Chapter 1

Ways to Choose

\[ O \text{ hell! to choose by another’s eye.} \]

\emph{A Midsummer Night’s Dream, WILLIAM SHAKESPEARE}

Aside from considerations of habit and fate, much of what we do appears to rest on deliberate and (one should hope) not infrequently intelligent choice. It is to this latter kind of choice that we devote our attention in these introductory pages.

The discussion we are about to undertake is intended primarily as a refresher. This is not to say that our level of presentation rests upon any formal prerequisites of much sophistication; it, rather, springs out of the usual dose of common wisdom that most of us share. The art of counting may rest primarily on the innate ability to discern patterns or systematic formations. When placed in the presence of well-developed mathematical technique this ability is usually strengthened, especially if the methods tend to assimilate...
well with the intuitive.

Discerning choice entails accurate assessment of chance which, in turn, requires good counting abilities. We shall be glad if the handful of techniques presented here enhance these abilities, even if only in marginal ways.

1.1 THE ESSENTIALS OF COUNTING

We begin with fundamental definitions and customary notation. A set is an assemblage of elements listed in any order without repetitions. The notation \( \{3, 1, 2\} \) describes a set with 1, 2, and 3 as elements; on the other hand \( \{1, 2, 2, 3\} \) is not a set. The number of elements in a set could be finite or infinite. All the sets in this book are finite, unless specific mention to the contrary is made. By \( |A| \) we denote the number of elements (or cardinality, or size) of the set \( A \). The notation \( a \in A \) indicates the fact that element \( a \) belongs to the set \( A \).

Given two sets \( A \) and \( B \), the symbol of inclusion \( A \subseteq B \) intimates the fact that all the elements of \( A \) could be found among those of \( B \); we say that \( A \) is a subset of \( B \). The intersection \( A \cap B \) consists of elements common to both \( A \) and \( B \). By \( A \cup B \), the union of the two sets, we denote the elements that are in either \( A \) or \( B \) (or in the intersection). When \( A \) is part of a larger set, \( \bar{A} \) denotes the elements not in \( A \) (but in that larger set); \( \bar{A} \) is called the complement of \( A \). The symbols introduced so far are related by the rules \( A \cap B = \bar{A} \cup \bar{B} \) and "dually" \( A \cup B = \bar{A} \cap \bar{B} \). Further, \( B - A \) denotes the elements of \( B \)
that are not in \( A \). One may think of \( B - A \) as the complement of \( A \) in \( B \). The symbol \( \emptyset \) is reserved for the empty set, that is, the set with no elements whatsoever. Due to its being vacuous, the empty set is somewhat illusory and thus difficult to grasp. Luckily its usefulness rests often in writing \( A \cap B = \emptyset \), which merely indicates that sets \( A \) and \( B \) have no common elements (we call such \( A \) and \( B \) disjoint). By \( \mathcal{P}(A) \) we denote the set of all subsets of \( A \) (including \( \emptyset \) and the set \( A \) itself). Two sets are equal if they consist of the same elements.

The Cartesian product of sets \( A \) and \( B \) is the set of ordered pairs defined as follows:

\[
A \times B = \{(a, b) : a \in A \text{ and } b \in B\}.
\]

And lastly, by \( N \), \( Q \), and \( R \) we indicate the (infinite) sets of natural, rational, and real numbers, respectively. (The natural numbers are sometimes called positive integers.)

### 1.1.2

A function from set \( A \) to set \( B \) is a rule by which we associate to each element of \( A \) a single element of \( B \). Let \( f : A \to B \) be a function. The most important implication in dealing with functions is that for two disjoint subsets \( B_1 \) and \( B_2 \) of \( B \) (i.e., \( B_1 \cap B_2 = \emptyset \)) we have disjoint preimages [i.e., \( f^{-1}(B_1) \cap f^{-1}(B_2) = \emptyset \)], where \( f^{-1}(C) \) denotes the subset \( \{a \in A : f(a) \in C\} \).

We call \( f \) injective (or one to one) if \( i \neq j \) implies \( f(i) \neq f(j) \).

We call \( f \) surjective (or onto) if for every \( b \) in \( B \) there exists \( a \) in \( A \) such that \( f(a) = b \).

The function \( f \) is said to be a bijection if it is both injective and surjective.

(Given an injection from \( A \) to \( B \) it is often helpful to effectively identify \( A \) with its
image through the injection. A surjection spreads, in a sense, $A$ over $B$, possibly several times.)

Though clear, it is important enough to state explicitly that

If $f$ is injective, then $|A| \leq |B|$.

If $f$ is surjective, then $|A| \geq |B|$.

If $f$ is bijective, then $|A| = |B|$.

1.1.3

With regard to Cartesian products we have

$$|A \times B| = |A||B|.$$ 

This is easy to see, since we count the number of ordered pairs with $|A|$ possibilities for the first entry and $|B|$ possibilities for the second:

(a totality of $|A||B|$ choices).

As an example, suppose your mate prepares for lunch three kinds of soup, five kinds of main course, six types of dessert, and five brands of drinks. How many choices for lunch do you have? The answer is $3 \cdot 5 \cdot 6 \cdot 5$, the cardinality of the Cartesian product of the sets of soup, main course, dessert, and drinks. (A lunch is understood to consist of precisely one choice out of each of the four available courses.)
1.1. THE ESSENTIALS OF COUNTING

1.1.4

A set with \( n \) elements has \( 2^n \) subsets. [Or, in terms of symbols, \( |\mathcal{P}(A)| = 2^{|A|} \).]

The simplest way to see this is to list the \( n \) elements of \( A \) as

\[(a_1, a_2, a_3, a_4, \ldots, a_{n-1}, a_n)\]

and then associate a vector of length \( n \) to each subset of \( A \) by placing a 1 in the position of the elements that occur in that subset and 0 in the remaining positions. The question now is how many vectors of length \( n \) with 0 or 1 as entries are there? Well, there are \( n \) positions to fill, with two choices (0 or 1) for each entry, so there are \( 2^n \) such vectors. We thus conclude that \( A (|A| = n) \) has \( 2^n \) subsets.

[Representing subsets as vectors is a nice little trick that we shall use again. Just to make sure the reader understands it we give a small example. If \( A = \{a_1, a_2, a_3, a_4, a_6, a_6\} \) is our set, then we associate as follows:

\[
\{a_1, a_2, a_4, a_6\} \leftrightarrow (1, 1, 0, 1, 0, 1).
\]

1.1.5

Let \( x \) be a real number or an indeterminate. Define \([x]_n\) to be \( x(x-1)(x-2)\ldots(x-n+1)\).

We set for convenience \([x]_0 = 1\). With \( n \leq m \) natural numbers, the number \([m]_n = m(m-1)(m-2)\ldots(m-n+1)\) has combinatorial meaning. Specifically:

* The number of ordered \( n \)-tuples with no repeated entries from a set of \( m \) symbols equals \([m]_n\).
* The number of sequences of length \( n \) with no repeated letters formed from \( m \) available (distinct) letters is \([m]_n\).

* The number of injections from \( A \) into \( B \) equals \([m]_n\), if \(|A| = n \) and \(|B| = m\).

In all these cases we have \( m \) choices at step 1, \( m - 1 \) choices at step 2, \( m - 2 \) choices at step 3, ..., and \( m - n + 1 \) choices at step \( n \). This gives a total of \( m(m - 1)(m - 2) \ldots (m - n + 1) = [m]_n \) possibilities:

\[
\begin{array}{cccccc}
\uparrow & \uparrow & \uparrow & \cdots & \uparrow \\
m \quad \text{choices} & m - 1 \quad \text{choices} & m - 2 \quad \text{choices} & \cdots & m - n + 1 \quad \text{choices}
\end{array}
\]

When \( m = n \) we denote \([n]_n\) by \( n! \) (\( n \) factorial). In addition 0! is by convention set to 1.

