(b) If C is the subset $\{(x, y, z) : 0 < x = y = z < 1\}$, compute $Q(C)$.

1.2.15. Let C denote the set $\{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$. Using spherical coordinates, evaluate

$$Q(C) = \int\int\int_C \sqrt{x^2 + y^2 + z^2}\, dxdydz.$$

1.2.16. To join a certain club, a person must be either a statistician or a mathematician or both. Of the 25 members in this club, 19 are statisticians and 16 are mathematicians. How many persons in the club are both a statistician and a mathematician?

1.2.17. After a hard-fought football game, it was reported that, of the 11 starting players, 8 hurt a hip, 6 hurt an arm, 5 hurt a knee, 3 hurt both a hip and an arm, 2 hurt both a hip and a knee, 1 hurt both an arm and a knee, and no one hurt all three. Comment on the accuracy of the report.

1.3 The Probability Set Function

Given an experiment, let $\mathcal{C}$ denote the sample space of all possible outcomes. As discussed in Section 1.1, we are interested in assigning probabilities to events, i.e., subsets of $\mathcal{C}$. What should be our collection of events? If $\mathcal{C}$ is a finite set, then we could take the set of all subsets as this collection. For infinite sample spaces, though, with assignment of probabilities in mind, this poses mathematical technicalities which are better left to a course in probability theory. We assume that in all cases, the collection of events is sufficiently rich to include all possible events of interest and is closed under complements and countable unions of these events. Using DeMorgan's Laws, Example 1.2.17, the collection is then also closed under countable intersections. We denote this collection of events by $\mathcal{B}$. Technically, such a collection of events is called a **σ-field** of subsets.

Now that we have a sample space, $\mathcal{C}$, and our collection of events, $\mathcal{B}$, we can define the third component in our probability space, namely a probability set function. In order to motivate its definition, we consider the relative frequency approach to probability.

Remark 1.3.1. The definition of probability consists of three axioms which we motivate by the following three intuitive properties of relative frequency. Let $\mathcal{C}$ be a sample space and let $C \subset \mathcal{C}$. Suppose we repeat the experiment N times. Then the relative frequency of C is $f_C = \#\{C\}/N$, where $\#\{C\}$ denotes the number of times C occurred in the N repetitions. Note that $f_C \geq 0$ and $f_{\mathcal{C}} = 1$. These are the first two properties. For the third, suppose that C_1 and C_2 are disjoint events. Then $f_{C_1 \cup C_2} = f_{C_1} + f_{C_2}$. These three properties of relative frequencies form the axioms of a probability, except that the third axiom is in terms of countable unions. As with the axioms of probability, the readers should check that the theorems we prove below about probabilities agree with their intuition of relative frequency. ■

Definition 1.3.1 (Probability). *Let $\mathcal{C}$ be a sample space and let $\mathcal{B}$ be the set of events. Let P be a real-valued function defined on $\mathcal{B}$. Then P is a* **probability set function** *if P satisfies the following three conditions:*

1. *$P(C) \geq 0$, for all $C \in \mathcal{B}$.*
2. *$P(\mathcal{C}) = 1$.*
3. *If $\{C_n\}$ is a sequence of events in $\mathcal{B}$ and $C_m \cap C_n = \phi$ for all $m \neq n$, then*

$$P\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} P(C_n).$$

A collection of events whose members are pairwise disjoint, as in (3), is said to be a **mutually exclusive** collection. The collection is further said to be **exhaustive** if the union of its events is the sample space, in which case $\sum_{n=1}^{\infty} P(C_n) = 1$. We often say that a mutually exclusive and exhaustive collection of events forms a **partition** of $\mathcal{C}$.

A probability set function tells us how the probability is distributed over the set of events, $\mathcal{B}$. In this sense we speak of a distribution of probability. We often drop the word "set" and refer to P as a probability function.

