1

Basic Counting

This chapter develops the basic counting techniques that form the foundation of enumerative combinatorics. We apply these techniques to study fundamental combinatorial structures such as words, permutations, subsets, functions, and lattice paths. The end of the chapter gives some applications of combinatorics to probability theory.

1.1 Review of Set Theory

We assume the reader is familiar with elementary aspects of logic and set theory, including proofs by induction. This material may be found in texts such as [34, 126]. Table 1.1 reviews the notation we will use from set theory. The word *iff* is defined to mean "if and only if."

TABLE 1.1
Review of notation from set theory.

Concept	Symbol	Meaning
membership	$x \in S$	x is an element of the set S.
set-building	$\{x : P(x)\}$	$y \in \{x : P(x)\}$ iff $P(y)$ is true.
subset	$A \subseteq B$	For all x, $x \in A$ implies $x \in B$.
set equality	$A = B$	For all x, $x \in A$ iff $x \in B$.
empty set	$\emptyset$	For all x, $x \notin \emptyset$.
cardinality	$\lvert A\rvert = n$	The set A has exactly n members.
union	$A \cup B$	$x \in A \cup B$ iff $x \in A$ or $x \in B$.
intersection	$A \cap B$	$x \in A \cap B$ iff $x \in A$ and $x \in B$.
set difference	$A \sim B$	$x \in A \sim B$ iff $x \in A$ and $x \notin B$.
ordered pair	(a, b)	$(a, b) = (c, d)$ iff $a = c$ and $b = d$.
Cartesian product	$A \times B$	$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}$.
finite union	$A_1 \cup \cdots \cup A_n$	$x \in A_1 \cup \cdots \cup A_n$ iff $x \in A_i$ for at least one $i \leq n$.
finite intersection	$A_1 \cap \cdots \cap A_n$	$x \in A_1 \cap \cdots \cap A_n$ iff $x \in A_i$ for all $i \leq n$.
ordered n-tuple	$(a_1, \ldots, a_n)$	$(a_1, \ldots, a_n) = (b_1, \ldots, b_n)$ iff $a_i = b_i$ for $1 \leq i \leq n$.
finite product	$A_1 \times \cdots \times A_n$	$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) : a_i \in A_i \text{ for } i \leq n\}$.

We use the notation $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$, $\mathbb{N}^+ = \{1, 2, 3, \ldots\}$, $\mathbb{Z}$ for the set of all integers, $\mathbb{Q}$ for the set of rational numbers, $\mathbb{R}$ for the set of real numbers, and $\mathbb{C}$ for the set of complex numbers. Informally, the notation $|A| = n$ means that A is a set consisting of n elements. We will give a more formal discussion of cardinality later (§1.6).

Two sets A and B are *disjoint* iff $A \cap B = \emptyset$. More generally, the sets $A_1, \ldots, A_n$ are called *pairwise disjoint* iff $A_i \cap A_j = \emptyset$ for all $i \neq j$. This means that no two sets in the given list overlap one another.

1.2 Sum Rule

The starting point for enumerative combinatorics is the following basic fact.

1.1. Counting Principle. If A and B are finite disjoint sets, then $|A \cup B| = |A| + |B|$.

The requirement that A and B be *disjoint* is certainly necessary. For example, if $A = \{1, 2, 3\}$ and $B = \{3, 5\}$, then $|A \cup B| = 4$, while $|A| + |B| = 3 + 2 = 5$. We will give a formal proof of 1.1 later (see 1.32). For now, let us deduce some consequences of this counting principle.

1.2. Sum Rule. If $A_1, \ldots, A_m$ are pairwise disjoint finite sets, then

$$|A_1 \cup \cdots \cup A_m| = |A_1| + \cdots + |A_m|.$$

Proof. We use induction on m. The case $m = 1$ is immediate, while the case $m = 2$ is true by 1.1. For $m > 2$, assume the result is known for $m-1$ sets. In 1.1, let $A = A_1 \cup \cdots \cup A_{m-1}$ and $B = A_m$. By induction hypothesis,

$$|A| = |A_1| + \cdots + |A_{m-1}|.$$

Since A_m does not intersect any A_j with $j < m$, we see that A and B are disjoint. So 1.1 gives

$$|A_1 \cup \cdots \cup A_m| = |A \cup B| = |A| + |B| = |A_1| + \cdots + |A_{m-1}| + |A_m|. \qquad \square$$

1.3. Difference Rule. If S and T are finite sets such that $T \subseteq S$, then $|S \sim T| = |S| - |T|$.

Proof. The set S is the union of the disjoint sets T and $S \sim T$. Therefore, 1.1 gives $|S| = |T| + |S \sim T|$. Subtracting the finite quantity $|T|$ from both sides gives the result. $\square$

We can generalize 1.1 to the case where the two sets in question are not disjoint, as follows.

1.4. Binary Union Rule. If A and B are arbitrary finite sets, then $|A \cup B| = |A| + |B| - |A \cap B|$.

Proof. Note that A is the disjoint union of $A \sim B$ and $A \cap B$; B is the disjoint union of $B \sim A$ and $A \cap B$; and $A \cup B$ is the disjoint union of $A \sim B$, $B \sim A$, and $A \cap B$. See Figure 1.1. Applying the sum rule repeatedly, we see that

$$|A| = |A \sim B| + |A \cap B|; \quad |B| = |B \sim A| + |A \cap B|; \quad |A \cup B| = |A \sim B| + |B \sim A| + |A \cap B|.$$

Using the first two equations to eliminate $|A \sim B|$ and $|B \sim A|$ in the third equation, we obtain the desired result. $\square$

The sum rule can also be extended to a formula for $|A_1 \cup \cdots \cup A_n|$, where $A_1, \ldots, A_n$ are arbitrary (not necessarily pairwise disjoint) finite sets. This formula is called the *inclusion-exclusion formula*; we will study it later (Chapter 4).

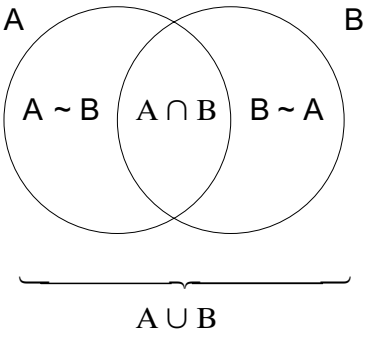


FIGURE 1.1
Proof of the binary union rule.

1.3 Product Rule

We can use the sum rule to compute the size of the Cartesian product of finite sets.

1.5. Product Rule for Sets. Suppose $S_1, \ldots, S_k$ are finite sets with $|S_i| = n_i$ for $1 \leq i \leq k$. Then

$$|S_1 \times S_2 \times \cdots \times S_k| = n_1 n_2 \cdots n_k.$$

Proof. We proceed by induction on k. There is nothing to prove when $k = 1$. Consider the case $k = 2$. Let $S_2 = \{x_1, x_2, \ldots, x_{n_2}\}$. The set $S_1 \times S_2$ is the disjoint union of the n_2 sets $S_1 \times \{x_1\}$, $S_1 \times \{x_2\}$, ..., $S_1 \times \{x_{n_2}\}$. Each of these sets has cardinality $|S_1| = n_1$. So, by the sum rule,

$$|S_1 \times S_2| = \sum_{i=1}^{n_2} |S_1 \times \{x_i\}| = \sum_{i=1}^{n_2} n_1 = n_1 n_2.$$

Next, let $k > 2$ and assume that the result is already known for products of $k - 1$ sets. We can regard $S_1 \times S_2 \times \cdots \times S_k$ as the Cartesian product $A \times B$, where $A = S_1 \times \cdots \times S_{k-1}$ and $B = S_k$. By induction, $|A| = n_1 n_2 \cdots n_{k-1}$. By the $k = 2$ case,

$$|S_1 \times S_2 \times \cdots \times S_k| = |A \times B| = |A| \cdot |B| = (n_1 n_2 \cdots n_{k-1}) n_k.$$

This completes the induction step. □

1.6. Example: License Plates. A California license plate consists of a digit, followed by three uppercase letters, followed by three more digits. Formally, we can view a license plate as an element of the set $S = D \times L \times L \times L \times D \times D \times D$, where $D = \{0, 1, 2, \ldots, 9\}$ and $L = \{A, B, C, \ldots, Z\}$. Thus,

$$|S| = 10 \times 26 \times 26 \times 26 \times 10 \times 10 \times 10 = 175,760,000.$$

1.7. Example: Phone Numbers. A phone number is a ten-digit sequence such that the first digit is not zero or one, while the second digit must be zero or one. Formally, we can view a phone number as an element of the set $S = \{2, 3, \ldots, 9\} \times \{0, 1\} \times D^8$, where the notation D^8 denotes the Cartesian product of 8 copies of the set $D = \{0, 1, \ldots, 9\}$. The number of phone numbers is

$$|S| = 8 \times 2 \times 10^8 = \text{ 1.6 billion.}$$

(To allow for more phone numbers, the restriction on the second digit of the area code was removed years ago.)

We will often be interested in finding the cardinality of a finite set S whose members are "structured objects." Frequently, we will be able to build up each object in S by making a sequence of choices. The next counting principle tells us how to compute $|S|$ in this situation.

1.8. Product Rule. Suppose each object x in a set S can be uniquely constructed by making a sequence of k choices. Suppose the first choice can be made in n_1 ways; the second choice can be made in n_2 ways (regardless of what the first choice was); and so on. In general, we suppose that the ith choice can be made in n_i ways, regardless of what happened in the first $i-1$ choices, for all $i \leq k$. Then

$$|S| = n_1 n_2 \cdots n_k.$$

The product rule is a consequence of 1.5, as we will explain in 1.34.

1.9. Example: Fraternity and Sorority Names. The name of a fraternity or sorority consists of any sequence of two or three uppercase Greek letters. (The Greek alphabet has 24 letters.) How many possible names are there? The set S of all such names is the disjoint union of S_2 and S_3, where S_k is the set of names of length k. Using the sum rule,

$$|S| = |S_2| + |S_3|.$$

We can calculate $|S_2|$ using the product rule. We build a typical word in S_2 by choosing the first letter (24 ways), then choosing the second letter (24 ways). By the product rule, $|S_2| = 24^2$. Similarly, $|S_3| = 24^3$, so $|S| = 24^2 + 24^3 = 14,400$. Note that we cannot directly use the product rule to calculate $|S|$, since the *number* of choices in a given application of the product rule must be fixed.

1.10. Example. How many three-digit odd numbers contain the digit 2 but not the digit 5? Let X be the set of all such numbers. We can write X as the disjoint union of three sets A, B, and C, where A consists of numbers in X with first and second digit 2, B consists of numbers in X with first digit 2 and second digit not 2, and C consists of numbers in X with second digit 2 and first digit not 2. To build a number in C, we choose the digits from left to right. There are seven choices for the first digit (we must avoid 0, 2, and 5), one choice for the second digit (it must be 2), and four choices for the third digit (which is odd and unequal to 5). By the product rule, $|C| = 7 \cdot 1 \cdot 4 = 28$. Similar reasoning shows that $|A| = 4$ and $|B| = 1 \cdot 8 \cdot 4 = 32$. Therefore, $|X| = |A| + |B| + |C| = 64$.

1.4 Words, Permutations, and Subsets

1.11. Definition: Words. Let A be a finite set. A *word* over the alphabet A is a sequence $w = w_1 w_2 \cdots w_k$, where each $w_i \in A$ and $k \geq 0$. The *length* of $w = w_1 w_2 \cdots w_k$ is k. Two words $w = w_1 w_2 \cdots w_k$ and $z = z_1 z_2 \cdots z_m$ are *equal* iff $k = m$ and $w_i = z_i$ for $1 \leq i \leq k$.

1.12. Example. Let $A = \{a, b, c, \ldots, z\}$ be the set of 26 lowercase letters in the English alphabet. Then *stop*, *opts*, and *stoops* are distinct words (of lengths 4, 4, and 6, respectively). If $A = \{0, 1\}$, the 8 words of length 3 over A are

$$000, \quad 001, \quad 010, \quad 011, \quad 100, \quad 101, \quad 110, \quad 111.$$

There is exactly one word of length zero, called the *empty word*. It is sometimes denoted by the special symbols $\cdot$ or ϵ.

1.13. Theorem: Enumeration of Words. If A is an n-letter alphabet and $k \geq 0$, then there are n^k words of length k over A.

Proof. We can uniquely construct a typical word $w = w_1 w_2 \cdots w_k$ by a sequence of choices. First, choose $w_1 \in A$ to be any of the n letters in A. Second, choose $w_2 \in A$ in any of n ways. Continue similarly, choosing $w_i \in A$ in any of n ways for $1 \leq i \leq k$. By the product rule, the number of words is $n \times n \times \cdots \times n$ (k factors), which is n^k. Note that the empty word is the unique word of length 0 over A, so our formula holds for $k = 0$ also. □

1.14. Definition: Permutations. Let A be an n-element set. A *permutation* of A is a word $w = w_1 w_2 \cdots w_n$ in which each letter of A appears exactly once. For example, the 6 permutations of $A = \{x, y, z\}$ are

$$xyz, \quad xzy, \quad yxz, \quad yzx, \quad zxy, \quad zyx.$$

1.15. Definition: Factorials. For each integer $n \geq 1$, *n-factorial* is

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1,$$

which is the product of the first n positive integers. We also define $0! = 1$.

1.16. Theorem: Enumeration of Permutations. There are $n!$ permutations of an n-letter alphabet A.

Proof. Build a typical permutation $w = w_1 w_2 \cdots w_n$ of A by making n choices. First, choose w_1 to be any of the n letters of A. Second, choose w_2 to be any of the $n-1$ letters of A different from w_1. Third, choose w_3 to be any of the $n-2$ letters of A different from w_1 and w_2. Proceed similarly; at the nth stage, choose w_n to be the unique letter of A that is different from $w_1, w_2, \ldots, w_{n-1}$. By the product rule, the number of permutations is $n \times (n-1) \times \cdots \times 1 = n!$. The result also holds when $n = 0$. □

1.17. Definition: k-Permutations. Let A be an n-element set. A *k-permutation* of A is a word $w = w_1 w_2 \cdots w_k$ consisting of k *distinct* letters in A. For example, the twelve 2-permutations of $A = \{a, b, c, d\}$ are

$$ab, \quad ac, \quad ad, \quad ba, \quad bc, \quad bd, \quad ca, \quad cb, \quad cd, \quad da, \quad db, \quad dc.$$

An n-permutation of A is the same as a permutation of A.

1.18. Theorem: Enumeration of k-Permutations. Suppose A is an n-letter alphabet. For $0 \leq k \leq n$, the number of k-permutations of A is

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

For $k > n$, there are no k-permutations of A.

Proof. Build a typical k-permutation $w = w_1 w_2 \cdots w_k$ of A by making k choices. First, choose w_1 to be any of the n letters of A. Second, choose w_2 to be any of the $n-1$ letters of A different from w_1. Continue similarly. When we choose w_i (where $1 \leq i \leq k$), we have already used the $i-1$ distinct letters $w_1, w_2, \ldots, w_{i-1}$. Since A has n letters, there are $n-(i-1) = n-i+1$ choices available at stage i. In particular, for the kth and final choice, there are $n-k+1$ ways to choose w_k. By the product rule, the number of k-permutations is $\prod_{i=1}^{k}(n-(i-1)) = n(n-1)\cdots(n-k+1)$. Multiplying this expression by $(n-k)!/(n-k)!$, we obtain the product of the integers 1 through n in the numerator, which is $n!$. Thus the answer is also given by the formula $n!/(n-k)!$. □

1.19. Definition: Power Set. For any set S, the *power set* $\mathcal{P}(S)$ is the set of all subsets of S. Thus, $T \in \mathcal{P}(S)$ iff $T \subseteq S$. For example, if $S = \{2, 5, 7\}$, then $\mathcal{P}(S)$ is the eight-element set

$$\{\emptyset, \{2\}, \{5\}, \{7\}, \{2, 5\}, \{2, 7\}, \{5, 7\}, \{2, 5, 7\}\}.$$

1.20. Theorem: Cardinality of Power Sets. An n-element set has 2^n subsets. In other words, if $|S| = n$, then $|\mathcal{P}(S)| = 2^n$.

Proof. Suppose $S = \{x_1, \ldots, x_n\}$ is an n-element set. We can build a typical subset T of S by making a sequence of n choices. First, decide whether $x_1 \in T$ or $x_1 \notin T$. This binary decision can be made in two ways. Second, decide whether $x_2 \in T$ or $x_2 \notin T$; again there are two possibilities. Continue similarly; decide in the ith choice whether $x_i \in T$ or $x_i \notin T$ (two possibilities). This sequence of choices uniquely determines which x_j's belong to T. Since T is a subset of S, this information uniquely determines the set T. By the product rule, the number of subsets is $2 \times 2 \times \cdots \times 2$ (n factors), which is 2^n. □

1.5 Functions

This section reviews the definitions of functions, injections, surjections, and bijections, which should already be familiar to the reader. We also enumerate the number of functions, injections, and bijections between two given finite sets. The enumeration of surjections is more subtle, and will be discussed later (§2.10).