Small values for factorials are: \( 1! = 1, 2! = 2, 3! = 6, 4! = 24, 5! = 120, 6! = 720 \).

Example. The number of sequences of length 2 made with the elements of the set \( \{a, b, c, d\} \) is \([4]_2 = 4 \cdot 3 = 12\).

Indeed, they are:

\[
\begin{align*}
\text{ab} & \quad \text{ba} & \quad \text{ca} & \quad \text{da} \\
\text{ac} & \quad \text{bc} & \quad \text{cb} & \quad \text{db} \\
\text{ad} & \quad \text{bd} & \quad \text{cd} & \quad \text{dc}
\end{align*}
\]

1.1.6 Binomial Numbers

For \( 0 \leq n \leq m \) we denote \([m]_n/n!\) by \( \binom{m}{n} \) (to be read ,”\( m \) choose \( n \)”); when \( n > m \) we define \( \binom{m}{n} \) to be 0.

The numbers \( \binom{m}{n} \) are called \textit{binomial numbers} and they have several possible combi-
natorial interpretations:

* The number of subsets with \( n \) elements of a set with \( m \) elements is \( \binom{m}{n} \).

* The number of sequences of length \( m \) with precisely \( n \) ones and \( (m - n) \) zeros is \( \binom{m}{n} \).

(Allowing 0 and 1 as sole possibilities explains the term “binomial” as nomenclature.)

* The number of nondecreasing paths of length \( m \) from \((0,0)\) to \((n, m - n)\) on the planar lattice of integral points equals \( \binom{m}{n} \).

**Example.**

A path is called nondecreasing if the sequence of coordinates of the points we successively select is nondecreasing in each of the coordinates.

**Example.** The number of subsets of a set with four elements is \( \frac{[4]_2}{2!} = \frac{4 \cdot 3}{2 \cdot 1} = 6 \).

Indeed, the six subsets are:

\[
\{a, b\} \\
\{a, c\} \{b, c\} \\
\{a, d\} \{b, d\} \{c, d\}.
\]
Let us prove the first assertion. Pick a subset of \( n \) elements \((n \leq m)\). One can make \( n! \) sequences with the elements of this fixed subset. There are \([m]_n\) sequences of length \( n \) in all. The quotient \([m]_n/n!\) counts therefore the number of subsets with \( n \) elements.

The second assertion is the same as the first upon identifying the 1’s in a sequence of length \( m \) with the subset of positions in which they occur.

The third statement is the same as the second if we attach a sequence to a path by writing a 1 whenever we move horizontally, and a 0 whenever we move vertically.

A short list of small values of \( \binom{m}{n} \) may be helpful: \( \binom{m}{0} = 1, \binom{m}{m} = 1, \binom{m}{1} = \binom{m}{m-1} = m, \binom{7}{2} = 21, \binom{9}{3} = 84.\)

**Properties of the Binomial Numbers**

(a) \( \binom{m}{n} = \binom{m-1}{n} + \binom{m-1}{n-1} \)

This property is known as Pascal’s triangle:

\[
\begin{array}{cccccc}
1 \\
1 & 1 \\
1 & 2 & 1 \\
1 & 3 & 3 & 1 \\
1 & 4 & 6 & 4 & 1 \\
1 & 5 & 10 & 10 & 5 & 1 \\
\ldots
\end{array}
\]

\( 10 = \binom{5}{3} = \binom{4}{3} + \binom{4}{2} = 6 + 4 \)

This property can be proved any number of ways. The number \( \binom{m}{n} \) of nondecreasing
paths between \((0,0)\) and \((n,m-n)\) equals the number of such paths between \((0,0)\) and \((n,m-1-n)\) [i.e., \(\binom{m-1}{n}\)] plus those between \((0,0)\) and \((n-1,m-n)\) [i.e., \(\binom{m-1}{n-1}\)]. This is one possible proof.

(b) \(\binom{m}{n} = \binom{m}{m-n}\).

One can see this by realizing that each time we pick a subset of size \(n\) we uniquely identify a subset of size \(m-n\), namely its complement.

(c) \(\binom{m}{n-1} \leq \binom{m}{n}\), for \(1 \leq n \leq (m+1)/2\).

Indeed,

\[
\binom{m}{n-1} = \frac{[m]_{n-1}}{(n-1)!} \leq \frac{[m]_n}{n!} = \binom{m}{n}
\]

if and only if \(n \leq m - n + 1\) (or \(n \leq (m+1)/2\)).

(d) The Vandermonde Convolution: \(\binom{m+n}{k} = \sum_{j=0}^{m} \binom{m}{j}\binom{n}{k-j}\).

One can see the proof at once from the following figure:

By coloring \(m\) of the objects red and \(n\) blue, a choice of \(k\) objects out of the \(m+n\) would
contain $j$ red objects and $k - j$ blue ones. Sorting out by the possible values of $j$ leads to the formula stated above.

*Liebnitz’s Formula*

A natural way in which the binomial numbers arise, apart from those already mentioned, is when computing higher order derivatives of a product of two functions.

Let $f$ and $g$ be functions differentiable any number of times. Denote by $D^n f$ the $n$th derivative of $f$. Conveniently write $f_n$ for $D^n f$ and understand also that $f_0 = f$ and $g_0 = g$. Applying the product rule we obtain

\[
D^0 f g = f_1 g_0 + f_0 g_1,
\]

\[
D^1 f g = f_2 g_0 + f_1 g_1 + f_1 g_0 + f_0 g_2 = f_2 g_0 + \binom{2}{1} f_1 g_1 + f_0 g_2,
\]

\[
D^2 f g = f_3 g_0 + \binom{3}{1} f_2 g_1 + \binom{3}{2} f_1 g_2 + f_0 g_3,
\]

and so on. Assume that $D^n f g = \sum_{k=0}^{n} \binom{n}{k} f_k g_{n-k}$. Then

\[
D^{n+1} f g = \sum_{k=0}^{n} \binom{n}{k} D^k f g_{n-k} = \sum_{k=0}^{n} \binom{n}{k} (f_{k+1} g_{n-k} + f_k g_{n-k+1})
\]

\[
= \sum_{k=0}^{n} \left( \binom{n}{k} + \binom{n}{k-1} \right) f_k g_{n-k+1} + \binom{n}{n} f_{n+1} g_0
\]

\[
= \sum_{k=0}^{n+1} \binom{n+1}{k} f_k g_{n+1-k},
\]

which completes our inductive proof.

We thus proved Liebnitz’s formula:

\[
D^n f g = \sum_{k=0}^{n} \binom{n}{k} f_k g_{n-k}.
\]
1.1.7 Stirling Numbers of the Second Kind

Subsets $A_1, A_2, \ldots, A_m$ of set $A$ form a *partition* of the set $A$ if $A_i \neq \emptyset$, for all $1 \leq i \leq m$,

$$A_i \cap A_j = \emptyset \quad \text{for } i \neq j,$$

and

$$A_1 \cup A_2 \cup \cdots \cup A_m = A.$$

The subsets $A_i$ are called the *classes* of the partition.

