Chapter 1

Basic Concepts

1.1 Dice, Coins, and the World Series

The subject of probability can be traced back to the 17th century when it arose out of the study of gambling games. As we will see the range of applications extends beyond games into business decisions, insurance, law, medical tests, and the social sciences. The stock market, "the largest casino in the world," cannot do without it. The telephone network, call centers, and airline companies with their randomly fulctuating loads could not have been economically designed without probability theory. To quote Pierre Simon, Marquis de Laplace from several hundred years ago:

"It is remarkable that this science, which originated in the consideration of games of chance, should become the most important object of human kowledge … The most important questions of life are, for the most part, really only problems of probability."

In order to address these applications, we need to develop a language for discussing them. Here and throughout the book **bold face type** indicates a term that is being defined. An **experiment** is an activity or procedure that produces distinct, well-defined possibilities called **outcomes**. For example, if our experiment is to roll one die, then there are six outcomes corresponding to the number that shows on the top. The set of all outcomes in this case is $\{1, 2, 3, 4, 5, 6\}$. It is called the **sample space** and is usually denoted by Ω. Symmetry dictates that all outcomes are equally likely so each has probability 1/6.

Things get a little more interesting when we roll two dice. If we suppose, for convenience, that they are red and green. then we can write the outcomes of this experiment as (m, n), where m is the number on the red die and n is the number on the green die. To visualize the set of outcomes it is useful to make a little table:

(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

There are $36 = 6 \cdot 6$ outcomes since there are 6 possible numbers to write in the first slot and for each number written in the first slot there are 6 possibilities for the second.

The goal of probability theory is to compute the probability of various events of interest. Intuitively, an event is a statement about the outcome of an experiment. The formal definition is: An **event** is a subset of the sample space. For example, "The sum is 8" $A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$. Since this event contains 5 of the 36 possible outcomes its probability $P(A) = 5/36$.

For a second example, consider $B =$ "there is at least one six." B consists of the last row and last column of the table so it contains 11 outcomes and hence has probability $P(B) = 11/36$. In general the probability of an event C concerning the roll of two dice is the number of outcomes in C divided by 36.

Abstractly, a **probability** is a function that assigns numbers to events, which satisfies:

(i) For any event A, $0 \le P(A) \le 1$.
(ii) If Ω is the sample space then $P(\Omega) = 1$.
(iii) if A and B are disjoint, i.e., $P(A \cap B) = \emptyset$ then

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

(iv) If $A_1, A_2, \ldots$ is an infinite sequence of pairwise disjoint events (i.e., $A_i \cap A_j \neq \emptyset$ when $i \neq j$)

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

These assumptions are motivated by the **frequency interpretation of probability**, which states that if we repeat an experiment a large number of times then the fraction of times the event A occurs will be close to $P(A)$. To be precise, if we let $N(A,n)$ be the number of times A occurs in the first n trials then

$$P(A) = \lim_{n \to \infty} N(A,n)/n \tag{1.1}$$

In Chapter 4 we will see this result is a theorem called the law of large numbers. For the moment we will use this interpretation of $P(A)$ to motivate the definition.

Given (1.1), (i) and (ii) are clear. The fraction of times that a given event A occurs must be between 0 and 1, and if Ω has been defined properly (recall that it is the set of ALL possible outcomes), the fraction of times something in Ω happens is 1. To explain (iii), note that if the A and B are disjoint then

$$N(A \cup B, n) = N(A,n) + N(B,n)$$

since $A \cup B$ occurs if either A or B occurs but it is impossible for both to happen. Dividing by n and letting $n \to \infty$, we arrive at (iii).

(iii) implies that (iv) holds for a finite number of sets but for infinitely many sets this is a new assumption and the last argument breaks down. Assumption (iv) is a little controversial. Some have argues passionately that it should not be imposed but without (iv) the theory of probability becomes much more difficult and less useful, so we will impose this assumption and having noted its existence not apologize further for it. In many cases the sample space is finite so (iv) is not relevant.

Example 1.1. *Suppose we pick a letter at random from the word TENNESSEE. What is the sample space Ω and what probabilities should be assigned to the outcomes?*

The sample space $\Omega = \{T, E, N, S\}$. To describe the probability it is enough to give the values for the individual outcomes, since (iii) implies that $P(A)$ is the sum of the probabilities of the outcomes in A. Since there are nine letters in TENNESSEE the probabilities are $P(\{T\}) = 1/9$, $P(\{E\}) = 4/9$, $P(\{N\}) = 2/9$, and $P(\{S\}) = 2/9$.

Astragali. Board games involving chance were known in Egypt, 3000 years before Christ. The elment of chance needed for these games was at first provided by tossing astragali, the ankle bones of sheep. These bones could come to rest on only four sides, the other two sides being rounded. The upper side of the bone, broad and slightly convex counted four; the opposite side broad and slightly concave counted three; the lateral side flat and narrow, one, and the opposite narrow lateral side, which is slightly hollow, six.

The outcomes of this experiment are $\Omega = \{1, 3, 4, 6\}$. There is no reason to suppose that all four sides have the same probability so our model will have probabilities for the four outcomes $p_1, p_3, p_4, p_6 \geq 0$ that have $p_1 + p_3 + p_4 + p_6 = 1$. To define the probability of an event A we let

$$P(A) = \sum_{i \in A} p_i$$

In words we add up the probabilities of the outcomes in A. With a little thought we see that any probability with a finite set of outcomes has this form.

Example 1.2. *In English language text the 26 letters in the alphabet oocur with the following frequencies*

E	13.0%	H	3.5%	W	1.6%
T	9.3%	L	3.5%	V	1.3%
N	7.8%	C	3.0%	B	0.9%
R	7.7%	F	2.8%	X	0.5%
O	7.4%	P	2.7%	K	0.3%
I	7.4%	U	2.7%	Q	0.3%
A	7.3%	M	2.5%	J	0.2%
S	6.3%	Y	1.9%	Z	0.1%
D	4.4 %	G	1.6%		

From this it follows that vowels (A,E,I,O,U) are used (7.3+13.0+7.4+7.4+2.7) = 37.8% of the time.

Having introduced a number of definitions, we will now derive some basic properties of probabilities and illustrate their use.

$$P(A) = 1 - P(A^c) \tag{1.2}$$

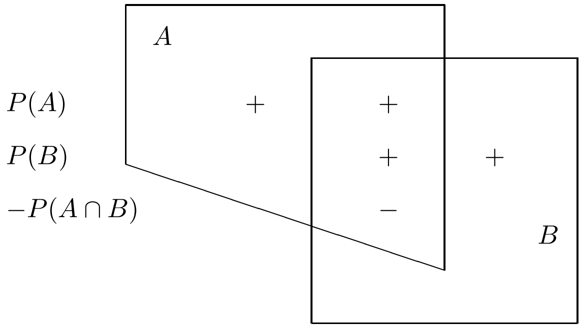
Proof. Let $A_1 = A$ and $A_2 = A^c$. Then $A_1 \cap A_2 = \emptyset$ and $A_1 \cup A_2 = \Omega$ so (iii) implies $P(A) + P(A^c) = P(\Omega) = 1$ by (ii). Subtracting $P(A)$ from each side of the equation gives the result. □

For an example consider A = at least one six. In this case A^c = "no six." There are $5 \cdot 5$ outcomes with no six so $P(A^c) = 25/36$ and $P(A) = 1 - 25/36 = 11/36$ as we computed before.

For any sets A and B,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.3}$$

Proof. We prove this by drawing a picture:



Intuitively, $P(A) + P(B)$ counts $A \cap B$ twice so we have to subtract $P(A \cap B)$ to make the net number of times $A \cap B$ is counted equal to 1. □

To illustrate this let A = red die shows six, B = green die shows six. In this case $A \cup B$ = "at least one 6," and $A \cap B = \{(6,6)\}$ so we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}$$

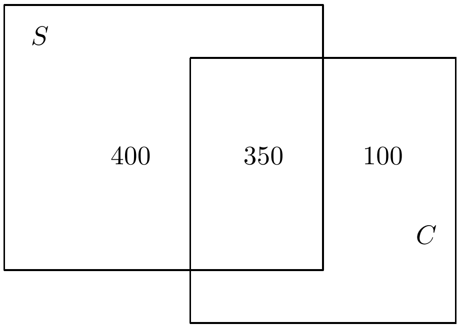
The same principle applies to counting outcomes in sets.

Example 1.3. *A survey of 1000 students revealed that 750 owned stereos, 450 owned cars, and 350 owned both. How many own either a car or a stereo?*

Letting $|S|$ denote the number of students with stereos, and $|C|$ the number with cars, the reasoning that led to (1.3) tells us that

$$|S \cup C| = |S| + |C| - |S \cap C| = 750 + 450 - 350 = 850$$

We can confirm this by drawing a picture:



Even simpler than rolling dice is flipping coins, which produces one of two outcomes, call "Heads" or "Tails." If we flip two coins there are four outcomes

	HH	HT TH	TT
heads	0	1	2
probability	1/4	1/2	1/4

Flipping three coins there are eight possibilities:

	HHH	HHT HTH THH	THH THT HHT	TTT
heads	0	1	2	3
probability	1/8	1/4	1/4	1/8

Our next problem concerns flipping four to seven coins:

Example 1.4 (The World Series, Stanley Cup, and NBA finals.). *All three of these sporting events have the format: the first team to win four games wins the championship. Obviously the series may last 4, 5, 6, or 7 games. However, a fan who wants to buy a ticket would like to know what are the probabilities of each of these outcomes.*

Ignoring potential complicating factors like the advantage of playing at home or psychological factors that make the outcome of one game effect the next one, we suppose that the games are decided by tossing a fair coin to determine if team A or team B wins.