1.21. Definition: Functions. Formally, a *function* f from X to Y is an ordered triple (X, Y, G), where G is a subset of $X \times Y$ such that for each $x \in X$ there is exactly one $y \in Y$ with $(x, y) \in G$. X is the *domain* of f, Y is the *codomain* of f, and G is the *graph* of f. We write $y = f(x)$ iff $(x, y) \in G$, and we write $f : X \to Y$ to signify that f is a function from X to Y. Let ${}^XY$ denote the set of all functions from X to Y.

Informally, we think of a function f as consisting of a rule that maps each $x \in X$ to a unique value $f(x) \in Y$. When X and Y are finite sets, it is convenient to visualize f by an *arrow diagram*. We obtain this diagram by drawing a dot for each element of X and Y, and drawing an arrow from x to y whenever $y = f(x)$. The definition of a function requires that each $x \in X$ have *exactly one* arrow emanating from it, and the arrow must point to an element of Y. On the other hand, an element $y \in Y$ may have zero, one, or more than one arrow hitting it. Figures 1.2, 1.3, 1.4, and 1.5 depict the arrow diagrams for some functions.

1.22. Theorem: Enumeration of Functions. Suppose $X = \{x_1, \ldots, x_n\}$ is an n-element set and $Y = \{y_1, \ldots, y_m\}$ is an m-element set. There are m^n functions from X to Y. In other words, $|{}^XY| = |Y|^{|X|}$.

Proof. To build a typical function $f \in {}^XY$, we make a sequence of n choices that uniquely determine the graph G of f. First, we choose $f(x_1)$ to be any of the m elements of Y. Second, we choose $f(x_2)$ to be any of the m elements of Y. Similarly, for each $i \leq n$, we choose $f(x_i)$ to be any of the m elements of Y. By the product rule, the number of functions we can build is $m \times m \times \cdots \times m$ (n factors), which is m^n. □

1.23. Definition: Injections. A function $g : X \to Y$ is an *injection* iff for all $x, x' \in X$, $x \neq x'$ implies $g(x) \neq g(x')$. Injective functions are also called *one-to-one* functions.

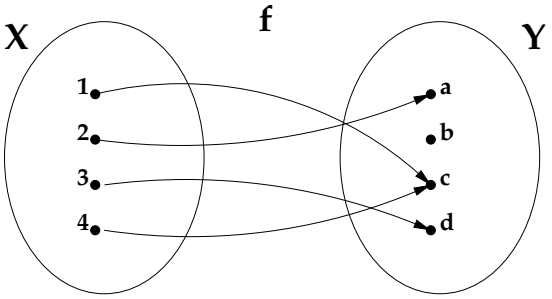


FIGURE 1.2
A function $f : \{1, 2, 3, 4\} \to \{a, b, c, d\}$.

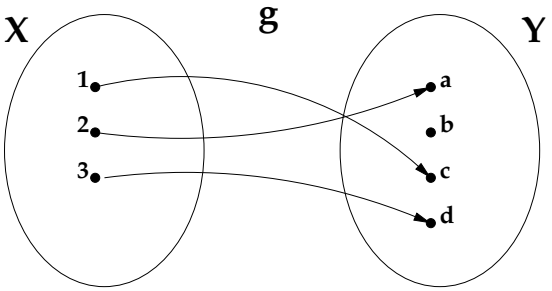


FIGURE 1.3
An injective function $g : \{1, 2, 3\} \to \{a, b, c, d\}$.

In the arrow diagram for an injective function, every $y \in Y$ has *at most one* arrow entering it. For example, the function f in Figure 1.2 is not injective, while the function g in Figure 1.3 is injective.

1.24. Theorem: Enumeration of Injections. Suppose $X = \{x_1, \ldots, x_n\}$ is an n-element set and $Y = \{y_1, \ldots, y_m\}$ is an m-element set. If $n \leq m$, the number of injections from X into Y is $m(m-1)(m-2)\cdots(m-n+1) = m!/(m-n)!$. If $n > m$, there are no injections from X to Y.

Proof. Assume first that $n \leq m$. As above, we construct a typical injection $g : X \to Y$ by choosing the n function values $g(x_i)$, for $1 \leq i \leq n$. For each $i \leq n$, we choose $g(x_i)$ to be an element of Y distinct from the elements $g(x_1), \ldots, g(x_{i-1})$ already chosen. Since the latter elements are pairwise distinct, we see that there are $m - (i-1) = m - i + 1$ alternatives for $g(x_i)$, no matter what happened in the first $i - 1$ choices. By the product rule, the number of injections is $m(m-1)\cdots(m-n+1) = m!/(m-n)!$.

On the other hand, suppose $n > m$. Try to build an injection g by choosing the values $g(x_1), g(x_2), \ldots$ as before. When we try to choose $g(x_{m+1})$, there are no elements of Y distinct from the previously chosen elements $g(x_1), \ldots, g(x_m)$. Since it is impossible to complete the construction of g, there are no injections from X to Y in this situation. □

1.25. Definition: Surjections. A function $h : X \to Y$ is a *surjection* iff for every $y \in Y$ there exists $x \in X$ with $y = f(x)$. Surjective functions are also said to be *onto* or to map *onto* the codomain Y.

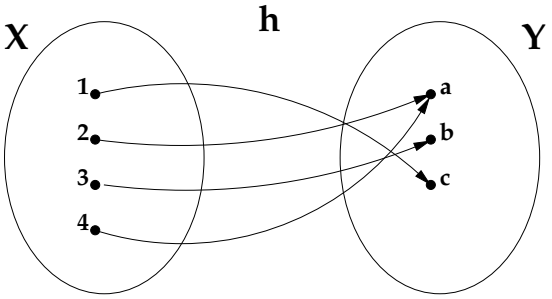


FIGURE 1.4
A surjective function $h : \{1, 2, 3, 4\} \to \{a, b, c\}$.

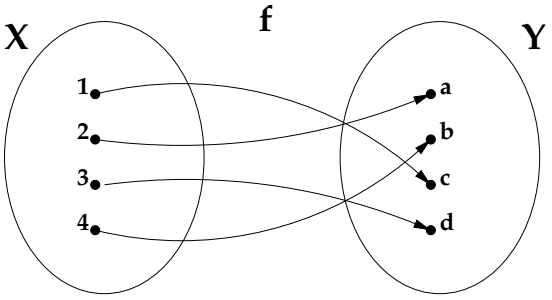


FIGURE 1.5
A bijective function $f : \{1, 2, 3, 4\} \to \{a, b, c, d\}$.

In the arrow diagram for a surjective function, every $y \in Y$ has *at least one* arrow entering it. For example, the functions f and g in Figures 1.2 and 1.3 are not surjective, while the function h in Figure 1.4 is surjective. Note that h is not injective. Counting surjections is harder than counting other classes of functions, so we defer discussion of this problem to a later section (§2.10).

1.26. Definition: Bijections. A function $f : X \to Y$ is a *bijection* iff f is both injective and surjective iff for every $y \in Y$ there exists a unique $x \in X$ with $y = f(x)$.

In the arrow diagram for a bijective function, every $y \in Y$ has *exactly one* arrow entering it. For example, the functions in Figures 1.2 through 1.4 are not bijective, while the function f in Figure 1.5 is bijective.

1.27. Theorem: Injectivity vs. Surjectivity. Suppose $f : X \to Y$ is a function. If X and Y are finite sets with the same number of elements, then f is injective iff f is surjective.

Proof. Suppose X and Y both have n elements, and write $X = \{x_1, \ldots, x_n\}$. Assume that $f : X \to Y$ is injective. Then the set $T = \{f(x_1), \ldots, f(x_n)\}$ is a subset of Y consisting of n *distinct* elements. Since Y has n elements, this subset must be all of Y. This means that every $y \in Y$ has the form $f(x_i)$ for some $x_i \in X$, so that f is surjective.

Conversely, assume that $f : X \to Y$ is not injective. Then there exist $i \neq j$ with $f(x_i) = f(x_j)$. It follows that the set $T = \{f(x_1), \ldots, f(x_n)\}$ contains fewer than n elements, since the displayed list of members of T contains at least one duplicate. Thus T is a proper subset of Y. Letting y be any element of $Y \sim T$, we see that y does not have the form $f(x)$ for any $x \in X$. Therefore f is not surjective. □

The previous result does not extend to infinite sets, as shown by the following examples. The function $f : \mathbb{N} \to \mathbb{N}$ defined by $f(n) = n+1$ is injective but not surjective. The function $g : \mathbb{N} \to \mathbb{N}$ defined by $g(2k) = g(2k+1) = k$ for all $k \geq 0$ is surjective but not injective. The function $\exp : \mathbb{R} \to \mathbb{R}$ defined by $\exp(x) = e^x$ is injective but not surjective. The function $h : \mathbb{R} \to \mathbb{R}$ defined by $h(x) = x(x-1)(x+1)$ is surjective but not injective.

1.28. Theorem: Enumeration of Bijections. Suppose X and Y are two n-element sets. Then there are $n!$ bijections from X to Y.

Proof. By 1.27, a function $f : X \to Y$ is injective iff f is surjective. Therefore, under the assumption that $|X| = |Y| = n$, f is injective iff f is bijective. We have already seen that the number of injections from X to Y is $n!/(n-n)! = n!$. The result follows. □

If X is an n-element set and Y is an m-element set and $m \neq n$, there are no bijections from X to Y (cf. the next section).

1.29. Remark. The reader may note the similarity between the formulas obtained here for functions and the formulas obtained earlier for words and permutations. This is not a coincidence. Indeed, we can formally define a word $w_1 w_2 \cdots w_k$ over an alphabet A as the *function* $w : \{1, 2, \ldots, k\} \to A$ defined by $w(i) = w_i$. The number of such words (functions) is $|A|^k$. The word $w_1 w_2 \cdots w_k$ is a k-permutation of A iff the w_i's are all distinct iff w is an *injective* function. The word $w_1 w_2 \cdots w_k$ is a permutation of A iff w is a *bijective* function. Finally, note that w is surjective iff every letter in the alphabet A occurs among the letters $w_1, \ldots, w_k$.

1.6 Bijections, Cardinality, and Counting

Bijections play a critical role in the theory of counting. Indeed, the very definition of cardinality is formulated in terms of bijections. In everyday life, we count the number of objects in a finite set S by pointing to each object in the set in turn and saying "one," "two," "three," etc. In essence, we are setting up a bijection between S and some set $\{1, 2, \ldots, n\}$ of natural numbers. This leads to the following definition, which provides a rigorous foundation for the informal notion of cardinality that we have used up to this point.

1.30. Definition: Cardinality. For any set A and any integer $n \geq 1$, we write $|A| = n$ iff there exists a bijection $f : A \to \{1, 2, \ldots, n\}$. We write $|A| = 0$ iff $A = \emptyset$. For any sets A and B, we write $|A| = |B|$ iff there exists a bijection $f : A \to B$. We write $|A| \leq |B|$ iff there exists an injection $g : A \to B$.

These definitions apply to infinite sets as well as finite sets, although we shall be mainly interested in finite sets. In the general case, one can prove the *Schröder-Bernstein Theorem:* $|A| \leq |B|$ and $|B| \leq |A|$ imply $|A| = |B|$ (see [125, p. 29] or 1.156 for a proof.) If A is nonempty and the axiom of choice is assumed, then $|A| \leq |B|$ is equivalent to the existence of a surjection $h : B \to A$. These properties are intuitively evident in the case of finite sets. For more discussion of the theory of cardinality for infinite sets, see [66] or [95].

Recall that if $f : X \to Y$ and $g : Y \to Z$ are functions, the *composition* of g and f is the function $g \circ f : X \to Z$ defined by $(g \circ f)(x) = g(f(x))$ for $x \in X$. We assume the reader is familiar with the following theorem, so we omit its proof.

1.31. Theorem: Properties of Bijections. Let X, Y, Z be any sets. (a) The identity map $\mathrm{id}_X : X \to X$, defined by $\mathrm{id}_X(x) = x$ for all $x \in X$, is a bijection. Hence, $|X| = |X|$.

(b) A function $f : X \to Y$ is bijective iff there exists a function $f' : Y \to X$ such that $f' \circ f = \mathrm{id}_X$ and $f \circ f' = \mathrm{id}_Y$. If such an f' exists, it is unique; we call it the *two-sided inverse* of f and denote it by f^{-1}. This inverse is also a bijection, and $(f^{-1})^{-1} = f$. Hence, $|X| = |Y|$ implies $|Y| = |X|$. (c) The composition of two bijections is a bijection. Hence, $|X| = |Y|$ and $|Y| = |Z|$ implies $|X| = |Z|$.

The definition of cardinality can be used to *prove* the basic counting principle 1.1.

1.32. Theorem. If $|A| = n$, $|B| = m$, and $A \cap B = \emptyset$, then $|A \cup B| = n + m$.

Proof. The assumption $|A| = n$ means that there is a bijection $f : A \to \{1, 2, \ldots, n\}$. The assumption $|B| = m$ means that there is a bijection $g : B \to \{1, 2, \ldots, m\}$. Define a function $h : A \cup B \to \{1, 2, \ldots, n + m\}$ by setting

$$h(x) = \begin{cases} f(x) & \text{if } x \in A; \\ g(x) + n & \text{if } x \in B. \end{cases}$$

The assumption that $A \cap B = \emptyset$ is needed to ensure that h is a well-defined (single-valued) function. Observe that h does map into the required codomain $\{1, 2, \ldots, n + m\}$. To see that h is a bijection, we display a two-sided inverse $h' : \{1, 2, \ldots, n + m\} \to A \cup B$. We define

$$h'(i) = \begin{cases} f^{-1}(i) & \text{if } 1 \leq i \leq n; \\ g^{-1}(i - n) & \text{if } n + 1 \leq i \leq n + m. \end{cases}$$

A routine case analysis verifies that $h \circ h'$ and $h' \circ h$ are identity maps. □

The product rule 1.8 can be phrased more formally in terms of bijections.

1.33. Formal Product Rule. Suppose there is a bijection

$$f : \{1, 2, \ldots, n_1\} \times \{1, 2, \ldots, n_2\} \times \cdots \times \{1, 2, \ldots, n_k\} \to S.$$

Then $|S| = n_1 n_2 \cdots n_k$.

Proof. S has the same cardinality as the product set $\{1, 2, \ldots, n_1\} \times \cdots \times \{1, 2, \ldots, n_k\}$, thanks to the bijection f. So the result follows from 1.5. □

1.34. Remark. Let us compare the formal product rule 1.33 to the informal version of the product rule given earlier (1.8). In informal applications of the product rule, we "build" objects in a set S by making a sequence of k choices, where there are n_i ways to make the ith choice. The input to the bijection f in the formal product rule is a k-tuple $(c_1, \ldots, c_k)$ where $1 \leq c_i \leq n_i$ for all $i \leq k$. Intuitively, c_i records which choice was made at the ith stage. In practice, the map f is described as an algorithm that tells us how to combine the choices c_i to build an object in S. The key point in the intuitive product rule is that each object in S can be constructed *in exactly one way* by making suitable choices. This corresponds to the requirement in the formal product rule that f be a *bijection* onto S. Most erroneous applications of the intuitive product rule occur when the underlying "construction map" f is not bijective (a point that is seldom checked explicitly when using the product rule).

1.35. Example. How many 4-letter words contain at least one E? One might try to construct such words by choosing a position that contains the E (4 choices), then filling the remaining positions from left to right with arbitrary letters (26 choices for each position). The product rule would then give $4 \times 26^3 = 70,304$ as the answer. However, this answer is incorrect. Our choice sequence implicitly defines a function

$$f : \{1, 2, 3, 4\} \times \{1, 2, \ldots, 26\}^3 \to X,$$

where X is the set of words under consideration. For example, $f(3,3,2,26) = \text{CBEZ}$. Our counting argument is flawed because the function f is surjective but not bijective. For instance, $f(1,1,5,1) = \text{EAEA} = f(3,5,1,1)$.

To obtain the correct answer, one can combine the product rule and the difference rule. There are 26^4 words of length 4, and there are 25^4 such words that do *not* contain the letter E. So the true answer is $26^4 - 25^4 = 66,351$. An alternative argument that is closer to our original attempt breaks X into the disjoint union $X_1 \cup X_2 \cup X_3 \cup X_4$, where X_i is the set of four-letter words where the *first* occurrence of E is at position i. A modification of the argument in the previous paragraph shows that $|X_i| = 25^{i-1}26^{4-i}$, so

$$|X| = 26^3 + 25 \cdot 26^2 + 25^2 \cdot 26 + 25^3 = 66,351.$$

1.36. Remark. One can give a *bijective* proof of the product rule (1.5), just as we gave a bijective proof of the sum rule (1.1) in 1.32. These bijective proofs have applications to the problems of listing, ranking, and unranking collections of combinatorial objects. These topics are discussed in Chapter 5.