A partition of a set with $n$ elements is said to be of type $1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n}$ ($\lambda_i \geq 0$) if it contains:

- $\lambda_1$ classes of cardinality 1
- $\lambda_2$ classes of cardinality 2
- $\vdots$
- $\lambda_n$ classes of cardinality $n$

We first observe that:

(a) *The number of partitions of type $1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n}$ of a set with $n$ elements* (where $n = \sum_{i=1}^{n} i \lambda_i$) *is* 

$$\frac{n!}{(1!)^{\lambda_1}(2!)^{\lambda_2} \cdots (n!)^{\lambda_n}(\lambda_1!)(\lambda_2!)\cdots(\lambda_n!)}.$$

*Proof.* Keeping in mind at all times the type of partition we seek (i.e., $1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n}$), we initially list all the $n!$ sequences we can possibly make with the $n$ elements of our set. The order of elements within any class being of no importance, we should divide $n!$ by 

$$(1!)^{\lambda_1}(2!)^{\lambda_2} \cdots (n!)^{\lambda_n}.$$ 

But classes of the same size can also be permuted among themselves
in any way whatever without changing the partition; we should, therefore, also divide by
$(\lambda_1!)(\lambda_2!)(\cdots)(\lambda_n!)$. This ends our proof, for any other permutation of elements would in
fact change the partition.

*Example.* Let $n = 14$ and the type be $1^22^33^24^0\cdots14^0$. That is:

```
{ | | | | | | | | | | | | | | | | }
\lambda_1

{ | | | | | | | | | | | | | | | | }
\lambda_2

{ | | | | | | | | | | | | | | | | }
\lambda_3
```

We have

\[
\frac{14!}{(1!)^2(2!)^3(3!)^2(2!)(3!)(2!)}
\]

partitions of this type.

The number of partitions of a set of $n$ objects into exactly $m$ classes is called the
*Stirling number of the second kind*. We denote this number by $S^m_n$.

In terms of the calculation performed in (a), we may write

\[
S^m_n = \sum \frac{n!}{(1!)^{\lambda_1}(2!)^{\lambda_2}(n!)^{\lambda_n}(\lambda_1!)(\lambda_2!)(\cdots)(\lambda_n!)}
\]

the sum extending over all vectors $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ with $\lambda_i$ nonnegative integers satisfying

$\sum_{i=1}^{m} \lambda_i = m$ and $\sum_{i=1}^{m} i\lambda_i = n$.

(b) The numbers $S^m_n$ satisfy the following recurrence relations:

\[
S^l_n = S^m_n = 1,
\]

and

\[
S^{m}_{n+l} = S^{m-1}_{n} + mS^{m}_{n}, \quad \text{for } 1 < m \leq n.
\]

(We define for convenience $S^0_n = 0$, and $S^m_n = 0$ for values of $m$ exceeding $n$.)
Proof. Consider the list of all partitions of \( n + 1 \) objects into \( m \) classes. There are \( S_{n+1}^m \) such partitions, by definition. Let \( w \) be an object among the \( n + 1 \). Separate the list of such partitions into two disjoint parts: those in which \( w \) is the sole element of a class and those in which any class containing \( w \) has cardinality 2 or more. There are \( S_{m-l}^n \) of the former kind and \( mS_n^m \) of the latter. Thus \( S_{n+1}^m = S_{m-l}^n + mS_n^m \).

To summarize graphically, denote by \( A \) the set of the \( n + 1 \) elements:

![Diagram](attachment:diagram.png)

This ends the proof.

The recurrence relations established in (b) allow us to write the small values of \( S_n^m \):
Example. We have $S_4^2 = 7$ partitions with two classes of a set with four elements. Let the set be \{a, b, c, d\}. The seven partitions are:

$$\{a\}\{b, c, d\},$$

$$\{b\}\{a, c, d\}, \{a, b\}\{c, d\},$$

$$\{c\}\{a, b, d\}, \{a, c\}\{b, d\},$$

$$\{d\}\{a, b, c\}, \{a, d\}\{b, c\}.$$  

For a real number $x$ we denote the product $x(x-1)(x-2)\ldots(x-k+1)$ by $[x]_k$. The powers of $x$ are related to the $[x]_k$'s via the Stirling numbers $S_n^k$. Specifically, we now prove that:

$$(c) \ x^n = \sum_{k=0}^{n} S_n^k [x]_k.$$  

Proof. The proof involves counting the number of functions from set $A$ to set $B$ in two different ways. Let $|A| = n$ and $|B| = m$. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>1 2 3 4 5 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>1 0 0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>1 1 0 0 0 0</td>
</tr>
<tr>
<td></td>
<td>1 3 1 0 0 0</td>
</tr>
<tr>
<td></td>
<td>1 7 6 1 0 0</td>
</tr>
<tr>
<td></td>
<td>1 15 25 10 1 0</td>
</tr>
<tr>
<td></td>
<td>1 31 90 65 15 1</td>
</tr>
</tbody>
</table>
First Way of Counting. Each element of $A$ can be mapped into any one of the $m$ elements of $B$. Hence there are $m \cdot m \cdot \ldots \cdot m = m^n$ functions in all between $A$ and $B$.

Second Way of Counting. Fix a subset $C$ of $B$ and count all the functions from $A$ with precisely $C$ as image (then sum over all nonempty subsets $C$ of $B$).

If $|C| = k$, then a function from $A$ onto $C$ can be identified with a partition of $A$ into precisely $k$ classes (a class being the preimage of a point in $C$). Conversely, and more importantly, a partition of $A$ into $k$ classes gives rise to $k!$ functions from $A$ onto $C$ (the possible ways of mapping the $k$ classes onto the elements of $C$). The number of functions from $A$ onto $C$ equals therefore $k!$ times the number of partitions of $A$ into $k$ classes, that is, it equals $k!S^k_n$.

Since there are $\binom{m}{k}$ subsets of size $k$ in $B$, we conclude that the number of all functions from $A$ to $B$ is $\sum_{k=1}^{m} \binom{m}{k} k! S^k_n = \sum_{k=1}^{m} S^k_n [m]_k$.

By the two ways of counting we conclude that

$$m^n = \sum_{k=1}^{m} S^k_n [m]_k \quad (1.1)$$

Let us now look at the polynomial

$$P(x) = x^n - \sum_{k=1}^{n} S^k_n [x]_k.$$  

By (1.1) it is clear that the numbers $1, 2, \ldots, n$ are all roots of $P$. But $0$ is also a root of $P$. We just finished exhibiting $n + 1$ distinct roots of $P$, and since $P$ is a polynomial of degree $n$ we conclude that $P$ must necessarily be the zero polynomial. That is,

$$x^n = \sum_{k=1}^{n} S^k_n [x]_k.$$
which (considering that $S_n^0$ is 0) is the content of (c). This ends our proof. (The technique employed to derive this result is a disguised form of Möbius inversion, which we study in detail in Chapter 9.)

Let us conclude Section 1.1.7 by proving the following formula:

(d) $S_{n+1}^m = \sum_{k=0}^{n} \binom{n}{k} S_k^{m-1}$ (when reading this formula the reader should recall that by definition $S_k^p = 0$ for $k < p$).

One may establish the above formula as follows. Consider the list of all partitions with $m$ classes of a set with $n + 1$ elements. (There are $S_{n+1}^m$ such partitions.)

Let us count these partitions in a different way. Fix an element, say $w$, of the $n + 1$ available elements. In each of the partitions eliminate the class containing $w$. We thus obtain partitions with precisely $m-1$ classes, on $k$ elements, where $k$ is at most $n$. Sorting out the partitions so obtained by the values of $k$ we obtain the formula stated in (d).

