Consider the diagonal starting at d_{11}. For each $j \geq 1$, choose e_j different from d_{jj}, 0, and 9. Then $y = .e_1e_2\cdots$ represents a real number in $[0,1)$ that is different from each x_i. But we assumed that the decimal representation of every real number in $[0,1)$ appears in the above array, and we have a contradiction. Our assumption that $[0,1)$ is countable leads to a contradiction.

EXERCISES 2.3

The last problem requires the use of the well-ordering property of the natural numbers N, which states that if $A \subset N$ and $A \neq \emptyset$, then A has a least element.

1. If $f = \{(x,y) : x \in R, y \in R, y = x^2, -1 \leq x \leq 1\}$, determine its domain and range.
2. If in the customary notation of the calculus $f(x) = \sqrt{1-x^4}$, describe f as a subset of $R \times R$ and determine its domain and range.
3. If in the customary notation of the calculus $f(x) = 1/\sqrt{1-x^2}$, describe f as a subset of $R \times R$ and determine its domain and range.
4. If $q \in N$ and $X = \{p/q : p \in N\}$, show that X is countable.
5. Let $X_1, X_2, \ldots, X_m$ be a finite sequence of countably infinite sets. Show that $X = X_1 \cup \cdots \cup X_m$ is countable.
6. Show that the set X of all infinite sequences of 0's and 1's is uncountable.
7. Show that $N \times N = \{(m,n) : m \in N, n \in N\}$ is countable by considering the collection of finite sets $A_k = \{(m,n) : m + n = k\}$.
8. Let A and B be countable sets. Show that $A \times B$ is countable.
9. Which of the following sets are countable?
   (a) The set of circles in the plane having centers with rational coordinates and rational radii.
   (b) The set of all polynomials $P(x) = a_nx^n + \cdots + a_1x + a_0$ having integer coefficients.
   (c) The set of all intervals $(a,b) \subset R$ having rational endpoints.
10. Let $X_1, X_2, \ldots$ be an infinite sequence of countably infinite sets. Show that $X = \cup_{n=1}^{\infty} X_n$ is countable.

2.4 AXIOMS

If A and B are disjoint collections of outcomes, we have seen that

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

More generally, if $A_1, \ldots, A_n$ are disjoint, it follows from the empirical law that

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j). \tag{2.4}$$

If the total number of outcomes is finite, there is no more to be said in regard to Equation 2.4. But what if $\{A_j\}$ is an infinite sequence of disjoint collections? An experiment with an infinite number of outcomes was discussed in Chapter 1; namely, flipping a coin until a head appears for the first time. In this case, it is possible to have an infinite sequence $\{A_j\}$ of disjoint collections of outcomes, and so it makes sense to ask if

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j). \tag{2.5}$$

This is a moot question for all the examples with finitely many outcomes and has an affirmative answer for the single model just described. Since Equation 2.5 is compatible with every example we have considered, can we assume that Equation 2.5, in addition to Equations 1.1, 1.2, and 1.3, is valid in a general model for probability theory? We can, but we cannot have everything we would like to have. It turns out that we cannot assume that Equation 2.5 is valid for all sequences $\{A_j\}$ of disjoint collections of outcomes and at the same time assume that $P(A)$ is meaningful for all possible A. We must give up one of the two assumptions. We will give up the latter, and so $P(A)$ may not be meaningful for some A.

Let Ω denote the collection of all outcomes for a given experiment. The following definitions are needed to limit the A for which $P(A)$ will be defined.

Definition 2.1 *A collection $\mathscr{A}$ of subsets of Ω is an* algebra *if*

1. $A, B \in \mathscr{A}$ *implies* $A \cup B \in \mathscr{A}$.
2. $A \in \mathscr{A}$ *implies* $A^c \in \mathscr{A}$.
3. $\Omega \in \mathscr{A}$. ■

That is, $\mathscr{A}$ is an algebra of subsets of Ω if it is closed under the operations of union and complementation and $\Omega \in \mathscr{A}$. A mathematical induction argument can be used to show that an algebra $\mathscr{A}$ is closed under finite unions; i.e., if $A_1, A_2, \ldots, A_n \in \mathscr{A}$, then $\cup_{j=1}^{n} A_j \in \mathscr{A}$.