Four games. There are two possible ways this can happen: A wins all four games or B wins all four games. There are $2 \cdot 2 \cdot 2 \cdot 2 = 16$ possible outcomes and these are 2 of them so $P(4) = 1/8$.

Five games. Here and in the next case we will compute the probability A wins in the specified number of games and then multiply by 2. There are four possible outcomes

$$BAAAA, \quad ABAAA, \quad AABAA, \quad AAABA$$

$AAAAB$ is not possible since in that case the series would have ended in four games. There are $2^5 = 32$ outcomes so $P(5) = 2 \cdot 4/32 = 1/4$.

Six games. In the next section we will learn systematic ways of doing this, but for now we will compute the probabilities by enumerating the possibilities.

$$\begin{array}{llll} BBAAAA & ABBAAA & AABBAA & AAABBA \\ BABAAA & ABABAA & AABABA & \\ BAABAA & ABAABA & & \\ BAAABA & & & \end{array}$$

The first column corresponds to outcomes in which B wins the first game, the second to outcomes in which the first game B wins is the second game, etc. We then move the remaining win for B through its possibilities. There are 10 outcomes out of $2^6 = 64$ total so remembering to multiply by 2 to account for the ways B can win in six games, $P(6) = 2 \cdot 10/64$.

Seven games. The analysis from the previous case becomes even messier here so we instead observe that the probabilities for the four possible outcomes must add up to one, so

$$P(7) = 1 - P(4) - P(5) - P(6) = 1 - \frac{2}{16} - \frac{4}{16} - \frac{5}{16} = \frac{5}{16}$$

As mentioned earlier we are ignoring things that many fans think are important to determining the outcomes of the games so our next step is to compare the probabilities just calculated to the observed frequencies.

Games	4	5	6	7
Probabilities	0.125	0.25	0.3125	0.3125
World Series (94)	0.181	0.224	0.224	0.372
Stanley Cup (74)	0.270	0.216	0.230	0.284
NBA finals (57)	0.122	0.228	0.386	0.263

To determine whether or not the data agrees with predictions, statisticians use a “chi-squared statistic:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

where o_i is the number of observations in category i and e_i is what the model predicts. The details of the test are beyond the scope of this book so we just quote the results: the Stanley Cup data is very unusual (has probability < 0.01) due to the larger than expected number of four game series. The World Series data does not fit the model well but is not very unusual (has probability > 0.05). On the other hand the NBA finals data looks like what we should expect to see. The excess of six game series can be due just to chance. We will have more to say about what the probabilities in parentheses mean later.

Example 1.5 (The problem of points). *This problem dates back to the 15th century. Its solution by Pascal and Fermat in 1654 is one of the earliest results in probability. Two teams are playing a competition in which the first term to win five games wins the match. They have each has bet $48 on the outcome but with team A leading team B 3 games to 1, rain forces the match to be canceled. How should they divide up the money that was bet?*

Earlier researchers had proposed specific formulas in terms of the number of games won or left to win. However, Pascal and Fermat came up with the reasonable idea that the fraction of the stakes a team should receive should be equal to the probability it would have won the match (assuming each team is equally likely to win each game). In the case under consideration it is easier to calculate the probability that team B wins. In order for team B to win by a score of 5-3 they must win four games in a row: $BBBB$, an event of probability 1/16. To win 5-4 they can lose one game but not the last game: $ABBBB$, $BABBB$, $BBABB$, $BBBAB$, which has probability 4/32. Adding the two we see that B wins with probability 3/16 and should get that fraction of the total wage, i.e., $(3/16)(96) = 18$.

Example 1.6 (The Birthday Problem). *There are 30 people at a party. Someone wants to bet you $10 that there are two people with exactly the same birthday. Should you take the bet?*

To pose a mathematical problem, we ignore Feb. 29 which only comes in leap years and suppose the each person at the party picks their birthday at random from the calendar. There are 365^{30} possible outcomes for that experiment. The number in which all the birthdays are different is

$$365 \cdot 364 \cdot 363 \cdots 336$$

since the second person must avoid the first person's birthday, the third the first two birthdays and so on until the 30th person must avoid the 29 previous birthdays. Let D be the event that all birthdays are different. Dividing we have

$$P(D) = \frac{365 \cdot 364 \cdot 363 \cdots 336}{365^{30}} = 0.293684$$

In words, only about 29% of the time all the birthdays are different, so you will lose the bet 71% of the time.

At first glance it is surprising that the probability of two people having the same birthday is so large, since there are only 30 people compared with 365 days on the calendar. Some of the surprise disappears if you realize that there are $(30 \cdot 29)/2 = 435$ pairs of people who are going to compare their birthdays. Let p_k be the probability that k people all have different birthdays. Clearly, $p_1 = 1$ and $p_{k+1} = p_k(365 - k)/365$. Using this recursion it is easy to generate a table of p_k for $1 \leq k \leq 40$:

1	1.00000	11	0.85886	21	0.55631	31	0.26955
2	0.99726	12	0.83298	22	0.52430	32	0.24665
3	0.99180	13	0.80559	23	0.49270	33	0.22503
4	0.98364	14	0.77690	24	0.46166	34	0.20468
5	0.97286	15	0.74710	25	0.43130	35	0.18562
6	0.95954	16	0.71640	26	0.40176	36	0.16782
7	0.94376	17	0.68499	27	0.37314	37	0.15127
8	0.92566	18	0.65309	28	0.34554	38	0.13593
9	0.90538	19	0.62088	29	0.31903	39	0.12178
10	0.88305	20	0.58856	30	0.29368	40	0.10877

1.2 Lotteries, Permutations and Combinations

As usual we begin with a question:

Example 1.7. *The New York State Lottery picks 6 numbers out of 54, or more precisely, a machine picks 6 numbered ping pong balls out of a set of 54. How many outcomes are there? The set of numbers chosen is all that is important. The order in which they were chosen is irrelevant.*

To work up to the solution we begin with something that is obvious but is a key step in some of the reasoning to follow.

Example 1.8. *A man has 4 pair of pants, 6 shirts, 8 pairs of socks, and 3 pairs of shoes. Ignoring the fact that some of the combinations may look ridiculous, he can get dressed in* $4 \cdot 6 \cdot 8 \cdot 3 = 576$ *ways.*

To explain why this is true we begin by noting that there are $4 \cdot 6 = 24$ possible combinations of pants and shirts. Each of these can be paired with one of 8 choices of socks, so there are $192 = 24 \cdot 8$ ways of putting on pants, shirt, and socks. Repeating the last argument one more time, we see that for each of these 192 combinations there are 3 choices of shoes and this gives the answer. The reasoning in the last solution can clearly be extended to more than four experiments, and does not depend on the number of choices at each stage, so we have

The multiplication rule. Suppose that m experiments are performed in order and that, no matter what the outcomes of experiments $1, \ldots, k-1$ are, experiment k has n_k possible outcomes. Then the total number of outcomes is $n_1 \cdot n_2 \cdots n_m$.

Example 1.9. *How many ways can 5 people stand in line?*

To answer this question, we think about building the line up one person at a time starting from the front. There are 5 people we can choose to put at the front of the line. Having made the first choice, we have 4 possible choices for the second position. (The set of people we have to choose from depends upon who was chosen first, but there are always 4 people to choose from.) Continuing, there are 3 choices for the third position, 2 for the fourth, and finally 1 for the last. Invoking the multiplication rule, we see that the answer must be

$$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

Generalizing from the last example we define n **factorial** to be

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 \tag{1.4}$$

To see that this gives the number of ways n people can stand in line, notice that there are n choices for the first person, $n-1$ for the second, and each subsequent choice reduces the number of people by 1 until finally there is only 1 person who can be the last in line.

Note that $n!$ grows very quickly since $n! = n \cdot (n-1)!$.

1!	1	7!	5,040
2!	2	8!	40,320
3!	6	9!	362,880
4!	24	10!	3,628,800
5!	120	11!	39,916,800
6!	720	12!	479,001,600

Example 1.10. *Twelve people belong to a club. How many ways can they pick a president, vice-president, secretary, and treasurer?*

Again we think of filling the offices one at a time in the order in which they were given in the last sentence. There are 12 people we can pick for president. Having made the first choice, there are always 11 possibilities for vice-president, 10 for secretary, and 9 for treasurer. So by the multiplication rule, the answer is

$$\frac{12}{P}\;\frac{11}{V}\;\frac{10}{S}\;\frac{9}{T}$$

Passing to the general situation, if we have k offices and n club members then the answer is

$$n \cdot (n-1) \cdot (n-2) \cdots (n-k+1)$$

To see this, note that there are n choices for the first office, $n-1$ for the second, and so on until there are $n-k+1$ choices for the last, since after the last person is chosen there will be $n-k$ left. Products like the last one come up so often that they have a name: the **permutation of n things taken k at a time**, or $P_{n,k}$ for short. Multiplying and dividing by $(n-k)!$ we have

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

which gives us a short formula,

$$P_{n,k} = n!/(n-k)! \tag{1.5}$$

The last formula would give us to the answer to the lottery problem if the order in which the numbers drawn was important. Our last step is to consider a related but slightly simpler problem.