[To pay attention to detail, if $w$ belongs to a class of size $j$, by eliminating the class containing $w$, we obtain a partition with $m-1$ classes of a set with $n-j+1$ elements. Each choice of (a class of) $j$ elements leads to $S_{n-j+1}^{m-1}$ partitions of the remaining $n-j+1$ elements. The element $w$ being always among the $j$ elements selected, we have $\binom{n}{j-1}$ possibilities to select a class with $j$ elements that contains $w$, a totality of $\binom{n}{j-1} S_{n-j+1}^{m-1}$ choices. Observe, in addition, that the largest value $j$ may take is $n+2-m$ (since we must have $m$ classes in the original partition). Summing up over $j$ we obtain

$$S_{n+1}^m = \sum_{j=1}^{n+2-m} \binom{n}{j-1} S_{n-j+1}^{m-1} = \sum_{k=0}^{n} \binom{n}{k} S_k^{m-1}.$$]
1.1. THE ESSENTIALS OF COUNTING

1.1.8 Stirling Numbers of the First Kind

The set of bijections from a set $A$ to itself is the object of study in this section. Let us formally denote

$$\text{Sym } A = \{ \sigma : A \to A, \; \sigma \text{ bijection} \}.$$ 

An element of Sym $A$ is called a permutation. If $A$ has $n$ elements, then the cardinality of Sym $A$ is $n!$.

To fix ideas, say $A = \{1, 2, 3, 4, 5\}$. Line up the elements of $A$ in some order, say 12345, and keep this order as reference at all times. The writing (or sequence) 34521 indicates the bijection $\sigma$:

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 \\
\downarrow \sigma \\
3 & 4 & 5 & 2 & 1 \\
\end{array}
\begin{pmatrix}
\sigma(1) & \sigma(2) & \sigma(3) & \sigma(4) & \sigma(5) \\
i.e., & || & || & || & ||
\end{pmatrix}
\]

(The reader now understands why $|\text{Sym } A| = n!$, if $|A| = n$; this is so simply because $|\text{Sym } A|$ counts the number of sequences on $n$ symbols with no repetitions, i.e., $|\text{Sym } A| = [n]_n = n!$.)

Although this sequential notation for a permutation comes in handy quite often, it is the decomposition of a permutation into disjoint cycles that we want to emphasize. The decomposition of a permutation $\sigma$ into disjoint cycles is carried out as follows: pick a symbol, say 1, and list the symbols obtained by repeated applications of $\sigma$, that is, $(1\sigma(1)\sigma(\sigma(1))\sigma(\sigma(\sigma(1)))\cdots)$. Close the cycle when we get back to 1 again. If any symbols are left, repeat the process until none are left.
For example, the permutation

\[
\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 \\
\downarrow & \sigma \\
3 & 4 & 5 & 2 & 1
\end{array}
\]

has cycle decomposition \((1\ 3\ 5)(2\ 4) = \sigma\).

More Examples.

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\downarrow & \sigma \\
6 & 3 & 5 & 4 & 2 & 1
\end{array}
\]

corresponds to \((1\ 6)(2\ 3\ 5)(4)\) and

\[
\begin{array}{cccccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\downarrow & \sigma \\
5 & 8 & 3 & 7 & 4 & 6 & 1 & 2
\end{array}
\]

corresponds to \((1\ 5\ 4\ 7)(2\ 8)(3)(6)\). (The reader should write the cyclic decompositions of the permutations 2 7 6 9 1 4 5 8 and 6 3 4 7 2 5 1.)

It is self-evident that each permutation has a unique decomposition into disjoint cycles (up to a rearrangement of the cycles).

A cycle decomposition of a permutation on \(n\) symbols is said to be of type \(1^{\lambda_1}2^{\lambda_2}\cdots n^{\lambda_n}\) (just notation), if it has

\[
\begin{align*}
\lambda_1 & \text{ cycles of length 1} \\
\lambda_2 & \text{ cycles of length 2} \\
& \vdots \\
\lambda_n & \text{ cycles of length } n
\end{align*}
\]
We shall count the number of permutations on \( n \) symbols with precisely \( m \) cycles. One should conveniently think of a permutation simply as a collection of disjoint cycles in a formal way. (If necessary we can interpret this symbolic writing as a bijection on a finite set, but this is seldom necessary.) It should, however, be clear to the reader that a same cycle of length \( n \) can be written out in \( n \) different ways: for example, \( (3, 2, 4, 1) = (1, 3, 2, 4) = (4, 1, 3, 2) = (2, 4, 1, 3) \). In other words the cyclic motion within a cycle gives the same cycle. And this cyclic motion is the only change within a cycle that preserves it.

Let us observe that:

(a) The number of permutations of type \( 1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n} \) of a set with \( n \) elements (or symbols)

\[
\frac{n!}{1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n}(\lambda_1!)(\lambda_2!) \cdots (\lambda_n)!}.
\]

(The proof is almost the same as that of statement (a) in Section 1.1.7; replace the word "class" with "cycle" and recall that, apart from the cyclic order, the order of the elements within a cycle is relevant.)

Let \((-1)^{n-m}s_n^m\) be the number of permutations with exactly \( m \) (disjoint) cycles on a set of \( n \) elements. The numbers \( s_n^k \) are called Stirling numbers of the first kind.

By (a) above,

\[
(-1)^{n-m}s_n^m = \sum \frac{n!}{1^{\lambda_1}2^{\lambda_2} \cdots n^{\lambda_n}(\lambda_1!)(\lambda_2!) \cdots (\lambda_n)!},
\]

where the sum extends over all vectors \((\lambda_1, \lambda_2, \ldots, \lambda_n)\), with \( \lambda_j \) positive integers satisfying \( \sum \lambda_j = m \) and \( \sum i\lambda_i = n \).

Comparing this last expression with its counterpart for \( S_n^m \) in Section 1.1.7 one observes
straightaway that $S_n^m \leq |s_n^m|$, simply because in the former we divide to $k!$ while in the latter we divide only to $k$.

Another helpful thing to notice is that:

(b) The numbers $s_n^m$ satisfy the following recurrence relations:

$$s_n^0 = 0, \quad s_n^n = 1,$$

and

$$s_{n+1}^m = s_n^{m-1} - ns_n^m, \quad \text{for } 1 \leq m \leq n$$

(for convenience we set $s_n^m = 0$ for $m$ larger than $n$).

**Proof.** The chain of arguments parallels that of the proof of (b) in Section 1.1.7. Consider the list of the $(-1)^{n+1-m}s_{n+1}^m$ permutations on the $n + 1$ elements of set $A$ with $m$ cycles. Fix an element $w$ and split this list into two parts as follows:
1.1. THE ESSENTIALS OF COUNTING

This concludes the proof.

A list of small values for $s_n^m$ is given below:

<table>
<thead>
<tr>
<th>$m$</th>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>-6</td>
<td>11</td>
<td>-6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>24</td>
<td>-50</td>
<td>35</td>
<td>-10</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>-120</td>
<td>274</td>
<td>-225</td>
<td>85</td>
<td>-15</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Example. There are $(-1)^4-2s_4^2 = 11$ permutations on four symbols with exactly two cycles. These are:

$(1)(234)$, $(3)(124)$, $(12)(34)$

$(1)(432)$, $(3)(421)$, $(13)(24)$

$(2)(134)$, $(4)(123)$, $(14)(23)$

$(2)(431)$, $(4)(321)$

We now prove the following identity:

(c) $[x]_n = \sum_{k=0}^{n} s_n^k x^k$.

