W = \bigcup_{i=1}^{26} (\{a_i\} \times B_i)$$

of the pairwise disjoint sets $\{a_1\} \times B_1, \{a_2\} \times B_2, \dots, \{a_{26}\} \times B_{26}$. From the Sum Rule (Theorem 12) and then the Basic Product Rule (Proposition 15),

$$\#(W) = \sum_{i=1}^{26} \#\left(\{a_i\} \times B_i\right) = \sum_{i=1}^{26} \#(\{a_i\}) \cdot \#(B_i).$$

Now $\#(\{a_i\}) = 1$ for each $i$, so that

$$\#(W) = \sum_{i=1}^{26} \#(B_i).$$

In this problem, it so happens that the sets $B_i$ all have the same number of elements, 25, no matter what $i$ is. Hence finally

$$\#(W) = \sum_{i=1}^{26} 25 = 26 \cdot 25 = 650.$$

Obviously the essential reason behind the counting in the preceding example is the same as that for the case of an actual cartesian product of two sets: the set to be counted is a union of pairwise disjoint sets to which the Sum Rule may be applied. For this reason, the principle for the kind of counting just done is often referred to as a "product rule". To make clear that what is being counted is not an actual cartesian product, we shall refer to this principle by qualifying the phrase "product rule" with the adjective "generalized". Here is a formal statement of the principle. The statement is formulated in a way that does not require actually indexing the elements of the first set $A$ that represents the initial choices.

**Proposition 20** (Basic Generalized Product Rule)**.** *Let $A$ be a finite set and for each $a \in A$ let $B_a$ be a finite set. Then the set $\bigcup_{a\in A} (\{a\} \times B_a)$ is also finite, and*

$$\#\left(\bigcup_{a\in A} (\{a\} \times B_a)\right) = \sum_{a\in A} \#(B_a).$$

*If $\#(B_a) = n$, the same number, for all $a \in A$, then*

$$\#\left(\bigcup_{a\in A} (\{a\} \times B_a)\right) = \#(A) \cdot n.$$

*Proof.* Exercise. (The proof is essentially the same as for the example of 2-letter words that have different letters.) □

**Example 21.** Set $S$ be a finite set with $\#(S) = n$. How many 2-element subsets does an $m$-element set have? Answer: $\frac{1}{2}n(n-1)$. Explanation: First count the set $P$ of ordered pairs $(x_1, x_2)$ consisting of *different* elements of $S$. According to the Basic Generalized Product Rule, $\#(P) = n \cdot (n-1)$. This is *not,* however, the desired answer, since each 2-element subset $\{x, y\}$ of $S$ has been counted twice—once in the ordered pair $(x, y)$ and then again in the ordered pair $(y, x)$. so the desired answer is $\frac{1}{2}n \cdot (n-1)$.

As an exercise, formalize the preceding argument about each 2-element subset of $S$ being counted twice by ordered pairs of distinct elements. (*Hint:* Think of forming two copies of each 2-element subset $\{x, y\}$ of $S$, say with one copy being painted red and the other copy being painted green.)

Next consider a situation of selecting ordered triples, where the choices available for the second entry depend upon what the first entry is, and in turn the choices available for the third entry depend upon what both the first and second entries are.

**Proposition 22** (Generalized Product Rule—3-stage case)**.** *Let $A$ be a finite set. For each $a \in A$, let $B_a$ be a finite set. For each ordered pair $(a, b)$ with $a \in A$ and $b \in B_a$, let $C_{(a,b)}$ be a finite set. Then the set $\bigcup_{a\in A} \bigcup_{b\in B_a} \{a\} \times B_a \times C_{(a,b)}$ is also finite, and*

$$\#\left(\bigcup_{a\in A} \bigcup_{b\in B_a} \{a\} \times B_a \times C_{(a,b)}\right) = \sum_{a\in A} \sum_{b\in B_a} \#(C_{(a,b)}).$$

*Proof.* Use the Basic Generalized Product Rule twice. □

The **Generalized Product Rule** is the generalization of Propositions 20 and 22 to the situation of forming ordered $k$-tuples, for an arbitrary $k \geq 2$, where the choice of each entry in a $k$-tuple depends upon the choice of all the preceding entries in that $k$-tuple. Just formulating this rule would be rather complicated, so it will not be attempted here. However, you should feel free to use the Generalized Product Rule!

Here is one special case of the Generalized Product Rule that arises frequently, where all the entries of the ordered $k$-tuples are from the same underlying set $A$.

**Proposition 23** (Ordered sampling without replacement)**.** *Let $B$ be a nonempty finite set with $n = \#(B)$. Let $r$ be an integer with $1 \leq r \leq n$. Let $P_r(B)$ denote the set of all ordered $r$-tuples whose entries are distinct elements of $B$. Then $P_r(B)$ is also finite, and*

$$\#\big(P_r(B)\big) = n(n-1)(n-2)\cdots(n-[r-1]) = n(n-1)(n-2)\cdots(n-r+1).$$

Although the preceding proposition can be derived immediately from the Generalized Product Rule (which we have not actually stated)—or by an easy induction from the Basic Generalized Product Rule (which we did state and prove)—another approach to the same conclusion will be taken below. That approach regards an ordered sample of size $r$ from a set $B$ without replacement as an injective function from $\{1, 2, \ldots, r\}$ to $B$.

## 2.4 Inclusion-Exclusion Principle

For two finite sets $A$ and $B$, the Union Rule (Proposition 14) gave the result

$$\#\,(A \cup B) = \#\,(A) + \#\,(B) - \#\,(A \cap B).$$

The total number is obtained by adding the counts of the elements that belong to $A$ and the elements that belong to $B$ and then subtracting out the doubly-counted elements, namely, those elements that belong to both $A$ and $B$.

For the union of three finite sets $A$, $B$, and $C$, the counting would start out the same way: Add the counts of the elements that belong to each of $A$, $B$, and $C$. Then subtract the doubly-counted elements—the elements that belong to some two of the three sets, that is, to $A \cap B$, to $A \cap C$, or to $B \cap C$. But then some elements that were doubly-counted have been subtracted twice—those that belong to all three of the sets, that is, to $A \cap B \cap C$. So the result should be given by the formula in the following proposition.

**Proposition 24** (Principle of Inclusion-Exclusion—Case of 3 sets)**.** *Let $A$, $B$, and $C$ be finite sets. Then their union $A \cup B \cup C$ is also finite, and*

$$\begin{aligned}&\#\,(A \cup B \cup C) = \\ &\quad \#\,(A) + \#\,(B) + \#\,(C) - \#\,(A \cap B) - \#\,(A \cap C) - \#\,(B \cap C) + - \#\,(A \cap B \cap C)\end{aligned}$$

*Proof.* Start with

$$A \cup B \cup C = (A \cup B) \cup C.$$

By the Union Rule (Proposition 14), the set $(A \cup B) \cup C$ is finite, and

$$\#\,\big((A \cup B) \cup C\big) = \#\,(A \cup B) + \#\,(C) - \#\,\big((A \cup B) \cap C\big). \qquad (*)$$

By the Union Rule again,

$$\#\,(A \cup B) = \#\,(A) + \#\,(B) - \#\,(A \cap B). \qquad (**)$$

Now

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C),$$

so by the Union rule again,

$$\#\,\big((A \cup B) \cap C\big) = \#\,(A \cap C) + \#\,(B \cap C) - \#\,\big((A \cap C) \cap (B \cap C)\big).$$

But

$$(A \cap C) \cap (B \cap C) = A \cap B \cap C,$$

so that

$$\#\,\big((A \cup B) \cap C\big) = \#\,(A \cap C) + \#\,(B \cap C) - \#\,A \cap B \cap C. \qquad (***)$$

Now substitute (**) and (***) into (*) to obtain the stated formula. ☐

A similar argument will establish the analogous result for the union of four finite sets. Here is the general result.

**Theorem 25** (Principle of Inclusion-Exclusion (PIE))**.** *Let $A_1$, $A_2$, ..., $A_n$ be a finite list of finite sets. Define*

$$\begin{aligned}
S_1 &= \sum_i \#(A_i), \\
S_2 &= \sum_{i \neq j} \#(A_i \cap A_j), \\
S_3 &= \sum_{i \neq j, i \neq k, j \neq k} \#(A_i \cap A_j \cap A_k), \\
&\vdots \\
S_n &= \#(A_1 \cap A_2 \cap \cdots \cap A_n),
\end{aligned}$$

*in other words, for each $r = 1, 2, \ldots, n$, the set $S_r$ is the sum of the $\binom{n}{r}$ numbers of elements in $r$-wise intersections of the sets $A_1, A_2, \ldots, A_n$. Then*

$$\#\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{r=1}^{n} (-1)^r S_r.$$

*Proof.* Use induction on $n$. □

In the preceding formula, the $\binom{n}{r}$ denotes the **binomial coefficient** "$n$ above $r$". By definition,

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

For more information, see Definition 36.

