6.2 Unions and Intersections

The probability of a union of events

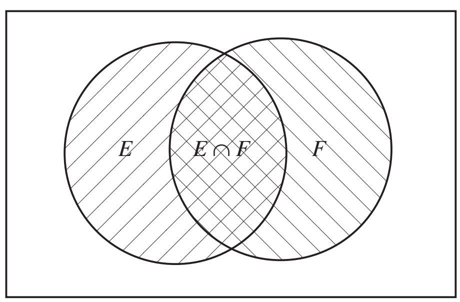
Exercise 6.2-1 If you roll two dice, what is the probability of an even sum or a sum of 8 or more?

Exercise 6.2-2 In Exercise 6.2-1, let E be the event "even sum" and let F be the event "8 or more." We found the probability of the union of the events E and F. Why isn't it the case that $P(E \cup F) = P(E) + P(F)$? What weights appear twice in the sum $P(E) + P(F)$? Find a formula for $P(E \cup F)$ in terms of the probabilities of E, F, and $E \cap F$. Apply this formula to Exercise 6.2-1. What is the value of expressing one probability in terms of three?

Exercise 6.2-3 What is $P(E \cup F \cup G)$ in terms of probabilities of the events E, F, and G and their intersections?

In the sum $P(E) + P(F)$ the weights of elements of $E \cap F$ each appear twice, while the weights of all other elements of $E \cup F$ each appear once. We can see this by looking at a diagram called a Venn Diagram, as in Figure 6.1. In a *Venn diagram*, the rectangle represents the sample space, and the circles represent the events. If we were to shade both E and F, we would wind

Figure 6.1: A Venn diagram for two events.



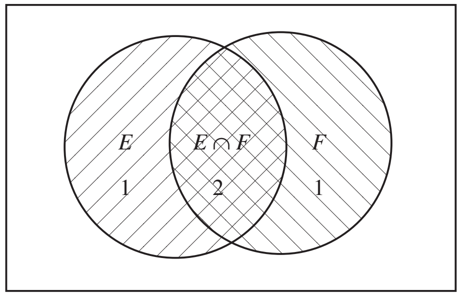
up shading the region $E \cap F$ twice. In Figure 6.2, we represent that by putting numbers in the regions, representing how many times they are shaded. This illustrates why the sum $P(E)+P(F)$ includes the probability weight of each element of $E \cap F$ twice. Thus to get a sum that includes the probability weight of each element of $E \cup F$ exactly once, we have to subtract the weight of $E \cup F$ from the sum $P(E) + P(F)$. This is why

$$P(E \cup F) = P(E) + P(F) - P(E \cap F) \tag{6.3}$$

We can now apply this to Exercise 6.2-1 by noting that the probability of an even sum is 1/2, while the probability of a sum of 8 or more is

$$\frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} = \frac{15}{36}.$$

Figure 6.2: If we shade each of E and F once, then we shade $E \cap F$ twice



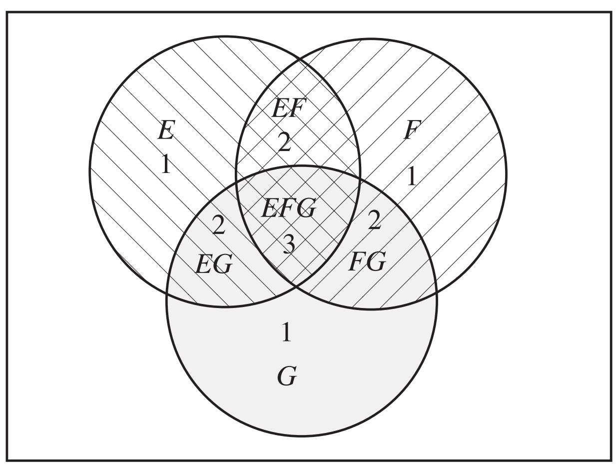
From a similar sum, the probability of an even sum of 8 or more is 9/36, so the probability of a sum that is even or is 8 or more is

$$\frac{1}{2} + \frac{15}{36} - \frac{9}{36} = \frac{2}{3}.$$

(In this case our computation merely illustrates the formula; with less work one could add the probability of an even sum to the probability of a sum of 9 or 11.) In many cases, however, probabilities of individual events and their intersections are more straightforward to compute than probabilities of unions (we will see such examples later in this section), and in such cases our formula is quite useful.