The important thing to remember about algebras is that by starting with a finite number of elements of the algebra and performing a finite number of union, intersection, and complementation operations on them, the result is still in the algebra.

EXAMPLE 2.10 Let $\mathscr{A}$ be an algebra of subsets of Ω and let A_1, A_2, A_3, A_4 be elements of the algebra with some or all having nonempty intersections. If we need to restructure the union $\cup_{i=1}^{4} A_i$ into a union of disjoint sets in $\mathscr{A}$, we could let $B_1 = A_1, B_2 = A_2 \cap A_1^c, B_3 = A_3 \cap (A_1 \cup A_2)^c$, and $B_4 = A_4 \cap (A_1 \cup A_2 \cup A_3)^c$. Then $B_j \subset A_j$, $1 \leq j \leq 4$, the B_j are disjoint, and $\cup A_j = \cup B_j$. ■

We need to postulate more if we want to deal with infinite sequences.

Definition 2.2 *A collection $\mathcal{F}$ of subsets of Ω is a* σ-algebra *if*

1. *$\mathcal{F}$ is an algebra.*
2. *If $\{A_j\}$ is an infinite sequence in $\mathcal{F}$, then $\cup A_j \in \mathcal{F}$.* ■

The important thing to remember about σ-algebras is that by starting with a sequence of elements of the σ-algebra and performing countably many union, intersection, and complementation operations on them, the result is still in the σ-algebra.

EXAMPLE 2.11 Let Ω be a finite set of outcomes and let $\mathcal{A}$ be the collection of all subsets of Ω. Then $\mathcal{A}$ is an algebra. ■

EXAMPLE 2.12 Let Ω be any set and let $\mathcal{F}$ be the collection of all subsets of Ω. Then $\mathcal{F}$ is a σ-algebra. Clearly, $\Omega \in \mathcal{F}$ since $\Omega \subset \Omega$. If $\{A_n\}$ is a finite or infinite sequence in $\mathcal{F}$, then $\cup A_n$ is a subset of Ω and therefore is in $\mathcal{F}$. If $A \in \mathcal{F}$, then $A^c \subset \Omega$ and $A^c \in \mathcal{F}$. ■

If $\mathcal{B}$ is any collection of subsets of a set Ω, then there is a "smallest σ-algebra," denoted by $\sigma(\mathcal{B})$, that contains $\mathcal{B}$. In discussing probability, we began with objects ω that were used to form collections A that have now been used to form σ-algebras $\mathcal{F}$. This process has taken us through three hierarchical levels of set theory, and to prove the result just stated would require going to a fourth hierarchical level. This fourth level is left to more advanced texts. For the time being, we have all the concepts needed to describe a general probability model. Ω will be a fixed collection of outcomes.

Definition 2.3 *A* probability space *is a triple $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is a nonempty σ-algebra of subsets of Ω and P is a mapping from $\mathcal{F}$ to R satisfying*

1. *$P(\Omega) = 1$.*
2. *$0 \leq P(A) \leq 1$ for all $A \in \mathcal{F}$.*
3. *If $\{A_j\}$ is a finite or infinite disjoint sequence in $\mathcal{F}$, then*

$$P(\cup A_j) = \sum P(A_j). \quad ■$$

All the simple games of chance described in Chapter 1 for which Ω is finite, $\mathcal{F}$ is the collection of all subsets of Ω, and P is defined as described there result in probability spaces $(\Omega, \mathcal{F}, P)$.

Definition 2.4 *If $(\Omega, \mathcal{F}, P)$ is a probability space, elements of $\mathcal{F}$ are called* events. ■

We return to a model discussed in Section 1.5.