1.7 Subsets, Binary Words, and Compositions

A fundamental method for counting a finite set A is to display a bijection between A and some other set B whose cardinality is already known. We illustrate this basic principle by revisiting the enumeration of subsets and binary words, and enumerating new combinatorial objects called compositions.

Let $X = \{x, y, z\}$, and consider the set $\mathcal{P}(X)$ of all subsets of X. We define a bijection $f : \mathcal{P}(X) \to \{0,1\}^3$ as follows:

$$\begin{array}{llll} f(\emptyset) = 000; & f(\{x\}) = 100; & f(\{y\}) = 010; & f(\{z\}) = 001; \\ f(\{x,y\}) = 110; & f(\{x,z\}) = 101; & f(\{y,z\}) = 011; & f(\{x,y,z\}) = 111. \end{array}$$

These values were computed by the following rule. Given $S \subseteq X$, we set $f(S) = w_1w_2w_3$ where $w_1 = 1$ if $x \in S$, $w_1 = 0$ if $x \notin S$, $w_2 = 1$ if $y \in S$, $w_2 = 0$ if $y \notin S$, $w_3 = 1$ if $z \in S$, and $w_3 = 0$ if $z \notin S$. We see by inspection that f is a bijection. Thus, $|\mathcal{P}(X)| = |\{0,1\}^3| = 2^3 = 8$. We now generalize this example to n-element sets. First, we introduce notation that will be used frequently throughout the text.

1.37. Definition: Truth Function. If P is any logical statement, we set $\chi(P) = 1$ if P is true, and $\chi(P) = 0$ if P is false.

1.38. Theorem: Subsets vs. Binary Words. Let X be an n-element set. For each ordering $x_1, \ldots, x_n$ of the elements of X, there is a bijection $f : \mathcal{P}(X) \to \{0,1\}^n$. Therefore, $|\mathcal{P}(X)| = 2^n$.

Proof. Given $S \subseteq X$, we define $f(S) = w_1w_2 \cdots w_n$, where $w_i = \chi(x_i \in S)$. To see that f is a bijection, define $f' : \{0,1\}^n \to \mathcal{P}(X)$ by setting

$$f'(w_1w_2 \cdots w_n) = \{x_i \in X : w_i = 1\}.$$

It is immediate that f' is the two-sided inverse of f. □

For example, if $X = \{1,2,3,4,5,6,7,8\}$ with the usual ordering, then $f(\{2,5,7,8\}) = 01001011$ and $f^{-1}(10000011) = \{1,7,8\}$.

1.39. Definition: Compositions. A *composition* of an integer $n > 0$ is a sequence $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, where each α_i is a positive integer and $\alpha_1 + \alpha_2 + \cdots + \alpha_k = n$. The *number of parts* of α is k. Let Comp_n be the set of all compositions of n.

1.40. Example. The sequences $(1,3,1,3,3)$ and $(3,3,3,1,1)$ are two distinct compositions of 11 with five parts. The four compositions of 3 are

$$(3), \quad (2,1), \quad (1,2), \quad (1,1,1).$$

1.41. Theorem: Enumeration of Compositions. For all $n > 0$, there are 2^{n-1} compositions of n.

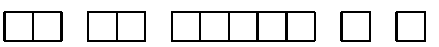
Proof. We define a bijection $g : \mathrm{Comp}_n \to \{0,1\}^{n-1}$. Given $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k) \in \mathrm{Comp}_n$, define

$$g(\alpha) = 0^{\alpha_1 - 1} 1 0^{\alpha_2 - 1} 1 \cdots 1 0^{\alpha_k - 1}.$$

Here, the notation 0^j denotes a sequence of j consecutive zeroes, and 0^0 denotes the empty word. For example, $g((3,1,3)) = 001100$. Since $\sum_{i=1}^{k} (\alpha_i - 1) = n - k$ and there are $k-1$ ones, we see that $g(\alpha) \in \{0,1\}^{n-1}$. Now define $g' : \{0,1\}^{n-1} \to \mathrm{Comp}_n$ as follows. We can uniquely write any word $w \in \{0,1\}^{n-1}$ in the form $w = 0^{b_1} 1 0^{b_2} 1 \cdots 1 0^{b_k}$ where $k \geq 1$, each $b_i \geq 0$, and $\sum_{i=1}^{k} b_i = (n-1) - (k-1) = n - k$ since there are $k-1$ ones. Define $g'(w) = (b_1 + 1, b_2 + 1, \ldots, b_k + 1)$, which is a composition of n. For example, $g'(100100) = (1,3,3)$. One may check that g' is the two-sided inverse of g, so g is a bijection. It follows that $|\,\mathrm{Comp}_n\,| = |\{0,1\}^{n-1}| = 2^{n-1}$. □

The bijections in the preceding proof are best understood pictorially. We represent an integer $i > 0$ as a sequence of i unit squares glued together. We visualize a composition $(\alpha_1, \ldots, \alpha_k)$ by drawing the squares for $\alpha_1, \ldots, \alpha_k$ in a single row, separated by gaps. For instance, the composition $(1,3,1,3,3)$ is represented by the picture

We now scan the picture from left to right and record what happens between each two successive boxes. If the two boxes in question are glued together, we record a 0; if there is a gap between the two boxes, we record a 1. The composition of 11 pictured above maps to the word $1001100100 \in \{0,1\}^{10}$. Going the other way, the word $0101000011 \in \{0,1\}^{10}$ leads first to the picture



and then to the composition $(2,2,5,1,1)$. One can check that the pictorial operations just described correspond precisely to the maps f and f' in the proof above. When $n = 3$, we have:

$$f((3)) = 00; \quad f((2,1)) = 01; \quad f((1,2)) = 10; \quad f((1,1,1)) = 11.$$

1.8 Subsets of a Fixed Size

We turn now to the enumeration of the k-element subsets of an n-element set. For example, there are ten 3-element subsets of $\{a,b,c,d,e\}$:

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

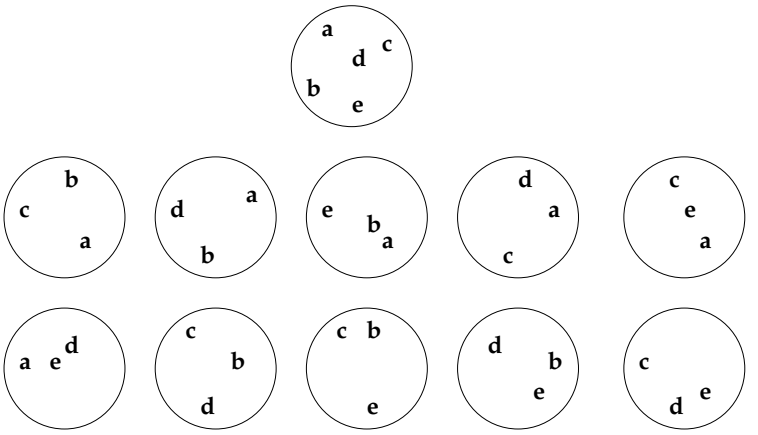


FIGURE 1.6
The set $\{a, b, c, d, e\}$ and its 3-element subsets.

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

In this example, we present a given set by listing its members between curly braces. This notation forces us to list the members of each set in a particular order (alphabetical in this case). If we reorder the members of the list, the underlying set does not change. For example, the sets $A_1 = \{a, c, d\}$ and $A_2 = \{c, d, a\}$ and $A_3 = \{d, c, a\}$ are all equal. This assertion follows from the very definition of set equality: $A = B$ means that for every x, $x \in A$ iff $x \in B$. In contrast, the ordering of elements in a sequence (or word) definitely makes a difference. For instance, the words *cad* and *dac* are unequal although they use the same three letters.

To emphasize that the members of a set do not come in any particular order, we often picture a finite set as a circle with the members of the set floating around in random positions inside the circle. For example, Figure 1.6 depicts the sets mentioned above.

Suppose we try to enumerate the k-element subsets of a given n-element set using the product rule. Recall that the product rule requires us to construct objects by making an *ordered sequence* of choices. We might try to construct a subset by choosing its first element in n ways, then its second element in $n-1$ ways, etc., which leads to the *incorrect* answer $n(n-1)\cdots(n-k+1)$. The trouble here is that there is no well-defined "first element" of a subset. In fact, our naive construction procedure generates each subset several times, once for each possible ordering of its members. There are $k!$ such orderings, so we obtain the correct answer by dividing the previous formula by $k!$. We make this argument more precise in the next theorem.

1.42. Theorem: Enumeration of k-element Subsets. For $0 \leq k \leq n$, the number of k-element subsets of an n-element set is

$$\frac{n!}{k!(n-k)!}.$$

Proof. Fix n and k with $0 \leq k \leq n$. Let A be an n-element set, and let x denote the number of k-element subsets of A. Let S be the set of all k-permutations of A. Recall that elements of S are *ordered* sequences $w_1 w_2 \cdots w_k$, where the w_i are distinct elements of A. We compute $|S|$ in two ways. First, we have already seen that $|S| = n!/(n-k)!$ by using the product rule — we choose w_1 in n ways, then choose w_2 in $n-1$ ways, etc., and finally

choose w_k in $n-k+1$ ways. On the other hand, here is a second way to construct a typical sequence $w_1 w_2 \cdots w_k$ in S. Begin by choosing a k-element subset of A in any of x ways. Then write down a permutation of this k-element subset in any of $k!$ ways. The result is an element of S. By the product rule,

$$x \cdot k! = |S| = n!/(n-k)!.$$

Solving for x, we obtain the desired formula. □

1.43. Definition: Binomial Coefficients. For $0 \leq k \leq n$, the *binomial coefficient* is

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

For $k < 0$ or $k > n$, we define $\binom{n}{k} = C(n,k) = 0$. Thus, for all $n \geq 0$ and all k, $\binom{n}{k}$ is the number of k-element subsets of an n-element set. In particular, $\binom{n}{k}$ is always an integer.

1.9 Anagrams

1.44. Definition: Anagrams. Suppose $a_1, \ldots, a_k$ are distinct letters from some alphabet A and $n_1, \ldots, n_k$ are nonnegative integers. Let $\mathcal{R}(a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k})$ denote the set of all words $w = w_1 w_2 \cdots w_n$ that are formed by rearranging n_1 copies of a_1, n_2 copies of a_2, ..., n_k copies of a_k (so that $n = n_1 + n_2 + \cdots + n_k$). Words in a given set $\mathcal{R}(a_1^{n_1} \cdots a_k^{n_k})$ are said to be *anagrams* or *rearrangements* of one another.

1.45. Example.

$$\mathcal{R}(0^2 1^3) = \{00111, 01011, 01101, 01110, 10011, 10101, 10110, 11001, 11010, 11100\};$$

$$\mathcal{R}(a^1 b^2 c^1 d^0) = \{abbc, abcb, acbb, babc, bacb, bbac, bbca, bcab, bcba, cabb, cbab, cbba\}.$$

1.46. Theorem: Enumeration of Anagrams. Suppose $a_1, \ldots, a_k$ are distinct letters, $n_1, \ldots, n_k$ are nonnegative integers, and $n = n_1 + \cdots + n_k$. Then

$$|\mathcal{R}(a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k})| = \frac{n!}{n_1! n_2! \cdots n_k!}.$$

Proof. We give two proofs of this result. *First Proof:* We use a technique similar to that used in the proof of 1.42. Define a new alphabet A consisting of n *distinct letters* by attaching distinct numerical superscripts to each copy of the given letters $a_1, \ldots a_k$:

$$A = \{a_1^{(1)}, a_1^{(2)}, \ldots, a_1^{(n_1)}, a_2^{(1)}, \ldots, a_2^{(n_2)}, \ldots, a_k^{(1)}, \ldots, a_k^{(n_k)}\}.$$

Let $x = |\mathcal{R}(a_1^{n_1} a_2^{n_2} \cdots a_k^{n_k})|$. Let S be the set of all permutations w of A. We count $|S|$ in two ways. On one hand, we already know that $|S| = n!$ (choose w_1 in n ways, then w_2 in $n-1$ ways, etc.). On the other hand, here is a different method for constructing each permutation of A exactly once. First, choose a word $v \in \mathcal{R}(a_1^{n_1} \cdots a_k^{n_k})$ in any of x ways. Second, attach the superscripts 1 through n_1 to the n_1 copies of a_1 in v in any of $n_1!$ ways. Third, attach the superscripts 1 through n_2 to the n_2 copies of a_2 in v in any of $n_2!$ ways. Continue similarly; at the last stage, we attach the superscripts 1 through n_k to the n_k copies of a_k in v in any of $n_k!$ ways. By the product rule,

$$x \cdot n_1! \cdot n_2! \cdot \ldots \cdot n_k! = |S| = n!.$$

Solving for x, we obtain the desired formula.

Second Proof. The second proof relies on 1.42 and an algebraic manipulation of factorials. We construct a typical object $w = w_1 w_2 \cdots w_n \in \mathcal{R}(a_1^{n_1} \cdots a_k^{n_k})$ by making the following sequence of k choices. Intuitively, we are going to choose the positions of the a_1's, then the positions of the a_2's, etc. First, choose any n_1-element subset S_1 of $\{1, 2, \ldots, n\}$ in any of $\binom{n}{n_1}$ ways, and define $w_i = a_1$ for all $i \in S_1$. Second, choose any n_2-element subset S_2 of $\{1, 2, \ldots, n\} \sim S_1$ in any of $\binom{n-n_1}{n_2}$ ways, and define $w_i = a_2$ for all $i \in S_2$. At the jth stage (where $1 \leq j \leq k$), we have already filled the positions in $S_1 \cup \cdots \cup S_{j-1} \subseteq \{1, 2, \ldots, n\}$, and there are $n - n_1 - n_2 - \cdots - n_{j-1}$ remaining positions in the word. We choose any n_j-element subset S_j of these remaining positions in any of $\binom{n-n_1-\cdots-n_{j-1}}{n_j}$ ways, and define $w_i = a_j$ for all $i \in S_j$. By the product rule, the number of rearrangements is

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-n_1-\cdots-n_{k-1}}{n_k} = \prod_{i=1}^{k} \frac{(n-n_1-\cdots-n_{i-1})!}{n_i!(n-n_1-\cdots-n_i)!}.$$

This is a telescoping product that simplifies to $n!/(n_1!n_2!\cdots n_k!)$. For instance, when $k = 4$, the product is

$$\frac{n!}{n_1!(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!(n-n_1-n_2-n_3)!} \cdot \frac{(n-n_1-n_2-n_3)!}{n_4!(n-n_1-n_2-n_3-n_4)!},$$

which simplifies to $n!/(n_1!n_2!n_3!n_4!)$. (Recall that $(n - n_1 - n_2 - \cdots - n_k)! = 0! = 1$.) □

1.47. Example. We now illustrate the constructions in each of the two preceding proofs. For the first proof, suppose we are counting $\mathcal{R}(a^3b^1c^4)$. The alphabet A in the proof consists of the eight distinct letters

$$A = \{a^{(1)}, a^{(2)}, a^{(3)}, b^{(1)}, c^{(1)}, c^{(2)}, c^{(3)}, c^{(4)}\}.$$

Let us build a specific permutation of A using the second counting method. First, choose an element of $\mathcal{R}(a^3b^1c^4)$, say $v = baccaacc$. Second, choose a labeling of the a's with superscripts, say $ba^{(3)}cca^{(1)}a^{(2)}cc$. Third, choose a labeling of the b's, say $b^{(1)}a^{(3)}cca^{(1)}a^{(2)}cc$. Finally, choose a labeling of the c's, say $b^{(1)}a^{(3)}c^{(1)}c^{(2)}a^{(1)}a^{(2)}c^{(4)}c^{(3)}$. We have now constructed a permutation of the alphabet A.

Next, let us see how to build the word 'baccaacc' using the method of the second proof. Start with an empty 8-letter word, which we denote - - - - - - - -. We first choose the 3-element subset $\{2, 5, 6\}$ of $\{1, 2, \ldots, 8\}$ and put a's in those positions, obtaining - a - - a a - -. We then choose the 1-element subset $\{1\}$ of $\{1, 3, 4, 7, 8\}$ and put a b in that position, obtaining b a - - a a - -. Finally, we choose the 4-element subset $\{3, 4, 7, 8\}$ of $\{3, 4, 7, 8\}$ and put c's in those positions, obtaining the word baccaacc.