The following theorems give us some other properties of a probability set function. In the statement of each of these theorems, $P(C)$ is taken, tacitly, to be a probability set function defined on the collection of events $\mathcal{B}$ of a sample space $\mathcal{C}$.

Theorem 1.3.1. *For each event $C \in \mathcal{B}$, $P(C) = 1 - P(C^c)$.*

Proof: We have $\mathcal{C} = C \cup C^c$ and $C \cap C^c = \phi$. Thus, from (2) and (3) of Definition 1.3.1, it follows that

$$1 = P(C) + P(C^c),$$

which is the desired result. ■

Theorem 1.3.2. *The probability of the null set is zero; that is, $P(\phi) = 0$.*

Proof: In Theorem 1.3.1, take $C = \phi$ so that $C^c = \mathcal{C}$. Accordingly, we have

$$P(\phi) = 1 - P(\mathcal{C}) = 1 - 1 = 0$$

and the theorem is proved. ■

Theorem 1.3.3. *If C_1 and C_2 are events such that $C_1 \subset C_2$, then $P(C_1) \leq P(C_2)$.*

Proof: Now $C_2 = C_1 \cup (C_1^c \cap C_2)$ and $C_1 \cap (C_1^c \cap C_2) = \phi$. Hence, from (3) of Definition 1.3.1,

$$P(C_2) = P(C_1) + P(C_1^c \cap C_2).$$

From (1) of Definition 1.3.1, $P(C_1^c \cap C_2) \geq 0$. Hence, $P(C_2) \geq P(C_1)$. ■

Theorem 1.3.4. *For each $C \in \mathcal{B}$, $0 \le P(C) \le 1$.*

Proof: Since $\phi \subset C \subset \mathcal{C}$, we have by Theorem 1.3.3 that

$$P(\phi) \le P(C) \le P(\mathcal{C}) \quad \text{or} \quad 0 \le P(C) \le 1,$$

the desired result. ■

Part (3) of the definition of probability says that $P(C_1 \cup C_2) = P(C_1) + P(C_2)$ if C_1 and C_2 are disjoint, i.e., $C_1 \cap C_2 = \phi$. The next theorem gives the rule for any two events.

Theorem 1.3.5. *If C_1 and C_2 are events in $\mathcal{C}$, then*

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$

Proof: Each of the sets $C_1 \cup C_2$ and C_2 can be represented, respectively, as a union of nonintersecting sets as follows:

$$C_1 \cup C_2 = C_1 \cup (C_1^c \cap C_2) \quad \text{and} \quad C_2 = (C_1 \cap C_2) \cup (C_1^c \cap C_2).$$

Thus, from (3) of Definition 1.3.1,

$$P(C_1 \cup C_2) = P(C_1) + P(C_1^c \cap C_2)$$

and

$$P(C_2) = P(C_1 \cap C_2) + P(C_1^c \cap C_2).$$

If the second of these equations is solved for $P(C_1^c \cap C_2)$ and this result substituted in the first equation, we obtain

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$

This completes the proof. ■

Remark 1.3.2 (Inclusion Exclusion Formula). It is easy to show (Exercise 1.3.9) that

$$P(C_1 \cup C_2 \cup C_3) = p_1 - p_2 + p_3,$$

where

$$\begin{aligned} p_1 &= P(C_1) + P(C_2) + P(C_3) \\ p_2 &= P(C_1 \cap C_2) + P(C_1 \cap C_3) + P(C_2 \cap C_3) \\ p_3 &= P(C_1 \cap C_2 \cap C_3). \end{aligned} \tag{1.3.1}$$

This can be generalized to the **inclusion exclusion formula**:

$$P(C_1 \cup C_2 \cup \cdots \cup C_k) = p_1 - p_2 + p_3 - \cdots + (-1)^{k+1} p_k, \tag{1.3.2}$$

where p_i equals the sum of the probabilities of all possible intersections involving i sets. It is clear in the case $k = 3$ that $p_1 \geq p_2 \geq p_3$, but more generally $p_1 \geq p_2 \geq \cdots \geq p_k$. As shown in Theorem 1.3.7,

$$p_1 = P(C_1) + P(C_2) + \cdots + P(C_k) \geq P(C_1 \cup C_2 \cup \cdots \cup C_k).$$