EXAMPLE 2.13 Let $\Omega = \{\omega_1, \omega_2, \ldots\}$ be countably infinite, $\mathcal{F}$ the σ-algebra of all subsets of Ω, and $p(\omega_i)$ a weight function as defined in Section 1.5. Define $P(A), A \in \mathcal{F}$, as in Section 1.5. Then $(\Omega, \mathcal{F}, P)$ is a probability space. To show this, we need only verify Item 3 of Definition 2.3. Let $\{A_j\}$ be a sequence of disjoint events in $\mathcal{F}$ and let $A = \cup A_j$. Suppose $A = \{\omega_{i_1}, \omega_{i_2}, \ldots\}$. The fact that the series $\sum_j p(\omega_{i_j})$ is a convergent series with sum $P(A)$ means that the terms of the series can be rearranged without affecting the sum of the series. This fact about absolutely convergent series is proved or at least discussed in most calculus books. We rearrange the terms of the series so that the terms $p(\omega_{i_j})$ with $\omega_{i_j} \in A_1$ come first, then the terms $p(\omega_{i_j})$ with $\omega_{i_j} \in A_2$ second, and so on, to obtain

$$\begin{aligned} P(A) &= \sum_j p(\omega_{i_j}) \\ &= \sum_{\omega_{i_j} \in A_1} p(\omega_{i_j}) + \sum_{\omega_{i_j} \in A_2} p(\omega_{i_j}) + \cdots \\ &= P(A_1) + P(A_2) + \cdots. \end{aligned}$$

Therefore,

$$P(\cup A_j) = \sum P(A_j),$$

and $(\Omega, \mathcal{F}, P)$ is a probability space. ■

We have previously encountered the following situation. Suppose a coin is flipped until a head appears for the first time with a maximum of n flips. Suppose n is a large positive integer. Let A be the event "the experiment terminates on the fifth flip." We can think of this experiment continuing through all n flips of the coin and simply ignoring what happens after the fifth flip. If $\omega \in A$, then the first four letters of ω are T, the fifth is H, and there are two choices for each of the remaining $n - 5$ letters. Thus, $|A| = 2^{n-5}$,

$$P(A) = \frac{2^{n-5}}{2^n} = \frac{1}{2^5},$$

and it appears that $P(A)$ does not depend upon n at all! This computation is based on Pascal's reasoning in which we think of the coin as continuing to be flipped beyond the fifth flip and simply ignoring everything beyond the fifth flip. Why bother to mention the number n at all? If we eliminate mentioning n at all, then we are confronted with a conceptual experiment in which a coin is continually flipped. We can, in fact, construct a probability space $(\Omega, \mathcal{F}, P)$ with Ω consisting of outcomes ω that are words of infinite length using the alphabet T, H, and probabilities for events such as A described above are calculated by fixing some large n. We will consider a more general model.

In the following example, Ω will denote the set of all infinite sequences $\{x_i\}_{i=1}^{\infty}$ where each x_i is a 1 or a 0. We can think of 1 and 0 as an encoding of H and T or S and F, respectively, where S stands for success and F stands for failure. This Ω is uncountable; i.e., not countable. (See Exercise 2.3.6.) Therefore, none of the models we have discussed pertain to Ω. The model depends upon a parameter p, called the *probability of success*, with $0 < p < 1$. The number $q = 1 - p$ is called the *probability of failure*. Whenever p and q appear, these conditions on p and q will be taken for granted without comment.

EXAMPLE 2.14 (Infinite Sequence of Bernoulli Trials) Fix $0 < p < 1$ and let $q = 1 - p$. Let Ω be the set described above and let $\mathcal{F}_0$ be the collection of subsets A of Ω of the form

$$A = \{\omega : \omega = \{x_i\}_{i=1}^{\infty}, x_{i_1} = \delta_1, \ldots, x_{i_n} = \delta_n\}, \tag{2.6}$$

where n is any positive integer, $1 \leq i_1 < i_2 < \cdots < i_n$, and each δ_i is a 0 or a 1. We think of the x_i as the results of successive trials. For $\mathcal{F}$ we take $\sigma(\mathcal{F}_0)$, the smallest σ-algebra containing $\mathcal{F}_0$. As an illustration of how probabilities are to be computed, consider the event "1 on the second trial, 0 on the fourth trial, and 1 on the eighth trial"; i.e., the event