1.48. Definition: Multinomial Coefficients. Suppose $n_1, \ldots, n_k$ are nonnegative integers and $n = n_1 + \cdots + n_k$. The *multinomial coefficient* is

$$\binom{n}{n_1, n_2, \ldots, n_k} = C(n; n_1, n_2, \ldots, n_k) = \frac{n!}{n_1!n_2!\cdots n_k!}.$$

This is the number of rearrangements of k letters where there are n_i copies of the ith letter. In particular, $\binom{n}{n_1, n_2, \ldots, n_k}$ is always an integer.

1.49. Theorem: Binomial vs. Multinomial Coefficients. For all nonnegative integers a and b, we have

$$\binom{a+b}{a} = \binom{a+b}{a, b}.$$

Proof. The result is immediate from the formulas for binomial coefficients and multinomial coefficients as quotients of factorials, but we want to give a bijective proof. Let U be the set of a-element subsets of $\{1, 2, \ldots, a+b\}$, and let $V = \mathcal{R}(1^a 0^b)$. We have already shown that $|U| = \binom{a+b}{a}$ and $|V| = \binom{a+b}{a,b}$. So we must define a bijection $f : U \to V$. Given $S \in U$, let $f(S) = w_1 w_2 \cdots w_{a+b}$, where $w_i = \chi(i \in S)$. Since S has a elements, $f(S)$ is a word consisting of a ones and b zeroes. The inverse of f is the map $f' : V \to U$ given by $f'(w_1 w_2 \ldots w_{a+b}) = \{i : w_i = 1\}$. (Note that these maps are the restrictions to U and V of the maps f and f' from the proof of 1.38.) □

1.50. Example: Compositions with k Parts. Let us determine the number of compositions $\alpha = (\alpha_1, \ldots, \alpha_k)$ of n that have exactly k parts. Recall the bijection $g : \mathrm{Comp}_n \to \{0,1\}^{n-1}$ from the proof of 1.41. Applying g to α produces a word with $k-1$ ones and $n-k$ zeroes. Conversely, any such word arises from a composition with k parts. Thus, g restricts to a bijection between the set of compositions of n with k parts and the set of words $\mathcal{R}(0^{n-k}1^{k-1})$. Consequently, the number of such compositions is $\binom{n-1}{n-k,k-1}$.

1.10 Lattice Paths

1.51. Definition: Lattice Paths. A *lattice path* in the plane is a sequence

$$P = ((x_0, y_0), (x_1, y_1), \ldots, (x_k, y_k)),$$

where the x_i's and y_i's are integers, and for each $i \geq 1$, either $(x_i, y_i) = (x_{i-1} + 1, y_{i-1})$ or $(x_i, y_i) = (x_{i-1}, y_{i-1} + 1)$. We say that P is a path *from* (x_0, y_0) *to* (x_k, y_k).

We often take (x_0, y_0) to be the origin $(0, 0)$. We represent P pictorially by drawing a line segment of length 1 from (x_{i-1}, y_{i-1}) to (x_i, y_i) for each i. For example, Figure 1.7 displays the ten lattice paths from $(0, 0)$ to $(2, 3)$.

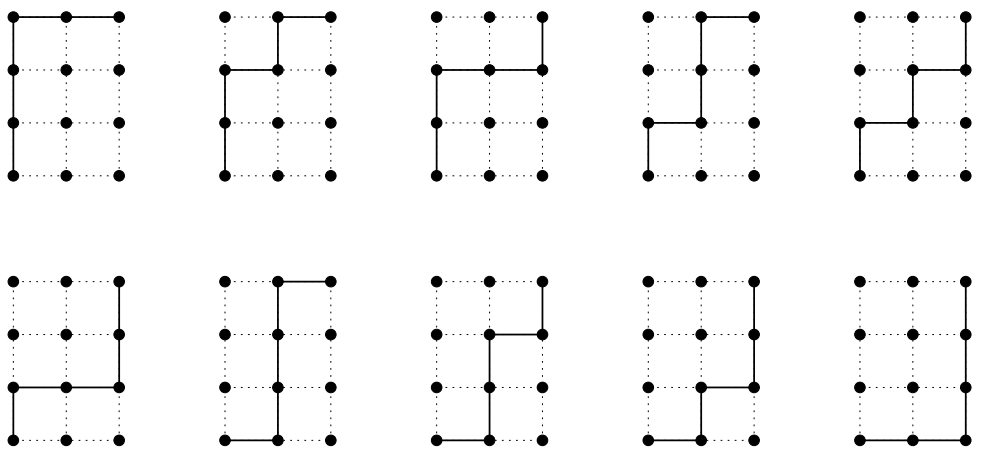


FIGURE 1.7
Lattice paths from $(0, 0)$ to $(2, 3)$.

1.52. Theorem: Enumeration of Lattice Paths in a Rectangle. For all integers $a, b \geq 0$, there are $\binom{a+b}{a,b} = \frac{(a+b)!}{a!b!}$ lattice paths from $(0, 0)$ to (a, b).

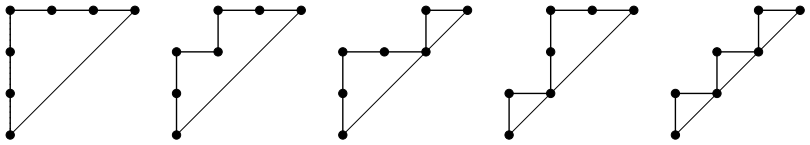


FIGURE 1.8
Dyck paths of order 3.

Proof. We can encode a lattice path P from $(0,0)$ to (a,b) as a word $w \in \mathcal{R}(E^a N^b)$ by setting $w_i = E$ if $(x_i, y_i) = (x_{i-1}+1, y_{i-1})$ and $w_i = N$ if $(x_i, y_i) = (x_{i-1}, y_{i-1}+1)$. Here, E stands for "east step," and N stands for "north step." Since the path ends at (a,b), w must have exactly a occurrences of E and exactly b occurrences of N. Thus we have a bijection between the given set of lattice paths and the set $\mathcal{R}(E^a N^b)$. Since $|\mathcal{R}(E^a N^b)| = \binom{a+b}{a,b}$, the theorem follows. □

For example, the paths shown in Figure 1.7 are encoded by the words

NNNEE, NNENE, NNEEN, NENNE, NENEN,
NEENN, ENNNE, ENNEN, ENENN, EENNN.

More generally, one can consider lattice paths in $\mathbb{R}^d$. Such a path is a sequence of points $(v_0, v_1, \ldots, v_k)$ in $\mathbb{Z}^d$ such that for each i, $v_i = v_{i-1} + e_j$ for some standard basis vector $e_j = (0, \ldots, 1, \ldots, 0) \in \mathbb{R}^d$ (the 1 occurs in position j).

1.53. Theorem: Enumeration of Lattice Paths in a d-dimensional Rectangle. For all integers $n_1, \ldots, n_d \geq 0$, the number of d-dimensional lattice paths from $(0, \ldots, 0)$ to $(n_1, \ldots, n_d)$ is

$$|\mathcal{R}(e_1^{n_1} e_2^{n_2} \cdots e_d^{n_d})| = \binom{n_1 + n_2 + \cdots + n_d}{n_1, n_2, \ldots, n_d}.$$

Proof. Encode a path P by the word $w_1 w_2 \cdots w_n$, where $n = n_1 + \cdots + n_d$ and $w_i = e_j$ iff $v_i = v_{i-1} + e_j$. □

Henceforth, we usually will make no distinction between a lattice path (which is a sequence of lattice points) and the word that encodes the lattice path.

We now turn to a more difficult enumeration problem involving lattice paths.

1.54. Definition: Dyck Paths. A *Dyck path of order n* is a lattice path from $(0,0)$ to (n,n) such that $y_i \geq x_i$ for all points (x_i, y_i) on the path. This requirement means that the path always stays weakly above the line $y = x$. For example, Figure 1.8 displays the five Dyck paths of order 3.

1.55. Definition: Catalan Numbers. For $n \geq 0$, the *nth Catalan number* is

$$C_n = \frac{1}{n+1}\binom{2n}{n,n} = \frac{1}{2n+1}\binom{2n+1}{n+1,n} = \frac{(2n)!}{n!(n+1)!} = \binom{2n}{n,n} - \binom{2n}{n+1,n-1}.$$

One may check that these expressions are all equal. For instance,

$$\binom{2n}{n,n} - \binom{2n}{n+1,n-1} = \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!} = \frac{(2n)!}{n!n!}\left[1 - \frac{n}{n+1}\right] = \frac{1}{n+1}\binom{2n}{n,n}.$$

The first few Catalan numbers are

$$C_0 = 1, \quad C_1 = 1, \quad C_2 = 2, \quad C_3 = 5, \quad C_4 = 14, \quad C_5 = 42, \quad C_6 = 132, \quad C_7 = 429.$$

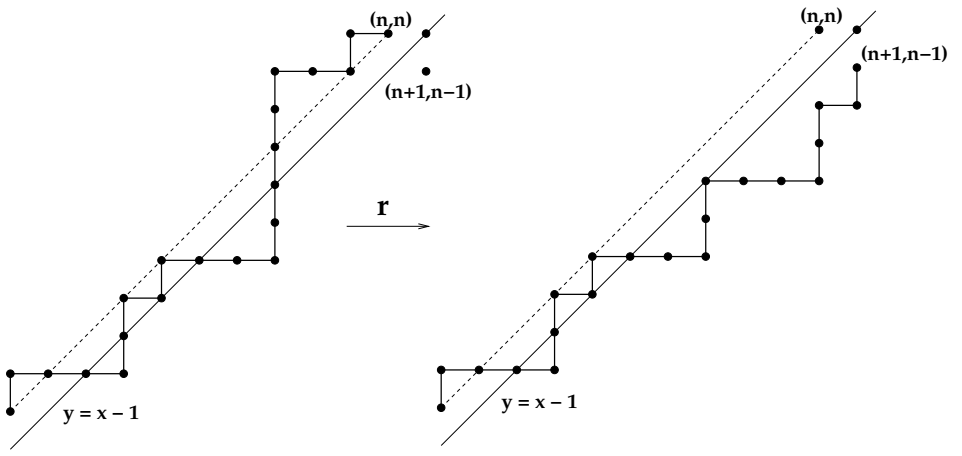


FIGURE 1.9
Example of the reflection map r.

1.56. Theorem: Enumeration of Dyck Paths. For $n \geq 0$, the number of Dyck paths of order n is the Catalan number $C_n = \binom{2n}{n,n} - \binom{2n}{n+1,n-1}$.

Proof. The following proof is essentially due to André [3]. Let A be the set of all lattice paths from $(0,0)$ to (n,n); let B be the set of all lattice paths from $(0,0)$ to $(n+1,n-1)$; let C be the set of all Dyck paths of order n; and let $D = A \sim C$ be the set of paths from $(0,0)$ to (n,n) that do go strictly below the line $y = x$. Since $C = A \sim D$, the difference rule gives

$$|C| = |A| - |D|.$$

We already know that $|A| = \binom{2n}{n,n}$ and $|B| = \binom{2n}{n+1,n-1}$. To establish the desired formula $|C| = C_n$, it therefore suffices to exhibit a bijection $r : D \to B$.

We define r as follows. Given a path $P \in D$, follow the path backwards from (n,n) until it goes below the diagonal $y = x$ for the first time. Let (x_i, y_i) be the first lattice point we encounter that is below $y = x$; this point must lie on the line $y = x - 1$. P is the concatenation of two lattice paths P_1 and P_2, where P_1 goes from $(0,0)$ to (x_i, y_i) and P_2 goes from (x_i, y_i) to (n,n). By choice of i, every lattice point of P_2 after (x_i, y_i) lies strictly above the line $y = x - 1$. Now, let P_2' be the path from (x_i, y_i) to $(n+1, n-1)$ obtained by reflecting P_2 in the line $y = x - 1$. Define $r(P)$ to be the concatenation of P_1 and P_2'. See Figure 1.9 for an example. Here, $(x_i, y_i) = (7,6)$, $P_1 =$ NEEENNENEEENN, $P_2 =$ NNNEENE, and $P_2' =$ EEENNEN. Note that $r(P)$ is a lattice path from $(0,0)$ to $(n+1, n-1)$, so $r(P) \in B$. Furthermore, (x_i, y_i) is the only lattice point of P_2' lying on the line $y = x - 1$.

The inverse map $r' : B \to D$ acts as follows. Given $Q \in B$, choose i maximal such that (x_i, y_i) is a point of Q on the line $y = x - 1$. Such an i must exist, since there is no way for a lattice path to reach $(n+1, n-1)$ from $(0,0)$ without passing through this line. Write $Q = Q_1 Q_2$, where Q_1 goes from $(0,0)$ to (x_i, y_i) and Q_2 goes from (x_i, y_i) to $(n+1, n-1)$. Let Q_2' be the reflection of Q_2 in the line $y = x - 1$. Define $r'(Q) = Q_1 Q_2'$, and note that this is a lattice path from $(0,0)$ to (n,n) which passes through (x_i, y_i), and hence lies in D. See Figure 1.10 for an example. Here, $(x_i, y_i) = (6,5)$, $Q_1 =$ NNENEEEENEN, $Q_2 =$ EEENENENN, and $Q_2' =$ NNNENENEE. From our observations about the point (x_i, y_i) in this paragraph and the last, one sees that r' is the two-sided inverse of r. □

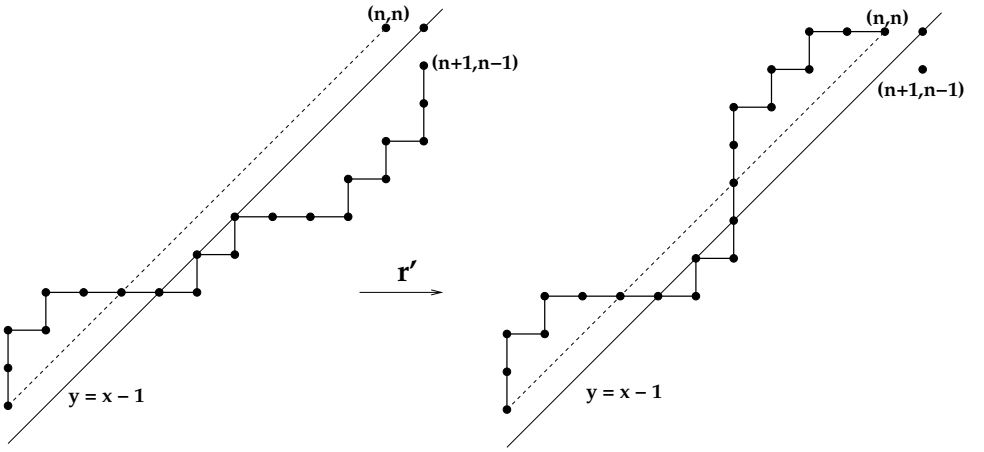


FIGURE 1.10
Example of the inverse reflection map.

The technique used in the preceding proof is called *André's reflection principle*. Another proof of the theorem, which leads directly to the formula $\frac{1}{2n+1}\binom{2n+1}{n+1,n}$, is given in §12.1. Yet another proof, which leads directly to the formula $\frac{1}{n+1}\binom{2n}{n,n}$, is given in §12.2.

1.11 Multisets

Recall that the concepts of *order* and *multiplicity* play no role when deciding whether two *sets* are equal. For instance, $\{1,3,5\} = \{3,5,1\} = \{1,1,1,5,5,3,3,3\}$ since all these sets have the same members. We now introduce the concept of a *multiset*, in which order still does not matter, but repetitions of a given element are significant.

1.57. Definition: Multisets. A *multiset* is an ordered pair $M = (S, m)$, where S is a set and $m : S \to \mathbb{N}^+$ is a function. For $x \in S$, the number $m(x)$ is called the *multiplicity* of x in M. The *number of elements of* M is $|M| = \sum_{x \in S} m(x)$.

In contrast, the number of *distinct* elements in M is $|S|$. We sometimes display a multiset as a list $[x_1, x_2, \ldots, x_k]$, where each $x \in S$ occurs exactly $m(x)$ times in the list. However, one must remember that the order of the elements in this list does not matter when deciding equality of multisets. For example, $[1,1,2,3,3] \neq [1,1,1,2,3,3] = [3,2,3,1,1,1]$. We often visualize M as a circle with the elements $x \in S$ appearing inside, each repeated $m(x)$ times. For example, Figure 1.11 displays the twenty 3-element multisets using letters in the alphabet $\{w,x,y,z\}$. The last circle in the first row represents the multiset $[w,x,x]$, which is formally the ordered pair $(\{w,x\},m)$ such that $m(w) = 1$ and $m(x) = 2$.