This is known as **Boole's inequality**. For $k = 2$, we have

$$1 \geq P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2),$$

which gives **Bonferroni's inequality**,

$$P(C_1 \cap C_2) \geq P(C_1) + P(C_2) - 1, \tag{1.3.3}$$

that is only useful when $P(C_1)$ and $P(C_2)$ are large. The inclusion exclusion formula provides other inequalities that are useful, such as

$$p_1 \geq P(C_1 \cup C_2 \cup \cdots \cup C_k) \geq p_1 - p_2$$

and

$$p_1 - p_2 + p_3 \geq P(C_1 \cup C_2 \cup \cdots \cup C_k) \geq p_1 - p_2 + p_3 - p_4. \blacksquare$$

Example 1.3.1. Let $\mathcal{C}$ denote the sample space of Example 1.1.2. Let the probability set function assign a probability of $\frac{1}{36}$ to each of the 36 points in $\mathcal{C}$; that is, the dice are fair. If $C_1 = \{(1,1), (2,1), (3,1), (4,1), (5,1)\}$ and $C_2 = \{(1,2), (2,2), (3,2)\}$, then $P(C_1) = \frac{5}{36}$, $P(C_2) = \frac{3}{36}$, $P(C_1 \cup C_2) = \frac{8}{36}$, and $P(C_1 \cap C_2) = 0$. ■

Example 1.3.2. Two coins are to be tossed and the outcome is the ordered pair (face on the first coin, face on the second coin). Thus the sample space may be represented as $\mathcal{C} = \{(H,H), (H,T), (T,H), (T,T)\}$. Let the probability set function assign a probability of $\frac{1}{4}$ to each element of $\mathcal{C}$. Let $C_1 = \{(H,H), (H,T)\}$ and $C_2 = \{(H,H), (T,H)\}$. Then $P(C_1) = P(C_2) = \frac{1}{2}$, $P(C_1 \cap C_2) = \frac{1}{4}$, and, in accordance with Theorem 1.3.5, $P(C_1 \cup C_2) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4}$. ■

Example 1.3.3 (Equilikely Case). Let $\mathcal{C}$ be partitioned into k mutually disjoint subsets $C_1, C_2, \ldots, C_k$ in such a way that the union of these k mutually disjoint subsets is the sample space $\mathcal{C}$. Thus the events $C_1, C_2, \ldots, C_k$ are **mutually exclusive and exhaustive**. Suppose that the random experiment is of such a character that it is reasonable to *assume* that each of the mutually exclusive and exhaustive events $C_i, i = 1, 2, \ldots, k$, has the same probability. It is necessary then that $P(C_i) = 1/k$, $i = 1, 2, \ldots, k$; and we often say that the events $C_1, C_2, \ldots, C_k$ are *equally likely*. Let the event E be the union of r of these mutually exclusive events, say

$$E = C_1 \cup C_2 \cup \cdots \cup C_r, \quad r \leq k.$$

Then

$$P(E) = P(C_1) + P(C_2) + \cdots + P(C_r) = \frac{r}{k}.$$

Frequently, the integer k is called the total number of ways (for this particular partition of $\mathcal{C}$) in which the random experiment can terminate and the integer r is called the number of ways that are favorable to the event E. So, in this terminology, $P(E)$ is equal to the number of ways favorable to the event E divided by the total number of ways in which the experiment can terminate. It should be emphasized that in order to assign, *in this manner*, the probability r/k to the event E, we must assume that each of the mutually exclusive and exhaustive events $C_1, C_2, \ldots, C_k$ has the same probability $1/k$. This assumption of equally likely events then becomes a *part* of our probability model. Obviously, if this assumption is not realistic in an application, the probability of the event E cannot be computed in this way. ■

In order to illustrate the equilikely case, it is helpful to use some elementary counting rules. These are usually discussed in an elementary algebra course. In the next remark, we offer a brief review of these rules.