# 3 The Pigeonhole Principle

Suppose that a flock of $r$ pigeons fly into $n$ pigeonholes to roost. If $r > n$, then at least one pigeonhole will contain more than one pigeon. (These can be real pigeons flying into real pigeonholes. Or the "pigeonholes" could be the little compartments in an old-fashioned desk that hold letters and other papers, and the "pigeons" could be letters that are distributed among the compartments.)

That statement seems "obvious". But why is it actually so? The mathematical justification is called the **Pigeonhole Principle** or the **Dirichlet drawer principle**. The former term will be used here.

## 3.1 The basic Pigeonhole Principle

The Pigeonhole Principle can be modeled mathematically either in terms of sets or in terms of functions. To start, here is the model in terms of sets. Instead of referring to pigeons and pigeonholes, let us speak a bit more abstractly of objects and boxes: If $r > n$ objects are going to be distributed among $n$ boxes, then at least one of the boxes must get more than one object.

To model this situation mathematically, represent the set of $r$ objects by an $r$-element set $A$ and the $n$ boxes by finite sets $A_1, A_2, \ldots, A_n$. Think, though, not of the boxes themselves, but of which of the objects from $A$ each of these boxes contain. In other words, in the mathematical representation:

- Each of the sets $A_j$ is a subset of $A$.

Since a particular object cannot be put into two different boxes at the same time, assume:

- The sets $A_j$ are pairwise disjoint.

That all the objects are distributed among the boxes means:

- The set $A$ is the union of its subsets $A_j$.

With this model, it is now easy to formulate the Pigeonhole Principle as a mathematical statement and to prove it.

**Theorem 26** (Pigeonhole Principle)**.** *Let $A_1, A_2, \ldots, A_n$ be* pairwise disjoint *subsets of a finite, $r$-element set $A$ with $A = A_1 \cup A_2 \cup A_n$. If $r > n$, then $\#(A_j) > 1$ for some $j \in \{1, 2, \ldots, n\}$.*

*Proof.* Assume $r > n$. Just suppose that $\#(A_j) \leq 1$ for each $j = 1, 2, \ldots, n$. By the Sum Rule (Theorem 12),

$$r = \#(A) = \sum_{j=1}^{n} \#(A_j) \leq n \cdot 1 = n.$$

This contradicts the assumption that $r > n$. □

In practice, the Pigeonhole Principle is often used informally in much the way it was originally formulated: one speaks of distributing so many pigeons among so many pigeonholes.

As simple-minded as it may seem, the Pigeonhole Principle has many fascinating and often unexpected consequences. The only one considered here will be the proof that a function from an $r$-element set to an $n$-element set cannot be injective when $r > n$. To prepare for the proof, the following function-theoretic model for the Pigeonhole principle is introduced.

Represent the set of pigeonholes—the set of "boxes" in the more abstract way of thinking about the situation—by a finite set $B$. Represent the set of pigeons—the objects to be distributed into the boxes—by a finite set $A$. Finally, represent the act of distributing the objects among the boxes as a function $f\colon A \to B$; for each object $a \in A$, the value $f(a) \in B$ represents the box into which the object $a$ is put. To say that some box must hold more than one of the objects is to say that some element of $B$ is the value of $f$ at more than one element of $A$.

**Corollary 27** (Pigeonhole Principle—function version)**.** *Let $f\colon A \to B$ be a function from an $r$-element finite set $A$ to an $n$-element finite set $B$. If $r > n$, then there exist elements $a, a' \in A$ with $a \neq a'$ such that $f(a) = f(a')$.*

*Proof.* Assume $r > n$. Write $B = \{b_1, b_2, \ldots, b_n\}$. For each $j = 1, 2, \ldots, n$, let $A_j$ be the subset of $A$ defined by

$$A_j = \{\, a \in A : f(a) = b_j \,\}.$$

Then the sets $A_1, A_2, \ldots, A_n$ are pairwise disjoint, and $A = A_1 \cup A_2 \cup \cdots \cup A_n$. By the Pigeonhole Principle (Theorem 26), there exists some $j$ with $\#(A_j) > 1$, that is $\#(A_j) \geq 2$. For such $j$, there exist $a, a' \in A_j$ with $a \neq a'$. But by definition of $A_j$, the value $f(a) = b_j = f(a')$. □

Note that in the preceding Corollary, the number $n$ must be strictly positive: If $r > n = 0$, there is no function whatsoever from an $r$-element set to the 0-element empty set.

Another way to state the preceding corollary is the following:

**Proposition 28.** *Let $f\colon A \to B$ be a function from an $r$-element finite set $A$ to an $n$-element finite set $B$. If $r > n$, then $f$ cannot be injective.*

## 3.2 The Generalized Pigeonhole Principle

There's another way to state the conclusion of the Pigeonhole Principle's function version (Prop. 27). It helps to introduce a bit of notation for this.

Again let $f\colon A \to B$ be a function. For an element $b \in B$, form the 1-element subset $\{b\}$ of $B$; the **inverse image** of $\{b\}$ is the subset of $A$ denoted by $f^{-1}\big(\{b\}\big)$ and defined by

$$f^{-1}\big(\{b\}\big) = \{\, a \in A : f(a) = b \,\}.$$

In other words, the inverse image $f^{-1}\big(\{b\}\big)$ consists of all those elements of $a$ of the domain $A$ that $f$ maps to that particular element $b$ of the codomain.

For example, take $A = \{1, 2, 3, 4\}, B = \{7, 8, 9\}$ and define $f\colon A \to B$ by $f(1) = 7, f(2) = 8, f(3) = 7, f(4) = 7$. Then $f^{-1}\big(\{7\}\big) = \{1, 3, 4\}$, $f^{-1}\big(\{8\}\big) = \{2\}$, and $f^{-1}\big(\{9\}\big) = \emptyset$.

In general, for $f\colon A \to B$ and $b \in B$, the function $f$ is constant on $f^{-1}\big(\{b\}\big)$.

The function version of the Pigeonhole Principle (Cor. 27) may now be stated: Let $f\colon A \to B$ be a function from an $r$-element set $A$ to an $n$-element set $B$. If $r > n$, then there is some $b \in B$ with $\#\left(f^{-1}\big(\{b\}\big)\right) \geq 2$.

In that statement, if you write the hypothesis $r > n$ in the form $r > k \cdot n$ with $k = 1$, then the conclusion claims the existence of some $b \in B$ with $\#\left(f^{-1}\big(\{b\}\big)\right) \geq k + 1$. This holds even when $k > 1$:

**Theorem 29** (Generalized Pigeonhole Principle)**.** *Let $f\colon A \to B$ be a function from an $r$-element set $A$ to an $n$-element set $B$. If $k$ is a positive integer for which $r > k \cdot n$, then there is some $b \in B$ for which $\#\left(f^{-1}\big(\{b\}\big)\right) \geq k + 1$.*

The proof is similar to the proof of the basic Pigeonhole Principle and so is left as an exercise.

**Example 30.** Each day a news digest web site displays an ad randomly selected from a bank of 30 ads. Then in any 100-day period, some ad must be displayed at least 4 times.

In fact, $100 > 3 \cdot 30$ and $4 = 3 + 1$. So the conclusion is an application of the Generalized Pigeonhole Principle, because we may take $A$ to be the set of 100 days, $B$ to be the set of 30 ads, and $f\colon A \to B$ to be the function defined by $f(a) =$ the ad displayed on day $a$.

There is another way to formulate the conclusion of the Generalized Pigeonhole Principle. For it, we use the following definition.

**Definition 31.** The **ceiling** $\lceil x \rceil$ of a real number $x$ is the least integer $m$ for which $x \leq m$.

For example, $\lceil 4.001 \rceil = 5$ and $\lceil 5 \rceil = 5$.

**Corollary 32.** *Let $f\colon A \to B$ be a function from an $r$-element set $A$ to an $n$-element set $B$. Then there is some $b \in B$ for which*

$$\#\left(f^{-1}\big(\{b\}\big)\right) \geq \left\lceil \frac{r}{n} \right\rceil.$$

*Proof.* In the theorem, take

$$k = \left\lceil \frac{r}{n} \right\rceil - 1.$$

Then

$$k \cdot n = \left(\left\lceil \frac{r}{n} \right\rceil - 1\right) \cdot n < \left(\left(\frac{r}{n} + 1\right) - 1\right) \cdot n = \frac{r}{n} \cdot n = n,$$

and so the hypothesis $r > k \cdot n$ of Theorem 29 holds. □

**Example 33.** Among 100 people, at least $\left\lceil \dfrac{100}{12} \right\rceil = 9$ were born in the same month.

# 4 Counting subsets: unordered selection with no repetitions

The combinatorial problem considered here is to count the number of ways to select a sample of $r$ objects from a population of $n$ objects, where no object may be selected twice and the order in which the objects are selected is of no interest.

Model the total population from which the objects are selected by a finite set $A$ having $m$ members. Then represent a sample of $r$ distinct objects from this population by an $r$-element subset of $A$. The question is: How many $r$-element subsets of $A$ are there?