1.58. Theorem: Enumeration of Multisets. The number of k-element multisets using letters from an n-letter alphabet is

$$\binom{k+n-1}{k,n-1} = \frac{(k+n-1)!}{k!(n-1)!}.$$

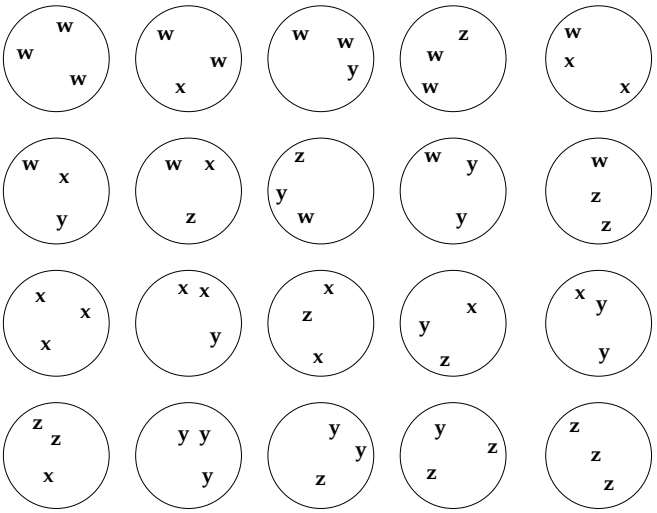


FIGURE 1.11
The 3-element multisets over the alphabet $\{w, x, y, z\}$.

Proof. We give two proofs of this result. *First Proof:* Let A be a fixed n-letter alphabet, and let U be the set of all k-element multisets using letters from A. Introduce the two symbols $\star$ ("star") and $|$ ("bar"), and let $V = \mathcal{R}(\star^k\,|^{n-1})$ be the set of all rearrangements of k stars and $n-1$ bars. We know that $|V| = \binom{k+n-1}{k, n-1}$. It therefore suffices to define a bijection $f : U \to V$.

Let $(a_1, a_2, \ldots, a_n)$ be a fixed ordering of the alphabet A. Let $M = (S, m)$ be a typical multiset in U. Set $m(a_i) = 0$ if $a_i \notin S$. Define

$$f(M) = \star^{m(a_1)}|\,\star^{m(a_2)}\,|\cdots|\,\star^{m(a_{n-1})}\,|\star^{m(a_n)} \in V.$$

In other words, we write a star for each occurrence of a_1 in M (if any), then a bar, then a star for each occurrence of a_2 in M (if any), then a bar, etc. There is no bar after the stars for a_n, so there are only $n-1$ bars total. Since M has k elements, there are k stars total. Thus $f(M)$ really is an element of V. For example, the multisets in the first column of Figure 1.11 are mapped to the following star-bar words:

$$f([w,w,w]) = \star\star\star|||, \quad f([w,x,y]) = \star|\star|\star|, \quad f([x,x,x]) = |\star\star\star||, \quad f([x,z,z]) = |\star||\star\star.$$

The multiset M is uniquely determined by $f(M)$. More precisely, define $f' : V \to U$ by letting $f'(\star^{m_1}|\,\star^{m_2}\,|\cdots|\star^{m_n})$ be the unique multiset that has m_i copies of a_i for $1 \le i \le n$ (here $m_i \ge 0$). Since $\sum_{i=1}^{n} m_i = k$, this is a k-element multiset using letters from A. For example, if $n = 6$, $k = 4$, and $A = \{1, 2, 3, 4, 5, 6\}$, then $f'(|| \star || \star | \star\star) = [3, 5, 6, 6]$. One may check that f' is the two-sided inverse of f.

Second Proof: We may assume (without loss of generality) that the alphabet A is $\{1, 2, \ldots, n\}$. As above, let U be the set of all k-element multisets using letters from A. Let W be the set of all k-element subsets of $B = \{1, 2, \ldots, k+n-1\}$. We know that $|W| = \binom{k+n-1}{k} = \binom{k+n-1}{k,n-1}$. So it suffices to define a bijection $g : U \to W$.

Given $M \in U$, we can write M uniquely in the form $M = [x_1, x_2, \ldots, x_k]$ by requiring that $x_1 \le x_2 \le \cdots \le x_k$. Now define

$$g(M) = g([x_1, x_2, \ldots, x_k]) = \{x_1 + 0, x_2 + 1, x_3 + 2, \ldots, x_i + (i-1), \ldots, x_k + (k-1)\}.$$

For example, if $n = 5$, $k = 5$, and $M = [1, 1, 4, 5, 5]$, then $g(M) = \{1, 2, 6, 8, 9\} \subseteq \{1, 2, \ldots, 9\}$. Notice that the elements of the *set* $g(M)$ all lie in $\{1, 2, \ldots, k + n - 1\}$ since $1 \leq x_i \leq n$ for all i. Also, the k displayed elements of $g(M)$ are pairwise distinct because, for any $i < j$, the assumption $x_i \leq x_j$ implies $x_i + (i - 1) < x_j + (j - 1)$. Thus, $g(M)$ is indeed a k-element subset of B.

Going the other way, define $g' : W \to U$ as follows. Given $S \in W$, we can write S uniquely in the form $S = \{y_1, y_2, \ldots, y_k\}$ where $y_1 < y_2 < \cdots < y_k$. Now define

$$g'(S) = g'(\{y_1, y_2, \ldots, y_k\}) = [y_1 - 0, y_2 - 1, \ldots, y_i - (i - 1), \ldots, y_k - (k - 1)].$$

For example, if $n = k = 5$ and $S = \{2, 3, 5, 7, 8\}$, then $g'(S) = [2, 2, 3, 4, 4] \in U$. Since every $y_j \geq 1$ and the y_i's form a strictly increasing sequence of integers, it follows that $i \leq y_i$ for all i. Similarly, since every $y_j \leq k + n - 1$ and there are $k - i$ entries that exceed y_i in the sequence (namely $y_{i+1}, \ldots, y_k$), we deduce that $y_i \leq (k + n - 1) - (k - i) = n + i - 1$ for all i. Subtracting $i - 1$, it follows that every element of the k-element multiset $g'(S)$ lies in the range $\{1, 2, \ldots, n\}$, so that $g'(S)$ really is an element of U. It is now routine to check that g' is the two-sided inverse of g. □

1.12 Probability

The basic techniques of counting can be applied to solve a number of problems from probability theory. This section introduces some fundamental concepts of probability and considers several examples.

1.59. Definition: Sample Spaces and Events. A *sample space* is a set S, whose members represent the possible outcomes of a "random experiment." In this section, we only consider *finite* sample spaces. An *event* is a subset of the sample space.

Intuitively, an event consists of the set of outcomes of the experiment that possess a particular property we are interested in.

1.60. Example: Coin Tossing. Suppose the experiment consists of tossing a coin five times. We could take the sample space for this experiment to be $S = \{H, T\}^5$, the set of all 5-letter words using the letters H (for heads) and T (for tails). The element HHHTH $\in S$ represents the outcome where the fourth toss was tails and all other tosses were heads. The subset $A = \{w \in S : w_1 = H\}$ is the event in which the first toss comes up heads. The subset $B = \{w \in S : w_1 \neq w_5\}$ is the event that the first toss is different from the last toss. The subset

$$C = \{w \in S : w_i = T \text{ for an odd number of indices } i\}$$

is the event that we get an odd number of tails.

1.61. Example: Dice Rolling. Suppose the experiment consists of rolling a six-sided die three times. The sample space for this experiment is $S = \{1, 2, 3, 4, 5, 6\}^3$, the set of all 3-letter words over the alphabet $\{1, 2, \ldots, 6\}$. The subset $A = \{w \in S : w_1 + w_2 + w_3 \in \{7, 11\}\}$ is the event that the sum of the three numbers rolled is 7 or 11. The subset $B = \{w \in S : w_1 = w_2 = w_3\}$ is the event that all three numbers rolled are the same. The subset $C = \{w \in S : w \neq (4, 1, 3)\}$ is the event that we do not see the numbers 4, 1, 3 (in that order) in the dice rolls.

1.62. Example: Lotteries. Consider the following random experiment. We put 49 white balls (numbered 1 through 49) into a machine that mixes the balls for awhile and then outputs a sequence of six distinct balls, one at a time. We could take the sample space here to be the set S' of all 6-letter words w consisting of six distinct letters from $A = \{1, 2, \ldots, 49\}$. In lotteries, the order in which the balls are drawn usually does not matter, so it is more common to take the sample space to be the set S of all 6-element subsets of A. (We will see later that using S instead of S' does not affect the probabilities we are interested in.) Suppose a lottery player picks a (fixed and known) 6-element subset T_0 of A. For $0 \leq k \leq 6$, define events $B_k = \{T \in S : |T \cap T_0| = k\} \subseteq S$. Intuitively, the event B_k is the set of outcomes in which the player has matched exactly k of the winning lottery numbers.

1.63. Example: Special Events. For any sample space S, $\emptyset$ and S are events. Intuitively, the event $\emptyset$ contains no outcomes, and therefore "never happens." On the other hand, the event S contains all the outcomes, and therefore "always happens." If A and B are *events* (i.e., subsets of S), note that $A \cup B$, $A \cap B$, $S \sim A$, and $A \sim B$ are also *events*. Intuitively, $A \cup B$ is the event that either A happens or B happens (or both); $A \cap B$ is the event that both A and B happen; $S \sim A$ is the event that A does not happen; and $A \sim B$ is the event that A happens but B does not happen.

Now we can formally define the concept of probability. Intuitively, for each event A, we want to define a number $P(A)$ that measures the probability or likelihood that A occurs. Numbers close to 1 represent more likely events, while numbers close to 0 represent less likely events. A probability-zero event is "impossible," while a probability-one event is "certain" to occur.

1.64. Definition: Probability. Assume S is a finite sample space. Recall that $\mathcal{P}(S)$ is the set of all subsets of S, i.e., the set of all events. A *probability measure* for S is a function $P : \mathcal{P}(S) \to [0, 1]$ such that $P(\emptyset) = 0$; $P(S) = 1$; and for any two *disjoint* events A and B, $P(A \cup B) = P(A) + P(B)$.

By induction, it follows that P satisfies the *finite additivity* property

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

for all pairwise disjoint sets $A_1, A_2, \ldots, A_n \subseteq S$.

1.65. Example: Classical Probability Spaces. Suppose S is a finite sample space in which all outcomes are equally likely. Then we must have $P(\{x\}) = 1/|S|$ for each outcome $x \in S$. For any event $A \subseteq S$, finite additivity gives

$$P(A) = \frac{|A|}{|S|} = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}}. \tag{1.1}$$

Thus the calculation of probabilities (in this classical setup) reduces to two counting problems: counting the number of elements in A and counting the number of elements in S. We can take equation (1.1) as the *definition* of our probability measure P. Note that the axiom $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$ is then a consequence of the sum rule. Also note that this probability model will only be appropriate if all the possible outcomes of the underlying random experiment are equally likely to occur.

1.66. Example: Coin Tossing. Suppose we toss a fair coin five times. The sample space is $S = \{H, T\}^5$, so that $|S| = 2^5 = 32$. Consider the event $A = \{w \in S : w_1 = H\}$ of getting a head on the first toss. By the product rule, $|A| = 1 \cdot 2^4 = 16$, so $P(A) = 16/32 = 1/2$. Consider the event $B = \{w \in S : w_1 \neq w_5\}$ in which the first toss differs from the last toss.

B is the disjoint union of $B_1 = \{w \in S : w_1 = H, w_5 = T\}$ and $B_2 = \{w \in S : w_1 = T, w_5 = H\}$. The product rule shows that $|B_1| = |B_2| = 2^3 = 8$, so that $P(B) = (8+8)/32 = 1/2$. Finally, consider the event

$$C = \{w \in S : w_i = T \text{ for an odd number of indices } i\}.$$

C is the disjoint union $C_1 \cup C_3 \cup C_5$, where (for $0 \leq k \leq 5$) C_k is the event of getting exactly k tails. We have $C_k = \mathcal{R}(T^k H^{5-k})$, so that $P(C_k) = \binom{5}{k}/2^5$. Therefore,

$$P(C) = \frac{\binom{5}{1} + \binom{5}{3} + \binom{5}{5}}{2^5} = 16/32 = 1/2.$$

1.67. Example: Dice Rolling. Consider the experiment of rolling a six-sided die twice. The sample space is $S = \{1,2,3,4,5,6\}^2$, so that $|S| = 6^2 = 36$. Consider the event $A = \{x \in S : x_1 + x_2 \in \{7, 11\}\}$ of rolling a sum of 7 or 11. By direct enumeration, we have

$$A = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1),(5,6),(6,5)\}; \qquad |A| = 8.$$

Therefore, $P(A) = 8/36 = 2/9$. Consider the event $B = \{x \in S : x_1 \neq x_2\}$ of getting two different numbers on the two rolls. The product rule gives $|B| = 6 \cdot 5 = 30$, so $P(B) = 30/36 = 5/6$.

1.68. Example: Balls in Urns. Suppose an urn contains n_1 red balls, n_2 white balls, and n_3 blue balls. Let the random experiment consist of randomly drawing a k-element subset of balls from the urn. What is the probability of drawing k_1 red balls, k_2 white balls, and k_3 blue balls, where $k_1 + k_2 + k_3 = k$? We can take the sample space S to be all k-element subsets of the set

$$\{1, 2, \ldots, n_1, n_1+1, \ldots, n_1+n_2, n_1+n_2+1, \ldots, n_1+n_2+n_3\}.$$

Here the first n_1 integers represent red balls, the next n_2 integers represent white balls, and the last n_3 integers represent blue balls. We know that $|S| = \binom{n_1+n_2+n_3}{k}$. Let A be the event where we draw k_1 red balls, k_2 white balls, and k_3 blue balls. To build a set $T \in A$, we choose a k_1-element subset of $\{1, 2, \ldots, n_1\}$, then a k_2-element subset of $\{n_1+1, \ldots, n_1+n_2\}$, then a k_3-element subset of $\{n_1+n_2+1, \ldots, n_1+n_2+n_3\}$. By the product rule, $|A| = \binom{n_1}{k_1}\binom{n_2}{k_2}\binom{n_3}{k_3}$. Therefore, the definition of the probability measure gives

$$P(A) = \frac{\binom{n_1}{k_1}\binom{n_2}{k_2}\binom{n_3}{k_3}}{\binom{n_1+n_2+n_3}{k_1+k_2+k_3}}.$$

This calculation can be generalized to the case where the urn has balls of more than three colors.

1.69. Example: Lotteries. Consider the lottery described in 1.62. Here the sample space S consists of all 6-element subsets of $A = \{1, 2, \ldots, 49\}$, so $|S| = \binom{49}{6} = 13,983,816$. Suppose a lottery player picks a (fixed and known) 6-element subset T_0 of A. For $0 \leq k \leq 6$, define events $B_k = \{T \in S : |T \cap T_0| = k\}$. B_k occurs when the player matches exactly k of the winning numbers. We can build a typical object $T \in B_k$ by choosing k elements of T_0 in $\binom{6}{k}$ ways, and then choosing $6-k$ elements of $A \sim T_0$ in $\binom{43}{6-k}$ ways. Hence,

$$P(B_k) = \frac{\binom{6}{k}\binom{43}{6-k}}{\binom{49}{6}}.$$

TABLE 1.2
Analysis of Virginia's "Lotto South" lottery.

Matches	Probability	Prize Value
3	0.01765 or 1 in 57	about \$5
4	0.0009686 or 1 in 1032	about \$75
5	0.00001845 or 1 in 54,201	about \$1000
6	7.15×10^{-8} or 1 in 13,983,816	jackpot

Table 1.2 shows the probability of matching k numbers, for $3 \leq k \leq 6$. The table also shows the amount of money one would win in the various cases. One can view this example as the special case of the previous example where the urn contains 6 balls of one color and 43 balls of another color.

In the lottery example, suppose we took the sample space to be the set S' of all *ordered* sequences of six distinct elements of $\{1, 2, \ldots, 49\}$. Let B'_k be the event that the player guesses exactly k numbers correctly (disregarding order, as usual). Let P' be the probability measure on the sample space S'. One may check that $|S'| = \binom{49}{6} \cdot 6!$ and $|B'_k| = \binom{6}{k}\binom{43}{6-k} \cdot 6!$, so that

$$P'(B'_k) = \frac{\binom{6}{k}\binom{43}{6-k}6!}{\binom{49}{6}6!} = P(B_k).$$

This confirms our earlier remark that the two sample spaces S and S' give the same probabilities for events that do not depend on the order in which the balls are drawn.

1.70. Example: Lattice Paths. Suppose we randomly choose a lattice path from $(0, 0)$ to (n, n). What is the probability that this path is a Dyck path? We know that there are $\frac{1}{n+1}\binom{2n}{n,n}$ Dyck paths and $\binom{2n}{n,n}$ lattice paths ending at (n, n). Therefore, the probability is $1/(n + 1)$. We discuss a remarkable generalization of this result, called the *Chung-Feller Theorem*, in §12.2.