Example 1.11. *A club has 23 members. How many ways can they pick 4 people to be on a committee to plan a party?*

To reduce this question to the previous situation, we imagine making the committee members stand in line, which by (1.5) can be done in $23 \cdot 22 \cdot 21 \cdot 20$ ways. To get from this to the number of committees, we note that each committee can stand in line 4! ways, so the number of committees is the number of lineups divided by 4! or

$$\frac{23 \cdot 22 \cdot 21 \cdot 20}{1 \cdot 2 \cdot 3 \cdot 4} = 23 \cdot 11 \cdot 7 \cdot 5 = 8,855$$

Passing to the general situation, suppose we want to pick k people out of a group of n. Our first step is to make the k people stand in line, which can be done in $P_{n,k}$ ways, and then to realize that each set of k people can stand in line $k!$ ways, so the number of ways to choose k people out of n is

$$C_{n,k} = \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} \tag{1.6}$$

by (1.6) and (1.4). Here, $C_{n,k}$ is short for the number of **combinations of** n **things taken** k **at a time**. $C_{n,k}$ is often written as $\binom{n}{k}$, a symbol that is read as “n choose k.” We are now ready for the

Answer to the Lottery Problem. We are choosing $k = 6$ objects out of a total of $n = 54$ when order is not important so the number of possibilities is

$$C_{54,6} = \frac{54!}{6!48!} = \frac{54 \cdot 53 \cdot 52 \cdot 51 \cdot 50 \cdot 49}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 25,827,165$$

You should consider this the next time you think about spending \$1 for a chance to win \$10 million. On the other hand when the probabilities of success are small it is not sensible to think in terms of how much you’ll win on the average.

World Series continued. Using (1.6) we can easily compute the probability that the series last seven games. For this to occur the score must be tied 3-3 after six games. The total number of outcomes for the first 6 games is $2^6 = 64$. The number that end in a 3-3 tie is

$$C_{6,3} = \frac{6!}{3!\,3!} = \frac{6 \cdot 5 \cdot 4}{1 \cdot 2 \cdot 3} = 20$$

since the outcome is determined by choosing the three games that team A will win. This gives us a probability of $20/64 = 5/16$ for the series to end in seven games. Returning to the calculation in the previous section, we see that the number of outcomes that lead to A winning in six games is the number of ways of picking two of the first five games for B to win or $C_{5,2} = 5!/(2!\,3!) = 5 \cdot 4/2 = 10$.

Example 1.12. *Suppose we flip five coins. Compute the probability that we get 0, 1, or 2 heads.*

There are $2^5 = 32$ total outcomes. There is only 1, $TTTTT$ that gives 0 heads. If we want this to fit into our previous formula we set $0! = 1$ (there is only one way for zero people to stand in line) so that

$$C_{5,0} = \frac{5!}{5!\,0!} = 1$$

There are 5 outcomes that have five heads. We can see this by writing out the possibilities: $HTTTT$, $THTTT$, $TTHTT$, $TTTHT$, and $TTTTH$, or note that the number of ways to pick 1 toss for the heads to occur is

$$C_{5,1} = \frac{5!}{4!\,1!} = 5$$

Extending the last reasoning to two heads, the number of outcomes is the number of ways of picking 2 tosses for the heads to occur or

$$C_{5,2} = \frac{5!}{3!\,2!} = \frac{5 \cdot 4}{2} = 10$$

By symmetry the number of outcomes for 3, 4, and 5 heads are 10, 5, and 1. In terms of binomial coefficients this says

$$C_{n,m} = C_{n,n-m} \tag{1.7}$$

The last equality is easy to prove: The number of ways of picking m objects out of n to take is the same as the number of ways of choosing $n - m$ to leave behind. Of course, one can also check this directly from the formula in (1.6).

Pascal's triangle. The number of outcomes for coin tossing problems fit together in a nice pattern:

$$\begin{array}{ccccccccccccccc}
 & & & & & & & 1 & & & & & & & \\
 & & & & & & 1 & & 1 & & & & & & \\
 & & & & & 1 & & 2 & & 1 & & & & & \\
 & & & & 1 & & 3 & & 3 & & 1 & & & & \\
 & & & 1 & & 4 & & 6 & & 4 & & 1 & & & \\
 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & & \\
 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 & \\
1 & & 7 & & 21 & & 35 & & 35 & & 21 & & 7 & & 1
\end{array}$$

Each number is the sum of the ones on the row above on its immediate left and right. To get the 1's on the edges to work correctly we consider the blanks to be zeros. In symbols

$$C_{n,k} = C_{n-1,k-1} + C_{n-1,k} \tag{1.8}$$

Proof. In picking k things out of n, which can be done in $C_{n,k}$ ways, we may or may not pick the last object. If we pick the last object then we must complete our set of k by picking $k - 1$ objects from the first $n - 1$, which can be done in $C_{n-1,k-1}$ ways. If we do not pick the last object then we must pick all k objects from the first $n - 1$, which can be done in $C_{n-1,k}$ ways. □

Binomial Theorem. The numbers in Pascal's triangle also arise if we take powers of $(x + y)$:

$$(x+y)^2 = x^2 + 2xy + y^2$$

$$(x+y)^3 = (x+y)(x^2 + 2xy + y^2) = x^3 + 3x^2y + 3xy^2 + y^3$$

$$\begin{aligned}
(x+y)^4 &= (x+y)(x^3 + 3x^2y + 3xy^2 + y^3) \\
&= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4
\end{aligned}$$

or in general

$$(x+y)^n = \sum_{m=0}^{n} C_{n,m} x^m y^{n-m} \tag{1.9}$$

To see this consider $(x+y)^5$ and write it as

$$(x+y)(x+y)(x+y)(x+y)(x+y)$$

Since we can choose x or y from each parenthesis, there are 2^5 terms in all. If we want a term of the form x^3y^2 then in 3 of the 5 cases we must pick x, so there are $C_{5,3} = (5 \cdot 4)/2 = 10$ ways to do this. The same reasoning applies to the other terms so we have

$$\begin{aligned}(x+y)^5 &= C_{5,5}x^5 + C_{5,4}x^4y + C_{5,3}x^3y^2 + C_{5,2}x^2y^3 + C_{5,1}xy^4 + C_{5,0}y^5 \\ &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5\end{aligned}$$

Poker. In the game of poker the following hands are possible; they are listed in increasing order of desirability. In the definitions the word *value* refers to A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, or 2. This sequence also describes the relative ranks of the cards, with one exception: an Ace may be regarded as a 1 for the purposes of making a straight. (See the example in (d), below.)

(a) *one pair:* two cards of equal value plus three cards with different values
$J\spadesuit$ $J\diamondsuit$ $9\heartsuit$ $Q\clubsuit$ $3\spadesuit$
(b) *two pair:* two pairs plus another card with a different value
$J\spadesuit$ $J\diamondsuit$ $9\heartsuit$ $9\clubsuit$ $3\spadesuit$
(c) *three of a kind:* three cards of the same value and two with different values
$J\spadesuit$ $J\diamondsuit$ $J\heartsuit$ $9\clubsuit$ $3\spadesuit$
(d) *straight:* five cards with consecutive values
$5\heartsuit$ $4\spadesuit$ $3\spadesuit$ $2\heartsuit$ $A\clubsuit$
(e) *flush:* five cards of the same suit
$K\clubsuit$ $9\clubsuit$ $7\clubsuit$ $6\clubsuit$ $3\clubsuit$
(f) *full house:* a three of a kind and a pair
$J\spadesuit$ $J\diamondsuit$ $J\heartsuit$ $9\clubsuit$ $9\spadesuit$
(g) *four of a kind:* four cards of the same value plus another card
$J\spadesuit$ $J\diamondsuit$ $J\heartsuit$ $J\clubsuit$ $9\spadesuit$
(h) *straight flush:* five cards of the same suit with consecutive values
$A\clubsuit$ $K\clubsuit$ $Q\clubsuit$ $J\clubsuit$ $10\clubsuit$

This example is called a *royal flush.*

To compute the probabilities of these poker hands we begin by observing that there are

$$C_{52,5} = \frac{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 2,598,960$$

ways of picking 5 cards out of a deck of 52, so it suffices to compute the number of ways each hand can occur. We will do three cases to illustrate the main ideas and then leave the rest to the reader.