Observe that the nature of the elements of $A$ is immaterial: If $A$ and $A'$ are two $n$-element finite sets, then the number of $r$-element subsets of $A$ is exactly the same as the number of $r$-element subsets of $A'$. This observation justifies the notation used in the following definition.

**Definition 34.** For nonnegative integers $n$ and $r$, denote by $C(n, r)$ the number of $r$-element subsets of an $n$-element set. The number $C(n, r)$ is called the **number of combinations of $n$ things taken $r$ at a time**. The notation $C(n, r)$ may be read as "$n$ choose $r$".

It is easy to determine $C(n, r)$ for some particular pairs of values of $n$ and $r$. Since the only 0-element subset of an $n$-element set is the empty set:

- $C(n, 0) = 1$

There is an obvious bijection between an $n$-element set $A$ and its 1-element subsets, namely, the function that assigns to each $a \in A$ the 1-element subset $\{a\}$ of $A$. Thus:

- $C(n, 1) = n$

To select an $(n-1)$-element subset of an $n$-element set $A$ is to leave 1 element of $A$ unselected, so that there is a one-to-one correspondence between the set of all $(n-1)$-element subsets of $A$ and the set of all 1-element subsets of $A$. Hence:

- $C(n, n-1) = n$

More generally:

- $C(n, n-r) = C(n, r)$

Since the only $n$-element subset of an $n$-element set is itself:

- $C(n, n) = 1$

**Exercise 35.** Derive the formula:

$$C(n, 2) = n(n-1)/2$$

(*Hint:* First count ordered 2-member samples, without repetition, from an $n$-element set.)

The values of $C(n, r)$ in the cases above agree with the corresponding coefficients of terms $a^r\, b^{n-r}$ in the expansion of a power $(a + b)^n$ of a binomial $a + b$. These coefficients are defined as follows.

**Definition 36.** Let $n$ and $r$ be nonnegative integers. The **binomial coefficient** $\binom{n}{r}$ is defined by

$$\binom{n}{r} = \begin{cases} \frac{n!}{n!\,(n-r)!} & \text{if } 0 \le r \le n, \\ 0 & \text{otherwise.} \end{cases}$$

The notation $\binom{n}{r}$ may be read as "$n$ above $r$".

Recalling that $0! = 1$ and $1! = 1$, easy computations establish the particular values:

$$\binom{n}{0} = 1 = \binom{n}{n},$$
$$\binom{n}{1} = n = \binom{n}{n-1},$$
$$\binom{n}{2} = \frac{n(n-1)}{2} = \binom{n}{n-2}.$$

In general, since $n - (n-r) = r$,

$$\binom{n}{n-r} = \binom{n}{r}.$$

Besides the preceding ones, there are many important formulas involving binomial coefficients. The following one will be needed to count the number of $r$-element subsets of a finite set.

**Proposition 37** (Addition Formula)**.** *Let $n$ and $r$ be integers with $0 \leq r < n+1$. Then*

$$\binom{n+1}{r} = \binom{n}{r-1} + \binom{n}{r}.$$

*Proof.* This is a fairly straightforward computation that is left as an exercise. Begin with the sum on the right side and combine the two terms to obtain the binomial coefficient on the left. □

**Theorem 38.** *For nonnegative integers $n$ and $r$, the number $C(n,r)$ of $r$-element subsets of an $n$-element set is the binomial coefficient $\binom{n}{r}$.*

*Proof.* If $r > n$, then the result is trivially true: from Proposition 8, $C(n,r) = 0$, and by definition $\binom{n}{r} = 0$.

Now use induction on $n$ to prove that, for all $n \geq 0$, the equality $C(n,k) = \binom{n}{k}$ holds for all $k$ with $0 \leq k \leq n$.

**Base step ($n = 0$):** The 0-element set (the empty set $\{\,\}$) has only one subset, namely, itself. Hence $C(0,0) = 1$. But $\binom{0}{0} = 0!/(0!0!) = 1/1 = 1$.

**Inductive step:** Let $n \geq 0$ and assume that $C(n,k) = \binom{n}{k}$ for all $k$ with $0 \leq k \leq n$.

Let $r$ be an integer with $0 \leq r \leq n+1$. It must be deduced that $C(n+1,r) = \binom{n+1}{r}$.

Let $A$ be any $(n+1)$-element set, and let $\mathcal{S}$ be the collection of all $r$-element subsets of $A$, so that

$$C(n+1,r) = \#(\mathcal{S}).$$

If $r = 0$, then $C(n+1,r) = 1$ as already noted [$(\mathcal{S})$ has only the empty set as a member]; and if $r = 0$, then $\binom{n+1}{r} = (n+1)!/(0!)(n+1-0)! = 1$ also. Thus $C(n+1,r) = \binom{n+1}{r}$ in case $r = 0$.

If $r = n+1$, then $C(n+1,n+1) = 1$ as already noted [$(\mathcal{S})$ has only the entire set $A$ as a member]; and if $r = n+1$, then $\binom{n+1}{r} = (n+1)!/(n+1)!(n+1-[n+1])! = 1$ also. Thus $C(n+1,r) = \binom{n+1}{r}$ in case $r = n+1$.

Now suppose $0 < r < n+1$, that is, $1 \leq r \leq n$.

Write $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$. For each $r$-element subset $S$ of $A$, either $a_{n+1} \in S$ or else $a_{n+1} \notin S$. So the collection $\mathcal{S}$ of all $r$-element subsets of $A$ may be written as the union of two disjoint collections—

$$\mathcal{S} = \mathcal{N} \cup \mathcal{Y}$$

—where

$$\mathcal{N} = \{\, S : S \subset A, \#(A) = r, \text{and } a_{n+1} \notin S \,\},$$
$$\mathcal{Y} = \{\, S : S \subset A, \#(A) = r, \text{and } a_{n+1} \in S \,\}.$$

(The notation was chosen to suggest "no" for $\mathcal{N}$ and "yes" for $\mathcal{Y}$.) By the Basic Sum Rule (Proposition 11),

$$\#(\mathcal{S}) = \#(\mathcal{N}) + \#(\mathcal{Y}),$$

and so

$$C(n+1, r) = \#(\mathcal{N}) + \#(\mathcal{Y}) \tag{*}$$

First count the sets belonging to $\mathcal{N}$. Each $S \in \mathcal{N}$ is an $r$-element subset of the $n$-element set $A \setminus \{a_{n+1}\} = \{a_1, a_2, \ldots, a_n\}$, and conversely. Hence $\#(\mathcal{N}) = C(n, r)$, so by the inductive assumption

$$\#(\mathcal{N}) = \binom{n}{r}. \tag{**}$$

Next count the sets belonging to $\mathcal{Y}$. Each $S \in \mathcal{Y}$ can be written uniquely in the form

$$S = S' \cup \{a_{n+1}\} \text{ where } S' \subset A \setminus \{a_{n+1}\} \text{ and } \#(S') = r - 1,$$

namely,

$$S' = S \setminus \{a_{n+1}\}.$$

And conversely, each subset $S'$ of $A \setminus \{a_{n+1}\}$ with $\#(S') = r - 1$ arises in this way from exactly one set $S \in \mathcal{Y}$, namely, from $S = S' \cup \{a_{n+1}\}$. Thus there is a one-to-one correspondence between sets $S$ belonging to $\mathcal{Y}$ and $(r-1)$-element subsets $S'$ of the $n$-element set $A \setminus \{a_{n+1}\}$. This means that

$$\#(\mathcal{Y}) = C(n, r-1).$$

By the inductive assumption again,

$$\#(\mathcal{Y}) = \binom{n}{r-1} \tag{**}$$

From (*), (**), and (***),

$$C(n+1, r) = \binom{n}{r} + \binom{n}{r-1}.$$

According to the Addition Formula for binomial coefficients (Proposition 37), the sum on the right is exactly $\binom{n+1}{r}$. □

Do you see why, in the proof of the inductive step above, it was necessary to treat separately the cases $r = 0$ and $r = n + 1$?

Every subset of an $n$-element finite set is also finite and has at most $n$ elements. In view of the relation $C(n, r) = \binom{n}{r}$, it follows that, for any $n$-element finite set $A$,

$$\#(A) = \sum_{r=o}^{n} \binom{n}{r}.$$

The sum on the right may be written in the more complicated form $\sum_{r=o}^{n} \binom{n}{r} 1^r \cdot 1^{n-r}$. This is a special case of the Binomial Formula:

**Theorem 39** (Binomial Formula)**.** *Let $a$ and $b$ be real (or complex) numbers and let $n$ be a nonnegative integer. Then*

$$\boxed{(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r\, b^{n-r}.}$$

*Proof.* The computations are easier when the binomial $a+b$ is of the special form $a+1$. To reduce to this special form, substitute $c = a/b$, so that

$$(a+b)^n = (c\,b+b)^n = b^n(c+1)^n.$$

Now prove the formula for the expansion of $(c+1)^n$ and multiply the result by $b^n$ to obtain the desired formula. What needs to be proved about $(c+1)^n$ is that

$$(c+1)^n = \sum_{r=0}^{n} \binom{n}{r} c^r$$

for every nonnegative integer $n$. To prove that, use induction on $n$. You will need to use the Addition Formula (Proposition 37) in the inductive step. $\square$

A consequence of the Binomial Formula is that

$$2^n = \sum_{r=o}^{n} \binom{n}{r}.$$

As observed above, the sum on the right is just the number of subsets of an $n$-element set. So this provides a proof that ***an $n$-element set has exactly $2^n$ subsets***. A different proof, not involving use of the Binomial Theorem, appears with Theorem 42, below.

