2

Combinatorial Probability

2.1 Permutations and combinations

As usual we begin with a question:

Example 2.1 The New York State Lottery picks 6 numbers out of 59, or more precisely, a machine picks 6 numbered ping-pong balls out of a set of 59. How many outcomes are there? The set of numbers chosen is all that is important. The order in which they are 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 2.2 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, in how many ways can he get dressed?

We begin by noting that there are $4 \cdot 6 = 24$ possible combinations of pants and shirts. Each of these can be paired with 1 of 8 choices of socks, so there are $24 \cdot 8 = 192$ ways of putting on pants, shirts, and socks. Repeating the last argument one more time, we see that for each of these 192 combinations there are 3 choices of shoes, so the answer is

$$4 \cdot 6 \cdot 8 \cdot 3 = 576 \text{ ways}$$

The reasoning in the last solution can clearly be extended to more than 4 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 $n_1 \cdot n_2 \cdot \cdots \cdot n_m$.

Example 2.3 How many ways can 5 people stand in line?

To answer this question, we think about building the line up 1 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 on 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

$$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 \tag{2.1}$$

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 and $n-1$ 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

The number of ways we can put the 22 volumes of an encyclopedia on a shelf is

$$22! = 1.24000728 \times 10^{21}$$

Here, we have used our TI-83. We typed in 22 and then used the MATH button to get to the PRB menu and scroll down to the fourth entry to get the !, which gives us 22! after we press ENTER.

The number of ways that cards in a deck of 52 can be arranged is

$$52! = 8.065817517 \times 10^{67}$$

Before there were calculators, people used **Stirling's formula**:

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n} \tag{2.2}$$

When $n = 52$, $52/e = 19.12973094$ and $\sqrt{2\pi n} = 18.07554591$, so

$$52! \approx (19.12973094)^{52} \cdot 18.07554591 = 8.0529 \times 10^{67}$$

Example 2.4 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

$$\frac{12}{P}\ \frac{11}{V}\ \frac{10}{S}\ \frac{9}{T} = 11{,}800$$

To compute $P_{12,4}$ with the TI-83 calculator, type 12, push the MATH button, move the cursor across to the PRB submenu, scroll down to nPr on the second row, and press ENTER. nPr appears on the screen after the 12. Now type 4 and press ENTER.

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

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

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 such as the last one come up so often that they have a name: the **number of permutations of** k **objects from a set of size** n, or $P_{n,k}$ for short. Multiplying and dividing by $(n-k)!$, we have

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

which gives us a short formula

$$P_{n,k} = \frac{n!}{(n-k)!} \tag{2.3}$$

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

Example 2.5 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 (2.3) 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

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

To compute $C_{23,4}$ with the TI-83 calculator, type 23, push the MATH button, move the cursor across to the PRB submenu, scroll down to nCr on the third row, and press ENTER. nCr appears on the screen after the 23. Now type 4 and press ENTER.

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

$$C_{n,k} = \frac{P_{n,k}}{k!} = \frac{n!}{k!(n-k)!} = \frac{n \cdot (n-1) \cdot \cdots \cdot (n-k+1)}{1 \cdot 2 \cdot \cdots \cdot k} \tag{2.4}$$

by (2.4) and (2.1). Here, $C_{n,k}$ is short for the **number of combinations of k things taken from a set of n**. $C_{n,k}$ 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, Example 2.1. We are choosing $k = 6$ objects out of a total of $n = 59$ when order is not important, so the number of possibilities is

$$\begin{aligned} C_{59,6} = \frac{59!}{6!53!} &= \frac{59 \cdot 58 \cdot 57 \cdot 56 \cdot 55 \cdot 54}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} \\ &= 59 \cdot 58 \cdot 19 \cdot 7 \cdot 11 \cdot 9 = 45{,}057{,}474 \end{aligned}$$

You should consider this the next time you think about spending \$1 for two chances to win a jackpot that starts at \$3 million and increases by \$1 million each week there is no winner.

Example 2.6 **World Series (continued).** Using (2.4) we can easily compute the probability that the series lasts 7 games. For this to occur the score must be tied 3–3 after 6 games. The total number of outcomes for the first 6 games is $2^6 = 64$. The number that ends 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 3 games, team A will win. This gives us a probability of $20/64 = 5/16$ for the series to end in 7 games. Returning to the calculation in the previous section, we see that the number of outcomes that lead to A winning in 6 games is the number of ways of picking two of the first 5 games for B to win or $C_{5,2} = 5!/(2!\,3!) = 5 \cdot 4/2 = 10$.

Example 2.7 Suppose we flip 5 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 head. 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 1 head. 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 2 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 numbers of outcomes for 3, 4, and 5 heads are 10, 5, and 1, or in general.

$$C_{n,m} = C_{n,n-m} \tag{2.5}$$

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 (2.4).

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 1's 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 0's. In symbols,

$$C_{n,k} = C_{n-1,k-1} + C_{n-1,k} \tag{2.6}$$

Verbal 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 things 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. □

Analytic proof. Using the definition (2.4),

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

Factoring out the parts common to the two fractions

$$= \frac{(n-1)!}{(n-k-1)!(k-1)!}\left(\frac{1}{n-k} + \frac{1}{k}\right)$$

$$= \frac{(n-1)!}{(n-k-1)!(k-1)!}\left(\frac{n}{(n-k)k}\right) = \frac{n!}{(n-k)!k!}$$

which proves (2.6). □

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

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

or in general

$$(x+y)^n = \sum_{m=0}^{n} C_{n,m} x^m y^{n-m} \tag{2.7}$$

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).)