Remark 1.3.3 (Counting Rules)**.** Suppose we have two experiments. The first experiment results in m outcomes, while the second experiment results in n outcomes. The composite experiment, first experiment followed by second experiment, has mn outcomes, which can be represented as mn ordered pairs. This is called the **multiplication rule** or the **mn-rule**. This is easily extended to more than two experiments.

Let A be a set with n elements. Suppose we are interested in k-tuples whose components are elements of A. Then by the extended multiplication rule, there are $n \cdot n \cdots n = n^k$ such a k-tuples whose components are elements of A. Next, suppose $k \leq n$ and we are interested in k-tuples whose components are distinct (no repeats) elements of A. There are n elements from which to choose for the first component, $n-1$ for the second component, $\ldots$, $n-(k-1)$ for the kth. Hence, by the multiplication rule, there are $n(n-1)\cdots(n-(k-1))$ such k-tuples with distinct elements. We call each such k-tuple a **permutation** and use the symbol P_k^n to denote the number of k permutations taken from a set of n elements. Hence, we have the formula

$$P_k^n = n(n-1)\cdots(n-(k-1)) = \frac{n!}{(n-k)!}. \tag{1.3.4}$$

Next, suppose order is not important, so instead of counting the number of permutations we want to count the number of subsets of k elements taken from A. We use the symbol $\binom{n}{k}$ to denote the total number of these subsets. Consider a subset of k elements from A. By the permutation rule it generates $P_k^k = k(k-1)\cdots 1$ permutations. Furthermore, all these permutations are distinct from permutations generated by other subsets of k elements from A. Finally, each permutation of k distinct elements drawn from A must be generated by one of these subsets. Hence, we have just shown that $P_k^n = \binom{n}{k}k!$; that is,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}. \tag{1.3.5}$$

We often use the terminology combinations instead of subsets. So we say that there are $\binom{n}{k}$ **combinations** of k things taken from a set of n things. Another common symbol for $\binom{n}{k}$ is C_k^n.

It is interesting to note that if we expand the binomial,

$$(a+b)^n = (a+b)(a+b)\cdots(a+b),$$

we get

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \tag{1.3.6}$$

because we can select the k factors from which to take a in $\binom{n}{k}$ ways. So $\binom{n}{k}$ is also referred to as a **binomial coefficient**. ■

Example 1.3.4 (Poker Hands). Let a card be drawn at random from an ordinary deck of 52 playing cards which has been well shuffled. The sample space $\mathcal{C}$ is the union of $k = 52$ outcomes, and it is reasonable to assume that each of these outcomes has the same probability $\frac{1}{52}$. Accordingly, if E_1 is the set of outcomes that are spades, $P(E_1) = \frac{13}{52} = \frac{1}{4}$ because there are $r_1 = 13$ spades in the deck; that is, $\frac{1}{4}$ is the probability of drawing a card that is a spade. If E_2 is the set of outcomes that are kings, $P(E_2) = \frac{4}{52} = \frac{1}{13}$ because there are $r_2 = 4$ kings in the deck; that is, $\frac{1}{13}$ is the probability of drawing a card that is a king. These computations are very easy because there are no difficulties in the determination of the appropriate values of r and k.