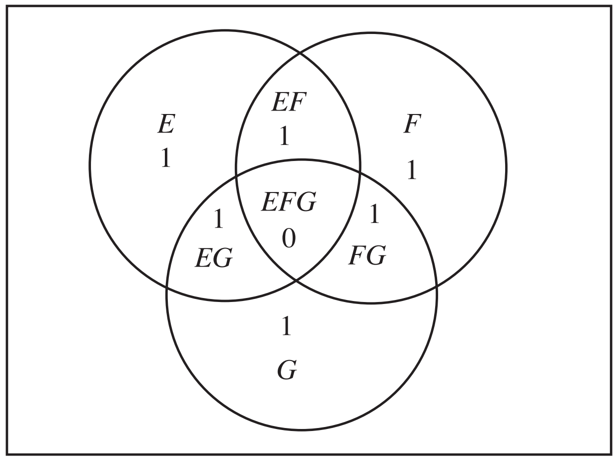
Now let's consider the case for three events and draw a Venn diagram and fill in the numbers for shading all E, F, and G. So as not to crowd the figure we use EF to label the region corresponding to $E \cap F$, and similarly label other regions. Doing so we get Figure 6.3. Thus we

Figure 6.3: The number of ways the intersections are shaded when we shade E, F, and G.



have to figure out a way to subtract from $P(E) + P(F) + P(G)$ the weights of elements in the regions labeled EF, FG and EG once, and the the weight of elements in the region labeled EFG twice. If we subtract out the weights of elements of each of $E \cap F$, $F \cap G$, and $E \cap G$, this does more than we wanted to do, as we subtract the weights of elements in EF, FG and EG once but the weights of elements in of EFG three times, leaving us with Figure 6.4. We then see that

Figure 6.4: The result of removing the weights of each intersection of two sets.



all that is left to do is to add weights of elements in the $E \cap F \cap G$ back into our sum. Thus we have that

$$P(E \cup F \cup G) = P(E) + P(F) + P(G) - P(E \cap F) - P(E \cap G) - P(F \cap G) + P(E \cap F \cap G).$$

Principle of inclusion and exclusion for probability

From the last two exercises, it is natural to guess the formula

$$P(\bigcup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i) - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(E_i \cap E_j) + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(E_i \cap E_j \cap E_k) - \ldots . \quad (6.4)$$

All the sum signs in this notation suggest that we need some new notation to describe sums. We are now going to make a (hopefully small) leap of abstraction in our notation and introduce notation capable of compactly describing the sum described in the previous paragraph. We use

$$\sum_{\substack{i_1,i_2,\ldots,i_k: \\ 1 \le i_1 < i_2 < \cdots < i_k \le n}} P(E_{i_1} \cap E_{i_2} \cap \cdots E_{i_k}) \quad (6.5)$$

to stand for the sum, over all sequences $i_1, i_2, \ldots i_k$ of integers between 1 and n of the probabilities of the sets $E_{i_1} \cap E_{i_2} \ldots \cap E_{i_k}$. More generally, $\sum_{\substack{i_1,i_2,\ldots,i_k: \\ 1 \le i_1 < i_2 < \cdots < i_k \le n}} f(i_1, i_2, \ldots, i_k)$ is the sum of $f(i_1, i_2, \ldots, i_k)$ over all increasing sequences of k numbers between 1 and n.

Exercise 6.2-4 To practice with notation, what is $\sum_{\substack{i_1,i_2,i_3: \\ 1 \le i_1 < i_2 < i_3 \le 4}} i_1 + i_2 + i_3$?