Proof. Let

$$[x]_n = x(x-1) \cdots (x-n+1) = c_n^0 + c_n^1 x + c_n^2 x^2 + \cdots + c_n^n x^n$$
(the symbol $i$ in $c_n^i$ being an index and not a power). Then
\[
\cdots + c_{n+1}^m x^m + \cdots = [x]_{n+1} = [x]_n (x - n)
\]
\[
= (\cdots + c_{n-1}^{m-1} x^{m-1} + c_n^m x^m + \cdots) (x - n)
\]
and we see that the numbers $c_n^m$ satisfy
\[
c_0^n = 0, \quad c_n^n = 1
\]
and
\[
c_{n+1}^m = c_{n}^{m-1} - n c_n^m.
\]
But we showed in (b) that the Stirling numbers $s_n^m$ satisfy these recurrence relations and
initial conditions. Thus $s_n^m = c_n^m$, and this ends our proof.

In the same way we proved the formula in (d), Section 1.1.7, we can prove the following:

(d) $|s_{n+1}^m| = \sum_{k=0}^{n} [n]_{n-k} |s_k^{m-1}|$

(e.g., $225 = |s_6^3| = [5][3]^2 |s_2^2| + [5][2]^2 |s_3^2| + [5][1]^2 |s_4^2| + [5][0]^2 |s_5^2|$.)

1.1.9 Bell Numbers

Let $B_n$ denote the number of partitions of a set with $n$ elements. The $B_n$’s are called Bell
numbers. By the definition of $B_n$ and the definition of $S_n^m$, the Stirling numbers of the
second kind, we have
\[
B_n = \sum_{m=1}^{n} S_n^m.
\]
As an example, let \( A = \{a, b, c, d\} \). Then the list of all the partitions of \( A \) is

\[
\{a, b, c, d\}, \{a\} \{b, c, d\}, \{a\} \{b\} \{c\} \{d\},
\{b\} \{a, c, d\}, \{a\} \{c\} \{b, d\},
\{c\} \{a, b, d\}, \{a\} \{d\} \{b, c\},
\{d\} \{a, b, c\}, \{b\} \{c\} \{a, d\},
\{a, b\} \{c, d\}, \{b\} \{d\} \{a, c\},
\{a, c\} \{b, d\}, \{c\} \{d\} \{a, b\},
\{a, d\} \{b, c\},
\]

(15 partitions in all).

Observe that

\[
B_{n+1} = \sum_{m+1}^{n+1} S_m^{n+1} = \sum_{m=1}^{n} \sum_{k=0}^{n} \binom{n}{k} S_k^{m-1}
= \sum_{k=0}^{n} \binom{n}{k} \sum_{m=1}^{n+1} S_k^{m-1} = \sum_{k=0}^{n} \binom{n}{k} B_k.
\]

The second equality sign is explained by (d), Section 1.1.7.

We proved the following:

(a) \( B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k \).

A formula that is appropriate to state here, though we postpone its proof until introducing generating functions in the next chapter (see specifically, Section 2.7, formula 3), is the following:

(b) \( B_n = e^{-1} \sum_{m=0}^{\infty} m^n / m! \). (This is Dobinski’s formula, where \( e \) is the Eulerian constant \( e \approx 2.71828 \ldots \)).

Observe, by the way, that the analog of the Bell number \( B_n \) for permutations is \( n! \), that is, \( n! = \sum_{m=1}^{n} |s_n^m| \). And since we know that \( S_n^m \leq |s_n^m| \), this tells us that \( B_n \leq n! \). A
more refined perception of the magnitude of $B_n$ can be cultivated upon reading Section 2.7.

1.2 OCCUPANCY

1.2.1

As need arises, we allow $A$ to be an assemblage of elements *distinguishable* from one another (in which case $A$ is a set) or have it consist of a number of *indistinguishable* elements, such as $A = \{1, 1, 1, 1\}$. The number of elements of $A$, distinguishable or not, is indicated by $|A|$.

Without specifying whether the elements of $A$ and $B$ are distinguishable or indistinguishable, by a *function* from $A$ to $B$ we understand a rule by which we assign to each element of $A$ a single element of $B$.

Compiled below are results regarding the numbers of functions, injections, surjections, and bijections from $A$ to $B$ ($|A| = n, |B| = m$).

<table>
<thead>
<tr>
<th>Elements of $A$</th>
<th>Elements of $B$</th>
<th>Functions</th>
<th>Injections</th>
<th>Surjections</th>
<th>Bijections</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$ Distinguish.</td>
<td>Distinguish.</td>
<td>$m^n$</td>
<td>$[m]_n$</td>
<td>$m!S^m_n$</td>
<td>$n!$</td>
</tr>
<tr>
<td>$\beta$ Indistinguish.</td>
<td>Distinguish.</td>
<td>$\binom{n+m-1}{m}$</td>
<td>$\binom{m}{m}$</td>
<td>$\binom{n-1}{m-1}$</td>
<td>1</td>
</tr>
<tr>
<td>$\gamma$ Distinguish.</td>
<td>Indistinguish.</td>
<td>$\sum_{k=1}^m S^k_n$</td>
<td>1</td>
<td>$S^m_n$</td>
<td>1</td>
</tr>
<tr>
<td>$\delta$ Indistinguish.</td>
<td>Indistinguish.</td>
<td>$\sum_{k=1}^m P_k(n)$</td>
<td>1</td>
<td>$P_m(n)$</td>
<td>1</td>
</tr>
</tbody>
</table>

In this table the entry $S^m_n$, for example, signifies the number of surjections from a set with
n distinguishable elements to a set with m indistinguishable elements. \(S_n^m\) is the Stirling number of the second kind.)

We devote the remainder of Section 2 to explaining the notation and proving the statements contained in this table.

\((\alpha)\) \(\{\text{Distinguishable}\} \xrightarrow{f} \{\text{Distinguishable}\}\)

**Functions.** \(A \xrightarrow{f} B; |A| = n, |B| = m.\) For each element in A we have m choices to map it to. A total of \(m^n\) choices.

**Injections.** \((n \leq m.)\) When defining an injection, we have m choices for the first element of A, \(m - 1\) choices for the second, \(m - 2\) for the third,..., \(m - n + 1\) for the nth. In all \([m]_n\) possibilities.

**Surjections.** \((m \leq n.)\) For each element \(b\) in B, consider the subset \(A_b = \{a \in A : f(a) = b\}\). Since \(f\) is a surjective function, the subsets \(\{A_b : b \in B\}\) form a partition of A with \(m\) classes. Conversely, to each partition of A with \(m\) classes, there correspond \(m!\) surjections from A to B. There are \(S_n^m\) partitions with \(m\) classes of A, and this gives a total of \(m!S_n^m\) surjections.

**Bijections.** \((n = m.)\) There are \([n]_n = n!\) of these.

\((\beta)\) \(\{\text{Indistinguishable}\} \xrightarrow{f} \{\text{Distinguishable}\}\)

**Bijections.** \((n = m.)\) There is only one bijection.

**Injections.** \((n \leq m.)\) Think of injecting the \(m\) (indistinguishable) elements of A into B. This corresponds to a choice of \(n\) elements out of the \(m\) available in B. The order in which we identify these \(n\) elements of B is of no consequence since the elements of A are indistinguishable. This shows that there are \(\binom{m}{n}\) injections (one associated to each subset
Functions. The sort of counting we perform here is quite relevant throughout the upcoming chapter. There is a little trick by which we do the counting, actually, which we hope the reader will appreciate. An entire section is now devoted to this.

1.2.2

We assert that the following four statements are effectively the same, and we substantiate this assertion by exhibiting explicit bijections between them:

(a) *The number of ways to distribute \( n \) indistinguishable balls into \( m \) distinguishable boxes is* \( \binom{n+m-1}{n} \).