However, instead of drawing only one card, suppose that five cards are taken, at random and without replacement, from this deck; i.e, a five card poker hand. In this instance, order is not important. So a hand is a subset of five elements drawn from a set of 52 elements. Hence, by (1.3.5) there are $\binom{52}{5}$ poker hands. If the deck is well shuffled, each hand should be equilikely; i.e., each hand has probability $1/\binom{52}{5}$. We can now compute the probabilities of some interesting poker hands. Let E_1 be the event of a flush, all five cards of the same suit. There are $\binom{4}{1} = 4$ suits to choose for the flush and in each suit there are $\binom{13}{5}$ possible hands; hence, using the multiplication rule, the probability of getting a flush is

$$P(E_1) = \frac{\binom{4}{1}\binom{13}{5}}{\binom{52}{5}} = \frac{4 \cdot 1287}{2598960} = 0.00198.$$

Real poker players note that this includes the probability of obtaining a straight flush.

Next, consider the probability of the event E_2 of getting exactly three of a kind, (the other two cards are distinct and are of different kinds). Choose the kind for the three, in $\binom{13}{1}$ ways; choose the three, in $\binom{4}{3}$ ways; choose the other two kinds, in $\binom{12}{2}$ ways; and choose one card from each of these last two kinds, in $\binom{4}{1}\binom{4}{1}$ ways. Hence the probability of exactly three of a kind is

$$P(E_2) = \frac{\binom{13}{1}\binom{4}{3}\binom{12}{2}\binom{4}{1}^2}{\binom{52}{5}} = 0.0211.$$

Now suppose that E_3 is the set of outcomes in which exactly three cards are kings and exactly two cards are queens. Select the kings, in $\binom{4}{3}$ ways, and select the queens, in $\binom{4}{2}$ ways. Hence, the probability of E_3 is

$$P(E_3) = \binom{4}{3}\binom{4}{2}\Big/\binom{52}{5} = 0.0000093.$$

The event E_3 is an example of a full house: three of one kind and two of another kind. Exercise 1.3.18 asks for the determination of the probability of a full house. ∎

Example 1.3.4 and the previous discussion allow us to see one way in which we can define a probability set function, that is, a set function that satisfies the requirements of Definition 1.3.1. Suppose that our space $\mathcal{C}$ consists of k distinct points, which, for this discussion, we take to be in a one-dimensional space. If the random experiment that ends in one of those k points is such that it is reasonable to assume that these points are equally likely, we could assign $1/k$ to each point and let, for $C \subset \mathcal{C}$,

$$\begin{aligned} P(C) &= \frac{\text{number of points in } C}{k} \\ &= \sum_{x \in C} f(x), \quad \text{where} \quad f(x) = \frac{1}{k}, \quad x \in \mathcal{C}. \end{aligned}$$

For illustration, in the cast of a die, we could take $\mathcal{C} = \{1, 2, 3, 4, 5, 6\}$ and $f(x) = \frac{1}{6}$, $x \in \mathcal{C}$, if we believe the die to be unbiased. Clearly, such a set function satisfies Definition 1.3.1.

The word *unbiased* in this illustration suggests the possibility that all six points might *not*, in all such cases, be equally likely. As a matter of fact, *loaded* dice do exist. In the case of a loaded die, some numbers occur more frequently than others in a sequence of casts of that die. For example, suppose that a die has been loaded so that the relative frequencies of the numbers in $\mathcal{C}$ *seem to stabilize* proportional to the number of spots that are on the *up* side. Thus we might assign $f(x) = x/21$, $x \in \mathcal{C}$, and the corresponding

$$P(C) = \sum_{x \in C} f(x)$$

would satisfy Definition 1.3.1. For illustration, this means that if $C = \{1, 2, 3\}$, then

$$P(C) = \sum_{x=1}^{3} f(x) = \frac{1}{21} + \frac{2}{21} + \frac{3}{21} = \frac{6}{21} = \frac{2}{7}.$$

Whether this probability set function is realistic can only be checked by performing the random experiment a large number of times.