The sum in Exercise 6.2-4 is $1+2+3+1+2+4+1+3+4+2+3+4 = 3(1+2+3+4) = 30$.

With this understanding of the notation in hand, we can now write down a formula that captures the idea in Equation 6.4 more concisely. Notice that in Equation 6.4 we include probabilities of single sets with a plus sign, probabilities of intersections of two sets with a minus sign, and in general, probabilities of intersections of any even number of sets with a minus sign and probabilities of intersections of any odd number of sets (including the odd number one) with a plus sign. Thus if we are intersecting k sets, the proper coefficient for the probability of the intersection of these sets is $(-1)^{k+1}$ (it would be equally good to use $(-1)^{k-1}$, and correct but silly to use $(-1)^{k+3}$). This lets us translate the formula of Equation 6.4 to Equation 6.6 in the theorem, called the *Principle of Inclusion and Exclusion for Probability*, that follows. We will give two completely different proofs of the theorem, one of which is a nice counting argument but is a bit on the abstract side, and one of which is straightforward induction, but is complicated by the fact that it takes a lot of notation to say what is going on.

Theorem 6.3 (Principle of Inclusion and Exclusion for Probability) *The probability of the union $E_1 \cup E_2 \cup \cdots \cup E_n$ of events in a sample space S is given by*

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{\substack{i_1,i_2,\ldots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} P\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}\right). \tag{6.6}$$

First Proof: Consider an element x of $\bigcup_{i=1}^{n} E_i$. Let $E_{i_1}, E_{i_2}, \ldots E_{i_k}$ be the set of all events E_i of which x is a member. Let $K = \{i_1, i_2, \ldots, i_k\}$. Then x is in the event $E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_m}$ if and only if $\{j_1, j_2 \ldots j_m\} \subseteq K$. Why is this? If there is a j_r that is not in K, then $x \notin E_j$ and thus $x \notin E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_m}$. Notice that every x in $\bigcup_{i=1}^{n} E_i$ is in at least one E_i, so it is in at least one of the sets $E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}$.

Recall that we define $P\left(E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_m}\right)$ to be the sum of the probability weights $p(x)$ for $x \in E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_m}$. Suppose we substitute this sum of probability weights for $P\left(E_{j_1} \cap E_{j_2} \cap \cdots \cap E_{j_m}\right)$ on the right hand side of Equation 6.6. Then the right hand side becomes a sum of terms each of with is plus or minus a probability weight. The sum of all the terms involving $p(x)$ on the right hand side of Equation 6.6 includes a term involving $p(x)$ for each nonempty subset $\{j_1, j_2, \ldots, j_m\}$ of K, and no other terms involving $p(x)$. The coefficient of the probability weight $p(x)$ in the term for the subset $\{j_1, j_2, \ldots, j_m\}$ is $(-1)^{m+1}$. Since there are $\binom{k}{m}$ subsets of K of size m, the sum of the terms involving $p(x)$ will therefore be

$$\sum_{m=1}^{k} (-1)^{m+1} \binom{k}{m} p(x) = \left(-\sum_{m=0}^{k} (-1)^m \binom{k}{m} p(x)\right) + p(x) = 0 \cdot p(x) + p(x) = p(x),$$

because $k \ge 1$ and thus by the binomial theorem, $\sum_{j=0}^{k} \binom{k}{j} (-1)^j = (1-1)^k = 0$. This proves that for each x, the sum of all the terms involving $p(x)$ after we substitute the sum of probability weights into Equation 6.6 is exactly $p(x)$. We noted above that for every x in $\cup_{i=1}^{n} E_i$ appears in at least one of the sets $E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}$. Thus the right hand side of Equation 6.6 is the sum of every $p(x)$ such that x is in $\cup_{i=1}^{n} E_i$. By definition, this is the left-hand side of Equation 6.6. ■