(b) *The number of vectors \((n_1, n_2, \ldots, n_m)\) with nonnegative integer entries satisfying* \( n_1 + n_2 + \cdots + n_m = n \)

*is* \( \binom{n+m-1}{n} \).

(c) *The number of ways to select \( n \) objects with repetitions from \( m \) different types of objects is* \( \binom{n+m-1}{n} \). (We assume that we have an unlimited supply of objects of each type and that the order of selection of the \( n \) objects is irrelevant.)

(d) *The number of functions from \( n \) indistinguishable elements to \( m \) distinguishable elements is* \( \binom{n+m-1}{n} \).

When we assert that these statements are the same we mean that there are bijective correspondences between them. In other words, it’s like meeting the same person on four different occasions, each time wearing a different attire. When appearances are ignored you do recognize the same person.
The disguise between (a) and (b) is pretty thin, actually. When placing the $n$ balls within the $m$ boxes we place $n_1$ in box 1, $n_2$ in box 2, ..., $n_m$ in box $m$ ($\sum_{i=1}^{m} n_i = n$). To this process we naturally attach the vector $(n_1, n_2, \ldots, n_m)$. Counting the number of such vectors is the same thing as counting the ways of placing the $n$ indistinguishable balls into the $m$ distinguishable boxes. [We prove shortly that the answer to this counting problem is $\binom{n+m-1}{n}$.

With regard to (a) and (c): Selecting $n$ objects with repetitions from $m$ distinct types of objects involves a selection of $n_1$ objects of type 1, $n_2$ of type 2, ..., $n_m$ of type $m$. This is the same as placing $n_1$ balls in box 1, $n_2$ in box 2, ..., $n_m$ in box $m$.

Statement (d) is surely the same as (a) upon thinking of each distinguishable element as a box and each indistinguishable element as a ball. A function is merely a way of assigning the indistinguishable balls to the distinguishable boxes.

(That all four statements are the same follows now by transitivity.)

The numerical answer shared by statements (a) through (d) is $\binom{n+m-1}{n}$. To see this we first solve the following problem:

* The number of ways of placing $n$ distinguishable balls into $m$ distinguishable boxes, paying attention to the order in which the balls are placed within the boxes, is equal to $[n+m−1]_n = m(m + 1)...(m + n - 1)$.

Proof. The trick is to visualize an assignment of balls to boxes as a sequence of $n$ balls and $m−1$ '/'s, as displayed below:
We now wish to count the number of distinct sequences of $n + m - 1$ objects consisting of the $n$ distinguishable balls and the $m - 1$ indistinguishable /'s. If the /'s were also distinguishable, we would have had $(n + m - 1)!$ sequences in all. Since they are not, we may permute the $m - 1$ available /'s among themselves in whichever positions they occur [and this can be done in $(m - 1)!$ ways] without changing the sequence. The number of distinct sequences is therefore $(n + m - 1)!/(m - 1)!$ or $[n + m - 1]_n$. This ends the proof.

When the $n$ objects are indistinguishable, we obtain the following consequence: The number of ways of placing $n$ indistinguishable objects into $m$ distinguishable boxes is

$$
\frac{[n + m - 1]_n}{n!} = \binom{n + m - 1}{n}.
$$

This, however, is statement (a).

*Example.* With two distinguishable balls $z$ and $w$, and three distinguishable boxes, the $[2 + 3 - 1]_2$ possibilities are:

- $zw//, /zw/, //zw, z/w/, z//w, /z/w,$
- $wz//, /wz/, //wz, w/z/, w//z, /w/z.$

If $z$ and $w$ are indistinguishable we have

$$
\frac{[2 + 3 - 1]_2}{2!} = \binom{2 + 3 - 1}{2}
$$
1.2. OCCUPANCY

possibilities:
\[ \cdot /, \quad / \cdot, \quad / \cdot /, \quad / \cdot /, \quad / \cdot \cdot. \]

Initially we wanted to count the number of functions from \( n \) indistinguishable elements to \( m \) distinguishable elements. By the counting we just did, and the equivalence of statements (a) and (d), we conclude that there are \( \binom{n+m-1}{n} \) functions.

**Surjections.** \( (m \leq n). \) Think of the elements of \( A \) being \( n \) indistinguishable balls and the elements of \( B \) being \( m \) distinguishable boxes. We then ask for the number of ways of placing the \( n \) balls into the \( m \) boxes with no box left empty.

We count as follows: place initially one ball within each box. There are \( n-m \) indistinguishable balls remaining, which we can now place into the \( m \) boxes without restrictions. The number of surjections from \( n \) indistinguishable elements to \( m \) distinguishable ones equals therefore the number of functions from the remaining \( n-m \) indistinguishable elements to the \( m \) distinguishable elements. We just finished counting the number of such functions. There are
\[
\binom{n-m+m-1}{n-m} = \binom{n-1}{n-m} = \binom{n-1}{m-1}
\]
of them.

Let us proceed with explaining the remaining entries in our table.

\((\gamma)\{\text{Distinguishable}\} \xrightarrow{f} \{\text{Indistinguishable}\}\)

**Bijections.** \( (n = m.) \) There is only one bijection.

**Injections.** \( (n \leq m.) \) There is only one injection.

**Surjections.** \( (n \geq m.) \) Given a surjection from \( A \) to \( B \), the preimages of the indistinguishable elements of \( B \) form a partition of \( A \) with \( m \) classes. Conversely, given a partition
with \(m\) classes of \(A\), one canonically constructs a surjection by mapping all the elements of a class to an (indistinguishable) element of \(B\). The number of surjections is therefore \(S^m_n\), the number of partitions of a set of \(n\) elements into \(m\) classes (a Stirling number).

**Functions.** We have \(\sum_{k=1}^m S^n_k\) functions from \(n\) distinguishable elements to \(m\) indistinguishable ones. This is apparent from the above way in which we counted the surjections, since any function is a surjection onto its image.

\[
(\delta) \{\text{Indistinguishable}\} \xrightarrow{L} \{\text{Indistinguishable}\}
\]

We have one bijection and one injection. Counting surjections is difficult. No closed formula is known. Let us denote by \(P_m(n)\) the number of surjections from \(n\) indistinguishable objects to \(m\) indistinguishable objects \((n \geq m)\). With this notation, the number of functions we seek is \(\sum_{k=1}^m P_k(n)\), again, because any function is a surjection on its image.

The task of specifying a surjection from a collection of \(n\) indistinguishable objects to another of \(m\) indistinguishable ones is equivalent to that of writing \(n\) as a sum of precisely \(m\) positive integers. The positive integers whose sum is \(n\) are commonly called the *parts* of \(m\). Hence \(P_m(n)\) can be interpreted as the number of ways of writing \(n\) as a sum of \(m\) (positive) parts. The order in which we list the parts is, of course, irrelevant here. We may as well then list the parts nonincreasingly, and specify \(P_m(n)\) as

\[
P_m(n) = \{ (\alpha_1, \alpha_2, \ldots, \alpha_m) : \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_m, \\
\sum_{i=1}^m \alpha_i = n, \text{ and } \alpha_i's \text{ are positive integers} \}.
\]

We refer the reader to Section 7 of Chapter 2 for more information on the numbers \(P_m(n)\), which we call the number of partitions of \(n\) with \(m\) parts.
1.3 MORE ON COUNTING

1.3.1 Multinomial Numbers

For \( n \) and \( n_i (1 \leq i \leq m) \) positive integers satisfying \( \sum_{i=1}^{m} n_i = n \), the numbers \( \frac{n!}{(n_1!n_2!\cdots n_m!)} \) have combinatorial meaning. [Observe that for \( m = 2 \), \( n_1 + n_2 = n \), we obtain the binomial numbers \( \binom{n}{n_1} \).] We call \( \frac{n!}{(n_1!n_2!\cdots n_m!)} \) multinomial numbers. We prove the following three statements:

* Given \( n \) distinguishable objects and \( m \) distinguishable boxes (labeled 1, 2, \ldots, \( m \)), the number of ways of placing \( n_1 \) objects in box 1, \( n_2 \) objects in box 2, \ldots, \( n_m \) objects in box \( m \) is \( \frac{n!}{(n_1!n_2!\cdots n_m!)} \) (the order in which the \( n_i \) objects are placed in box \( i \) is irrelevant) \( (\sum_{i=1}^{m} n_i = n) \).