**Exercise 40.** Give a direct proof, using induction, that the number of subsets of an $n$-element set is $2^n$.

## 5 Sets of functions between finite sets

Suppose you sample in order $r$ times, with replacement, from a set $B$. Such an ordered sample may be modelled by a function $f\colon \{1, 2, \ldots, r\} \to B$, where for each $i = 1, 2, \ldots, r$, the value $f(i)$ is the element of $B$ chosen at the $i$th selection. Then the set of all length $r$ **ordered samples, with replacement**, from $B$ may be modelled by the set of all functions from the $r$-element set $\{1, 2, \ldots, r\}$ to $B$. (Another model for the same thing was formulated as a special case of the Product Rule—see page 7.)

**Theorem 41** (Number of functions)**.** *Let $A$ and $B$ be finite sets. Then the set $\mathcal{F}$ of all functions from $A$ to $B$ is also finite, and*

$$\#(\mathcal{F}) = \#(B)^{\#(A)}.$$

*Proof.* Let $r = \#(A)$ and write $A = \{a_1, a_2, \ldots, a_r\}$. Each function $f\colon A \to B$ may then be represented by the ordered $m$-tuple $\big(f(a_1), f(a_2), \ldots, f(a_r)\big)$ of its values, which is just an element of the cartesian product $\prod_{i=1}^{r} B_i$ where $B_i = B$ for each $i$. Conversely, each ordered $r$-tuple $(b_1, b_2, \ldots, b_r) \in \prod_{i=1}^{r} B_i$ has the form $(b_1, b_2, \ldots, b_r) = \big(f(a_1), f(a_2), \ldots, f(a_r)\big)$ for a unique function $f\colon A \to B$, namely, the function defined by $f(a_i) = b_i$ for each $i = 1, 2, \ldots, r$.

Thus there is a bijection between the set $\mathcal{F}$ of all functions from $A$ to $B$ and the set $\prod_{i=1}^{r} B_i$ of all ordered $r$-tuples of elements of $B$. Hence $\mathcal{F}$ is finite, and

$$\#\,(\mathcal{F}) = \#\left(\prod_{i=1}^{r} B_i\right).$$

By the Product Rule (Theorem 17),

$$\#\left(\prod_{i=1}^{r} B_i\right) = \prod_{i=1}^{r} \#\,(B_i).$$

Since $B_i = B$ for every $i = 1, 2, \ldots, r$, the product on the right is just $\#\,(B)^r = \#\,(B)^{\#\,(A)}$. Hence $\#\,(\mathcal{F}) = \#\,(B)^{\#\,(A)}$. $\square$

Because of the preceding formula, the set of all functions from a set $A$ to a set $B$ is often denoted by $B^A$. Then when both $A$ and $B$ are finite, the formula in the preceding proposition takes the pleasing form

$$\#\left(B^A\right) = \#(B)^{\#(A)}.$$

The preceding proof shows that our two models for ordered sampling with replacement from a finite set are essentially the same. These two models are:

1. the cartesian product $B_1 \times B_2 \times \cdots B_r$ with $B_1 = B_2 = \cdots B_r = B$ for an $n$-element set $B$; and

2. the set $B^A$ of all functions from an $r$-element set $A$ to an $n$-element set $B$.

One important application of Theorem 41 is to count the number of subsets of a finite set. The set of all subsets of a set $A$ is called the **power set** of $A$ and is denoted by $\mathcal{P}(A)$. For counting the number of elements in $\mathcal{P}(A)$, a representation of subsets of $A$ by certain functions will be used.

Let $A$ be a finite set. For a subset $S$ of $A$, the **characteristic function of $S$ (in $A$)** is the function

$$c_S \colon A \to \{0,1\}$$

defined by

$$c_S(a) = \begin{cases} 1 & \text{if } a \in S, \\ 0 & \text{if } a \notin S. \end{cases}$$

Thus the characteristic function $c_S$ of a subset $S$ of $A$ is a certain member of the set $\{0,1\}^A$ of all functions from $A$ to $\{0,1\}$.

**Theorem 42.** *Let $A$ be a finite set. Then the power set $\mathcal{P}(A)$ is also finite, and*

$$\#\big(\mathcal{P}(A)\big) = 2^{\#(A)}.$$

*Proof.* Define a function

$$\phi \colon \mathcal{P}(A) \to \{0,1\}^A$$

by

$$\phi(A) = c_A \qquad (\text{for all } A \in \mathcal{P}(A)).$$

It is easy to show that $\phi$ is a one-to-one correspondence. Since $\#(\{0,1\}) = 2$, the result now follows from Theorem 41 □

In short: a set with $n$ elements has exactly $2^n$ subsets.

**Remark 43.** The number of elements of the set $B^A$ of all functions from a finite set $A$ to a finite set $B$ depends only on $\#(A)$ and $\#(B)$, not on the sets $A$ and $B$ themselves. That is, if $\#(A') = \#(A)$ and $\#(B') = \#(B)$, then the number of functions from $A$ to $B$ is the same as the number of functions from $A'$ to $B'$.

The reason this is so is that there is a one-to-one correspondence between the function sets $B^A$ and $B'^{A'}$. (As an exercise, construct such a one-to-one correspondence. Begin with bijections $\phi \colon A \to A'$ and $\psi \colon B \to B'$.)

## 5.1 Permutations

This section concerns the problem of counting **ordered arrangements** of a finite set.

**Example 44.** Problem: How many ways can you arrange in a row the 3 letters `D`, `O`, `G` of the word `DOG`? That is, how many "strings" can you form using all of these three letters, using each letter only once? Solution: 6, because the set of all such strings is:

$$\{\mathtt{DOG}, \mathtt{DGO}, \mathtt{ODG}, \mathtt{OGD}, \mathtt{GDO}, \mathtt{GOD}\}$$

Observe that one of the six arrangements is the original string `DOG`. Nonetheless, all six of the displayed strings are often referred to as "rearrangements" of the letters of `DOG`. But "arrangements" would doubtless be a better term.

Here's a way to think about the preceding example that will lead to a mathematical model for ordered arrangements. Each of the arrangements of the 3 letters may be specified by: which letter takes the first position, originally occupied by `D`; which letter takes the second position, originally occupied by `O`; and which takes the third position, originally occupied by `G`.

To say which one of the 3 letters from the set $\{\mathtt{D}, \mathtt{O}, \mathtt{G}\}$ takes the position originally occupied by `D`, which one takes the position originally occupied by `O`, and which one takes the position originally occupied by `G` is to describe a *bijective function* from the set $A = \{\mathtt{D}, \mathtt{O}, \mathtt{G}\}$ to itself. For example, the arrangement `ODG` is represented by the bijection $f\colon A \to A$ given by $f(\mathtt{D}) = \mathtt{O}$, $f(\mathtt{O}) = \mathtt{D}$, $f(\mathtt{G}) = \mathtt{G}$.

**Definition 45.** A **permutation** of a set $A$ is a bijective function from $A$ to $A$. The set of all permutations of $A$ is denoted by $\mathcal{S}(A)$.

Thus the model for ordered arrangements of a set $A$ is the set $\mathcal{S}(A)$ of all permutations of $A$. Now suppose, as in the problem about `DOG`, that $A$ is finite. Since the set of all permutations of $A$—by definition the set of all bijections from $A$ to $A$— is a subset of the set of all functions from $A$ to $A$, it follows from Theorem 41 and Proposition 8 that the set of all permutations of the finite set $A$ is finite. (More generally, the set of all bijective functions from a finite set to a finite set is itself finite.) The next theorem will state how many elements $\mathcal{S}(A)$ has.

One tool needed in the proof of that theorem is the following analog for bijective functions between two sets of Remark 43 about all functions between two sets.

**Remark 46.** The number of bijections from a finite set $A$ to a finite set $B$ depends only on $\#(A)$ and $\#(B)$, not on the sets $A$ and $B$ themselves. That is, if $\#(A') = \#(A)$ and $\#(B') = \#(B)$, then the number of bijections from $A$ to $B$ is the same as the number of bijections from $A'$ to $B'$. (Proof: Exercise.)

Other "invariance properties" concerning sets formed from finite sets, akin to that stated in the preceding remark, will be used implicitly in subsequent proofs.

**Theorem 47** (Number of permutations)**.** *If $A$ is an $n$-element finite set, then the number of permutations of $A$ is $n!$.*

*Proof.* Use induction on $n$.