$$A = \{\omega : \omega = \{x_i\}_{i=1}^{\infty}, x_2 = 1, x_4 = 0, x_8 = 1\}.$$

Then

$$P(A) = p^2 q = p^2(1-p).$$

Note that

$$P(A) = p^{x_2+x_4+x_8} q^{3-(x_2+x_4+x_8)}.$$

For an event A of the type described in Equation 2.6, its probability is defined to be

$$P(A) = p^{(\sum_{j=1}^{n} \delta_j)} q^{(n - \sum_{j=1}^{n} \delta_j)}. \tag{2.7}$$

Note that $\sum_{j=1}^{n} \delta_j$ is the number of 1's in the trials numbered $i_1, i_2, \ldots, i_n$ and $n - \sum_{j=1}^{n} \delta_j$ is the number of 0's in the same trials. It cannot be done here, but it is possible to extend the definition of P so that $P(A)$ is defined for all $A \in \mathcal{F}$. Any set of outcomes that can be expressed in terms of events placing restrictions on only a finite number of trials will also be an event. Consider the event A described by "a 1 eventually appears in the outcome ω"; i.e.,

$$A = \{\omega : \omega = \{x_j\}_{j=1}^{\infty}, \sum_{j=1}^{\infty} x_j \geq 1\}.$$

If we let A_j be the event "1 appears for the first time on the jth trial," then

$$A = \bigcup A_j \in \mathcal{F}$$

for the reasons just cited, and the A_j are disjoint. Since $P(A_j) = q^{j-1}p$ and the latter is the general term of a geometric series,

$$P(A) = \sum_{j=1}^{\infty} q^{j-1}p = p\sum_{j=1}^{\infty} q^{j-1} = \frac{p}{1-q} = 1. \blacksquare$$

There is no reason to limit the number of results of each trial to just the 0 and 1 of the preceding example. We can allow the possibility that each trial results in one of k possibilities $r_1, r_2, \ldots, r_k$ with associated weights $p_1, p_2, \ldots, p_k$, where $0 \le p_i \le 1, i = 1, \ldots, k$. Suppose $n \ge 1$, $1 \le i_1 < i_2 < \cdots < i_n$, and $\delta_1, \ldots, \delta_n \in \{r_1, \ldots, r_k\}$. For the event

$$A = \{\omega : \omega = \{x_j\}_{j=1}^{\infty}, x_{i_1} = \delta_1, \ldots, x_{i_n} = \delta_n\},$$

we can define

$$P(A) = p_1^{m_1} \times p_2^{m_2} \times \cdots \times p_k^{m_k},$$

where m_i is the number of trials resulting in r_i, $1 \le i \le k$. This model is applicable to an unending sequence of throws of a die where the result of each throw is one of the integers 1, 2, 3, 4, 5, 6 with weight 1/6 associated with each.

EXERCISES 2.4

1. Using mathematical induction, write out a formal proof that an algebra $\mathcal{A}$ is closed under finite unions; i.e., for every $n \ge 1$,

$$A_1, \ldots, A_n \in \mathcal{A} \text{ implies } \bigcup_{j=1}^{n} A_j \in \mathcal{A}.$$

2. If $\mathcal{A}$ is an algebra of subsets of Ω, show that $\mathcal{A}$ is closed under finite intersections.
3. Let $\mathcal{F}$ be a σ-algebra of subsets of Ω. Show that $\mathcal{F}$ is closed under countable intersections; i.e., if $\{A_j\}$ is a finite or infinite sequence in $\mathcal{F}$, then $\cap A_j \in \mathcal{F}$.
4. Let Ω be an uncountable set and let $\mathcal{F}$ be the collection of subsets A of Ω such that either A is countable or A^c is countable. Show that $\mathcal{F}$ is a σ-algebra.
5. Consider an infinite sequence of Bernoulli trials with probability of success p. What is the probability that a success (or 1) will occur for the first time on an even-numbered trial?