We end this section with an additional property of probability which proves useful in the sequel. Recall in Exercise 1.2.8 we said that a sequence of events

$\{C_n\}$ is a nondecreasing sequence if $C_n \subset C_{n+1}$, for all n, in which case we wrote $\lim_{n\to\infty} C_n = \cup_{n=1}^{\infty} C_n$. Consider $\lim_{n\to\infty} P(C_n)$. The question is: can we interchange the limit and P? As the following theorem shows, the answer is yes. The result also holds for a decreasing sequence of events. Because of this interchange, this theorem is sometimes referred to as the continuity theorem of probability.

Theorem 1.3.6. *Let $\{C_n\}$ be a nondecreasing sequence of events. Then*

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P\left(\bigcup_{n=1}^{\infty} C_n\right). \tag{1.3.7}$$

Let $\{C_n\}$ be a decreasing sequence of events. Then

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P\left(\bigcap_{n=1}^{\infty} C_n\right). \tag{1.3.8}$$

Proof. We prove the result (1.3.7) and leave the second result as Exercise 1.3.19. Define the sets, called rings, as $R_1 = C_1$ and for $n > 1$, $R_n = C_n \cap C_{n-1}^c$. It follows that $\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} R_n$ and that $R_m \cap R_n = \phi$, for $m \neq n$. Also, $P(R_n) = P(C_n) - P(C_{n-1})$. Applying the third axiom of probability yields the following string of equalities:

$$\begin{aligned} P\left[\lim_{n\to\infty} C_n\right] &= P\left(\bigcup_{n=1}^{\infty} C_n\right) = P\left(\bigcup_{n=1}^{\infty} R_n\right) = \sum_{n=1}^{\infty} P(R_n) = \lim_{n\to\infty} \sum_{j=1}^{n} P(R_j) \\ &= \lim_{n\to\infty} \left\{P(C_1) + \sum_{j=2}^{n} [P(C_j) - P(C_{j-1})]\right\} = \lim_{n\to\infty} P(C_n). \end{aligned} \tag{1.3.9}$$

This is the desired result. ■

Another useful result for arbitrary unions is given by

Theorem 1.3.7 (Boole's Inequality). *Let $\{C_n\}$ be an arbitrary sequence of events. Then*

$$P\left(\bigcup_{n=1}^{\infty} C_n\right) \leq \sum_{n=1}^{\infty} P(C_n). \tag{1.3.10}$$

Proof: Let $D_n = \bigcup_{i=1}^{n} C_i$. Then $\{D_n\}$ is an increasing sequence of events which go up to $\bigcup_{n=1}^{\infty} C_n$. Also, for all j, $D_j = D_{j-1} \cup C_j$. Hence, by Theorem 1.3.5,

$$P(D_j) \leq P(D_{j-1}) + P(C_j),$$

that is,

$$P(D_j) - P(D_{j-1}) \leq P(C_j).$$

In this case, the C_is are replaced by the D_is in expression (1.3.9). Hence, using the above inequality in this expression and the fact that $P(C_1) = P(D_1)$, we have

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} C_n\right) &= P\left(\bigcup_{n=1}^{\infty} D_n\right) = \lim_{n\to\infty}\left\{P(D_1) + \sum_{j=2}^{n}[P(D_j) - P(D_{j-1})]\right\} \\ &\le \lim_{n\to\infty}\sum_{j=1}^{n} P(C_j) = \sum_{n=1}^{\infty} P(C_n). \quad \blacksquare \end{aligned}$$

EXERCISES

1.3.1. A positive integer from one to six is to be chosen by casting a die. Thus the elements c of the sample space $\mathcal{C}$ are $1, 2, 3, 4, 5, 6$. Suppose $C_1 = \{1, 2, 3, 4\}$ and $C_2 = \{3, 4, 5, 6\}$. If the probability set function P assigns a probability of $\frac{1}{6}$ to each of the elements of $\mathcal{C}$, compute $P(C_1)$, $P(C_2)$, $P(C_1 \cap C_2)$, and $P(C_1 \cup C_2)$.