* Given

\[
\begin{align*}
n_1 & \text{ indistinguishable objects of color 1} \\
n_2 & \text{ indistinguishable objects of color 2} \\
& \quad \vdots \\
n_m & \text{ indistinguishable objects of color } m
\end{align*}
\]

the number of distinct sequences of length \( n \) possible to make with these objects is \( \frac{n!}{(n_1!n_2!\cdots n_m!)} \). (How many sequences of length 11 can there be made with the letters of the word Mississippi?)

* The number of nondecreasing paths from \((0, 0, \ldots, 0)\) to \((n_1, n_2, \ldots, n_m)\) (with \( \sum_{i=1}^{m} n_i = n \)) on the \( m \)-dimensional lattice of integral points is \( \frac{n!}{(n_1!n_2!\cdots n_m!)} \). (A path is understood to be nondecreasing if the coordinates of the points successively selected form
nondecreasing sequences in each of the coordinates.)

Let us prove these three statements. The first assertion is proved as follows: Box 1 can be filled in \( \binom{n}{n_1} \) ways; with box 1 full, box 2 can be filled in \( \binom{n-n_1}{n_2} \) ways; with boxes 1 and 2 full, box 3 can be filled in \( \binom{n-n_1-n_2}{n_3} \) ways, and so on. The total number of ways is therefore

\[
\binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \cdots \binom{n}{m} = \frac{n!}{n_1! n_2! \cdots n_m!}.
\]

Statement two is also quite easy to prove. If the \( n \) objects were all distinguishable we would have had \( n! \) sequences. Objects of color 1, however, can be permuted among themselves in \( n_1! \) ways (in whichever positions they occur). The same is true for the other colors. We are thus left with \( n!/(n_1!n_2!\cdots n_m!) \) distinct sequences.

One can easily see that the last statement is effectively the same as the second. Indeed, color coordinate \( i \) with color \( i \). Then attach a sequence of colors to a nondecreasing path by marking down the color of the coordinate that changes at each step of the way. (Through this process the notions of a nondecreasing path and a sequence of colors become identical.) By statement two we conclude that there are \( n!/(n_1!n_2!\cdots n_m!) \) such paths.

The last of the three combinatorial interpretations of the multinomial numbers appears to be the richest and most convenient to rely on, chiefly because of its geometrical appeal. One good example of this is the proof of the *multinomial Vandermonde convolution*:

\[
\frac{(p + q)!}{n_1! n_2! \cdots n_m!} = \sum \frac{p!}{k_1! k_2! \cdots k_m!} \frac{q!}{(n_1 - k_1)! (n_2 - k_2)! \cdots (n_m - k_m)!}
\]

where the sum is over all points \((k_1, k_2, \ldots, k_m)\), with \( 0 \leq k_i \leq n_i \), \( k_i \) integral, and \( \sum_{i=1}^{m} k_i = p \).
Indeed, \((p + q!/(n_1! \cdots n_m!))\) counts the number of paths of length \(p + q\) between 
\((0, \ldots, 0)\) and \((n_1 \cdots n_m)\). We now count these paths in a different way. Any path of 
length \(p + q\) ending at \((n_1, \ldots, n_m)\) is composed of a path of length \(p\) ending, say, at 
\((k_1, \ldots, k_m)\), followed by a path of length \(q\) from \((k_1, \ldots, k_m)\) to \((n_1, \ldots, n_m)\). By placing 
the origin at \((k_1, \ldots, k_m)\) momentarily, the latter path of length \(q\) can be envisioned as 
a path between \((0, \ldots, 0)\) and \((n_1 - k_1, \ldots, n_m - k_m)\). There are \(p!/(k_1! \cdots k_m!)(n_1 - k_1)!(n_2 - k_2)\cdots(n_m - k_m)!\) choices 
for the initial path of length \(p\) and \(q!/(n_1 - k_1)! \cdots (n_m - k_m)!\) for its follow-up path of 
length \(q\). With the point \((k_1, \ldots, k_m)\) fixed we thus have 

\[
\frac{p!}{k_1!k_2! \cdots k_m!} \frac{q!}{(n_1 - k_1)!(n_2 - k_2)\cdots(n_m - k_m)!}
\]

paths passing through it. Summing up over all intermediate points \((k_1, \ldots, k_m)\) \(p\) steps 
away from the origin, we obtain the multinomial Vandermonde convolution. (Throughout 
the proof of this formula we understand by a path a nondecreasing path, of course.)

### 1.3.2 Lah Numbers

The numbers \((n!/m!)(n-1\choose m-1)\) have combinatorial meaning, specifically as follows:

* The number of ways of placing \(n\) distinguishable objects into \(m\) indistinguishable boxes 
  with no box left empty, paying attention to the order in which they are put inside the 
  boxes, is equal to \((n!/m!)(n-1\choose m-1)\).

To see this, think initially of the boxes being distinguishable and symbolized by the 
slashes, /.

```
Box 1  Box 2  Box m
/    /   ...   /
```
(We have \(m - 1\) slashes in all.)

Take any sequence formed by the \(n\) objects

\[
a_1 \wedge a_2 \wedge a_3 \wedge a_4 \wedge a_5 \wedge \cdots \wedge a_{n-1} \wedge a_n.
\]

The \(n - 1\) hats "\(\wedge\)" indicate the possible places where slashes can be inserted. No consecutive slashes are permitted. We have to insert \(m - 1\) slashes in \(n - 1\) places (indicated by the hats), a task that can be accomplished in \(\binom{n-1}{m-1}\) ways. Since there are \(n!\) sequences possible to make with the \(n\) distinguishable objects, we obtain a total of \(n!\binom{n-1}{m-1}\) ways of assigning the \(n\) distinguishable objects to the \(m\) distinguishable boxes with no box left empty. But the \(m\) boxes are in fact indistinguishable, thus leaving us with \((n!/m!)/\binom{n-1}{m-1}\) ways.

We define numbers \(L^m_n\) by setting,

\[
L^m_n = (-1)^n \frac{n!}{m!} \binom{n-1}{m-1}
\]

and call them Lah numbers. They have many nice properties, some of which are described in Section 2.9 of Chapter 2.

Example. \(n = 4, m = 3\). The \((4!/3!)\binom{4-1}{3-1}\) = 12 ways are indicated below:

\[
1/2/34 1/3/24 1/4/23
1/2/43 1/3/42 1/4/32
2/3/14 2/4/13 3/4/12
\]
1.4 EXERCISES

1. Is it more likely to obtain a sum of 9 when rolling two fair dice or when rolling three?

2. Are there as many subsets of even cardinality as there are of odd cardinality? (Exhibit a “constructive” bijection, if possible.)