(d) *straight:* $10 \cdot 4^5$
A straight must start with a card that is 5 or higher, 10 possibilities. Once the values are decided on, suits can be assigned in 4^5 ways. This counting regards a straight flush as a straight. If you want to exclude straight flushes, suits can be assigned in $4^5 - 4$ ways.

(f) *full house:* $13 \cdot C_{4,3} \cdot 12 \cdot C_{4,2}$
We first pick the value for the three of a kind (which can be done in 13 ways), then assign suits to those three cards ($C_{4,3}$ ways), then pick the value for the pair (12 ways), then we assign suits to the last two cards ($C_{4,2}$ ways).

(a) *one pair:* $13 \cdot C_{4,2} \cdot C_{12,3} \cdot 4^3$
We first pick the value for the pair (13 ways), next pick the suits for the pair ($C_{4,2}$ ways), then pick three values for the other cards ($C_{12,3}$ ways) and assign suits to those cards (in 4^3 ways).

A common incorrect answer to this question is $13 \cdot C_{4,2} \cdot 48 \cdot 44 \cdot 40$. The faulty reasoning underlying this answer is that the third card must not have the same value as the cards in the pair (48 choices), the fourth must be different from the third and the pair (44 choices), ... However, this reasoning is flawed since it counts each outcome $3! = 6$ times. (Note that $48 \cdot 44 \cdot 40/3! = C_{12,3} \cdot 4^3$.)

The numerical values of the probabilities of all poker hands are given in the next table.

(a) *one pair*	.422569
(b) *two pair*	.047539
(c) *three of a kind*	.021128
(d) *straight*	.003940
(e) *flush*	.001981
(f) *full house*	.001441
(g) *four of a kind*	.000240
(h) *straight flush*	.000015

The probability of getting none of these hands can be computed by summing the values for (a) through (g) (recall that (d) includes (h)) and subtracting the result from 1. However, it is much simpler to observe that we have nothing if we have five different values that do not make a straight or a flush. So the number of nothing hands is $(C_{13,5} - 10) \cdot (4^5 - 4)$ and the probability of a nothing hand is 0.501177.

More than two categories: the multinomial. We defined $C_{n,k}$ as the number of ways of picking k things out of n. To motivate the next generalization we would like to observe that $C_{n,k}$ is also the number of ways we can divide n objects into two groups, the first one with k objects and the second with $n - k$. To connect this observation with the next problem, think of it as asking: "How many ways can we divide 12 objects into three numbered groups of sizes 4, 3, and 5?"

Example 1.13. *A house has 12 rooms. We want to paint 4 yellow, 3 purple, and 5 red. In how many ways can this be done?*

This problem can be solved using what we know already. Wefirst pick 4 of the 12 rooms to be painted yellow, which can be done in $C_{12,4}$ ways, and then pick 3 of the remaining 8 rooms to be painted purple, which can be done in $C_{8,3}$ ways. (The 5 unchosen rooms will be painted red.) The answer is:

$$C_{12,4}C_{8,3} = \frac{12!}{4!\,8!} \cdot \frac{8!}{3!\,5!} = \frac{12!}{4!\,3!\,5!} = 27,720.$$

A second way of looking at the problem, which gives the last answer directly is to first decide the order in which the rooms will be painted, which can be done in 12! ways, then paint the first 4 on the list yellow, the next 3 purple, and the last 5 red. One example is

$$\frac{9}{Y}\frac{6}{Y}\frac{11}{Y}\frac{1}{Y}\frac{8}{P}\frac{2}{P}\frac{10}{P}\frac{5}{R}\frac{3}{R}\frac{7}{R}\frac{12}{R}\frac{4}{R}$$

Now, the first four choices can be rearranged in 4! ways without affecting the outcome, the middle three in 3! ways, and the last five in 5! ways. Invoking the multiplication rule, we see that in a list of the 12! possible permutations each possible painting thus appears 4! 3! 5! times. Hence the number of possible paintings is

$$\frac{12!}{4!\,3!\,5!}$$

The second computation is a little more complicated than the first, but makes it easier to see

If we have a group of n objects to be divided into m groups of size $n_1, \dots, n_m$ with $n_1 + \cdots + n_m = n$ this can be done in

$$\frac{n!}{n_1!\,n_2!\cdots n_m!} \text{ ways.} \tag{1.10}$$

The formula may look complicated but it is easy to use.

Example 1.14. *Four people play a card game in which each gets 13 cards. How many possible deals are there?*

$$52!/(13!)^4$$

Example 1.15. *There are 37 students in a class. In how many ways can a professor give out 9 A's, 13 B's, 12 C's, and 5 F's?*

$$37!/(9!\,13!\,12!\,5!)$$

1.3 Card Games and Other Urn Problems

A number of problems in probability have the following form.

Example 1.16. *Suppose we pick 4 balls out of an urn with 12 red balls and 8 black balls. What is the probability of B = "We get two balls of each color"?*

Almost by definition, there are

$$C_{20,4} = \frac{20 \cdot 19 \cdot 18 \cdot 17}{1 \cdot 2 \cdot 3 \cdot 4} = 5 \cdot 19 \cdot 3 \cdot 17 = 4,845$$

ways of picking 4 balls out of the 20. To count the number of outcomes in B, we note that there are $C_{12,2}$ ways to choose the red balls and $C_{8,2}$ ways to choose the black balls, so the multiplication rule implies

$$|B| = C_{12,2}C_{8,2} = \frac{12 \cdot 11}{1 \cdot 2} \cdot \frac{8 \cdot 7}{1 \cdot 2} = 6 \cdot 11 \cdot 4 \cdot 7 = 1,848$$

It follows that $P(B) = 1848/4845 = 0.3814$.

Our next four examples are practical applications of drawing balls out of urns.

Example 1.17 (Bridge). *In the game of bridge there are four players called North, West, South, and East according to their positions at the table. Each player gets 13 cards. The game is somewhat complicated so we will content ourselves to analyze one situation that is important in the play of the game. Suppose that North and South have a total of eight Hearts. What is the probability that West will have 3 and East will have 2?*

Even though this is not how the cards are usually dealt, we can imagine that West randomly draws 13 cards from the 26 that remain. This can be done in

$$C_{26,13} = \frac{26!}{13!\,13!} = 10,400,600 \text{ ways}$$

North and South have 8 hearts and 18-non hearts so in the 26 that remain there are $13-8=5$ hearts and $39-18=21$ non-hearts. To construct a hand for West with 3 hearts and 10 non-hearts we must pick 3 of the 5 hearts, which can be done in $C_{5,3}$ ways and 10 of the 21 non-hearts in $C_{21,10}$. The multiplication rule then implies that the number of outcomes for West with 3 hearts is $C_{5,3} \cdot C_{21,10}$ and the probability of interest is

$$\frac{C_{5,3} \cdot C_{21,10}}{C_{26,13}} = 0.339$$

Multiplying by 2 gives the probability that one player will have 3 cards and the other 2, something called a $3-2$ split. Repeating the reasoning gives that an $i-j$ split $(i+j=5)$ has probability

$$2 \cdot \frac{C_{5,i} \cdot C_{21,13-i}}{C_{26,13}}$$

This formula tells us that the probabilities are

3-2	0.678
4-1	0.282
5-0	0.039

Thus while a 3-2 split is the most common, one should not ignore the possibility of a 4-1 split. Similar calculations show that if North and South have 9 hearts then the probabilities are

2-2	0.406
3-1	0.497
4-0	0.095

In this case the uneven 3-1 split is more common than the 2-2 split since it can occur two ways, i.e., West might have 3 or 1.

Example 1.18 (Disputed elections). *In a close election in a small town, 2,656 people voted for candidate A compared to 2,594 who voted for candidate B, a margin of victory of 62 votes. An investigation of the election, instigated no doubt by the loser, found that 136 of the people who voted in the election should not have. Since this is more than the margin of victory, should the election results be thrown out even though there was no evidence of fraud on the part of the winner's supporters?*

Like many problems that come from the real world (*DeMartini v. Power*, 262 NE2d 857), this one is not precisely formulated. To turn this into a probability problem we suppose that all the votes were equally likely to be one of the 136 erroneously cast and we investigate what happens when we remove 136 balls from an urn with 2,656 white balls and 2,594 black balls. Now the probability of removing exactly m white and $136 - m$ black balls is

$$\frac{\binom{2{,}656}{m}\binom{2{,}594}{136-m}}{\binom{5{,}250}{136}}$$

In order to reverse the outcome of the election, we must have

$$2,656 - m \leq 2,594 - (136 - m) \quad \text{or} \quad m \geq 99$$

With the help of a short computer program we can sum the probability above from $m = 99$ to 136 to conclude that the probability of the removal of 136 randomly chosen votes reversing the election is 7.492×10^{-8}. This computation supports the Court of Appeals decision to overturn a lower court ruling that voided the election in this case.

Exercise. What do you think should have been done in *Ipolito v. Power*, 241 NE2d 232, where the winning margin was 1,422 to 1,405 but 101 votes had to be thrown out?