Second Proof: The proof is simply an application of mathematical induction using Equation 6.3. When $n = 1$ the formula is true because it says $P(E_1) = P(E_1)$. Now suppose inductively that for any family of $n-1$ sets $F_1, F_2, \ldots, F_{n-1}$

$$P\left(\bigcup_{i=1}^{n-1} F_i\right) = \sum_{k=1}^{n-1}(-1)^{k+1} \sum_{\substack{i_1,i_2,i\ldots,i_k: \\ 1\le i_1<i_2<\cdots<i_k\le n-1}} P(F_{i_1} \cap F_{i_2} \cap F_{i_k}) \tag{6.7}$$

If in Equation 6.3 we let $E = E_1 \cup \ldots \cup E_{n-1}$ and $F = E_m$, we may apply Equation 6.3 to to compute $P\left(\cup_{i=1}^n E_i\right)$ as follows:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P(E_n) - P\left(\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n\right). \tag{6.8}$$

By the distributive law,

$$\left(\bigcup_{i=1}^{n-1} E_i\right) \cap E_n = \bigcup_{i=1}^{n-1}\left(E_i \cap E_n\right),$$

and substituting this into Equation 6.8 gives

$$P\left(\bigcup_{i=1}^{n} E_i\right) = P\left(\bigcup_{i=1}^{n-1} E_i\right) + P\left(E_n\right) - P\left(\bigcup_{i=1}^{n-1}(E_i \cap E_n)\right).$$

Now we use the inductive hypothesis (Equation 6.7) in two places to get

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} E_i\right) &= \sum_{k=1}^{n-1}(-1)^{k+1} \sum_{\substack{i_1,i_2,i\ldots,i_k: \\ 1\le i_1<i_2<\cdots<i_k\le n-1}} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) \\ &+ P(E_n) \\ &- \sum_{k=1}^{n-1}(-1)^{k+1} \sum_{\substack{i_1,i_2,i\ldots,i_k: \\ 1\le i_1<i_2<\cdots<i_k\le n-1}} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k} \cap E_n). \end{aligned} \tag{6.9}$$

The first summation on the right hand side sums $(-1)^{k+1}P\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}\right)$ over all lists $i_1, i_2, \ldots, i_k$ that *do not* contain n, while the $P(E_n)$ and the second summation work together to sum $(-1)^{k+1}P\left(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}\right)$ over all lists $i_1, i_2, \ldots, i_k$ that *do* contain n. Therefore,

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{i_1,i_2,i\ldots,i_k: \\ 1\le i_1<i_2<\cdots<i_k\le n}} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}).$$

Thus by the principle of mathematical induction, this formula holds for all integers $n > 0$. ■

Exercise 6.2-5 At a fancy restaurant n students check their backpacks. They are the only ones to check backpacks. A child visits the checkroom and plays with the check tickets for the backpacks so they are all mixed up. If there are 5 students named Judy, Sam, Pat, Jill, and Jo, in how many ways may the backpacks be returned so that Judy gets her own backpack (and maybe some other students do, too)? What is the probability that this happens? What is the probability that Sam gets his backpack (and maybe some other students do, too)? What is the probability that Judy and Sam both get their own backpacks (and maybe some other students do, too)? For any particular two element set of students, what is the probability that these two students get their own backpacks (and maybe some other students do, too)? What is the probability that at least one student gets his or her own backpack? What is the probability that no students get their own backpacks? What do you expect the answer will be for the last two questions for n students? This classic problem is often stated using hats rather than backpacks (quaint, isn't it?), so it is called the *hatcheck problem*. It is also known as the *derangement problem*; a *derangement* of a set being a one-to-one function from a set onto itself (i.e., a bijection) that sends each element to something not equal to it.