3. How many sequences of length five with 0 or 1 as entries contain at least two consecutive 0’s? (Attempt to generalize and answer this question for sequences of length \( n \) that contain runs of \( k \) or more 0’s.)

4. A basket contains four red, five yellow, and seven green apples. Pick six apples at random. What is the chance that two are red, three yellow, and one green?

5. How many subsets of three (distinct) integers between 1 and 90 are there whose sum is: (a) even, (b) divisible by 3, (c) divisible by 9?

6. Give a combinatorial interpretation to the inequality

\[
\binom{m}{n-1} \leq \binom{m}{n}, \quad \text{for } 1 \leq n \leq \frac{m+1}{2}.
\]

[Hint: Construct a bijection between two Cartesian products.]

7. Show that:

(a) \( \sum_{k=m}^{n} \binom{k}{m} = \binom{n+1}{m+1} \).

(b) \( \sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n} \).

(c) \( \sum_{k=0}^{p} \binom{p}{k} \binom{q}{k} \binom{n+k}{p+q} = \binom{n}{p} \binom{n}{q} \).

(Hint: Use nondecreasing lattice paths.)
8. What is the probability that a random hand of eight cards (out of a usual deck of 52 cards) has:

(a) Two cards of each suit?

(b) The same number of hearts and spades?

(c) Two hearts and at least three spades, and the values of the spades are all strictly greater than the values of the hearts?

9. What is the chance that two or more people have the same birthdate (i.e., same month and day) among a random group of \( n \) people?

10. How many ways are there to invite one of three friends over for dinner on six successive evenings such that no friend is invited more than three times?

11. If you flip a fair coin 20 times and get 12 heads and 8 tails, what is the chance that there are no consecutive tails?

12. How many sequences formed with the letters in the word ”Mississippi” have no consecutive i’s and no consecutive s’s?

13. Let \( p \) be a prime. For \( 1 \leq k \leq p-2 \) show that \( p \) divides \( \binom{p}{k} \). Show also that \( \binom{2p}{p} = 2 \) (modulo \( p \)). Prove that \( \binom{pm}{nm} = \binom{n}{m} \) (modulo \( p \)), where \( m \) and \( n \) are nonnegative integers.

14. How many triangles can be drawn all of whose vertices are among those of a given \( n \)-gon and all of whose sides are diagonals (but not sides) of the \( n \)-gon?

15. Like the binomial numbers, the Stirling numbers \( S_n^0, S_n^1, \ldots, S_n^k, \ldots, S_n^n \) (and \( |s_n^0|, |s_n^1|, \ldots, |s_n^n| \)) start off by increasing up to a certain value of \( k \) and then strictly
decrease. The maximum, however, is not achieved at the middle terms, as is the case with the binomial numbers. Prove this.

16. Show that the (absolute value of the) Stirling number of the first kind, \(|s^k_n|\), is the sum of all products of \(n - k\) different integers taken from \(\{1, 2, \ldots, n - 1\}\). For example,

\[ s^3_5 = 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 4 + 2 \cdot 3 + 2 \cdot 4 + 3 \cdot 5. \]

Show that \(|s^k_n| < (n - k)!\binom{n-1}{k-1}^2\).

17. Show that the Stirling number of the second kind, \(S^k_n\), is the sum of all products of \(n - k\) not necessarily distinct integers taken from \(\{1, 2, \ldots, k\}\). For example,

\[ S^3_5 = 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 + 2 \cdot 2 + 2 \cdot 3 + 3 \cdot 3. \]

Prove also that \(S^k_n < k^{n-k}\binom{n-1}{k-1}\).

18. Prove that \(\sum_{k=m}^{n} \binom{k}{m}|s^k_n| = |s^{m+1}_n|\).

19. Prove that \(S^n_m = (1/ml) \sum_{k=0}^{n-1} (-1)^k \binom{m}{k}(m-k)^n\).

20. Prove that the number of sequences of integers \(x_1x_2\cdots x_r\) \((1 \leq x_i \leq n)\) containing less than \(i\) entries less than or equal to \(i\) (for \(i = 1, 2, \ldots, n\)) is \((n-r)n^{r-1}; 1 \leq r \leq n\).

21. Persons \(A\) and \(B\) independently proofread a book. \(A\) finds \(a\) errors, \(B\) finds \(b\) errors, while \(c\) errors were spotted by both \(A\) and \(B\). Estimate the number of remaining errors in the book.

22. In how many ways can \(r\) rooks be placed on a chessboard with \(m\) rows and \(n\) columns such that no rook can attack another?
23. Show that
\[ \sum |A_1 \cup A_2 \cup \cdots \cup A_k| = n(2^k - 1)^{k(n-1)}, \]
where the sum is taken over all ordered selections of subsets \( A_1, A_2, \ldots, A_k \) of a set with \( n \) elements.

24. What is the largest number of subsets noncomparable with respect to inclusion of a set with \( n \) elements? (Noncomparable with respect to inclusion means \( A_i \not\subseteq A_j \) for \( i \neq j \).)

25. Prove that among a set of \( n + 1 \) positive integers, none of which exceed \( 2n \), at least one divides another.

26. Select the minimal element in each of the \( \binom{n}{r} \) subsets of \( \{1, 2, \ldots, n\} \). Show that the average of the numbers so obtained is \( (n + 1)(r + 1)^{-1} \).

27. Show that the number of permutations \( \sigma \) of the set \( \{1, 2, \ldots, n\} \) that have the property that there exist exactly \( k \) elements \( j \) for which \( \sigma(j) > \sigma(i) \), for every \( i < j \), is equal to \( |s_n^k| \).

1.5 HISTORICAL NOTE

Fascination with producing configurations with remarkable properties, or counting the exact number of available choices goes back in history a rather long way. The square
has the "magical" property that the sum of dots along each of the four edges and the two diagonals is always 15. Legend tells [1] that it was observed on the back of a divine tortoise that emerged from the River Lo (in China) some 4000 years ago.

River Ho was known to shelter a no less remarkably decorated tortoise, displaying a magic figure which in present day notation takes the following form:

\[ \begin{array}{c}
6 \\
1 \\
10 \\
9 & 4 & 5 & 3 & 8 \\
10 \\
2 \\
7 
\end{array} \]

The reader will observe the symmetry displayed with respect to the center, for example

\[ 5 + 4 = 9; \ 4 + 10 + 1 = 9 + 6; \ 2 + 10 + 4 = 9 + 7. \]
Written documentation on binomial numbers dates back to the twelfth century’s Indian school of arithmetic led by Bhaskra. The binomial recurrence that we call "Pascal’s triangle" was actually known to Nasir-Ad-Din [2], a Persian philosopher of the thirteenth century. In connection with games of chance the binomial numbers were rediscovered by Pascal and Fermat about the middle of the seventeenth century, to become commonplace in the works of Euler and Laplace.

James Stirling is responsible for bringing to attention the two species of numbers that now bear his name. Descendant of one of the oldest landed families of Scotland, he was educated at Oxford and was in correspondence with most of the elite mathematicians of his day, such as Euler, Newton, Bernoulli, and MacLaurin (see [4]). The numbers in question appear in his major work [3], along with the more famous Stirling formula that approximates factorials. Taking not quite the form we recognize today, his original formula was an expansion of the logarithm of the factorial. A few generations earlier, Napier, the inventor of the logarithms (and a neighboring landowner), had married Elizabeth, a Stirling. Intermarriage between the two families has in fact occurred more than once.

We recommend [5] and [6] for more information on the topics covered in this chapter.

1.6 REFERENCES