1.3.2. A random experiment consists of drawing a card from an ordinary deck of 52 playing cards. Let the probability set function P assign a probability of $\frac{1}{52}$ to each of the 52 possible outcomes. Let C_1 denote the collection of the 13 hearts and let C_2 denote the collection of the 4 kings. Compute $P(C_1)$, $P(C_2)$, $P(C_1 \cap C_2)$, and $P(C_1 \cup C_2)$.

1.3.3. A coin is to be tossed as many times as necessary to turn up one head. Thus the elements c of the sample space $\mathcal{C}$ are H, TH, TTH, $TTTH$, and so forth. Let the probability set function P assign to these elements the respective probabilities $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}$, and so forth. Show that $P(\mathcal{C}) = 1$. Let $C_1 = \{c : c \text{ is } H, TH, TTH, TTTH, \text{ or } TTTTH\}$. Compute $P(C_1)$. Next, suppose that $C_2 = \{c : c \text{ is } TTTTH \text{ or } TTTTTH\}$. Compute $P(C_2)$, $P(C_1 \cap C_2)$, and $P(C_1 \cup C_2)$.

1.3.4. If the sample space is $\mathcal{C} = C_1 \cup C_2$ and if $P(C_1) = 0.8$ and $P(C_2) = 0.5$, find $P(C_1 \cap C_2)$.

1.3.5. Let the sample space be $\mathcal{C} = \{c : 0 < c < \infty\}$. Let $C \subset \mathcal{C}$ be defined by $C = \{c : 4 < c < \infty\}$ and take $P(C) = \int_C e^{-x}\,dx$. Show that $P(\mathcal{C}) = 1$. Evaluate $P(C)$, $P(C^c)$, and $P(C \cup C^c)$.

1.3.6. If the sample space is $\mathcal{C} = \{c : -\infty < c < \infty\}$ and if $C \subset \mathcal{C}$ is a set for which the integral $\int_C e^{-|x|}\,dx$ exists, show that this set function is not a probability set function. What constant do we multiply the integrand by to make it a probability set function?

1.3.7. If C_1 and C_2 are subsets of the sample space $\mathcal{C}$, show that

$$P(C_1 \cap C_2) \le P(C_1) \le P(C_1 \cup C_2) \le P(C_1) + P(C_2).$$

1.3.8. Let C_1, C_2, and C_3 be three mutually disjoint subsets of the sample space $\mathcal{C}$. Find $P[(C_1 \cup C_2) \cap C_3]$ and $P(C_1^c \cup C_2^c)$.

1.3.9. Consider Remark 1.3.2.

(a) If C_1, C_2, and C_3 are subsets of $\mathcal{C}$, show that

$$P(C_1 \cup C_2 \cup C_3) = P(C_1) + P(C_2) + P(C_3) - P(C_1 \cap C_2) - P(C_1 \cap C_3) - P(C_2 \cap C_3) + P(C_1 \cap C_2 \cap C_3).$$

(b) Now prove the general inclusion exclusion formula given by the expression (1.3.2).

Remark 1.3.4. In order to solve Exercises (1.3.10)-(1.3.18), certain reasonable assumptions must be made. ■

1.3.10. A bowl contains 16 chips, of which 6 are red, 7 are white, and 3 are blue. If four chips are taken at random and without replacement, find the probability that: (a) each of the four chips is red; (b) none of the four chips is red; (c) there is at least one chip of each color.

1.3.11. A person has purchased 10 of 1000 tickets sold in a certain raffle. To determine the five prize winners, five tickets are to be drawn at random and without replacement. Compute the probability that this person wins at least one prize. *Hint:* First compute the probability that the person does not win a prize.

1.3.12. Compute the probability of being dealt at random and without replacement a 13-card bridge hand consisting of: (a) 6 spades, 4 hearts, 2 diamonds, and 1 club; (b) 13 cards of the same suit.