For Exercise 6.2-5, let E_i be the event that person i on our list gets the right backpack. Thus E_1 is the event that Judy gets the correct backpack and E_2 is the event that Sam gets the correct backpack. The event $E_1 \cap E_2$ is the event that Judy *and* Sam get the correct backpacks (and maybe some other people do). In Exercise 6.2-5, there are 4! ways to pass back the backpacks so that Judy gets her own, as there are for Sam or any other single student. Thus $P(E_1) = P(E_i) = \frac{4!}{5!}$. For any particular two element subset, such as Judy and Sam, there are 3! ways that these two people may get their backpacks back. Thus, for each i and j, $P(E_i \cap E_j) = \frac{3!}{5!}$. For a particular k students the probability that each one of these k students gets his or her own backpack back is $(5-k)!/5!$. If E_i is the event that student i gets his or her own backpack back, then the probability of an intersection of k of these events is $(5-k)!/5!$ The probability that at least one person gets his or her own backpack back is the probability of $E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5$. Then by the principle of inclusion and exclusion, the probability that at least one person gets his or her own backpack back is

$$P(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5) = \sum_{k=1}^{5} (-1)^{k+1} \sum_{\substack{i_1, i_2, \ldots, i_k: \\ 1 \le i_1 < i_2 < \cdots < i_k \le 5}} P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}). \tag{6.10}$$

As we argued above, for a set of k people, the probability that all k people get their backpacks back is $(5-k)!/5!$. In symbols, $P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = \frac{(5-k)!}{5!}$. Recall that there are $\binom{5}{k}$ sets of k people chosen from our five students. That is, there are $\binom{5}{k}$ lists $i_1, i_2, \ldots i_k$ with $1 < i_1 < i_2 < \cdots < i_k \le 5$. Thus, we can rewrite the right hand side of the Equation 6.10 as

$$\sum_{k=1}^{5} (-1)^{k+1} \binom{5}{k} \frac{(5-k)!}{5!}.$$

This gives us

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3 \cup E_4 \cup E_5) &= \sum_{k=1}^{5} (-1)^{k-1} \binom{5}{k} \frac{(5-k)!}{5!} \\ &= \sum_{k=1}^{5} (-1)^{k-1} \frac{5!}{k!(5-k)!} \frac{(5-k)!}{5!} \\ &= \sum_{k=1}^{5} (-1)^{k-1} \frac{1}{k!} \\ &= 1 - \frac{1}{2} + \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!}. \end{aligned}$$

The probability that nobody gets his or her own backpack is 1 minus the probability that someone does, or

$$\frac{1}{2} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!}$$

To do the general case of n students, we simply substitute n for 5 and get that the probability that at least one person gets his or her own backpack is

$$\sum_{i=1}^{n}(-1)^{i-1}\frac{1}{i!} = 1 - \frac{1}{2} + \frac{1}{3!} - \cdots + \frac{(-1)^{n-1}}{n!}$$

and the probability that nobody gets his or her own backpack is 1 minus the probability above, or

$$\sum_{i=2}^{n}(-1)^{i}\frac{1}{i!} = \frac{1}{2} - \frac{1}{3!} + \cdots + \frac{(-1)^{n}}{n!} \qquad (6.11)$$

Those who have had power series in calculus may recall the power series representation of e^x, namely

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{i=0}^{\infty}\frac{x^i}{i!}.$$

Thus the expression in Equation 6.11 is the approximation to e^{-1} we get by substituting -1 for x in the power series and stopping the series at $i = n$. Note that the result depends very "lightly" on n; so long as we have at least four or five people, no matter how many people we have, the probability that no one gets their hat back remains at roughly e^{-1}. Our intuition might have suggested that as the number of students increases, the probability that *someone* gets his or her own backpack back approaches 1 rather than $1 - e^{-1}$. Here is another example of why it is important to use computations with the rules of probability instead of intuition!

The Principle of Inclusion and Exclusion for Counting

Exercise 6.2-6 How many functions are there from an n-element set N to a k-element set $K = \{y_1, y_2, \ldots y_k\}$ that map nothing to y_1? Another way to say this is if I have n distinct candy bars and k children Sam, Mary, Pat, etc., in how ways may I pass out the candy bars so that Sam doesn't get any candy (and maybe some other children don't either)?