6. An experiment consists of tossing a pair of dice until a score of 8 is observed for the first time, whereupon the experiment is terminated. What is the probability that it will terminate on an odd number of tosses of the dice?
7. A bowl contains w white chips, r red chips, and b black chips. Chips are successively selected at random from the bowl with replacement. What is the probability that a white chip will appear before a black chip?
8. If $\mathcal{F}$ is a σ-algebra of subsets of Ω and $\{A_j\}_{j=1}^{\infty}$ is an increasing sequence in $\mathcal{F}$ (i.e., $A_n \subset A_{n+1}$ for all $n \geq 1$), show that there is a disjoint sequence $\{B_j\}_{j=1}^{\infty}$ in $\mathcal{F}$ such that $\cup_{j=1}^{\infty} A_j = \cup_{j=1}^{\infty} B_j$.
9. If $\mathcal{F}$ is a σ-algebra of subsets of Ω, $\{A_j\}_{j=1}^{\infty}$ is a decreasing sequence in $\mathcal{F}$ (i.e., $A_{n+1} \subset A_n$ for all $n \geq 1$), and $A = \cap_{j=1}^{\infty} A_j$, show that there is a decreasing sequence $\{B_j\}_{j=1}^{\infty}$ in $\mathcal{F}$ such that $A_n = A \cup B_n, A \cap B_n = \varnothing$ for all $n \geq 1$ and $\cap_{j=1}^{\infty} B_j = \varnothing$.

2.5 PROPERTIES OF PROBABILITY FUNCTIONS

Throughout this section, $(\Omega, \mathcal{F}, P)$ will be a fixed probability space as described in Definition 2.3. We will now deduce several properties of the probability function P from the axioms listed in Definition 2.3.

Consider two events $A, B \in \mathcal{F}$. Since $\Omega = A \cup A^c$, intersecting both sides of this equation with B we obtain

$$B = B \cap \Omega = B \cap (A \cup A^c) = (B \cap A) \cup (B \cap A^c);$$

i.e., we can decompose B into two parts according to whether or not an outcome in B is in A or not in A. Since $B \cap A$ and $B \cap A^c$ are disjoint, by Item 3 of Definition 2.3,

$$P(B) = P(B \cap A) + P(B \cap A^c) \text{ for all } A, B \in \mathcal{F}. \tag{2.8}$$

If we put $B = \Omega$ and use Item 1 of Definition 2.3, then $1 = P(\Omega) = P(\Omega \cap A) + P(\Omega \cap A^c)$, so that

$$P(A^c) = 1 - P(A) \text{ for all } A \in \mathcal{F}. \tag{2.9}$$

In particular, $P(\varnothing) = 1 - P(\Omega) = 0$.

EXAMPLE 2.15 Consider n flips of a coin and let A be the event "the outcome ω has one or more heads." Calculating $P(A)$ directly is complicated, but calculating $P(A^c)$ is easily done because A^c consists of just one outcome having a label of n T's. Since each outcome has probability $1/2^n$, $P(A) = 1 - P(A^c) = 1 - (1/2^n)$. ■

Suppose now that $A, B \in \mathcal{F}, A \subset B$. Then $A \cap B = A$, and so Equation 2.8 becomes $P(B) = P(A) + P(B \cap A^c)$. Since $P(B \cap A^c) \geq 0$ by Item 2 of Definition 2.3,

$$P(A) \leq P(B) \text{ whenever } A, B \in \mathcal{F}, A \subset B. \tag{2.10}$$

If A and B are any two events, then $A \cup B = (A \cap B^c) \cup (A \cap B) \cup (A^c \cap B)$; i.e., $A \cup B$ can be split into three parts: (1) those outcomes in A but not in B, (2) those outcomes in both A and B, and (3) those outcomes in B but not in A. Thus,

$$P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B).$$

Applying Equation 2.8 to the first and third terms on the right and simplifying, we obtain