Example 1.19 (Quality control). *A shipment of 50 precision parts including 4 that are defective is sent to an assembly plant. The quality control division selects 10 at random for testing and rejects the entire shipment if 1 or more are found defective. What is the probability this shipment passes inspection?*

There are $C_{50,10}$ ways of choosing the test sample, and $C_{46,10}$ ways of choosing all good parts so the probability is

$$\begin{aligned}\frac{C_{46,10}}{C_{50,10}} &= \frac{46!/36!10!}{50!/40!/10!}\frac{46\cdot 45\cdots 37}{50\cdot 49\cdots 41} \\ &= \frac{40\cdot 39\cdot 38\cdot 37}{50\cdot 49\cdot 48\cdot 47} = 0.396\end{aligned}$$

Using almost identical calculations a company can decide on how many bad units they will alow in a shipment and design a testing program with a given probability of success.

Example 1.20 (Capture-recapture experiments). *An ecology graduate student goes to a pond and captures* $k = 60$ *water beetles, marks each with a dot of paint, and then releases them. A few days later she goes back and captures another sample of* $r = 50$*, finding* $m = 12$ *marked beetles and* $r - m = 38$ *unmarked. What is her best guess about the size of the population of water beetles?*

To turn this into a precisely formulated problem, we will suppose that no beetles enter or leave the population between the two visits. With this assumption, if there were N water beetles in the pond, then the probability of getting m marked and $r - m$ unmarked in a sample of r would be

$$p_N = \frac{C_{k,m}\, C_{N-k,r-m}}{C_{N,r}}$$

To estimate the population we will pick N to maximize p_N, the so-called **maximum likelihood estimate**. To find the maximizing N, we note that

$$C_{j-1,i} = \frac{(j-1)!}{(j-i-1)!i!} \quad \text{so} \quad C_{j,i} = \frac{j!}{(j-i)!i!} = \frac{jC_{j-1,i}}{(j-i)}$$

and it follows that

$$p_N = p_{N-1}\cdot\frac{N-k}{N-k-(r-m)}\cdot\frac{N-r}{N}$$

Now $p_N/p_{N-1} \geq 1$ if and only if

$$(N-k)(N-r) \geq N(N-k-r+m)$$

that is,

$$N^2 - kN - rN + kr \geq N^2 - kN - rN + mN$$

or equivalently if $kr \geq Nm$ or $N \leq kr/m$. Thus the value of N that maximizes the probability p_N is the largest integer $\leq kr/m$. This choice is reasonable since when $N = kr/m$ the proportion of marked beetles in the population $k/N = m/r$, the proportion of marked beetles in the sample. Plugging in the numbers from our example, $kr/m = (60\cdot 50)/12 = 250$, so the probability is maximized when $N = 250$.

1.4 Probabilities of Unions, Joe DiMaggio

In Section 1.1, we learned that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. In this section we will extend this formula to $n > 2$ events. We begin with $n = 3$ events:

$$\begin{aligned} P(A \cup B \cup C) \;=\; & P(A) + P(B) + P(C) \\ & -P(A \cap B) - P(A \cap C) - P(B \cap C) \\ & +P(A \cap B \cap C) \end{aligned} \tag{1.11}$$

Proof. As in the proof of the formula for two sets, we have to convince ourselves that the net number of times each part of $A \cup B \cup C$ is counted is 1. To do this, we make a table that identifies the areas counted by each term and note that the net number of pluses in each row is 1:

	A	B	C	$A \cap B$	$A \cap C$	$B \cap C$	$A \cap B \cap C$
$A \cap B \cap C$	+	+	+	−	−	−	+
$A \cap B \cap C^c$	+	+		−			
$A \cap B^c \cap C$	+		+		−		
$A^c \cap B \cap C$		+	+			−	
$A \cap B^c \cap C^c$	+						
$A^c \cap B \cap C^c$		+					
$A^c \cap B^c \cap C$			+				

Example 1.21. *Suppose we roll three dice. What is the probability that we get at least one 6?*

Let A_i = "We get a 6 on the ith die." Clearly,

$$\begin{aligned} & P(A_1) = P(A_2) = P(A_3) = 1/6 \\ & P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = 1/36 \\ & P(A_1 \cap A_2 \cap A_3) = 1/216 \end{aligned}$$

So plugging into (1.11) gives

$$P(A_1 \cup A_2 \cup A_3) = 3 \cdot \frac{1}{6} - 3 \cdot \frac{1}{36} + \frac{1}{216} = \frac{108 - 18 + 1}{216} = \frac{91}{216}$$

To check this answer, we note that $(A_1 \cup A_2 \cup A_3)^c$ = "no 6" $= A_1^c \cap A_2^c \cap A_3^c$ and $|A_1^c \cap A_2^c \cap A_3^c| = 5 \cdot 5 \cdot 5 = 125$ since there are five "non-6's" that we can get on each roll. Since there are $6^3 = 216$ outcomes for rolling three dice, it follows that $P(A_1^c \cap A_2^c \cap A_3^c) = 125/216$ and $P(A_1 \cup A_2 \cup A_3) = 1 - P(A_1^c \cap A_2^c \cap A_3^c) = 91/216$.

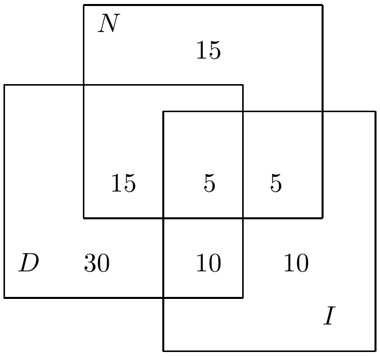
The same reasoning applies to sets.

Example 1.22. *In a freshman dorm, 60 students read the Cornell Daily Sun, 40 read the New York Times and 30 read the Ithaca Journal. 20 read the Daily Sun and the NY Times, 15 read the Daily Sun and the Ithaca Journal, 10 read the NY Times and the Ithaca Journal, and 5 read all three. How many read at least one newspaper.*

Using our formula the answer is

$$60 + 40 + 30 - 20 - 15 - 10 + 5 = 90$$

To check this we can draw picture using D, N, and I for the three newspapers



To figure out the number of students in each category we work out from the middle. $D \cap N \cap I$ has 5 students and $D \cap N$ has 20 so $D \cap N \cap I^c$ has 15. In the same way we compute that $D \cap N^c \cap I$ has $15 - 5 = 10$ students and $D^c \cap N \cap I$ has $10 - 5 = 5$ students. Having found that 30 of students in D read at least one other newspaper, the number that read only D is $60 - 30 = 30$. In a similar way we compute that there are $40 - 25 = 15$ students that only read N and $30 - 20 = 10$ students that only read I. Adding up the numbers in the seven regions gives a total of 90 as we found before.

The general formula for n events, called the **inclusion-exclusion formula**, is

$$\begin{aligned} P(\cup_{i=1}^n A_i) &= \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ &\quad \cdots + (-1)^{n+1} P(A_1 \cap \ldots \cap A_n) \qquad (1.12) \end{aligned}$$

In words, we take all possible intersections of one, two, … n sets and the signs of the sums alternate.

Proof. A point that is in k sets is counted k times by the first sum, $C_{k,2}$ by the second, $C_{k,3}$ be the third and so on until it is counted 1 time by the kth term. The net result is

$$C_{k,1} - C_{k,2} + C_{k,3} \ldots + (-1)^{k+1} 1$$

To show that this adds up to 1, we recall the Binomial theorem

$$(a + b)^k = a^k + C_{k,1} a^{k-1} b + C_{k,2} a^{k-2} b^2 + \cdots + b^k$$

Setting $a = 1$ and $b = -1$ we have

$$0 = 1 - C_{k,1} + C_{k,2} - C_{k,3} \dots - (-1)^{k+1} 1$$

which proves the desired result. □

Example 1.23. *You pick 7 cards out of deck of 52. What is the probability that you have exactly three cards of some denomination (e.g., three Kings or three 7's)?*

Let A_i for $1 \le i \le 13$ be the event you have three cards of type i where 1 is Ace, 11 is Jack, 12 is Queen, and 13 is King. It is impossible for three of these events to occur so

$$P(\cup_{i=1}^{13} A_i) = 13 P(A_1) - C_{13,2} P(A_1 \cap A_2)$$

A_1 can occur in $C_{4,3} C_{48,4} = 778,320$ ways, $A_1 \cap A_2$ can occur in $(C_{4,3})^2 \cdot 44 = 704$ ways so the answer is

$$\frac{13 \cdot 778,320 - 78 \cdot 704}{C_{52,7}} = \frac{10,118,160 - 54,912}{133,784,560} = 0.075219$$

Notice that the first term gives most of the answer and the second is only a small correction to account for the rare event of have two three of a kinds.