Exercise 6.2-7 How many functions map nothing to a j-element subset J of K? Another way to say this is if I have n distinct candy bars and k children Sam, Mary, Pat, etc., in how ways may I pass out the candy bars so that some particular j-element subset of the children don't get any (and maybe some other children don't either)?

Exercise 6.2-8 What is the number of functions from an n-element set N to a k element set K that map nothing to at least one element of K? Another way to say this is if I have n distinct candy bars and k children Sam, Mary, Pat, etc., in how ways may I pass out the candy bars so that some child doesn't get any (and maybe some other children don't either)?

Exercise 6.2-9 On the basis of the previous exercises, how many functions are there from an n-element set onto a k element set?

The number of functions from an n-element set to a k-element set $K = \{y_1, y_2, \dots y_k\}$ that map nothing to y_1 is simply $(k-1)^n$ because we have $k-1$ choices of where to map each of our n elements. Similarly the number that map nothing to a particular set J of j elements will be $(k-j)^n$. This warms us up for Exercise 6.2-8.

In Exercise 6.2-8 we need an analog of the principle of inclusion and exclusion for the size of a union of k sets (set i being the set of functions that map nothing to element i of the set K). Because we can make the same argument about the size of the union of two or three sets that we made about probabilities of unions of two or three sets, we have a very natural analog. That analog is the *Principle of Inclusion and Exclusion for Counting*

$$\left|\bigcup_{i=1}^{n} E_i\right| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{i_1,i_2,\dots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} \left|E_{i_1}\cap E_{i_2}\cap\cdots\cap E_{i_k}\right|. \tag{6.12}$$

In fact, this formula is proved by induction or a counting argument in virtually the same way. Applying this formula to the number of functions from N that map nothing to at least one element of K gives us

$$|\bigcup_{i=1}^{k} E_i| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{i_1,i_2,\dots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} |E_{i_1}\cap E_{i_2}\cap\cdots\cap E_{i_k}| = \sum_{j=1}^{k}(-1)^{j-1}\binom{k}{j}(k-j)^n.$$

This is the number of functions from N that map nothing to at least one element of K. The total number of functions from N to K is k^n. Thus the number of onto functions is

$$k^n - \sum_{j=1}^{k}(-1)^{j-1}\binom{k}{j}(k-j)^n = \sum_{j=0}^{k}(-1)^{j}\binom{k}{j}(k-j)^n,$$

where the second equality results because $\binom{k}{0}$ is 1 and $(k-0)^n$ is k^n.

Important Concepts, Formulas, and Theorems

1. *Venn diagram.* To draw a *Venn diagram*, for two or three sets, we draw a rectangle that represents the sample space, and two or three mutually overlapping circles to represent the events.

2. *Probability of a union of two events.* $P(E\cup F) = P(E) + P(F) - P(E\cap F)$

3. *Probability of a union of three events.* $P(E\cup F\cup G) = P(E) + P(F) + P(G) - P(E\cap F) - P(E\cap G) - P(F\cap G) + P(E\cap F\cap G)$.

4. A summation notation. $\displaystyle\sum_{\substack{i_1,i_2,\dots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} f(i_1, i_2, \dots, i_k)$ is the sum of $f(i_1, i_2, \dots, i_k)$ over all increasing sequences of k numbers between 1 and n.

5. *Principle of Inclusion and Exclusion for Probability.* The probability of the union $E_1\cup E_2\cup \cdots\cup E_n$ of events in a sample space S is given by

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{i_1,i_2,\dots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} P\left(E_{i_1}\cap E_{i_2}\cap\cdots\cap E_{i_k}\right).$$

6. *Hatcheck Problem.* The *hatcheck problem* or *derangement problem* asks for the probability that a bijection of an n element set maps no element to itself. The answer is
$$\sum_{i=2}^{n}(-1)^i\frac{1}{i!} = \frac{1}{2} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!},$$
the result of truncating the power series expansion of e^{-1} at the $\frac{(-1)^n}{n!}$. Thus the result is very close to $\frac{1}{e}$, even for relatively small values of n.