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{2.11}$$

EXAMPLE 2.16 A card is selected at random from a deck of 52 cards. What is the probability that the card selected will be a king or a spade? Let A be the event "the outcome ω is a king" and let B be the event "the outcome ω is a spade." The required probability is $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 1/13 + 1/4 - 1/52 = 4/13$. ■

If A, B, and C are any three events, then

$$\begin{aligned} P(A \cup B \cup C) = &\ P(A) + P(B) + P(C) - P(A \cap B) \\ &- P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \end{aligned}$$

More generally, if $A_1, \ldots, A_N$ are any events, then

$$\begin{aligned} P(A_1 \cup \cdots \cup A_N) &= \sum_{i=1}^{N} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq N} P(A_{i_1} \cap A_{i_2}) \\ &+ \sum_{1 \leq i_1 < i_2 < i_3 \leq N} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) \\ &- \cdots \pm P(A_1 \cap A_2 \cap \cdots \cap A_N) \\ &= \sum_{r=1}^{N} (-1)^{r-1} \sum_{1 \leq i_1 < \cdots < i_r \leq N} P(A_{i_1} \cap \cdots \cap A_{i_r}). \end{aligned} \tag{2.12}$$

This result goes by the name *inclusion/exclusion principle* and can be proved using mathematical induction.

Returning to Equation 2.11,

$$P(A \cup B) \leq P(A) + P(B) \text{ for all } A, B \in \mathcal{F}$$

since $P(A \cap B) \geq 0$ by Item 2 of Definition 2.3. This inequality is a special case of a more general inequality whose proof will require the following lemma.

Lemma 2.5.1 *If $\{A_j\}_{j=1}^{\infty}$ is a sequence of events, then there is a disjoint sequence $\{B_j\}_{j=1}^{\infty}$ of events such that $B_j \subset A_j$ for all $j \geq 1$, $\cup_{j=1}^{n} A_j = \cup_{j=1}^{n} B_j$ for all $n \geq 1$, and $\cup B_j = \cup A_j$.*

PROOF: Let $B_1 = A_1$ and $B_j = A_j \cap (\cup_{i=1}^{j-1} A_i)^c$ for $j \geq 2$. Clearly, $B_j \subset A_j$ for all $j \geq 1$. For $1 \leq i \leq j-1$,

$$B_i \cap B_j \subset A_i \cap B_j \subset \left(\bigcup_{i=1}^{j-1} A_i \right) \cap B_j = \varnothing.$$

Thus, $B_i \cap B_j = \varnothing$ whenever $1 \leq i \leq j-1$, $j \geq 1$. This means that the B_j are disjoint. Clearly, $\cup_{j=1}^{n} B_j \subset \cup_{j=1}^{n} A_j$. Suppose $\omega \in \cup_{j=1}^{n} A_j$. Then there is a smallest integer $k \leq n$ such that $\omega \in A_k$. Thus, $\omega \in A_k \cap (\cup_{i=1}^{k-1} A_i)^c = B_k \subset \cup_{k=1}^{n} B_k$, and it follows that $\cup_{j=1}^{n} A_j \subset \cup_{j=1}^{n} B_j$ and therefore that the two are equal. The proof of the last assertion is essentially the same. ■

Theorem 2.5.2 (Boole's Inequality) *If $\{A_j\}$ is any sequence of events, then $P(\cup A_j) \leq \sum P(A_j)$.*

PROOF: By Lemma 2.5.1, there is a disjoint sequence of events $\{B_j\}$ such that $B_j \subset A_j$, $j \geq 1$, and $\cup B_j = \cup A_j$. By Inequality 2.10, $P(B_j) \leq P(A_j)$, $j \geq 1$. Since the B_j are disjoint,

$$P(\cup A_j) = P(\cup B_j) = \sum P(B_j) \leq \sum P(A_j). \quad ■$$

Theorem 2.5.3 *Let $\{A_j\}_{j=1}^{\infty}$ be a sequence of events.*

- *(i) If $A_1 \subset A_2 \subset \cdots$ is an increasing sequence and $A = \cup_{j=1}^{\infty} A_j$, then $P(A) = \lim_{n\to\infty} P(A_n)$.*
- *(ii) If $A_1 \supset A_2 \supset \cdots$ is a decreasing sequence and $A = \cap_{j=1}^{\infty} A_j$, then $P(A) = \lim_{n\to\infty} P(A_n)$.*