**Base step ($n = 0$):** In this case, $A$ is the empty set $\{\,\}$. Then the unique function from $A$ to $A$ is certainly bijective. ( I dare you to find some element of $A$ that is not a value of that function; and I dare you to find two different elements of $A$ at which that function takes different values!) Thus $\#(\mathcal{S}(A)) = 1$. But $0! = 1$ by definition of the factorial function. Hence $\#(\mathcal{S}(A)) = n!$ in this case.

**Inductive step:** Let $n \geq 0$. Assume that $\#(\mathcal{S}(A)) = n!$ for every finite set $A$ with $\#(A) = n$.

Let $A$ be a finite set with $\#(A) = n + 1$. What must be deduced is that $\#(\mathcal{S}(A)) = (n + 1)!$.

Write $A = \{a_1, a_2, \ldots, a_n, a_{n+1}\}$. For each $j = 1, 2, \ldots, n + 1$, let

$$\mathcal{S}_j = \{\, f \in \mathcal{S}(A) : f(a_{n+1}) = a_j \,\}.$$

Then the sets $\mathcal{S}_j$ are pairwise disjoint, and $\mathcal{S}(A) = \bigcup_{j=1}^{n+1} \mathcal{S}_j$. By the Sum Rule (Theorem 12),

$$\#\,(\mathcal{S}(A)) = \sum_{j=1}^{n+1} \#\,(\mathcal{S}_j). \tag{*}$$

Let $1 \le j \le n+1$ and count the set $\mathcal{S}_j$. This set consists of those bijective functions $f\colon A \to A$ for which $f(a_{n+1}) = a_j$. If $f$ is such a bijection, then $f(a_i) \in A \setminus \{a_j\}$ whenever $i \neq n+1$, and each $a_k \in A \setminus \{a_j\}$ is the value $f(a_i)$ for a unique $i \neq n+1$. Then the restriction of $f$ to $A \setminus \{a_{n+1}\}$ defines a bijection

$$f_j\colon A \setminus \{a_{n+1}\} \to A \setminus \{a_j\}.$$

Conversely, if $g\colon A \setminus \{a_{n+1}\} \to A \setminus \{a_j\}$ is bijection, then $g = f_j$ for a unique permutation $f$ of $A$ for which $f(a_{n+1}) = a_j$, namely, the function $f\colon A \to A$ defined by $f(a_i) = g(a_i)$ for each $i \neq n+1$ and $f(a_{n+1}) = a_j$. Thus there is a one-to-one correspondence between the set of all permutations $f \in \mathcal{S}_j$, on the one hand, and the set of all bijections from $A \setminus \{a_{n+1}\}$ to $A \setminus \{a_j\}$, on the other hand.

Each of the sets $A \setminus \{a_{n+1}\}$ and $A \setminus \{a_j\}$ has exactly $n$ elements. According to Remark 46, the set of all bijections from the first of these sets to the second has the same number of elements as the set of all bijections from an $n$-element set to itself. Hence $\#\,(\mathcal{S}_j)$ is the number of permutations of an $n$-element set. By the inductive assumption, this number is $n!$. Thus

$$\#\,(\mathcal{S}_j) = n!. \tag{**}$$

From (*) and (**), it follows that

$$\#\,(\mathcal{S}(A)) = \sum_{j=1}^{n+1} n! = (n+1)\,n! = (n+1)!,$$

as was to be deduced. □

The model used above for the set of all arrangements of a finite set $A$ is the set of all permutations of that set. Sometimes either of two other, equivalent, models is useful. [The tacit assumption, at least right now, is that $n = \#\,(A) > 0$.]

An ordered arrangement of $A$ may be regarded as a particular numbering $a_1, a_2, \ldots, a_n$ of the $n$ distinct elements of $A$. And such a numbering may be modelled by a function $f\colon \{1, 2, \ldots, n\} \to A$ where, for each $i = 1, 2, \ldots, n$, the value $f(i) = a_i$. In an ordered arrangement of $A$:

- every element of $A$ must be numbered, that is, $A = \{a_1, a_2, \ldots, a_n\}$; and
- no two distinct elements of $A$ can be numbered the same, that is, $a_i \neq a_j$ whenever $i \neq j$.

Thus such a function $f\colon \{1, 2, \ldots, n\} \to A$ that represents an ordered arrangement of $A$ must be *bijective*. Thus **the set of all ordered arrangements of the $n$-element set $A$ may be modelled by *the set of all bijections from* $\{1, 2, \ldots, n\}$ *to* $A$.**

For an $n$-element set $A$, each bijection $f\colon \{1, 2, \ldots, n\} \to A$ may be represented uniquely by an ordered $n$-tuple $(a_1, a_2, \ldots, a_n)$ of *distinct* elements of $A$, namely, the ordered $n$-tuple

$$\bigl(f(1), f(2), \ldots, f(n)\bigr). \tag{*}$$

And conversely, each ordered $n$-tuple $(a_1, a_2, \ldots, a_n)$ of distinct elements of $A$ has the form (*) for a unique bijection $f\colon \{1, 2, \ldots, n\} \to A$, namely, the function $f$ defined by $f(j) = a_j$ for each $j = 1, 2, \ldots, n$. Thus **the set of all arrangements of an $n$-element set $A$ may be modelled by *the set of all ordered $n$-tuples of distinct elements of* $A$.**

## 5.2 Injections

As explained in the preceding section, the set of all arrangements of an $n$-element set $A$ may be modelled by the set of all ordered $n$-tuples of *distinct* elements of $A$. Instead of ordered $n$-tuples of distinct elements of an $n$-element set, consider now the more general situation of ordered $r$-tuples of *distinct* elements of an $n$-element set, where $r$ is not necessarily the same as $n$.

As in the preceding section, such an ordered $r$-tuple of distinct elements of $A$ may also be regarded as a function $f\colon \{1, 2, \ldots, r\} \to A$ that is *injective*—an injective function from an $r$-element set to an $n$-element set.

An ordered $r$-tuple of *distinct* elements of an $n$-element set $A$—or, equivalently, an injection from an $r$-element set to $A$—serves as the model for an **ordered sample, without replacement, of size $r$ from a set** $A$. (In some texts, such an ordered sample without replacement is called a "permutation", but here that term is reserved for the case, considered above, that $r = n$.) Thus to count the number ordered samples of size $r$, without replacement, from a finite set $A$ is to count the number of injections from an $r$-element set to an $n$-element set.

Suppose $A$ and $B$ are finite sets. According to Theorem 41, the set of all functions from $A$ to $B$ is finite. Hence the set of all *injective* functions from $A$ to $B$ is also finite. The number of elements of this set of functions depends only on $\#(A)$ and $\#(B)$, not on the sets $A$ and $B$ themselves. That is, if $\#(A') = \#(A)$ and $\#(B') = \#(B)$, then the number of injective functions from $A$ to $B$ is the same as the number of injective functions from $A'$ to $B'$. This justifies the following notation in the following definition.

**Definition 48.** For nonnegative integers $n$ and $r$, the integer $P(n, r)$ is defined to be the number of *injective* functions from an $r$-element set to an $n$-element set.

There is exactly one function from the empty set to the empty set, and this function is injective (I dare you to find two elements of the empty set!). Thus:

$$P(0, 0) = 1$$

Since there are no functions whatsoever from an $n$-element set to an $r$ element set when $r = 0$ but $n > 0$, there are no injective functions. Thus:

$$P(n, 0) = 0 \quad \text{if } n > 0$$

A consequence of the functional version of the Pigeonhole Principle (see Proposition 27) is that a function from an $r$-element set to an $n$ element set cannot be injective when $r > n$. Thus:

$$P(n, r) = 0 \quad \text{if } r > n > 0$$

From this and the preceding formula, for $P(n, 0)$:

$$\boxed{P(n, r) = 0 \quad \text{if } r > n \geq 0}$$

**Theorem 49.** *Let $n$ and $r$ be nonnegative integers with $0 \leq r \leq n$. Then the number $P(n, r)$ of injective functions from an $r$-element set to an $n$-element set is given by:*

$$\boxed{P(n, r) = \frac{n!}{(n - r)!}.}$$

*Proof.* Let $A$ and $B$ be finite sets with $\#(A) = r$ and $\#(A) = n$. The idea of the proof is that each injective function $f\colon A \to B$ can be formed in two stages: First, choose the $r$-element subset of $B$ that will be the set of values of $f$; there are $C(n, r)$ such sets. Second, for a given $r$-element $S$ subset of $B$, choose the arrangement of that subset that will determine which element of $S$ will be the value of $f$ at which element of $A$; there are $r!$ such arrangements.

By the Generalized Product Theorem, the number of injective functions from $A$ to $B$ is therefore $C(n,r)r!$. By Theorem 38,

$$C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Hence the number of injective functions from $A$ to $B$ is

$$\frac{n!}{r!(n-r)!}r! = \frac{n!}{(n-r)!},$$

as stated. □

Take $r = n$ in the just-proved formula for $P(n,r)$ to obtain

$$P(n,n) = n!.$$

This result says that the number of injective functions from an $n$-element set $A$ to an $n$-element set $B$ (possibly $B = A$) is the same as the number of bijective functions from $A \to B$. And that is hardly a coincidence in view of the following proposition.