7. *Principle of Inclusion and Exclusion for Counting.* The *Principle of inclusion and exclusion for counting* says that
$$\left|\bigcup_{i=1}^{n} E_i\right| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{\substack{i_1,i_2,\ldots,i_k:\\ 1\le i_1<i_2<\cdots<i_k\le n}} \left|E_{i_1}\cap E_{i_2}\cap\cdots\cap E_{i_k}\right|.$$

Problems

1. Compute the probability that in three flips of a coin the coin comes heads on the first flip or on the last flip.

2. The eight kings and queens are removed from a deck of cards and then two of these cards are selected. What is the probability that the king or queen of spades is among the cards selected?

3. Two dice are rolled. What is the probability that we see a die with six dots on top?

4. A bowl contains two red, two white and two blue balls. We remove two balls. What is the probability that at least one is red or white? Compute the probability that at least one is red.

5. From an ordinary deck of cards, we remove one card. What is the probability that it is an Ace, is a diamond, or is black?

6. Give a formula for the probability of $P(E\cup F\cup G\cup H)$ in terms of the probabilities of E,F, G, and H, and their intersections.

7. What is
$$\sum_{\substack{i_1,i_2,i_3:\\ 1\le i_1<i_2<i_3\le 4}} i_1 i_2 i_3 \ ?$$

8. What is
$$\sum_{\substack{i_1,i_2,i_3:\\ 1\le i_1<i_2<i_3\le 5}} i_1 + i_2 + i_3 \ ?$$

9. The boss asks the secretary to stuff n letters into envelopes forgetting to mention that he has been adding notes to the letters and in the process has rearranged the letters but not the envelopes. In how many ways can the letters be stuffed into the envelopes so that nobody gets the letter intended for him or her? What is the probability that nobody gets the letter intended for him or her?

10. If we are hashing n keys into a hash table with k locations, what is the probability that every location gets at least one key?

11. From the formula for the number of onto functions, find a formula for $S(n, k)$ which is defined in Problem 12 of Section 1.4. These numbers are called *Stirling numbers (of the second kind).*

12. If we roll 8 dice, what is the probability that each of the numbers 1 through 6 appear on top at least once? What about with 9 dice?

13. Explain why the number of ways of distributing k identical apples to n children is $\binom{n+k-1}{k}$. In how many ways may you distribute the apples to the children so that Sam gets more than m? In how many ways may you distribute the apples to the children so that no child gets more than m?

14. A group of n married couples sits a round a circular table for a group discussion of marital problems. The counselor assigns each person to a seat at random. What is the probability that no husband and wife are side by side?

15. Suppose we have a collection of m objects and a set P of p "properties," an undefined term, that the objects may or may not have. For each subset S of the set P of all properties, define $N_a(S)$ (a is for "at least") to be the number of objects in the collection that have at least the properties in S. Thus, for example, $N_a(\emptyset) = m$. In a typical application, formulas for $N_a(S)$ for other sets $S \subseteq P$ are not difficult to figure out. Define $N_e(S)$ to be the number of objects in our collection that have exactly the properties in S. Show that

$$N_e(\emptyset) = \sum_{K : K \subseteq P} (-1)^{|K|} N_a(K).$$

Explain how this formula could be used for computing the number of onto functions in a more direct way than we did it using unions of sets. How would this formula apply to Problem 9 in this section?

16. In Problem 14 of this section we allow two people of the same sex to sit side by side. If we require in addition to the condition that no husband and wife are side by side the condition that no two people of the same sex are side by side, we obtain a famous problem known as the *mènage* problem. Solve this problem.

17. In how many ways may we place n distinct books on j shelves so that shelf one gets at least m books? (See Problem 7 in Section 1.4.) In how many ways may we place n distinct books on j shelves so that no shelf gets more than m books?

18. In Problem 15 in this section, what is the probability that an object has none of the properties, assuming all objects to be equally likely? How would this apply the Problem 6-10?