1.71. Example: General Probability Measures on a Finite Sample Space. We now extend the previous discussion to the case where not all outcomes of the random experiment are equally likely. Let S be a finite sample space and let $p : S \to [0, 1]$ be a map such that $\sum_{x \in S} p(x) = 1$. Intuitively, $p(x)$ is the probability that the outcome x occurs. Now p is not a probability measure, since its domain is S instead of $\mathcal{P}(S)$. We build a probability measure from p by defining $P(A) = \sum_{x \in A} p(x)$. The axioms for a probability measure may be routinely verified. A similar construction works in the case where S is a countably infinite sample space. (Recall that a set S is *countably infinite* iff there exists a bijection $f : \mathbb{N} \to S$.)

1.72. Remark. In this section, we used counting techniques to solve basic probability questions. It is also possible to use probabilistic arguments to help solve counting problems. Examples of such arguments appear in §12.4 and §12.10.

1.13 Games of Chance

In this section, we use counting techniques to analyze two popular games of chance: powerball lotteries and five-card poker.

TABLE 1.3
Analysis of the Powerball lottery.

Matches	Probability	Prize Value
0 white, 1 red	0.0145 or 1 in 69	\$3
1 white, 1 red	0.00788 or 1 in 127	\$4
2 white, 1 red	0.00134 or 1 in 745	\$7
3 white, 0 red	0.00344 or 1 in 291	\$7
3 white, 1 red	0.0000838 or 1 in 11,927	\$100
4 white, 0 red	0.0000702 or 1 in 14,254	\$100
4 white, 1 red	0.000001711 or 1 in 584,432	\$10,000
5 white, 0 red	2.81×10^{-7} or 1 in 3.56 million	\$200,000
5 white, 1 red	6.844×10^{-9} or 1 in 146 million	jackpot

1.73. Example: Powerball. A *powerball lottery* has two kinds of balls: white balls (numbered $1, \ldots, M$) and red balls (numbered $1, \ldots, R$). Each week, one red ball and a set of n distinct white balls are randomly chosen. Lottery players guess what the n white balls will be, and they also guess the red ball (called the "power ball"). Players win prizes based on how many balls they guess correctly. Players always win a prize for matching the red ball, even if they incorrectly guess all the white balls.

To analyze this lottery, let the sample space be

$$S = \{(T, x) : T \text{ is an } n\text{-element subset of } \{1, 2, \ldots, M\} \text{ and } x \in \{1, 2, \ldots, R\}\}.$$

Let (T_0, x_0) be a fixed and known element of S representing a given player's lottery ticket. For $0 \le k \le n$, let A_k be the event $\{(T, x) \in S : |T \cap T_0| = k, x \neq x_0\}$ in which the player matches exactly k white balls but misses the power ball. Let B_k be the event $\{(T, x) \in S : |T \cap T_0| = k, x = x_0\}$ in which the player matches exactly k white balls and also matches the power ball. We have $|S| = \binom{M}{n}R$ by the product rule. To build a typical element in A_k, we first choose k elements of T_0, then choose $n - k$ elements of $\{1, 2, \ldots, M\} \sim T_0$, then choose $x \in \{1, 2, \ldots, R\} \sim \{x_0\}$. Thus, $|A_k| = \binom{n}{k}\binom{M-n}{n-k}(R-1)$, so

$$P(A_k) = \frac{\binom{n}{k}\binom{M-n}{n-k}(R-1)}{\binom{M}{n}R}.$$

Similarly,

$$P(B_k) = \frac{\binom{n}{k}\binom{M-n}{n-k} \cdot 1}{\binom{M}{n}R}.$$

In one version of this lottery, we have $M = 55$, $R = 42$, and $n = 5$. The probabilities of certain events A_k and B_k are shown in Table 1.3 together with the associated prize amounts.

Now we turn to an analysis of five-card poker.

1.74. Definition: Cards. A *suit* is an element of the 4-element set Suits $= \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. A *value* is an element of the 13-element set Values $= \{2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A\}$, where J, Q, K, and A stand for "jack," "queen," "king," and "ace," respectively. A *card* is an element of the set Deck = Values $\times$ Suits.

Note that $|\,\text{Deck}\,| = 13 \cdot 4 = 52$, by the product rule. For instance, $(A, \spadesuit) \in$ Deck. We often abbreviate this notation to $A\spadesuit$, and similarly for other cards.

1.75. Definition: Poker Hands. A *(five-card) poker hand* is a 5-element subset of Deck. Given such a hand H, let $V(H)$ be the set of values that appear among the cards of H, and let $S(H)$ be the set of suits that appear among the cards of H. For each $x \in$ Values, let $n_x(H)$ be the number of cards in H with value x.

1.76. Example. $H = \{A\diamondsuit, 3\clubsuit, 3\diamondsuit, J\diamondsuit, K\clubsuit\}$ is a five-card poker hand with $V(H) = \{A, 3, J, K\}$, $S(H) = \{\diamondsuit, \clubsuit\}$, $n_3(H) = 2$, $n_A(H) = n_J(H) = n_K(H) = 1$, and $n_x(H) = 0$ for all $x \notin V(H)$.

We now study the sample space X consisting of all five-card poker hands. We know that $|X| = \binom{52}{5} = 2,598,960$. In poker, certain hands in X play a special role. We define these hands now.

1.77. Definition: Special Card Hands. Let H be a five-card poker hand.

- H is a *four-of-a-kind hand* iff there exists $x \in$ Values with $n_x(H) = 4$.
- H is a *full house* iff there exist $x, y \in$ Values with $n_x(H) = 3$ and $n_y(H) = 2$.
- H is a *three-of-a-kind hand* iff there exist $x, y, z \in$ Values with $y \neq z$, $n_x(H) = 3$, and $n_y(H) = n_z(H) = 1$.
- H is a *two-pair hand* iff there exist $x, y, z \in$ Values with $y \neq z$, $n_x(H) = 1$, and $n_y(H) = n_z(H) = 2$.
- H is a *one-pair hand* iff there exist distinct $w, x, y, z \in$ Values with $n_w(H) = 2$ and $n_x(H) = n_y(H) = n_z(H) = 1$.
- H is a *straight* iff $V(H)$ is one of the following sets:

$$\{A, 2, 3, 4, 5\} \text{ or } \{i, i+1, i+2, i+3, i+4\} \text{ for some } i \text{ with } 2 \leq i \leq 6$$

$$\text{or } \{7, 8, 9, 10, J\} \text{ or } \{8, 9, 10, J, Q\} \text{ or } \{9, 10, J, Q, K\} \text{ or } \{10, J, Q, K, A\}.$$

- H is a *flush* iff $|S(H)| = 1$.
- H is a *straight flush* iff H is a straight and a flush.
- H is an *ordinary hand* iff H satisfies none of the above conditions.

1.78. Example: Card Hands.

- $\{5\spadesuit, 8\clubsuit, 5\diamondsuit, 5\clubsuit, 5\heartsuit\}$ is a four-of-a-kind hand.
- $\{J\spadesuit, 9\clubsuit, J\diamondsuit, J\clubsuit, 9\heartsuit\}$ is a full house.
- $\{J\spadesuit, 2\clubsuit, J\diamondsuit, J\clubsuit, 9\heartsuit\}$ is a three-of-a-kind hand.
- $\{2\spadesuit, 9\clubsuit, K\diamondsuit, 2\clubsuit, 9\heartsuit\}$ is a two-pair hand.
- $\{9\clubsuit, 10\diamondsuit, 10\heartsuit, A\clubsuit, 4\clubsuit\}$ is a one-pair hand.
- $\{7\clubsuit, 6\diamondsuit, 3\heartsuit, 5\clubsuit, 4\clubsuit\}$ is a straight that is not a flush.
- $\{10\heartsuit, 3\heartsuit, Q\heartsuit, J\heartsuit, 8\heartsuit\}$ is a flush that is not a straight.

TABLE 1.4
Probability of five-card poker hands.

Card Hand	Number	Probability
straight flush	40	1.54×10^{-5}
four-of-a-kind	624	0.00024
full house	3744	0.00144
flush (not straight)	5108	0.001965
straight (not flush)	10,200	0.00392
three-of-a-kind	54,912	0.02113
two pair	123,552	0.04754
one pair	1,098,240	0.42257
none of the above	1,302,540	0.50117
TOTAL	2,598,960	1.00000

- $\{10\spadesuit, J\spadesuit, Q\spadesuit, K\spadesuit, A\spadesuit\}$ is a straight flush. (A straight flush such as this one, which "starts at 10 and ends at A," is called a *royal flush*. There are four royal flushes, one for each suit.)
- $\{9\clubsuit, 10\diamondsuit, 7\heartsuit, A\clubsuit, 4\clubsuit\}$ is an ordinary hand.

We now compute the probability of the various five-card poker hands. This amounts to enumerating the hands of each type and dividing these counts by $|X| = \binom{52}{5} = 2,598,960$. Our results are summarized in Table 1.4. In each case, the desired counting result will follow from careful applications of the product rule. Less frequently occurring poker hands are more valuable in the game. So, for instance, a flush beats a straight. A full house beats both a straight and a flush separately, but is beaten by a straight flush.

- *Four-of-a-kind hands.* To build a typical four-of-a-kind hand H, first choose the value x that occurs 4 times in any of $|\,\text{Values}\,| = 13$ ways. All four cards of this value must belong to H. Second, choose the fifth card of H in any of $52 - 4 = 48$ ways. This gives $13 \times 48 = 624$ four-of-a-kind hands. The sample hand above was constructed by choosing the value 5 followed by the card $8\clubsuit$.
- *Full house hands.* To build a typical full house H, first choose a value $x \in$ Values to occur 3 times. This can be done in 13 ways. Second, choose 3 of the 4 cards of value x to appear in the hand. This can be done in $\binom{4}{3} = 4$ ways. Third, choose a value $y \in$ Values $\sim \{x\}$ to occur twice in H. This can be done in 12 ways. Fourth, choose 2 of the 4 cards of value y to appear in the hand. This can be done in $\binom{4}{2} = 6$ ways. The total is $13 \cdot 4 \cdot 12 \cdot 6 = 3744$ full house hands. The sample hand above was constructed by choosing the value J, then the three cards $\{J\spadesuit, J\diamondsuit, J\clubsuit\}$, then the value 9, then the two cards $\{9\clubsuit, 9\heartsuit\}$.
- *Three-of-a-kind hands.* To build a typical three-of-a-kind hand H, first choose a value $x \in$ Values to occur 3 times. This can be done in 13 ways. Second, choose 3 of the 4 cards of value x to appear in the hand. This can be done in $\binom{4}{3} = 4$ ways. Third, choose a set of 2 values $\{y, z\} \subseteq$ Values $\sim \{x\}$ that will occur once each in H. This can be done in $\binom{12}{2} = 66$ ways. Let the notation be such that $y < z$ (where $10 < J < Q < K < A$). Fourth, choose one of the 4 cards of value y to be in the hand in any of 4 ways. Fifth, choose one of the 4 cards of value z to be in the hand in any of 4 ways. The total is $13 \cdot 4 \cdot 66 \cdot 4 \cdot 4 = 54,912$. The sample hand above was constructed by choosing the value

J, then the three cards $\{J\spadesuit, J\diamondsuit, J\clubsuit\}$, then the values $\{2, 9\}$, then the card $2\clubsuit$, and then the card $9\heartsuit$.

- *Two-pair hands.* To build a typical two-pair hand H, first choose a set of two values $\{x, y\} \in$ Values to occur twice each. This can be done in $\binom{13}{2} = 78$ ways. Let the notation be such that $x < y$. Second, choose a set of two cards of value x in any of $\binom{4}{2} = 6$ ways. Third, choose a set of two cards of value y in any of $\binom{4}{2} = 6$ ways. Fourth, choose the last card in the hand. Since this card cannot have value x or y, the number of possibilities here is $52 - 8 = 44$. The total is $78 \cdot 6 \cdot 6 \cdot 44 = 123,552$. The sample hand above was constructed by choosing the values $\{2, 9\}$, then the cards $\{2\spadesuit, 2\clubsuit\}$, then the cards $\{9\clubsuit, 9\heartsuit\}$, then the card $K\diamondsuit$.

- *One-pair hands.* To build a typical one-pair hand H, first choose a value w to occur twice in the hand. This can be done in 13 ways. Second, choose a set of two cards of value w in any of $\binom{4}{2} = 6$ ways. Third, choose a set $\{x, y, z\} \subset$ Values $\sim \{x\}$ (where $x < y < z$) in any of $\binom{12}{3} = 220$ ways. Fourth, choose a card of value x in 4 ways. Fifth, choose a card of value y in 4 ways. Sixth, choose a card of value z in 4 ways. The total is $13 \cdot 6 \cdot 220 \cdot 4 \cdot 4 \cdot 4 = 1,098,240$. The sample hand above was constructed by choosing the value $w = 10$, then the cards $\{10\diamondsuit, 10\heartsuit\}$, then the values $\{4, 9, A\}$, then the card $4\clubsuit$, then the card $9\clubsuit$, then the card $A\clubsuit$.

- *Straight hands.* To build a typical straight H, first choose one of the ten allowable sets $V(H)$ in the definition of a straight. Then, for each of the five distinct values in $V(H)$, taken in increasing order, choose a suit for the card of that value. This can be done in 4 ways for each value. The total is $10 \cdot 4^5 = 10,240$. The sample hand above was constructed by choosing the value set $V(H) = \{3, 4, 5, 6, 7\}$, then the suit $\heartsuit$ for the 3, then the suit $\clubsuit$ for the 4, then the suit $\clubsuit$ for the 5, then the suit $\diamondsuit$ for the 6, and then the suit $\clubsuit$ for the 7. In the table entry for straights, we subtract the number of straight flushes (namely 40, as shown below) so that the entries in the table will be pairwise disjoint subsets of X.

- *Flush hands.* To build a typical flush H, first choose the one-element set $S(H)$ in any of $\binom{4}{1} = 4$ ways. Then choose the five-element set $V(H)$ in any of $\binom{13}{5}$ ways. H is now completely determined since all cards in H have the same suit. The total is therefore $4 \cdot \binom{13}{5} = 5148$. The sample hand above was constructed by choosing $S(H) = \{\heartsuit\}$, then $V(H) = \{3, 8, 10, J, Q\}$. In the table entry for flushes, we subtract the number of straight flushes (namely 40, as shown below) so that the entries in the table will be pairwise disjoint subsets of X.

- *Straight flushes.* To build a typical straight flush H, first choose one of the ten allowable sets $V(H)$ in the definition of a straight. Then choose one of the four suits to be the common suit of all cards in H. The total is $10 \cdot 4 = 40$. The sample hand above was constructed by choosing $V(H) = \{10, J, Q, K, A\}$ and then $S(H) = \{\spadesuit\}$.

- *Ordinary hands.* To count ordinary hands, one can subtract the total of the preceding counts from $|X|$. However, the answer can also be obtained directly from the product rule as follows. To build an ordinary hand H, first choose the value set $V(H)$. We must have $|V(H)| = 5$ to avoid hands such as two-pair, full house, etc. Also we must avoid the ten special choices of $V(H)$ in the definition of straight (all of which are five-element sets). We conclude that $V(H)$ can be chosen in $\binom{13}{5} - 10 = 1277$ ways. Write $V(H) = \{v_1, v_2, v_3, v_4, v_5\}$, where $v_1 < v_2 < v_3 < v_4 < v_5$. For each v_i in turn, choose the suit for the card of that value in any of 4 ways. This would give 4^5 choices, but we must avoid the four choice sequences in which all v_i's are assigned the same suit (which

would lead to a flush). So there are only $4^5 - 4 = 1020$ ways to assign suits to the chosen values. The hand is now completely determined, so the total number of ordinary hands is $1277 \cdot 1020 = 1,302,540$. The sample hand above was constructed by choosing $V(H) = \{4, 7, 9, 10, A\}$, then ♣ as the suit for the 4, ♡ as the suit for the 7, ♣ as the suit for the 9, ♢ as the suit for the 10, and ♣ as the suit for the ace.

1.14 Conditional Probability and Independence

Suppose that, in a certain random experiment, we are told that a particular event has occurred. Given this additional information, we can recompute the probability of other events occurring. This leads to the notion of conditional probability.

1.79. Definition: Conditional Probability. Suppose A and B are events in some sample space S such that $P(B) > 0$. The *conditional probability of A given B*, denoted $P(A|B)$, is defined by setting

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

In the case where S is a finite set of equally likely outcomes, we have $P(A|B) = |A \cap B|/|B|$. This conditional probability need not have any relation to the *unconditional* probability of A, which is $P(A) = |A|/|S|$.