**Proposition 50.** *Let $A$ be a* finite *set. Then every injective function $f\colon A \to A$ is bijective and every surjective function $f\colon A \to A$ is bijective.*

*Proof.* Omitted. □

Since a bijective function $f\colon A \to A$ is both injective and surjective, it follows that: **for a function from a finite set to itself, the properties of being injective, surjective, and bijective are equivalent**.

You might expect the next topic in this section to be the number of *surjective* functions between two finite sets. However, additional tools concerning partitions of a set, developed in Section 7, will be required for that.

## 5.3 Derangements

Let $A$ be a finite set and order $A$ in some way. A **derangement** of $A$ is a rearrangement of the elements of $A$ that leaves *no* element in its original order. In other words, a derangement of $A$ is a permutation $f$ of $A$ for which $f(a) \neq a$ for each $a \in A$.

The number of all derangements of a set with $n$ elements is denoted by $d_n$.

**Theorem 51.** *The number $d_n$ of derangements of an $n$-element set satisfies the relations:*

$$d_1 = 0, \quad d_2 = 1,$$
$$d_n = (n-1)(d_{n-1} + d_{n-2})$$

*Proof.* The only permutation of a 1-element set is the identity, and that is not a derangement; thus $d_1 = 0$. The permutation of a 2-element set are the identity and the permutation that interchanges the two elements, and only the latter is a derangement; thus $d_2 = 1$.

Now suppose $n \geq 3$. Consider the standard $n$-element set $A = \{1, 2, \ldots, n\}$. A derangement $f$ of $A$ has, in particular, the property that $f(n) \neq n$. But since a derangement is a permutation, then

$$f(k) = n$$

for a unique $k$ with $1 \leq k \leq n-1$. Of course there are exactly $n-1$ possibilities for the value of $k$.

The derangement $f$ of $A$ can be one of two types:

**Type (i): $f(n) = k$.** In this case, the derangement $f$ of $A$ just interchanges $n$ and $k$. The restriction $f_k$ of $f$ to the set

$$A_k = A \setminus \{k, n\}$$

is then a derangement of $A_k$. And conversely, if $g$ is a derangement of $A_k$, then $g = f_k$ for exactly one derangement $f$ of $A$, namely, the function $f\colon A \to A$ given by $f(j) = g(j)$ for $j \in A_k$, $f(k) = n$, and $f(n) = k$.

[To follow the reasoning for a type (i) $f$, consider, for example, $n = 5$ and

$$f(1) = 4, f(2) = 5, f(3) = 1, f(4) = 3, f(5) = 2.$$

Then $k = 2$, $A_k = \{1, 3, 4\}$, and the function $f_k$ is given by

$$f_k(1) = 4, f_k(3) = 1, f_k(4) = 3,$$

so that $f_k$ is a derangement of $A_k$.]

Thus when $f$ is of type (i), for a particular $k$, there is a one-to-one correspondence between derangements $f$ of type (i), on the one hand, and derangements of the $(n-2)$-element set $A_k$, on the other hand. By definition, there are $d_{n-2}$ such derangements of $A_k$. Hence for each particular $k$, there are $d_{n-2}$ derangements $f$ of $A$ that are of type (i).

**Type (ii):** $f(n) \neq k$**.** Then the restriction of $f$ to the set

$$B_k = A \setminus \{k\}$$

defines a bijection

$$h_k\colon B_k \to C$$

where

$$C = A \setminus \{n\} = \{1, 2, \ldots, n-1\}.$$

Still $h_k(i) \neq i$ for all $i \in B_k$, and $h_k(n) \neq k$. Now alter $h_k$ as follows. From its domain $B_k = A \setminus \{k\}$ remove $n$ and in its place insert $k$ so as to obtain the set $C$; define the function

$$g_k\colon C \to C$$

by

$$\begin{aligned} g_k(i) &= h_k(i) = f(i) \qquad \text{for } i \neq k, n, \\ g_k(k) &= h_k(n) = f(n). \end{aligned}$$

This new function $g_k$ is a derangement of $C$. And conversely, each derangement of $C$ can be obtained in this way as $g_k$ for a unique derangement $f$ of $A$ that is of type (ii) [you should verify that].

[To follow the reasoning for a type (ii) $f$, consider again, for example, $n = 5$ but now and

$$f(1) = 4, f(2) = 1, f(3) = 5, f(4) = 3, f(5) = 2.$$

Then $k = 3$, $B_k = \{1, 2, 4, 5\}$, and

$$h_k(1) = 4, h_k(2) = 1, h_k(4) = 3, h_k(5) = 2.$$

Now $C = \{1, 2, 3, 4\}$, and the function $g_k$ of $C$ is given by

$$g_k(1) = 4, g_k(2) = 1, g_k(3) = 2, g_k(4) = 3,$$

so that $g_k$ is a derangement of $C$.]

Thus when $f$ is of type (ii), for a particular $k$, there is a one-to-one correspondence between derangements $f$ of type (ii), on the one hand, and derangements of the $(n-1)$-element set $C$, on the other hand. By definition, there are $d_{n-1}$ such derangements of $C$. Hence for each particular $k$ there are $d_{n-1}$ derangements $f$ of $A$ that are of of type (ii).

By the Basic Sum Rule, for each $k = 1, 2, \ldots, n-1$, there are $d_{n-2} + d_{n-1}$ derangements $f$ of $A$ with $f(k) = n$. By the Sum Rule, there are in all $(n-1)(d_{n-1} + d_{n-2})$ derangements of $A$. But the number of derangements of $A$ is, by definition, $d_n$. This completes the proof. $\square$

**Example 52.** Your dorm has a bank of 100 mailboxes. One day the postal carrier arrives there with 100 letters from the registrar, one addressed to each of you. The carrier, having forgotten to wear his glasses that day, cannot make out the names on the envelopes. So he randomly puts one letter in each of the mailboxes. What is the probability that nobody receives the correct letter?

*Answer.* The number of ways the postal carrier can distribute the 100 letters among the mailboxes is 100!. The number of ways he can do that in such a way that nobody gets the correct letter is the derangement number $d_{100}$. So the desired probability is $d_{100}/100!$. (What is the numerical value of that ratio?)

# 6 Indistinguishable objects

So far, all the examples dealt with sets whose are "distinguishable" objects, that is, different elements. This section treats three special models that help in conceptualizing counting problems where the objects are indistinguishable from one another. As we shall see, each model allows us to solve the problem by just counting $k$-element subsets of a finite set. For that reason, this entire section may be studied immediately after Section 4.

## 6.1 MISSISSIPPI models

**Example 53.** How many different "words" can be formed by rearranging the letters of `MISSISSIPPI`? By a "word" we really mean just a string of letters, not necessarily a real word such as you would find in a dictionary. For example, `ISSIPIIMSPS` is such a word.

Our starting word `MISSISSIPPI` has 11 letters. But this is *not* a problem of counting the permutations of those 11 letters, because some of the letters appear multiple times (while others appear only once). In other words, some of the letters are indistinguishable from each other (while others are distinguishable).

*Solution.* To begin, tally the number of appearances of each letter in `MISSISSIPPI`:

```
M I S P
  I S P
  I S
  I S
```

There are 4 `I`s, 4 `S`s, 2 `P`s, and 1 `M`.

Next, create a "blank" 11-letter word where the letters of `MISSISSIPPI` will go:

-- -- -- -- -- -- -- -- -- -- --

Whereas the multiple instances of the letter `I`, for example, are indistinguishable, the *locations* in this 11-letter blank word *are* distinguishable. Indeed, imagine numbering the 11 locations left-to-right as 1, 2, 3, ..., 11.

Now apply a 4-step process: first, place the 4 `I`s; next, second, place the 4 `S`s; third, place the 2 `P`s; and fourth, place the `M`. (Of course you could arrange the steps in a different order, e.g., first place the single `M`, next place the `S`s, etc.)

Count the number of ways to carry out each step:

- There are $\binom{11}{4}$ ways to select the 4 locations for the 4 `I`s; once these 4 locations have been selected, only $11 - 4 = 7$ locations remain to be filled with the other letters.
- There are $\binom{7}{4}$ ways to select the 4 locations for the 4 `S`s; once these locations have been selected, only $7 - 4 = 3$ locations remain.
- There are $\binom{3}{2}$ ways to select the 2 locations for the 2 `P`s; once these have been selected, only $3 - 2 = 1$ location remains.

- There is $\binom{1}{1}$ way to select the 1 remaining location for the 1 `M`.