Example 1.24. *Suppose we roll a die 15 times. What is the probability that we do not see all 6 numbers at least once?*

Let A_i be the event that we never see i. $P(A_i) = 5^{15}/6^{15}$ since there are 6^{15} outcomes in all but only 5^{15} that contain no i's. $5^{15}/6^{15} = 0.064905$, so

$$\sum_{i=1}^{n} P(A_i) = 6(0.064905) = 0.389433$$

Turning to the second, we note that for any $i < j$, we have $P(A_i \cap A_j) = 4^{15}/6^{15} = 0.002284$ and there are $C_{6,2} = (6 \cdot 5)/2 = 15$ choices for $i < j$ so

$$\sum_{i<j} P(A_i \cap A_j) = 15(0.002284) = 0.03426$$

To the third term we note that for any $i < j < k$, we have $P(A_i \cap A_j \cap A_k) = 3^{15}/6^{15} = 3.05 \times 10^{-5}$ and there are $C_{6,3} = (6 \cdot 5 \cdot 4)/3! = 20$ choices for $i < j < k$ so

$$P\left(\cup_{i=1}^{6} A_i\right) = 20(3.05 \times 10^{-5}) 0.00061$$

At this point the pattern should be clear:

$$\begin{aligned}&C_{6,1}(5/6)^{15} - C_{6,2}(4/6)^{15} + C_{6,3}(3/6)^{15} - C_{6,4}(2/6)^{15} + C_{6,5}(1/6)^{15}\\ &= 0.389433 - 0.03426 + 6.1 \times 10^{-4} - 1.045 \times 10^{-6} + 1.276 \times 10^{-11}\end{aligned}$$

$$= 0.355787$$

Even better than the inclusion-exclusion formula are the associated

The Bonferroni Inequalities. In brief, if you stop the inclusion-exclusion formula with a + term you get an upper bound; if you stop with a − term you get a lower bound.

$$P(\cup_{i=1}^n A_i) \le \sum_{i=1}^n P(A_i) \tag{1.13}$$

$$\ge \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) \tag{1.14}$$

$$\le \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \tag{1.15}$$

To explain the usefulness of these inequalities, we note that in the previous example they imply that the probaility of interest is

$$\begin{aligned} &\le \; 0.389433 \\ &\ge \; 0.389433 - 0.03426 = 0.355178 \\ &\le \; 0.389433 - 0.03426 + 6.1 \times 10^{-4} = 0.355738 \end{aligned}$$

so we have a very accurate result after three terms.

Proof. The first inequality is obvious since the right-hand side counts each outcome in $\cup_{i=1}^n A_i$ at least once. To prove the second consider an outcome that is in k sets. If $k = 1$ the first term will count it once and the second not at all. If $k = 2$ the first term counts it 2 times and the second 1 for a net total of 1. If $k \ge 3$ the first term counts it k times and the second $C_{k,2} = k(k-1)/2 > k$ times so the net number of coutings is < 0. The third formula is similar but more complicated so we leave its proof to the reader. □

Example 1.25 (The Streak). *In the summer of 1941, Joe DiMaggio had what many people consider the greatest record in sports, in which he had at least one hit in each of 56 games. What is the probability of this event?*

To compute the probability of this event we will introduce three assumptions that are somewhat questionable: (i) a player gets exactly four at bats per game (during the streak, DiMaggio averaged 3.98 at bats per game), (ii) the outcomes of different at bats are independent with the probability of a hit being 0.325, Joe DiMaggio's lifetime batting average, and (iii) the outcomes for successive games are independent. From assumptions (i) and (ii) it follows that the probability p of a hit during a game is $1 - (0.675)^4 = 0.7924$.

Assuming a 162-game season, we could let A_i be the probability that a player got hits in games $i, i+1, \ldots i+55$ for $1 \le i \le 107$. Using (1.13) it follows that the probablity of the streak is

$$\le 107(0.7924)^{56} = 2.345 \times 10^{-4}$$

As we will see below this 4.86 times the actual answer. The problem is that if A_i occurs, it becomes much easier for A_{i+1}, A_{i-1}, and other "nearby" events to occur. To avoid this problem, we will let B_i be the event the player gets hits in games $i, i+1, i+2, \ldots, i+55$ but no hit in game $i+56$, where $1 \le i \le 106$. The event of interest $S = \cup_{i=1}^{106} B_i$ so

$$P(S) \le \sum_{i=1}^{106} P(B_i) = 106(0.7924)^{56}(0.2076) = 4.824 \times 10^{-5} \tag{1.16}$$

To compute the second bound we begin by noting $B_i \cap B_j = \emptyset$ if $i < j \le i+56$ since B_i requires a loss in game $i+56$ while B_j requires a win. If $56+i < j \le 106$ then $P(B_i \cap B_j) = P(B_i)P(B_j)$. There are $49+48+\cdots+2+1 = (50 \cdot 49)/2 = 1225$ terms so

$$\sum_{1 \le i < j \le 106} P(B_i \cap B_j) = 1225(0.7924)^{112}(0.2076)^2 = 32.537 \times 10^{-10}$$

Since this is the number we have to subtract from the upper bound to get the lower bound, our upper bound is extremely accurate.

To interpret our result, note that the probability in (1.16) is roughly 1/200,000, so even if there were 200 players with 0.325 batting averages, it would take 1000 years for this to occur again.

Example 1.26 (A less famous streak). *Sports Illustrated reports that a high school football team in Bloomington, Indiana lost 21 straight pre-game coin flips before finally winning one. Taking into account the fact that there are approximately 15,000 high school and college football teams, is this really surprising?*

We will first compute the probability that this happens to one team some time in the decade 1995-2004 assuming that the team plays 10 games per year. Taking a lesson from the previous example, we let B_i be the event that the team loses coin flips in games $i, i+1, \ldots i+20$ and finally wins the one for game $i+21$, where $1 \le i \le 79$.

$$P(S) \le \sum_{i=1}^{79} P(B_i) = 79(1/2)^{22} = 1.883 \times 10^{-5}$$

To compute the second bound we begin by noting $B_i \cap B_j = \emptyset$ if $i < j \le i+21$ since B_i requires a wining coin flip $i+21$ while B_j requires a loss. If $21+i < j \le 79$ then $P(B_i \cap B_j) = P(B_i)P(B_j)$. There are $58+57+\cdots+2+1 = (58 \cdot 57)/2 = 1653$ terms so

$$\sum_{1 \le i < j \le 79} P(B_i \cap B_j) = 1653(1/2)^{44} = 9.396 \times 10^{-11}$$

Again, since this is the number we have to subtract from the upper bound to get the lower bound, our upper bound is extremely accurate.

What we have computed is the probability that one particular team will have this type of bad luck some time in the last decade. The probability that none of the 15,000 teams will do this is

$$(1 - 1.883 \times 10^{-5})^{15{,}000} = 0.7539$$

i.e., with probability 0.2461 some team will have this happen to them. As a check on the last calculation, note that (1.13) gives an upper bound of

$$15,000 \times 1.883 \times 10^{-5} = 0.2825$$

1.5 Blackjack

In this book we will analyze craps and roulette, casino games where the player has a substantial disadvantage. In the case of blackjack, a little strategy, which we will explain in this section, can make the game almost even. To begin we will describe the rules and the betting.

In the game of blackjack a King, Queen, or Jack counts 10, an Ace counts 1 or 11, and the other cards count the numbers that are shown on them (e.g., a 5 counts 5). The object of the game is to get as close to 21 as you can without going over. You start with 2 cards and draw cards out of the deck until either you are happy with your total or you go over 21, in which case you "bust."

If your initial two cards total 21, this is a blackjack, and if the dealer does not have one, you win 1.5 times your original bet. If you bust then you immediately lose your bet. This is the main source of the casino advantage since if the dealer busts later you have already lost. If you stop with 21 or less and the dealer busts you win. If you and the dealer both end with 21 or less, then the one with higher hand wins. In the case of a tie no money changes hands.

In casino blackjack the dealer plays by a simple rule: He draws a card if his total is ≤ 16, otherwise he stops. The first step in analyzing blackjack is to compute the probability the dealer's ending total is k when he has a total of j. To deal with the complication that an Ace can count as 1 or 11, we introduce $b(j,k) =$ the probability that the dealer's ending total is k when he has a total of j including one Ace that is being counted as 11. Such hands are called "soft" because even if you draw a 10 you will not bust. We define $a(j,k) =$ the probability the dealer's ending total is k when he has a hard total of j, i.e., a hand in which any Ace is counted as 1.

We start by observing that $a(j,j) = b(j,j) = 1$ when $j \geq 17$ and then start with 16 and work down. Let $p_i = 1/13$ for $1 \leq i < 9$ and $p_{10} = 4/13$. If $11 \leq j \leq 16$ then a new Ace must count as 1 so

$$a(j,k) = p_1 a(j+1,k) + \sum_{m=2}^{10} p_m a(j+m,k)$$

When $2 \leq j \leq 10$ a new Ace counts as 11 and produces a soft hand:

$$a(j,k) = p_1 b(j+11,k) + \sum_{m=2}^{10} p_m a(j+m,k)$$

For soft hands, an Ace counts as 11, so there are no soft hands with totals of less than 12. If the card we draw takes us over 21 then we have to change the Ace from counting 11 to counting 1, producing a hard hand, so

$$b(j,k) = p_1 b(j+1,k) + \sum_{m=2}^{21-j} p_m b(j+m,k) + \sum_{m=22-j}^{10} p_m a(j+m-10,k)$$

When $j = 11$ the second sum runs from 11 to 10 and is considered to be 0.