1.80. Example: Dice Rolling. Consider the experiment of rolling a fair die twice. What is the probability of getting a sum of 7 or 11, given that the second roll comes up 5? Here, the sample space is $S = \{1, 2, 3, 4, 5, 6\}^2$. Let A be the event of getting a sum of 7 or 11, and let B be the event that the second die shows 5. We have $P(B) = 1/6$, and we saw earlier that $P(A) = 2/9$. Listing outcomes, we see that $A \cap B = \{(2, 5), (6, 5)\}$, so $P(A \cap B) = 2/36 = 1/18$. Therefore, the required conditional probability is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1/18}{1/6} = 1/3 > 2/9 = P(A).$$

On the other hand, let C be the event that the second roll comes up 4. Here $A \cap C = \{(3, 4)\}$, so

$$P(A|C) = \frac{1/36}{6/36} = 1/6 < 2/9 = P(A).$$

Next, let D be the event that the first roll is an odd number. Then

$$A \cap D = \{(1, 6), (3, 4), (5, 2), (5, 6)\},$$

so

$$P(A|D) = \frac{4/36}{18/36} = 2/9 = P(A).$$

These examples show that the conditional probability of A given some other event can be greater than, less than, or equal to the unconditional probability of A.

1.81. Example: Balls in Urns. Suppose an urn contains r red balls and b blue balls, where $r, b \geq 2$. Consider an experiment in which two balls are drawn from the urn in succession, without replacement. What is the probability that the first ball is red, given

that the second ball is blue? We take the sample space to be the set S of all words w_1w_2, where $w_1 \neq w_2$ and

$$w_1, w_2 \in \{1, 2, \ldots, r, r+1, \ldots, r+b\}.$$

Here, the numbers 1 through r represent red balls and the numbers $r+1$ through $r+b$ represent blue balls. The event of drawing a red ball first is the subset

$$A = \{w_1w_2 : 1 \leq w_1 \leq r\}.$$

The event of drawing a blue ball second is the subset

$$B = \{w_1w_2 : r+1 \leq w_2 \leq r+b\}.$$

By the product rule, $|S| = (r+b)(r+b-1)$, $|A| = r(r+b-1)$, $|B| = b(r+b-1)$, and $|A \cap B| = rb$. The conditional probability of A given B is

$$P(A|B) = P(A \cap B)/P(B) = r/(r+b-1).$$

In contrast, the unconditional probability of A is

$$P(A) = |A|/|S| = r/(r+b).$$

The conditional probability is slightly higher than the unconditional probability; intuitively, we are more likely to have gotten a red ball first if we know the second ball was not red. The probability that the second ball is blue, given that the first ball is red, is

$$P(B|A) = P(B \cap A)/P(A) = b/(r+b-1).$$

Note that $P(B|A) \neq P(A|B)$ (unless $r = b$).

1.82. Example: Card Hands. What is the probability that a 5-card poker hand is a full house, given that the hand is void in clubs (i.e., no card in the hand is a club)? Let A be the event of getting a full house, and let B be the event of being void in clubs. We have $|B| = \binom{39}{5} = 575,757$ since we must choose a five-element subset of the $52 - 13 = 39$ non-club cards. Next, we must compute $|A \cap B|$. To build a full house hand using no clubs, make the following choices: first, choose a value to occur three times (13 ways); second, choose the suits for this value (1 way, as clubs are forbidden); third, choose a value to occur twice (12 ways); fourth, choose the suits for this value ($\binom{3}{2} = 3$ ways). By the product rule, $|A \cap B| = 13 \cdot 1 \cdot 12 \cdot 3 = 468$. Accordingly, the probability we want is $P(A|B) = 468/575,757 \approx 0.000813$.

Next, what is the probability of getting a full house, given that the hand has at least two cards of the same value? Let C be the event that at least two cards in the hand have the same value; we seek $P(A|C) = P(A \cap C)/P(C) = |A \cap C|/|C|$. The numerator here can be computed quickly: since $A \subseteq C$, we have $A \cap C = A$ and hence $|A \cap C| = |A| = 3744$ (see Table 1.4). To compute the denominator, let us first enumerate $X \sim C$, where X is the full sample space of all five-card poker hands. Note that $X \sim C$ occurs iff all five cards in the hand have different values. Choose these values ($\binom{13}{5}$ ways), and then choose suits for each card (4 ways each). By the product rule, $|X \sim C| = 1,317,888$. So

$$|C| = |X| - |X \sim C| = 1,281,072.$$

The desired conditional probability is

$$P(A|C) = \frac{3744}{1,281,072} \approx 0.00292.$$

In some situations, the knowledge that a particular event D occurs does not change the probability that another event A will occur. For instance, events D and A in the dice rolling example 1.80 have this property because $P(A|D) = P(A)$. Writing out the definition of $P(A|D)$ and multiplying by $P(D)$, we see that the stated property is equivalent to $P(A \cap D) = P(A)P(D)$ (assuming $P(D) > 0$). This suggests the following definition, which is valid even when $P(D) = 0$.

1.83. Definition: Independence of Two Events. Two events A and D are called *independent* iff

$$P(A \cap D) = P(A)P(D).$$

Unlike the definition of conditional probability, this definition is symmetric in A and D. So, A and D are independent iff D and A are independent. As indicated above, when $P(D) > 0$, independence of A and D is equivalent to $P(A|D) = P(A)$. Similarly, when $P(A) > 0$, independence of A and D is equivalent to $P(D|A) = P(A)$. So, when considering two independent events of positive probability, knowledge that either event has occurred gives us no new information about the probability of the other event occurring.

1.84. Definition: Independence of a Collection of Events. Suppose $A_1, \ldots, A_n$ are events. This list of events is called *independent* iff for all choices of indices $i_1 < i_2 < \cdots < i_k \leq n$,

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k}).$$

1.85. Example. Let $S = \{a, b, c, d\}$, and suppose each outcome in S occurs with probability $1/4$. Define events $B = \{a, b\}$, $C = \{a, c\}$, and $D = \{a, d\}$. One verifies immediately that B and C are independent; B and D are independent; and C and D are independent. However, the triple of events B, C, D is *not* independent, because

$$P(B \cap C \cap D) = P(\{a\}) = 1/4 \neq 1/8 = P(B)P(C)P(D).$$

1.86. Example: Coin Tossing. Suppose we toss a fair coin 5 times. Take the sample space to be $S = \{H, T\}^5$. Let A be the event that the first and last toss agree; let B be the event that the third toss is tails; let C be the event that there are an odd number of heads. Routine counting arguments show that $|S| = 2^5 = 32$, $|A| = 2^4 = 16$, $|B| = 2^4 = 16$, $|C| = \binom{5}{1} + \binom{5}{3} + \binom{5}{5} = 16$, $|A \cap B| = 2^3 = 8$, $|A \cap C| = 2(\binom{3}{1} + \binom{3}{3}) = 8$, $|B \cap C| = \binom{4}{1} + \binom{4}{3} = 8$, and $|A \cap B \cap C| = 4$. It follows that

$$P(A \cap B) = P(A)P(B); \quad P(A \cap C) = P(A)P(C); \quad P(B \cap C) = P(B)P(C);$$

$$P(A \cap B \cap C) = P(A)P(B)P(C).$$

Thus, the triple of events (A, B, C) is independent.

We often assume that unrelated physical events are independent (in the mathematical sense) to help us construct a probability model. The next example illustrates this process.

1.87. Example: Tossing an Unfair Coin. Consider a random experiment in which we toss an unbalanced coin n times in a row. Suppose that the coin comes up heads with probability q and tails with probability $1 - q$, and that successive coin tosses are unrelated to one another. Let the sample space be $S = \{H, T\}^n$. Since the coin is unfair, it is not appropriate to assume that every point of S occurs with equal probability. Given an outcome $w = w_1 w_2 \cdots w_n \in S$, what should the probability $p(w)$ be? Consider an example where $n = 5$ and $w = \text{HHTHT}$. Consider the five events $B_1 = \{z \in S : z_1 = H\}$, $B_2 = \{z \in S : z_2 = H\}$, $B_3 = \{z \in S : z_3 = T\}$, $B_4 = \{z \in S : z_4 = H\}$, and $B_5 = \{z \in S : z_5 = T\}$. Our physical assumptions suggest that $B_1, \ldots, B_5$ should be independent events (since different

tosses of the coin are unrelated), $P(B_1) = P(B_2) = P(B_4) = q$, and $P(B_3) = P(B_5) = 1-q$. Since $B_1 \cap B_2 \cap B_3 \cap B_4 \cap B_5 = \{w\}$, the definition of independence leads to

$$p(w) = P(B_1 \cap \cdots \cap B_5) = P(B_1)P(B_2)\cdots P(B_5) = qq(1-q)q(1-q) = q^3(1-q)^2.$$

Similar reasoning shows that if $w = w_1 w_2 \cdots w_n \in S$ is an outcome consisting of k heads and $n-k$ tails (arranged in one particular order), then we should define $p(w) = q^k(1-q)^{n-k}$. Next, define $P(A) = \sum_{w \in A} p(w)$ for every event $A \subseteq S$. For example, let A_k be the event that we get k heads and $n-k$ tails (in any order). Note that $|A_k| = |\mathcal{R}(H^k T^{n-k})| = \binom{n}{k}$, and $p(w) = q^k(1-q)^{n-k}$ for each $w \in A_k$. It follows that

$$P(A_k) = \binom{n}{k} q^k (1-q)^{n-k}.$$

We have not yet checked that $P(S) = 1$, which is one of the requirements in the definition of a probability measure. This fact can be deduced from the binomial theorem (discussed in §2.2), as follows. Since S is the disjoint union of $A_0, A_1, \ldots, A_n$, we have

$$P(S) = \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k}.$$

By the binomial theorem 2.14, the right side is $(q + [1-q])^n = 1^n = 1$.

Summary

We end each chapter by summarizing some of the main definitions and results discussed in the chapter.

- *Notation.* Factorials: $0! = 1$ and $n! = n \times (n-1) \times \cdots \times 1$.
  Binomial coefficients: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ for $0 \le k \le n$; $\binom{n}{k} = 0$ otherwise.
  Multinomial coefficients: Given $n_1, \ldots, n_k \ge 0$ and $N = n_1 + \cdots + n_k$, $\binom{N}{n_1,\ldots,n_k} = \frac{N!}{n_1! n_2! \cdots n_k!}$.
  Rearrangements: $\mathcal{R}(a_1^{n_1} \cdots a_k^{n_k})$ is the set of all words consisting of n_i copies of a_i.

- *Basic Counting Rules.* Sum Rule: If $A_1, \ldots, A_k$ are pairwise disjoint finite sets, then $|A_1 \cup \cdots \cup A_k| = |A_1| + \cdots + |A_k|$.
  Union Rule: If A and B are arbitrary finite sets, then $|A \cup B| = |A| + |B| - |A \cap B|$.
  Difference Rule: If $A \subseteq B$ and B is finite, then $|B \sim A| = |B| - |A|$.
  Product Rule: If $A_1, \ldots, A_k$ are arbitrary finite sets, then $|A_1 \times \cdots \times A_k| = |A_1| \cdot \ldots \cdot |A_k|$.
  Bijection Rule: If there is a bijection $f : A \to B$, then $|A| = |B|$.

- *Counting Words.* Let A be an n-letter alphabet.
  There are n^k words of length k using letters from A.
  If the letters must be distinct, there are $n!/(n-k)!$ words of length $k \le n$.
  There are $n!$ permutations of all the letters in A.
  There are $\binom{n_1+\cdots+n_k}{n_1,\ldots,n_k}$ words in $\mathcal{R}(a_1^{n_1} \cdots a_k^{n_k})$.

- *Counting Sets and Multisets.*
  The number of k-element subsets of an n-element set is the binomial coefficient $\binom{n}{k}$.
  The total number of subsets of an n-element set is 2^n.
  The number of k-element multisets using n available objects is $\binom{k+n-1}{k,n-1}$.

- *Counting Functions.* Let $|X| = a$ and $|Y| = b$.
  There are b^a functions mapping X into Y.
  For $a \leq b$, there are $b!/(b-a)!$ injections from X to Y.
  If $a = b$, there are $a!$ bijections from X onto Y.

- *Counting Lattice Paths.* There are $\binom{a+b}{a,b}$ lattice paths from $(0,0)$ to (a,b).
  There are $\binom{n_1+\cdots+n_d}{n_1,\ldots,n_d}$ lattice paths in $\mathbb{R}^d$ from the origin to $(n_1, n_2, \ldots, n_d)$.
  The number of paths from $(0,0)$ to (n,n) that never go below $y = x$ is the Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{2n+1}\binom{2n+1}{n} = \binom{2n}{n,n} - \binom{2n}{n+1,n-1}.$$

  This can be proved using a reflection bijection to convert paths ending at (n,n) that do go below $y = x$ to arbitrary paths from $(0,0)$ to $(n+1, n-1)$.

- *Compositions.* A composition of n is an ordered sequence $(\alpha_1, \ldots, \alpha_k)$ of positive integers that sum to n. There are 2^{n-1} compositions of n. There are $\binom{n-1}{k-1}$ compositions of n with k parts.

- *Probability Definitions.* A *sample space* is the set S of outcomes for some random experiment. An *event* is a subset of the sample space. When all outcomes in S are equally likely, the *probability* of an event A is $P(A) = |A|/|S|$. The *conditional probability of A given B* is $P(A|B) = P(A \cap B)/P(B)$, when $P(B) > 0$. Events A and B are *independent* iff $P(A \cap B) = P(A)P(B)$.

Exercises

1.88. (a) How many numbers between 1 and 1000 are divisible by 5 or 7? (b) How many such numbers are divisible by 5 or 7, but not both?

1.89. How many three-digit numbers: (a) do not contain the digits 5 or 7; (b) contain the digits 5 and 7; (c) contain the digits 5 or 7; (d) contain 5 or 7, but not both?

1.90. How many seven-digit phone numbers do not begin with one of the prefixes 1, 911, 411, or 555?

1.91. How many n-letter words over the alphabet $\{0,1\}$ use both the symbols 0 and 1?

1.92. (a) How many four-letter words w using an n-letter alphabet satisfy $w_i \neq w_{i+1}$ for $i = 1, 2, 3$? (b) How many of the words in (a) also satisfy $w_4 \neq w_1$?

1.93. A key for the *DES encryption system* is a binary word of length 56. A key for a *permutation cipher* is a permutation of the 26-letter English alphabet. Which encryption system has more keys?

1.94. A key for the *AES encryption system* is a binary word of length 128. Suppose we try to decrypt an AES message by exhaustively trying every possible key. Assume six billion computers are running in parallel, where each computer can test one trillion keys per second. Estimate the number of years required for this attack to search the entire space of keys.

1.95. A pizza shop offers ten toppings. How many pizzas can be ordered with: (a) three different toppings; (b) up to three different toppings; (c) three toppings, with repeats allowed; (d) four different toppings, but pepperoni and sausage cannot be ordered together?

1.96. How many lattice paths from $(0,0)$ to $(7,5)$ pass through the point $(2,3)$?

1.97. How many n-letter words contain: (a) only vowels; (b) no vowels; (c) at least one vowel; (d) alternating vowels and consonants; (e) two vowels and $n-2$ consonants? (The vowels are A, E, I, O, and U.)

1.98. How many four-digit even numbers contain the digit 5 but not the digit 2?

1.99. A *palindrome* is a word $w = w_1 w_2 \cdots w_k$ that reads the same in reverse, i.e., $w_1 w_2 \cdots w_k = w_k \cdots w_2 w_1$. Count the number of k-letter palindromes using letters from an n-letter alphabet.

1.100. Explicitly list the following objects: (a) all 4-letter words using the alphabet $\{0,1\}$; (b) all permutations of $\{a,b,c,d\}$; (c) all 2-permutations of $\{u,v,w,x,y\}$; (d) all words in $\mathcal{R}(x^2y^2z^1)$.

1.101. Explicitly list the following objects: (a) all bijections from $\{1,2,3\}$ to $\{i,j,k\}$; (b) all surjections from $\{1,2,3\}$ to $\{0,1\}$; (c) all injections from $\{a,b\}$ to $\{c,d,e,f\}$.

1.102. Explicitly list the following objects: (a) all subsets of $\{0,1,2\}$; (b) all three-element subsets of $\{1,2,3,4,5\}$; (c) all three-element multisets using the alphabet $\{a,b,c\}$.

1.103. Explicitly list the following objects: (a) all compositions of 4; (b) all compositions of 7 with exactly three parts; (c) all lattice paths from $(0,0)$ to $(4,2)$; (d) all Dyck paths of order 4.

1.104. Draw pictures of all compositions of 5. For each composition, determine the associated word in $\{0,1\}^4$ constructed in the proof of 1.41.

1.105. How many lattice paths start at $(0,0)$ and end on the line $x+y=n$?

1.106. Let r be the bijection in the proof of 1.56. Compute

$$r(\text{NNEEEENNNNEEEENN}) \text{ and } r^{-1}(\text{NENEENNEEENEEENN}).$$

1.107. Draw all the non-Dyck lattice paths from $(0,0)$ to $(3,3)$ and compute their images under the reflection map r from the proof of 1.56.