Finally, apply the Generalized Product Rule (2.3): the total number of words that can be formed is:

$$\binom{11}{4} \cdot \binom{7}{4} \cdot \binom{3}{2} \cdot \binom{1}{1} = \frac{11!}{4! \times 7!} \cdot \frac{7!}{4! \times 3!} \cdot \frac{3!}{2! \times 1!} \cdot \frac{1!}{1! \times 0!} = \frac{11!}{4! \times 4! \times 2! \times 1!} = 34650$$

Notice that solving the problem boiled down to using a $k$-step process, where $k = 4$ and wherein the $j$th step you count the number of $r_j$-element subsets of a given $n_j$-element set. The initial $n_1 = n$ is the total number of locations and $r_1$ the number of repetitions of the first letter; then $n_2 = n_1 - r_1$ and $r_2$ is the number of repetitions of the second letter; etc. Thus

$$n_1 = n, \quad n_2 = n_1 - r_1, \quad n_3 = n_2 - r_2, \ldots n_k = n_{k-1} - r_{k-1}.$$

Moreover, since all locations must be filled,

$$r_1 + r_2 + \ldots r_k = n.$$

The form of the final count is then:

$$\binom{n_1}{r_1} \cdot \binom{n_2}{r_2} \cdot \cdots \cdot \binom{n_k}{r_k} = \binom{n}{r_1} \cdot \binom{n - r_1}{r_2} \cdot \cdots \cdot \binom{n - r_1 - r_2 - \cdots - r_{k-1}}{r_k}$$
$$= \frac{n!}{r_1! \times r_2! \times \cdots \times r_k!} = \frac{r_1 + r_2 + \cdots + r_k}{r_1! \times r_2! \times \cdots \times r_k!}$$

A number of this form is sometimes denoted by $(r_1, r_2, \ldots, r_k)!$ and is called a **multinomial coefficient**. If the $n$th power $(x_1 + x_2 + \cdots + x_k)^n$ of a multinomial $x_1 + x_2 + \cdots + x_k$ is expanded, then the coefficient of the term having the power product $x_1^{r_1} x_2^{r_2} \ldots x_k^{r_k}$ is precisely $(r_1, r_2, \ldots, r_k)!$. Can you see why?

We advise against memorizing the formula for a multinomial coefficient as the solution of a `MISSISSIPI`-type problem. Rather, just work out the solution the way it was done above.

## 6.2 "Stars-and-bars" models

**Example 54.** A market is stocked with apples, bananas, cherries, and durians. You will choose a dozen pieces of this fruit to buy. How many different choices can you make?

*Interpretation.* The intention of the problem is that one apple is as good as another—that for purposes of this problem, all the apples are indistinguishable—and similarly with the other three kinds of fruit. But, of course, you can distinguish apples from bananas, etc.

*Solution.* Imagine a particular choice of a dozen pieces of fruit you'll buy. Line them up in a row; we'll represent each piece by a "star" (an asterisk, really):

$$* \; * \; * \; * \; * \; * \; * \; * \; * \; * \; * \; *$$

Now we need to say which kind of fruit each of the twelve is. So put the apples first (if you chose any), the bananas second (if you chose any), the cherries third (if you chose any), and the durians last (if you chose any). For example, you might have chosen two apples, five bananas, one cherry, and four durians. Imagine moving each kind of fruit that you chose apart from the other kinds; we'll indicate this by inserting a "bar" (a vertical line segment) between each two different kinds of fruit:

$$* \; * \mid * \; * \; * \; * \; * \mid * \mid * \; * \; * \; *$$

Remember, the understanding is that the order, from left to right, is: apples, bananas, cherries, durians. Of course you might have chosen no cherries whatsoever, so your selection might look like this:

$$* \; * \mid * \; * \; * \; * \; * \; * \mid \mid * \; * \; * \; *$$

Or, you might have chosen no apples—

$$| * * * * * * * | * | * * * *$$

—or, instead, no durians:

$$* * | * * * * * | * * * * * |$$

So your selection of the 12 pieces selected from the $r = 4$ kinds of fruit is indicated by placing $r-1 = 3$ bars into position relative to the $n = 12$ stars. In all there are $n+(r-1) = 12+3 = 15$ symbols, and a particular selection of fruit is indicated by where among these $n+r-1 = 15$ symbols the $r-1 = 3$ bars go.

So the problem boils down to this: We have a set of distinguishable objects after all, namely, the set of $n+r-1$ positions for the $n+r-1$ symbols. And a particular selection of $r$ pieces of fruit is represented by an $r-1$-element subset of that set, namely, the positions for the bars.

Then the answer to the problem is simply the number of 3-element subsets of a 15-element set. This number is, of course:

$$\binom{15}{3} = \frac{15 \cdot 14 \cdot 13}{3 \cdot 2 \cdot 1} = 455$$

In general, suppose there are $r$ different "types" or "kinds" of objects, with the objects of each type indistinguishable from one another. And suppose we are going to select $n$ objects in all (assuming there are an unlimited number of each type). Then such a selection is represented by identifying the $r-1$ bars among $n+r-1$ symbols, so the total number of such selections will be:

$$\binom{n+r-1}{r-1}$$

Thus a problem represented by such a "stars-and-bars" model boil down to counting $(r-1)$ subsets of an $(n+r-1)$-element set.

Don't memorize the preceding formula for this kind of problem! Just do the representation in terms of stars-and-bars and think through what it is that you must count.

## 6.3 "Xs-and-wedges" models

**Example 55.** Six men and ten women are to be seated in a long row so that no two men are sitting next to one another. How many different seatings are possible?

*Interpretation.* As in the preceding fruit example (Example 54), likewise here the intention is that the men are indistinguishable and the women are indistinguishable. This is a fair interpretation of the question, since there was no indication that it matters which particular man is where or which particular woman is where, just that no two of the men sit next to one another.

*Solution.* Represent the 10 women by 10 Fs (for "female") in a row:

```
F F F F F F F F F F
```

Represent the 6 men by 6 Ms, where we need to place these Ms in that row. That no two men should be seated next to one another means that no two (or more) Ms may be placed between two adjacent Fs; nor may two (or more) Ms be placed to the left of all the Fs or to the right of all the Fs. And *each* individual one of the 6 Ms must go either between two adjacent Fs or else to the left of all the Fs or to the right of all the Fs. Note that some pairs of adjacent Fs may get no Ms between them: it's OK for two women to be seated next to one another.

Indicate the possible positions for the Ms by "wedges" (carets, really):

```
^F^F^F^F^F^F^F^F^F^F^
```

Since there are $n = 10$ women, we have $n+1 = 11$ wedges. Where the $r = 6$ men are seated is represented merely by selecting $r = 6$ of these wedges. Hence the number of seatings meeting the stated requirement is:

$$\binom{11}{6} = \frac{11 \cdot 10 \cdot 9 \cdot 8 \cdot 7}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 462$$

More generally, suppose we have $n$ indistinguishable objects of a first kind and $r \leq n+1$ indistinguishable objects of a second kind. The objects of the first kind are arranged in a row. The problem is to place the objects of the second kind in that same row so that no two of them are adjacent to one another. If we represent the objects of the first kind by `X`s and the possible locations of objects of the second kind by wedges, then as in the preceding example the number of ways to distribute the objects of the second kind is given by:

$$\binom{n+1}{r}$$

Thus a problem represented by such an "`X`s-and-wedges" model boils down to counting $r$-element subsets of an $(n + 1)$-element set.

Don't memorize the preceding formula for this kind of problem! Just do the representation in terms of `X`s-and-wedges and think through what it is that you must count.

# 7 Partitions

Start with a simple, motivating example.

**Example 56.** All 8 students in a combinatorics class are going to work on a certain project. They are going to be split up into 3 separate work groups, with each group working on that same project. In how many ways can these work groups be formed?

We may represent the set of students by the set $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Then we may represent each work group as a subset of $A$—a *nonempty* subset, of course, since a work group consisting of no members doesn't get any work done! That the class is split up into separate work groups means that each of the 8 students is going to belong to exactly one such work group. Then a way to split up the class may be represented by a collection of such nonempty subsets of $A$ such that each member of $A$ belongs to exactly one of these subsets. For example, $\{\,\{1, 3, 4, 7\}, \{2, 6, 8\}, \{5\}\,\}$ would be such a collection. Such a collection is a *partition* of the set $A$.

In general, by a **partition** of a set $A$ we mean a collection $\mathcal{P}$ such that:

- each member of $\mathcal{P}$ is a subset of $A$;
- no member of $\mathcal{P}$ is empty; and
- each $x \in A$ belongs to exactly one of the members of $\mathcal{P}$.

Said more tersely, a partition of $A$ is *a collection $\mathcal{P}$ of nonempty, pairwise disjoint subsets of $A$ whose union is $A$.*

The members of a partition of $A$ are called the **blocks** of that partition. For example, the partition $\{\,\{1, 3, 4, 7\}, \{2, 6, 8\}, \{5\}\,\}$ of $A = \{1, 2, 3, 4, 5, 6, 7, 8\}$ consists of 3 blocks, namely, $\{1, 3, 4, 7\}$, $\{2, 6, 8\}$, and $\{5\}$.

Thus the answer to our question in Example 56 is: the number of partitions of an 8-element set into 3 blocks. But what is this number? We shall develop a method for calculating it.

When $A$ is an $n$-element finite set with $n > 0$, then each block of a partition of $A$ is necessarily a finite set, with at most $n$-elements (and with strictly less than $n$ elements unless the partition has just one block, namely, the entire set $A$). Since by definition the blocks of a partition are nonempty, then a partition consisting of $r$ blocks exists only when $r \geq 1$.