The last three formulas are too complicated tow rok with by hand but are easy to manipulate using a computer. The next table gives the probabilities of the various results for the dealer conditional on the value of his first card. We have broken things down this way because when blackjack is played in a casino, we can see one of the dealer's two cards.

	17	18	19	20	21	bust
2	.13981	.13491	.12966	.12403	.11799	**.35361**
3	.13503	.13048	.12558	.12033	.11470	**.37387**
4	.13049	.12594	.12139	.11648	.11123	**.39447**
5	.12225	.12225	.11770	.11315	.10825	**.41640**
6	.16544	.10627	.10627	.10171	.09716	**.42315**
7	**.36857**	.13780	.07863	.07863	.07407	.26231
8	.12857	**.35934**	.12857	.06939	.06939	.24474
9	.12000	.12000	**.35076**	.12000	.06082	.22843
10	.11142	.11142	.11142	**.34219**	.11142	.21211
ace	.13079	.13079	.13079	.13079	**.36156**	.11529

You should note that when the dealer's upcard is 2, 3, 4, 5, or 6, her most likely outcome is to bust, but when her first card is $k = 7, 8, 9, 10$, or Ace $= 11$, her most likely total is $10 + k$. To make this clear we have given the most likely probabilities in boldface.

The analysis of the player's options is even more complicated than that of the dealer's so we will not attempt it here. The first analysis was performed in the mid-'50s (see Baldwin et al. in *J. Amer. Stat. Assoc.* 51, 429–439) and has been redone by a number of other people since that time. To describe the optimal strategy in a few words we use "stand on n" as short for "take a card if your total is $< n$ but not if it is $\geq n$."

Hard Hands

Stand on 17 if the dealer shows 7, 8, 9, 10, or A
Stand on 12 if the dealer shows 2, 3, 4, 5, or 6
Exception: Draw to 12 if the dealer shows 2 or 3

Soft Hands

Stand on 18
Exception: Draw to 18 if the dealer has 9 or 10

To help remember the rules for hard hands, observe that with two exceptions the strategy there is a combination of "mimic the dealer" and "never bust" (that is, "only take a card if you have 11 or less"), and it is exactly what we would do if the dealer's down card was a 10. If her up card is 7, 8, 9, 10, or A, then we must get to 17 to have a chance of winning. If her upcard is 2, 3, 4, 5, or 6 then we don't draw and hope that she busts.

Using these rules, the probability you will win is about 0.49, close enough to even if you are only looking for an evening's entertainment. You can reduce the house edge even further by learning about "doubling down" and splitting pairs.

Doubling Down. In this move you turn up your two cards, double your bet, and ask for one card to be dealt down to you. You are not allowed to ask for a second card if you don't like the first one. Double down

- if your total is 11
- if your total is 10 and the dealer's up card is 9 or less
- if your total is 9 and the dealer's up card is 3 through 6
- if you have A and 2 thhrough 7 and the dealer's up card is 4, 5, or 6

Again the doubling down rules can be explained by assuming that we are going to get a 10. Some Nevade casinos only allow doubling down on 11 or 10.

Splitting Pairs. If you have a pair you can split them, an extra card is dealt to each one, you place another bet on the table so there is one on each hand, and then play two hands separately.

- Always splits A's or 8's
- Never split 4's, 5's, or 10's
- Split 2's, 3's 6's, and 7's when the dealer's up card is 3 through 7.
- Split 9's when the dealer's up card is 2 through 6, 8 or 9.

The reason for splitting aces should be obvious. It is such a good play that it has on occasion been forbidden. Some casino rules do not allow further drawing after aces are split and if a 10 lands on the ace it is not a Blackjack. To see why 8's are singled out for splitting, note that $8 + 8 = 16$ which wins only if the dealer busts, while an 8 paired with a 10 produces an 18.

Counting cards. Edward Thorp's book *Beat the Dealer*, which astonished the world in 1962 by demonstrating that by "counting cards" (i.e., by keeping track of the difference between the numbers of cards you have seen that count 10 and those that count 2 through 6) and by adjusting your betting you can make money from blackjack. Before the reader plans a trip to Las Vegas or Atlantic City, we would like to point out that playing this strategy requires hard work, that making money with it requires a lot of capital, and that casinos are allowed to ask you to leave if they think you are playing it.

1.6 Exercises

Basic definitions

1. A man receives presents from his three children, Allison, Betty, and Chelsea. To avoid disputes he opens the presents in a random order. What are the possible outcomes?

2. Suppose we pick a number at random from the phone book and look at the last digit. (a) What is the set of outcomes and what probability should be assigned to each outcome? (b) Would this model be appropriate if we were looking at the first digit?

3. Two students arrive late for a math final exam with the excuse that their car had a flat tire. Suspicious, the professor says "each one of you write down on a piece of paper which tire was flat. What is the probability that both students pick the same tire?

4. Suppose we roll a red die and a green die. What is the probability the number on the red die is larger (>) than the number on the green die?

5. Two dice are rolled. What is the probability (a) the two numbers will differ by 1 or less, (b) the maximum of the two numbers will be 5 or larger?

6. If we flip a coin 5 times, what is the probability that the number of heads is an even number (i.e., divisible by 2)?

7. The 1987 World Series was tied at two games a piece before the St. Louis Cardinals won the fifth game. According to the Associated Press, "The number of history support the Cardinals and the momentum they carry. Whenever the series has been tied 2-2 the team that won the fifth game won the series 71% of the time." If momentum is not a factor and each team has a 50% chance of winning each game, what the probaiblity that the game 5 winner will win the series?

8. Two boys are repeatedly playing a game that they each have probability 1/2 of winning. The first person to win five games wins the match. What is the probability that Al will win if (a) he has won 4 games and Bobby has won 3; (b) he leads by a score of 3 games to 2?

9. Two red cards and two black cards are lying face down on the table. You pick two cards and turn them over. What is the probability that the two cards are different colors?

10. 20 families live in a neighborhood. 4 have 1 child, 8 have 2 children, 5 have 3 children, and 3 have 4 children. If we pick a child at random what is the probaiblity they come from a family with 1, 2, 3, 4 children.

11. In Galileo's time people thought that when three dice were rolled, a sum of 9 and a sum of 10 had the same probability since each could be obtained in 6 ways:

$$9 : 1+2+6,\ 1+3+5,\ 1+4+4,\ 2+2+5,\ 2+3+4,\ 3+3+3$$

$$10 : 1+3+6,\ 1+4+5,\ 2+4+4,\ 2+3+5,\ 2+4+4,\ 3+3+4$$

Compute the probabilities of these sums and show that 10 is a more likely total than 9.

12. Suppose we roll three dice. Compute the probability that the sum is (a) 3, (b) 4, (c) 5, (d) 6, (e) 7, (f) 8.

13. In a group of students, 25% smoke cigarettes, 60% drink alcohol, and 15% do both. What fraction of students have at least one of these bad habits?

14. In a group of 320 high school graduates, only 160 went to college but 100 of the 170 men did. How many women did not go to college?

15. In the freshman class, 62% of the students take math, 49% take science, and 38% take both science and math. What percentage takes at least one science or math course?

16. 24% of people have American Express Cards, 61% have VISA cards and 8% have both. What percentage of people have at least one credit card?

17. Suppose $\Omega = \{a, b, c\}$, $P(\{a, b\}) = 0.7$, and $P(\{b, c\}) = 0.6$. Compute the probabilities of $\{a\}$, $\{b\}$, and $\{c\}$.

18. Suppose A and B are disjoint with $P(A) = 0.3$ and $P(B) = 0.5$. What is $P(A^c \cap B^c)$?

19. Given two events A and B with $P(A) = 0.4$ and $P(B) = 0.7$. What are the maximum and minimum possible values for $P(A \cap B)$?

Permutations and Combinations

20. How many possible batting orders are there for nine baseball players?

21. A tire manufacturer wants to test four different types of tires on three different types of roads at five different speeds. How many tests are required?

22. 16 horses race in the Kentucky Derby. How many possible results are there for win, place, and show (first, second, and third)?

23. A school gives awards in five subjects to a class of 30 students but no one is allowed to win more than one award. How many outcomes are possible?

24. A tourist wants to visit six of America's ten largest cities. In how many ways can she do this if the order of her visits is (a) important, (b) not important?

25. Five businessmen meet at a convention. How many handshakes are exchanged if each shakes hands with all the others?

26. A commerical for Glade Plug-ins says that by using inserting two of a choice of 11 scents you can make more than 50 combinations. If we include the boring choice of two of the same scent how many possibilities are there?

27. In a class of 19 students, 7 will get A's. In how many ways can this set of students be chosen?

28. (a) How many license plates are possible if the first three places are occupied by letters and the last three by numbers? (b) Assuming all combinations are equally likely, what is the probability the three letters and the three numbers are different?

29. How many four-letter "words" can you make if no letter is used twice and each word must contain at least one vowel (A, E, I, O or U)?

30. Assuming all phone numbers are equally likely, what is the probability that all the numbers in a seven-digit phone number are different?

31. A domino is an ordered pair (m, n) with $0 \leq m \leq n \leq 6$. How many dominoes are in a set if there is only one of each?