1.3.13. Three distinct integers are chosen at random from the first 20 positive integers. Compute the probability that: (a) their sum is even; (b) their product is even.

1.3.14. There are five red chips and three blue chips in a bowl. The red chips are numbered 1, 2, 3, 4, 5, respectively, and the blue chips are numbered 1, 2, 3, respectively. If two chips are to be drawn at random and without replacement, find the probability that these chips have either the same number or the same color.

1.3.15. In a lot of 50 light bulbs, there are 2 bad bulbs. An inspector examines five bulbs, which are selected at random and without replacement.

(a) Find the probability of at least one defective bulb among the five.

(b) How many bulbs should be examined so that the probability of finding at least one bad bulb exceeds $\frac{1}{2}$?

1.3.16. If $C_1, \ldots, C_k$ are k events in the sample space $\mathcal{C}$, show that the probability that at least one of the events occurs is one minus the probability that none of them occur; i.e.,

$$P(C_1 \cup \cdots \cup C_k) = 1 - P(C_1^c \cap \cdots \cap C_k^c). \quad (1.3.11)$$

1.3.17. A secretary types three letters and the three corresponding envelopes. In a hurry, he places at random one letter in each envelope. What is the probability that at least one letter is in the correct envelope? *Hint:* Let C_i be the event that the ith letter is in the correct envelope. Expand $P(C_1 \cup C_2 \cup C_3)$ to determine the probability.

1.3.18. Consider poker hands drawn from a well-shuffled deck as described in Example 1.3.4. Determine the probability of a full house, i.e, three of one kind and two of another.

1.3.19. Prove expression (1.3.8).

1.3.20. Suppose the experiment is to choose a real number at random in the interval $(0, 1)$. For any subinterval $(a, b) \subset (0, 1)$, it seems reasonable to assign the probability $P[(a, b)] = b - a$; i.e., the probability of selecting the point from a subinterval is directly proportional to the length of the subinterval. If this is the case, choose an appropriate sequence of subintervals and use expression (1.3.8) to show that $P[\{a\}] = 0$, for all $a \in (0, 1)$.

1.3.21. Consider the events C_1, C_2, C_3.

(a) Suppose C_1, C_2, C_3 are mutually exclusive events. If $P(C_i) = p_i$, $i = 1, 2, 3$, what is the restriction on the sum $p_1 + p_2 + p_3$?

(b) In the notation of part (a), if $p_1 = 4/10$, $p_2 = 3/10$, and $p_3 = 5/10$, are C_1, C_2, C_3 mutually exclusive?

For the last two exercises it is assumed that the reader is familar with σ-fields.

1.3.22. Suppose $\mathcal{D}$ is a nonempty collection of subsets of $\mathcal{C}$. Consider the collection of events

$$\mathcal{B} = \cap\{\mathcal{E} : \mathcal{D} \subset \mathcal{E} \text{ and } \mathcal{E} \text{ is a } \sigma\text{-field}\}.$$

Note that $\phi \in \mathcal{B}$ because it is in each σ-field, and, hence, in particular, it is in each σ-field $\mathcal{E} \supset \mathcal{D}$. Continue in this way to show that $\mathcal{B}$ is a σ-field.

1.3.23. Let $\mathcal{C} = R$, where R is the set of all real numbers. Let $\mathcal{I}$ be the set of all open intervals in R. The Borel σ-field on the real line is given by

$$\mathcal{B}_0 = \cap\{\mathcal{E} : \mathcal{I} \subset \mathcal{E} \text{ and } \mathcal{E} \text{ is a } \sigma\text{-field}\}.$$

By definition, $\mathcal{B}_0$ contains the open intervals. Because $[a, \infty) = (-\infty, a)^c$ and $\mathcal{B}_0$ is closed under complements, it contains all intervals of the form $[a, \infty)$, for $a \in R$. Continue in this way and show that $\mathcal{B}_0$ contains all the closed and half-open intervals of real numbers.