**Definition 57.** Let $n$ and $r$ be integers with $0 \leq r \leq n$ or $0 = r < n$. Then the number of partitions of an $n$-element set consisting of $r$ blocks is called a **Stirling number of the second kind** and is denoted by $S(n, r)$.

**Theorem 58.** *The Stirling numbers of the second kind satisfy the relations:*

$$\begin{aligned} S(n,0) &= 0 \qquad (n \geq 1), \\ S(n,n) &= 1 \qquad (n \geq 0), \\ S(n+1,r) &= S(n,r-1) + rS(n,r) \qquad (1 \leq r \leq n) \end{aligned}$$

*Proof.* The first formula, $S(n, 0) = 0$, for $n \geq 1$, holds because a nonempty set cannot be partitioned in a collection of no nonempty sets.

The second formula $S(n, n) = 1$, holds when $n \geq 1$ because an $n$-element set $A$ has only one partition with $n$ blocks, namely, the partition $\{ \{a\} : a \in A \}$ of $A$ into one-element subsets. And the second formula $S(n, n) = 1$ holds when $n = 0$ because the empty set $\emptyset$ has exactly one partition, namely, the empty collection of subsets of itself. (Yes, the empty collection is a legitimate partition in this case, when the set $A$ is empty—and only in that case!)

To establish the third, recursive formula, fix $r$ and $n$ with $1 \leq r \leq n$. Let $A = \{1, 2, \ldots, n, n+1\}$ and let $B = A \setminus \{n+1\} = \{1, 2, \ldots, n\}$.

Suppose $\mathcal{P}$ is an partition of $A$ with exactly $r$ blocks. There is exactly one member $P \in \mathcal{P}$ for which $n + 1 \in P$. Then the partition $\mathcal{P}$ must be one of two types.

**Type (i):** $P = \{n + 1\}$**.** In this case, each other member of the partition $\mathcal{P}$ of $A$ is in fact a subset of $B$, so that $\mathcal{P} \setminus \{P\}$ is a partition of an $n$-element set into $(r - 1)$ blocks. Thus there are $S(n, r - 1)$ partitions of $A$ that are of type (i).

**Type (ii):** $P \neq \{n + 1\}$**.** In this case, there are some elements of $A$ other than $n + 1$ that belong to $P$. Then $\mathcal{P}$ can be constructed as follows: Choose some partition $\mathcal{Q}$ of $B$ with exactly $r$ blocks; there are $S(n, r)$ such partitions. Choose some one of those blocks and insert $n + 1$ into it to form the set $P$; there are $r$ such choices. By the Generalized Product Rule, there are thus $r\,S(n, r)$ partitions of $A$ that are of type (ii).

It now follows from the Sum Rule that the total number $S(n + 1, r)$ of partitions of $A$ with exactly $r$ blocks is $S(n, r - 1) + r\,S(n, r)$. □

The numerical answer to the question posed in Example 56 is $S(8, 3)$. You should use the preceding recursive formula for $S$ to find this number.

## 7.1 Surjections

Here is a variant of Example 56.

**Example 59.** The 8 students in a combinatorics class are going to be split into 3 separate workgroups to work on three different projects. In how many ways can this be done?

First model: The set of *ordered* partitions of an 8-element set into 3 subsets. The sets in the partition says who is in which workgroup. The ordering of the three workgroups determines which group works on which of the 3 problems. So the answer to the question is the number of ordered partitions of an 8-element set into 3 subsets. Can you calculate the number already?

Second model: Let $A = \{s_1, s_2, \ldots, s_8\}$ represent the set of 8 students and let $B = \{p_1, p_2, p_3\}$ represent the set of the 3 projects. To split up the class in the way stated is to say to which project each of the students is assigned; in other words, to prescribe a function $f\colon A \to B$ that is *surjective*. So the answer to the question is also the number of surjective functions from an 8-element set to a 3-element set.

The two models are, of course, just two equivalent ways of looking at the same thing. And, indeed, that equivalence lies at the heart of the proof of Theorem 61, below.

Already we know the following:

- there are exactly $n^r$ functions in all from an $r$-element set to an $n$-element set;
- there are exactly $P(n,r) = \frac{n!}{(n-r)!}$ injective functions from an $r$-element set to an-$n$ element set when $r \leq n$ (and none when $r > n$); and in particular,
- there are exactly $n!$ bijective functions from an $n$-element set to an $n$-element set.

Now, finally, we determine the number of surjective functions between two finite sets. Of course the set of surjective functions from a finite set $A$ to a finite set $B$ is finite, because it is a subset of the set of all functions from $A$ to $B$.

**Definition 60.** Let $\operatorname{Onto}(n,r)$ denote the number of surjective functions from an $n$-element set to an $r$ element set.

Notice that the roles of $n$ and $r$ are reversed in the notation $\operatorname{Onto}(n,r)$ as contrasted to the notation $P(n,r)$: the notation $\operatorname{Onto}(n,r)$ denotes the number of surjective functions from an $n$-element set to an $r$-element set, whereas the notation $P(n,r)$ denotes the number of injective functions from an $r$-element set to an $n$-element set.

Suppose $r = 0$. Then the only value of $n$ for which a function from an $n$-element set to the $r$-element set can be surjective is that $n = 0$ as well; in this case, the unique function from the empty set to itself is surjective (I dare you to find a member of the codomain $\{\}$ of that function that is not a value of the function!).Thus

$$\operatorname{Onto}(0,0) = 1, \qquad \operatorname{Onto}(n,0) = 0 \text{ if } n > 0.$$

According to Proposition 50, a surjective function from an $n$-element set to an $n$-element set is necessarily bijective. Since there are exactly $n!$ bijections from an $n$-element set to an $n$-element set, it follows that:

$$\operatorname{Onto}(n,n) = n!.$$

Suppose $n < r$. If $f\colon A \to B$ is a function from and $n$-element set to an $r$-element set, then the image $f(A)$ of $A$—the set of all values of $f$—has at most $n$ different elements (and it may have fewer than $n$ in case $f$ is not injective). Since $f(A)$ is a subset of the $r$-element set $B$, then $f(A) \neq B$. This means that $f$ is not surjective. Thus

$$\operatorname{Onto}(n,r) = 0 \text{ if } n < r.$$

**Theorem 61.** *Let $0 < r \leq n$. Then the number $\operatorname{Onto}(n,r)$ of surjective functions from an $n$-element set to an $r$-element set is given by*

$$\boxed{\operatorname{Onto}(n,r) = r!\,S(n,r),}$$

*where $S(n,r)$ is the Stirling number of the second kind.*

*Proof.* The idea of the proof is that each surjective function is determined by two things: (i) the partition of its domain into subsets on each of which the function is constant, and (ii) the constant values that the function takes on those subsets of its domain. (In following the proof, you may want to make things concrete by thinking of $n = 3$ and $r = 2$.)

Let $A = \{1, 2, \ldots, n\}$ and $B = \{1, 2, \ldots, r\}$.

Let $f\colon A \to B$ be a surjective function. For each $j = 1, 2, \ldots, r$, let

$$A_j - \{\, i \in A : f(i) = j \,\},$$

in other words, $A_j$ is the subset of the domain $A$ of $f$ on which $f$ takes the constant value $j$. Since $f$ is surjective, each $A_j$ is nonempty. Since $f$ is a function, the sets $A_j$ are pairwise disjoint. Since the value of $f$ at an element of $A$ is some element $j$ of $B$, the union of the sets $A_j$ is $A$. Thus the sets $A_1, A_2, \ldots, A_r$ form a partition of $A$.

In this way, each surjection $f\colon A \to B$ gives rise to a corresponding *ordered* $r$-tuple of sets $(A_1, A_2, \ldots, A_r)$ for which $\{A_1, A_2, \ldots, A_r\}$ is an $r$-block partition of $A$.

Conversely, let $(P_1, P_2, \ldots, P_r)$ be an ordered $r$-tuple of sets for which $\{P_1, P_2, \ldots, P_n\}$ is an $r$-block partition of $A$. Then there is a unique surjective function $f\colon A \to B$ for which, in the notation above, $A_j = P_j$ for each $j = 1, 2, \ldots, 4$, namely, the function $f$ defined by:

$$f(a) = j \text{ where } j \text{ is the unique integer for which } a \in P_j.$$

Thus there is a one-to-one correspondence between the set of all surjective functions $f\colon A \to B$, on the one hand, and the set of all *ordered* $r$-tuples of subsets of $A$ that form an $r$-block partition $A$, on the other hand. Hence the number $\mathrm{Onto}(n, r)$ is the same as the all orderings of all $r$-block partitions of the $n$-element set $A$.

The number of all $r$-block partitions of the $n$-element set $A$ is the Stirling number of the second kind $S(n, r)$. And for each $r$-block partition of $A$, there are $r!$ possible orderings (permutations) of that partition. By the Product Rule, the number of all orderings of all $r$-block partitions of the $n$-element set $A$ is the product $r!\,S(n, r)$. Hence $\mathrm{Onto}(n, r)$ is this product. □