1.108. A *bit* is one of the symbols 0 or 1. Find the minimum k such that every printable character on a standard computer keyboard can be encoded by a distinct bit string of length exactly k. Does the answer change if we allow nonempty bit strings of length *at most* k?

1.109. Ten lollipops are to be distributed to four children. All lollipops of the same color are considered identical. How many distributions are possible if (a) all lollipops are red; (b) all lollipops have different colors; (c) there are four red and six blue lollipops? (d) What are the answers if each child must receive at least one lollipop?

1.110. Given a positive integer n, let the prime factorization of n be $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$, where each $e_j > 0$ and the p_j are distinct primes. How many positive divisors does n have? How many divisors does n have in $\mathbb{Z}$?

1.111. (a) Given k and N, count the number of weakly increasing sequences $(i_1 \le i_2 \le \cdots \le i_k)$ with $1 \le i_j \le N$ for all j. (b) Count the number of strictly decreasing sequences $(i_1 > i_2 > \cdots > i_k)$ with $1 \le i_j \le N$ for all j. (c) For a fixed choice of k, count the number of permutations w of N objects such that

$$w_1 < w_2 < \cdots < w_k > w_{k+1} < w_{k+2} < \cdots < w_N. \tag{1.2}$$

(d) How many permutations satisfy (1.2) for some $k < N$?

1.112. Euler's ϕ Function. For each $n \ge 1$, let $\Phi(n)$ be the set of integers k between 1 and n such that $\gcd(k, n) = 1$, and let $\phi(n) = |\Phi(n)|$. (a) Compute $\Phi(n)$ and $\phi(n)$ for $1 \le n \le 12$. (b) Compute $\phi(p)$ for p prime. (c) Compute $\phi(p^e)$ for p prime and $e > 1$. (Exercise 1.150 shows how to compute $\phi(n)$ for any n.)

1.113. (a) How many 4-element subsets of $\{1, 2, \ldots, 11\}$ contain no two consecutive integers? (b) Given d, k, n, how many k-element subsets S of $\{1, 2, \ldots, n\}$ are such that any two distinct elements of S differ by at least d?

1.114. (a) How many anagrams of 'MISSISSIPPI' are there? (b) How many of these anagrams begin and end with P? (c) In how many of these anagrams are the two P's adjacent? (d) In how many of these anagrams are no two I's adjacent?

1.115. A *two-to-one function* is a function $f : X \to Y$ such that for every $y \in Y$, there exist *exactly two* elements $x_1, x_2 \in X$ with $f(x_1) = y = f(x_2)$. How many two-to-one functions are there from a $2n$-element set to an n-element set?

1.116. A *monomial* in N variables is a term of the form $x_1^{k_1} x_2^{k_2} \cdots x_N^{k_N}$, where each $k_i \ge 0$. The *degree* of this monomial is $k_1 + k_2 + \cdots + k_N$. How many monomials in N variables have degree (a) exactly d; (b) at most d?

1.117. How many multisets (of any size) can be formed from an n-letter alphabet if each letter can appear at most k times in the multiset?

1.118. Two fair dice are rolled. Find the probability that: (a) the same number appears on both dice; (b) the sum of the numbers rolled is 8; (c) the sum of the numbers rolled is divisible by 3; (d) the two numbers rolled differ by 1.

1.119. In blackjack, you have been dealt two cards from a shuffled 52-card deck: 9♡ and 6♣. Find the probability that drawing one more card will cause the sum of the three card values to go over 21. (Here, an ace counts as 1 and other face cards count as 10.)

1.120. Find the probability that a random 5-letter word: (a) has no repeated letters; (b) contains no vowels; (c) is a palindrome.

1.121. A company employs ten men (one of whom is Bob) and eight women (one of whom is Alice). A four-person committee is randomly chosen. Find the probability that the committee: (a) consists of all men; (b) consists of two men and two women; (c) does not have both Alice and Bob as members.

1.122. A fair coin is tossed ten times. (a) Find the probability of getting exactly seven heads. (b) Find the probability of getting at least two heads. (c) Find the probability of getting exactly seven heads, given that the number of heads was prime.

1.123. A fair die is tossed ten times. What is the probability that, in these ten tosses, 1 comes up 5 times, 3 comes up 2 times, and 6 comes up 3 times?

1.124. Ten balls are drawn (without replacement) from an urn containing 40 red, 30 blue, and 30 white balls. (a) What is the probability that no blue balls are drawn? (b) What is the probability of getting 4 red, 3 blue, and 3 white balls? (c) What is the probability that all ten balls have the same color? (d) Answer the same questions assuming the balls are drawn with replacement.

1.125. Urn A contains two red balls and three black balls. Urn B contains one red ball and four black balls. Urn C contains four red balls and one black ball. A ball is randomly chosen from each of the three urns. Find the probability that all three balls are the same color.

1.126. Consider the three urns from 1.125 (with five balls in each urn). An urn is selected at random, and then one ball is selected from that urn. What is the probability that: (a) the ball is black, given that urn B was chosen; (b) the ball is black; (c) urn B was chosen, given that the ball was black?

1.127. A fair coin is tossed three times. (a) Describe the sample space. (b) Consider the following events. A: second toss is tails; B: first and last tosses agree; C: all tosses are the same; D: the number of heads is odd. Describe each event as a subset of the sample space. (c) Which pairs of events from $\{A, B, C, D\}$ are independent? (d) Is the triple of events A, B, D independent? Explain.

1.128. Let the prime factorization of $n!$ be $p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$. Prove that $e_i = \sum_{k=1}^{\infty} \lfloor n/p_i^k \rfloor$. (The notation $\lfloor x \rfloor$ denotes the greatest integer not exceeding the real number x.) Hence determine the number of trailing zeroes in the decimal notation for 100!.

1.129. Find a bijection on Comp_n that maps compositions with k parts to compositions with $n + 1 - k$ parts for all k.

1.130. (a) How many numbers between one and one million contain the digit 7? (b) If one writes down the numbers from one to one million, how often will one write the digit 7? (c) What are the answers to (a) and (b) if 7 is replaced by 0?

1.131. A *relation from X to Y* is any subset of $X \times Y$. Suppose X has n elements and Y has k elements. (a) How many relations from X to Y are there? (b) How many relations R satisfy the following property: for each $y \in Y$, there exists at most one $x \in X$ with $(x, y) \in R$?

1.132. Suppose we play five-card poker using a 51-card deck in which the queen of spades has been removed. Compute the probabilities of the poker hands in Table 1.4 relative to this deck.

1.133. Suppose we play five-card poker using two identical decks mixed together. Compute the probabilities of the poker hands in Table 1.4 in this situation. Also compute the probability of a "five-of-a-kind" hand, which is a poker hand H such that $|V(H)| = 1$.

1.134. Consider a five-card poker hand dealt from a 52-card deck. (a) What is the probability that the hand contains only red cards (i.e., hearts and diamonds)? (b) What is the probability that the hand contains exactly two eights? (c) What is the probability that the hand contains only numerical cards (i.e., ace, jack, queen, and king may not appear)?

1.135. Consider a five-card poker hand dealt from a 52-card deck. (a) What is the probability that the hand is a flush, given that the hand contains no clubs? (b) What is the probability that the hand contains at least one card from each of the four suits? (c) What is the probability of getting a two-pair hand, given that at least two cards in the hand have the same value?

1.136. Let K be the event that a five-card poker hand contains the card $K\heartsuit$. Find the conditional probability of each event in Table 1.4, given K. Which of these events are independent of K?

1.137. Texas Hold 'em. In a popular version of poker, a player is dealt an *ordered* sequence of seven distinct cards from a 52-card deck. We model this situation using the sample space

$$S = \{(C_1, C_2, \ldots, C_7) : C_i \in \text{Deck},\ C_i \neq C_j \text{ for } i \neq j\}.$$

(The last five cards in this sequence are "community cards" shared with other players. In this exercise we concentrate on a single player, so we ignore this aspect of the game.) The player uses these seven cards to form the best possible five-card poker hand (cf. Table 1.4). For example, if we were dealt the hand

$$(4\heartsuit, 7\clubsuit, 3\diamondsuit, 9\clubsuit, 5\clubsuit, 6\clubsuit, Q\clubsuit),$$

we would have a flush (the five club cards) since this beats the straight (3,4,5,6,7 of various suits). (a) Compute $|S|$. (b) What is the probability of getting 4-of-a-kind? (c) What is the probability of getting a flush? (d) What is the probability of getting 4-of-a-kind, given $C_1 = 3\heartsuit$ and $C_2 = 3\spadesuit$? (e) What is the probability of getting a flush, given $C_1 = 5\diamondsuit$ and $C_2 = 9\diamondsuit$?

1.138. Prove that the following conditions are equivalent for any sets A and B: (a) $A \subseteq B$; (b) $A \cap B = A$; (c) $A \cup B = B$; (d) $A \sim B = \emptyset$.

1.139. Prove that if A and B are unequal nonempty sets, then $A \times B \neq B \times A$.

1.140. Use the binary union rule 1.4 to prove that for all finite sets X, Y, Z,

$$|X \cup Y \cup Z| = |X| + |Y| + |Z| - |X \cap Y| - |X \cap Z| - |Y \cap Z| + |X \cap Y \cap Z|.$$

1.141. (a) For fixed k, prove that $\lim_{n\to\infty} \frac{n!}{(n-k)!n^k} = 1$. (b) Give a probabilistic interpretation of this result.

1.142. Let $f : \mathcal{P}(X) \to \{0,1\}^n$ be the bijection in 1.38. Given two words $v, w \in \{0,1\}^n$, define words $v \wedge w$, $v \vee w$, and $\neg v$ by setting $(v \wedge w)_i = \min(v_i, w_i)$, $(v \vee w)_i = \max(v_i, w_i)$, and $(\neg v)_i = 1 - v_i$ for all $i \leq n$. Prove that for all $S, T \subseteq X$, $f(S \cap T) = f(S) \wedge f(T)$, $f(S \cup T) = f(S) \vee f(T)$, $f(X \sim S) = \neg f(S)$, $f(\emptyset) = 00\cdots0$, and $f(X) = 11\cdots1$.

1.143. Let A, B, C be events in a probability space S. Assume A and C are independent, and B and C are independent. (a) Give an example where $A \cup B$ and C are not independent. (b) Prove that $A \cup B$ and C are independent if A and B are disjoint. (c) Must $A \cap B$ and C be independent? Explain.

1.144. Properties of Injections. Prove the following statements about injective functions. (a) If $f : X \to Y$ and $g : Y \to Z$ are injective, then $g \circ f$ is injective. (b) If $g \circ f$ is injective, then f is injective but g may not be. (c) $f : X \to Y$ is injective iff for all W and all $g, h : W \to X$, $f \circ g = f \circ h$ implies $g = h$.

1.145. Properties of Surjections. Prove the following statements about surjective functions. (a) If $f : X \to Y$ and $g : Y \to Z$ are surjective, then $g \circ f$ is surjective. (b) If $g \circ f$ is surjective, then g is surjective but f may not be. (c) $f : X \to Y$ is surjective iff for all Z and all $g, h : Y \to Z$, $g \circ f = h \circ f$ implies $g = h$.

1.146. Sorting by Comparisons. Consider a game in which player 1 picks a permutation w of n letters, and player 2 must determine w by asking player 1 a sequence of yes/no questions. (Player 2 can choose later questions in the sequence based on the answers to earlier questions.) Let $K(n)$ be the minimum number such that, no matter what w player 1 chooses, player 2 can correctly identify w after at most $K(n)$ questions. (a) Prove that $(n/2)\log_2(n/2) \leq \lceil \log_2(n!) \rceil \leq K(n)$. (b) Prove that $K(n) = \lceil \log_2(n!) \rceil$ for $n \leq 5$. (c) Prove that (b) still holds if we restrict player 2 to ask only questions of the form "is $w_i < w_j$?" at each stage. (d) What does (a) imply about the length of time needed to sort n distinct elements using an algorithm that makes decisions by comparing two data elements at a time?

1.147. (a) You are given twelve seemingly identical coins and a balance scale. One coin is counterfeit and is either lighter or heavier than the others. Describe a strategy that can be used to identify which coin is fake in only three weighings. (b) If there are thirteen coins, can the fake coin always be found in three weighings? Justify your answer. (c) If there are N coins (one of which is fake), derive a lower bound for the number of weighings required to find the fake coin.

1.148. Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}^+$ by $f(a, b) = 2^a(2b + 1)$. Prove that f is a bijection.

1.149. Define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $f(a, b) = ((a + b)^2 + 3a + b)/2$. Prove that f is a bijection.

1.150. Chinese Remainder Theorem. In this exercise, we write "a mod k" to denote the unique integer b in the range $\{1, 2, \ldots, k\}$ such that k divides $(a - b)$. Suppose m and n are fixed positive integers. Define a map

$$f : \{1, 2, \ldots, mn\} \to \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \text{ by setting } f(z) = (z \bmod m, z \bmod n).$$

(a) Show that $f(z) = f(w)$ iff $\text{lcm}(m, n)$ divides $z - w$. (b) Show that f is injective iff $\gcd(m, n) = 1$. (c) Deduce that f is a bijection iff $\gcd(m, n) = 1$. (d) Prove that for $\gcd(m, n) = 1$, f maps $\Phi(mn)$ bijectively onto $\Phi(m) \times \Phi(n)$, and hence $\phi(mn) = \phi(m)\phi(n)$. (See 1.112 for the definition of Φ and ϕ.) (e) Suppose n has prime factorization $p_1^{e_1} \cdots p_k^{e_k}$. Prove that $\phi(n) = n \prod_{i=1}^{k} (1 - 1/p_i)$.

1.151. Bijective Product Rule. For any positive integers m, n, define

$$g : \{0, 1, \ldots, m - 1\} \times \{0, 1, \ldots, n - 1\} \to \{0, 1, \ldots, mn - 1\}$$

by setting $g(i, j) = ni + j$. Carefully prove that g is a bijection.

1.152. Bijective Laws of Algebra. (a) For all sets X, Y, Z, prove that $X \cup Y = Y \cup X$, $(X \cup Y) \cup Z = X \cup (Y \cup Z)$, and $X \cup \emptyset = X = \emptyset \cup X$. (b) For all sets X, Y, Z, define bijections $f : X \times Y \to Y \times X$, $g : (X \times Y) \times Z \to X \times (Y \times Z)$, and (for Y, Z disjoint) $h : X \times (Y \cup Z) \to (X \times Y) \cup (X \times Z)$. (c) Use (a), (b), and counting rules to deduce the algebraic laws $x + y = y + x$, $(x + y) + z = x + (y + z)$, $x + 0 = x = 0 + x$, $xy = yx$, $(xy)z = x(yz)$, and $x(y + z) = xy + xz$, valid for all integers $x, y, z \geq 0$.

1.153. Bijective Laws of Exponents. (a) If X, Y, Z are sets with $Y \cap Z = \emptyset$, define a bijection from ${}^{Y \cup Z}X$ to ${}^{Y}X \times {}^{Z}X$. (b) If X, Y, Z are any sets, define a bijection from ${}^{Z}({}^{Y}X)$ to ${}^{Y \times Z}X$. (c) By specializing to finite sets, deduce the laws of exponents $x^{y+z} = x^y x^z$ and $(x^y)^z = x^{yz}$ for all integers $x, y, z \geq 0$.

1.154. Let X be any set (possibly infinite). Prove that there exists an injection $g : X \to \mathcal{P}(X)$, but there exists no surjection $f : X \to \mathcal{P}(X)$. Conclude that $|X| < |\mathcal{P}(X)|$, and in particular $n < 2^n$ for all $n \geq 0$.

1.155. Show that the set of functions ${}^{\mathbb{N}}\{0,1\}$ (which can be viewed as the set of infinite sequences of zeroes and ones) is uncountably infinite.

1.156. Suppose X and Y are sets (possibly infinite), $f : X \to Y$ is any function, and $g : Y \to X$ is an injective function. (a) Show that there exist sets A, B, C, D such that X is the disjoint union of A and B, Y is the disjoint union of C and D, $C = f[A] = \{f(x) : x \in A\}$, and $B = g[D] = \{g(y) : y \in D\}$. (Let $Z = X \sim g[Y]$ and $h = g \circ f$; then let A be the intersection of all subsets U of X such that $Z \cup h[U] \subseteq U$.) (b) Deduce the Schröder-Bernstein Theorem from (a).

1.157. A sample space S consists of 25 equally likely outcomes. Suppose we randomly choose an ordered pair (A, B) of events in S. (a) Find the probability that A and B are disjoint. (b) Find the probability that A and B are independent events.

Notes

General treatments of combinatorics may be found in the textbooks [1, 10, 13, 16, 21, 23, 26, 60, 113, 115, 127, 131, 134]. For elementary accounts of probability theory, see [68, 93]. Two advanced probability texts that include measure theory are [11, 30]. More information on the theory of cardinality for infinite sets may be found in [66, 95].