32. A person has 12 friends and will invite 7 to a party. (a) How many choices are possible if Al and Bob are fueding and will not both go to the party? (b) How many choices are possible if Al and Betty insist that they both go or neither one goes?

33. A basketball team has 5 players over six feet tall and 6 who are under six feet. How many ways can they have their picture taken if the 5 taller players stand in a row behind the 6 shorter players who are sitting on a row of chairs?

34. Six students, three boys and three girls, lineup in a random order for a photograph. What is the probability that the boys and girls alternate?

35. Seven people sit at a round table. How many ways can this be done if Mr. Jones and Miss Smith (a) must sit next to each other, (b) must not sit next to each other? (Two seating patterns that differ only by a rotation of the table are considered the same).

36. How many ways can four rooks be put on a chessboard so that no rook can capture any other rook? Or, what is the same: How many ways can eight markers be placed on an 8×8 grid of squares so that there is at most one in each row or column?

Multinomial

37. How many different ways can the letters in the following words be arranged: (a) money, (b) banana, (c) statistics, (d) mississippi?

38. Twelve different toys are to be divided among 3 children so that each one gets 4 toys. How many ways can this be done?

39. A club with 50 members is going to form two committees, one with 8 members and the other with 7. How many ways can this be done (a) if the committees must be disjoint? (b) if they can overlap?

40. If seven dice are rolled, what is the probability that each of the six numbers will appear at least once?

41. How many ways can 5 history books, 3 math books, and 4 novels be arranged on a shelf if the books of each type must be together?

42. Suppose three runners from team A and three runners from team B have a race. If all six runners have equal ability, what is the probability that the three runners from team A will finish first, second, and fourth?

43. Four men and four women are shipwrecked on a tropical island. How many ways can they (a) form four male-female couples, (b) get married if we keep track of the order in which the weddings occur, (c) divide themselves into four unnumbered pairs, (d) split up into four groups of two to search the North, East, West, and South shores of the island, (e) walk single-file up the ramp to the ship when they are rescued, (f) take a picture to remember their ordeal if all eight stand in a line but each man stands next to his wife?

Urn problems

44. Four people are chosen at random from 5 couples. What is the probability two men and two women are selected?

45. You pick 5 cards out of a deck of 52. What is the probability you get exactly 2 spades?

46. Seven students are chosen at random from a class with 17 boys and 13 girls. What is the probability that 4 boys and 3 girls are selected?

47. In a carton of 12 eggs, 2 are rotten. If we pick 4 eggs to make an omelet, what is the probability we do not get a rotten egg?

48. A scrabble set contains 54 consonants, 44 vowels, and 2 blank tiles. Find the probability that your initial draw contains 5 consonants and 2 vowels.

49. (a) How many ways can can we pick 4 students from a group of 40 to be on the math team? (b) Suppose there are 18 boys and 12 girls. What is the probability the team will have 2 boys and 2 girls.

50. The following probability problem arose in a court case concerning possible discrimation against black nurses. 26 white nurses and 9 black nurses took an exam. All the white nurses passed but only 4 of the black nurses did. What is the probability we would get this outcome if the five nurses who failed were chosen at random?

51. A closet contains 8 pairs of shoes. You pick out 5. What is the probability of (a) no pair, (b) exactly one pair, (c) two pairs?

52. A drawer contains 10 black, 8 brown, and 6 blue socks. If we pick two socks at random, what is the probability they match?

53. A dance class consists of 12 men and 10 women. Five men and five women are chosen and paired up to dance. In how many ways can this be done?

54. Suppose we pick 5 cards out of a deck of 52. What is the probability we get at least one card of each suit?

55. A bridge hand in which there is no card higher than a 9 is called a *Yarborough* after the Earl who liked to bet at 1000 to 1 that your bridge hand would have a

card that was 10 or higher. What is the probability of a Yarborough when you draw 13 cards out of a deck of 52.

56. Two cards are a blackjack if one is an A and the other is a K, Q, J, or 10. (a) If you pick two cards out of a deck, what is the probability you will get a blackjack? (b) Suppose you are playing blackjack against the dealer with a freshly shuffled deck. What is the probability that you or the dealer will get a black jack.

57. In Keno the casino picks 20 balls from a set of 80 numbered 1 through 80. Before this drawing is done, you pick 10 numbers. What is the probability that exactly 5 of your numbers will be in the 20 selected?

58. A student studies 12 problems from which the professor will randomly choose 6 for a test. If the student can solve 9 of the problems, what is the probability she can solve at least 5 of the problems on the test?

59. A football team has 16 seniors, 12 juniors, 8 sophomores, and 4 freshmen. If we pick 5 players at random, what is the probability we will get 2 seniors and 1 from each of the other 3 classes?

60. In a kindergarten class of 20 students, one child is picked each day to help serve the morning snack. What is the probability that in one week five different children are chosen?

61. An investor picks 3 stocks out of 10 recommended by his broker. Of these, six will show a profit in the next year. What is the probability the investor will pick (a) 3 (b) 2 (c) 1 (d) 0 profitable stocks?

62. Four red cards (i.e., hearts and diamonds) and four black cards are face down on the table. A psychic who claims to be able to locate the four black cards turns over 4 cards and gets 3 black cards and 1 red card. What is the probability he would do this if he were guessing?

63. A town council considers the question of closing down an “adult” theatre. The five men on the council all vote against this and the three women vote in favor. What is the probability we would get this result (a) if the council members determined their votes by flipping a coin? (b) if we assigned the five “no” votes to council members chosen at random?

64. An urn contains white balls numbered 1 to 15 and black balls also numbered 1 to 15. Suppose you draw 4 balls. What is the probability that (a) no two have the same number? (b) you get exactly one pair with the same number? (c) you get two pair with the same numbers?

65. A town has four TV reparimen. In the first week of September four TV sets break and their owners call repairmen chosen at random. Find the probability that the number of repairmen who have jobs is 1, 2, 3, 4.

66. **Poker dice.** Compute the probabilities of the following poker hands when we roll five six sided dice.

(a) five of a kind	0.000771
(b) four of a kind,	0.019290
(c) a full house,	0.038580
(d) three of a kind	0.154320
(e) two pair	0.231481
(f) one pair	0.462962
(g) no pair	0.092592

67. In seven-card stud you receive seven cards and use them to make the best poker hand you can. Ignoring the possibility of a straight or a flush the probability that the best hand you can make with your cards is

	seven cards	five cards
(a) four of a kind,	0.001680	0.000240
(b) a full house,	0.025968	0.001441
(c) three of a kind	0.049254	0.021128
(d) two pair	0.240113	0.047539
(e) one pair	0.472839	0.422569
(f) no pair	0.210150	0.507082

Verify the probabilities for seven card stud. Hint: For full house you need to consider hand patterns: 3-3-1 and 3-2-2 in addition to the more likely 3-2-1-1. For two pair you also have to consider the possibility of three pair.

Inclusion-exclusion

68. Six high school teams play each other in the Southern Tier division. Each team plays all of the other teams once. What is the probability some team has a perfect $5-0$ season?

69. Suppose you draw seven cards out of a deck of 52. What is the probability you will have (a) exactly five cards of one suit? (b) at least five cards of one suit?

70. In a certain city 60% of the people subscribe to newspaper A, 50% to B, 40% to C, 30% to A and B, 20% to B and C, and 10% to A and C, but no one subscribes to all three. What percentage subscribe to (a) at least one newspaper, (b) exactly one newspaper?

71. Santa Claus has 45 drums, 50 cars, and 55 baseball bats in his sled. 15 boys will get a drum and a car, 20 a drum and a bat, 25 a bat and a car, and 5 will get three presents. (a) How many boys will receive presents? (b) How many boys will get just a drum?

72. Use the inclusion-exclusion formula to compute the probability that a randomly chosen number between 0000 and 9999 contains at least one 1. Check this by computing the probability there is no 1.

73. Ten people call an electrician and ask him to come to their houses on randomly chosen days of the work week (Monday through Friday). What is the probability of $A =$ "He has at least one day with no jobs"?

74. We pick a number between 0 and 999, then a computer picks one at random from that range. Use (1.11) to compute the probability at least two of our digits will match the computer's number. (Note: We include any leading zeros, so 017 and 057 have two matching digits.)

75. You pick 13 cards out of a deck of 52. What is the probability that you will not get a card from every suit?

76. You pick 13 cards out of a deck of 52. Let $A =$ "You have exactly six cards in at least one suit" and $B =$ "You have exactly six spades." The first Bonferroni inequality says that $P(A) \leq 4P(B)$. Compute $P(A)$ and $P(A)/P(B)$.

77. Use the first two Bonferroni inequalities to compute an upper and a lower bound on the probability that in a group of 60 people, at least 3 were born on the same day.

78. Suppose we roll two dice 6 times. Use the first three Bonferroni inequalities to compute bounds on the probability of $A =$ "We get at least one double 6." Compare the bounds with the exact answer $1 - (35/36)^6$.

79. Suppose we try 20 times for an event with probability 0.01. Use the first three Bonferroni inequalities to compute bounds on the probability of one success.