PROOF: (*i*) Let $\{A_j\}_{j=1}^{\infty}$ be an increasing sequence of events and let $A = \cup_{j=1}^{\infty} A_j \in \mathcal{F}$. Note that $\cup_{j=1}^{n} A_j = A_n$. By Lemma 2.5.1, there is a disjoint sequence of events $\{B_j\}_{j=1}^{\infty}$ such that $B_j \subset A_j$, $j \geq 1$, $\cup_{j=1}^{n} A_j = \cup_{j=1}^{n} B_j$, and $\cup A_j = \cup B_j$. By Item 3 of Definition 2.3,

$$\begin{aligned} P(A) &= P\left(\bigcup_{j=1}^{\infty} A_j\right) = P\left(\bigcup_{j=1}^{\infty} B_j\right) = \sum_{j=1}^{\infty} P(B_j) \\ &= \lim_{n\to\infty} \sum_{j=1}^{n} P(B_j) = \lim_{n\to\infty} P\left(\bigcup_{j=1}^{n} B_j\right) = \lim_{n\to\infty} P(A_n). \end{aligned}$$

(*ii*) Let $\{A_j\}_{j=1}^{\infty}$ be a decreasing sequence of events and let $A = \bigcap_{j=1}^{\infty} A_j \in \mathcal{F}$. Then $\{A_j^c\}_{j=1}^{\infty}$ is an increasing sequence of events, and $A^c = (\bigcap_{j=1}^{\infty} A_j)^c = \bigcup_{j=1}^{\infty} A_j^c$. By the first part of the proof,

$$\begin{aligned} 1 - P(A) = P(A^c) &= \lim_{n\to\infty} P(A_n^c) \\ &= \lim_{n\to\infty} (1 - P(A_n)) = 1 - \lim_{n\to\infty} P(A_n), \end{aligned}$$

and so $P(A) = \lim_{n\to\infty} P(A_n)$. ■

EXERCISES 2.5

1. In manufacturing brass cylindrical sleeves, 5 percent are defective because the outer diameter is too small and 3 percent are defective because the inner diameter is too large. What is the best you can say about the probability that a sleeve selected at random from a lot will be defective?
2. If A, B, and C are any three events, show that $P(A \cup B \cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$.
3. Consider three events A, B, C for which $P(A) = 1/3$, $P(B) = 1/4$, $P(C) = 1/2$, $P(A \cap B) = 1/8$, $P(A \cap C) = 1/8$, $P(B \cap C) = 3/16$, and $P(A \cap B \cap C) = 1/32$. Calculate $P(A \cup B \cup C)$.
4. An integer is chosen at random between 0000 and 9999. (a) Use the inclusion/exclusion principle to calculate the probability that at least one 1 will appear in the number. (b) Calculate the same probability assuming that the experiment is that of four Bernoulli trials.
5. Show that the probability that one and only one of the events A and B will occur is
$$P(A) + P(B) - 2P(A \cap B).$$
6. The mid–seventeenth century gambler Chavalier de Méré thought that the probability of getting at least one ace with the throw of four dice is equal to the probability of getting at least one double ace in 24 throws of two dice. Was de Méré correct?
7. Consider an infinite sequence of Bernoulli trials with probability of success p. If ω_0 is any outcome, show that $P(\{\omega_0\}) = 0$. (Note: There is no significance to the fact that each outcome has probability 0 whereas the aggregate of all outcomes has probability 1! After all, points in the interval $[0, 1]$ have zero length, but the aggregate $[0, 1]$ has length 1.)
8. If $P(A) = .8$ and $P(B) = .75$, show that $P(A \cap B) \geq .55$. More generally, show that if A and B are any two events, then
$$\min(P(A), P(B)) \geq P(A \cap B) \geq P(A) + P(B) - 1.$$