(a) *One pair:* Two cards of equal value plus three cards with different values
J♠ J♢ 9♡ Q♣ 3♠
(b) *Two pair:* Two pairs plus another card with a different value
J♠ J♢ 9♡ 9♣ 3♠
(c) *Three of a kind:* Three cards of the same value and two with different values
J♠ J♢ J♡ 9♣ 3♠
(d) *Straight:* Five cards with consecutive values
5♡ 4♠ 3♠ 2♡ A♣
(e) *Flush:* Five cards of the same suit
K♣ 9♣ 7♣ 6♣ 3♣
(f) *Full house:* A three of a kind and a pair
J♠ J♢ J♡ 9♣ 9♠
(g) *Four of a kind:* Four cards of the same value plus another card
J♠ J♢ J♡ J♣ 9♠
(h) *Straight flush:* Five cards of the same suit with consecutive values
A♣ K♣ Q♣ J♣ 10♣

The last 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), and then 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), and 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*	0.422569
(b) *Two pair*	0.047539
(c) *Three of a kind*	0.021128
(d) *Straight*	0.003940
(e) *Flush*	0.001981
(f) *Full house*	0.001441
(g) *Four of a kind*	0.000240
(h) *Straight flush*	0.000015

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 5 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.

2.1.1 More than two categories

We defined $C_{n,k}$ as the number of ways of picking k objects 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 2.8 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. We first 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, and then paint the first 4 on the list yellow, the next 3 purple, and the last 5 red. One example is

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

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

Theorem 2.1. *The number of ways a group of n objects to be divided into m groups of size $n_1, \ldots, n_m$ with $n_1 + \cdots + n_m = n$ is*

$$\frac{n!}{n_1!\,n_2! \cdots n_m!} \tag{2.8}$$

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

Example 2.9 There are 39 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?

$$\frac{39!}{9!\,13!\,12!\,5!} = 1.57 \times 10^{22}$$

Example 2.10 **Bridge.** Four people play a card game in which each gets 13 cards. How many possible deals are there?

$$\frac{52!}{(13!)^4} = 5.364473777 \times 10^{28}$$

Example 2.11 Suppose we draw 13 cards from a deck. How many outcomes are there? How many lead to hands with 4 spades, 3 hearts, 3 diamonds, and 3 clubs? 3 spades, 5 hearts, 2 diamonds, and 3 clubs?

$$C_{52,13} = 6.350135596 \times 10^{11}$$
$$C_{13,4}C_{13,3}C_{13,3}C_{13,3} = 715 \cdot (286)^3 = 16,726,464,040$$
$$C_{13,3}C_{13,5}C_{13,2}C_{13,3} = 286 \cdot 1,287 \cdot 78 \cdot 286 = 8,211,173,256$$

Example 2.12 **Suit distributions.** The last bridge hand in the previous example is said to have a 5–3–3–2 distribution. Here, we have listed the number cards in the longest suit first and continued in decreasing order. Permuting the four numbers we see that the example 3 spades, 5 hearts, 2 diamonds, and 3 clubs is just one of $4!/2!$ possible ways of assigning the numbers to suits, so the probability of a 5–3–3–2 distribution is

$$\frac{12 \cdot 8{,}211{,}173{,}256}{6.350135596 \times 10^{11}} = 0.155$$

Similar computations lead to the results in the next table. We have included the number of different permutations of the pattern to help explain why slightly unbalanced distributions have larger probability than 4–3–3–3.

Distribution	Probability	Permutations
4–4–3–2	0.216	12
5–3–3–2	0.155	12
5–4–3–1	0.129	24
5–4–2–2	0.106	12
4–3–3–3	0.105	4
6–3–2–2	0.056	12

2.2 Binomial and multinomial distributions

Example 2.13 Suppose we roll 6 dice. What is the probability of A = "we get exactly two 4's"?

One way that A can occur is

$$\frac{\times}{1}\,\frac{4}{2}\,\frac{\times}{3}\,\frac{4}{4}\,\frac{\times}{5}\,\frac{\times}{6}$$

where $\times$ stands for "not a 4." Since the six events "die 1 shows $\times$," "die 2 shows 4," ..., "die 6 shows $\times$" are independent, the indicated pattern has probability

$$\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^4$$

Here, we have been careful to say "pattern" rather than "outcome" since the given pattern corresponds to 5^4 outcomes in the sample space of 6^6 possible outcomes for 6 dice. Each pattern that results in A corresponds to a choice of 2 of the 6 trials on which a 4 will occur, so the number of patterns is $C_{6,2}$. When we write out the probability of each pattern, there will be two 1/6s and four 5/6s, so each pattern has the same probability and

$$P(A) = C_{6,2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^4$$

Generalizing from the last example, suppose we perform an experiment n times and on each trial an event we call "success" has probability p. (Here and in what follows, when we repeat an experiment, we assume that the outcomes of the various trials are independent.) Then the probability of k successes is

$$C_{n,k}\, p^k (1-p)^{n-k} \tag{2.9}$$

This is called the **binomial(n, p) distribution**. Taking $n = 6, k = 2,$ and $p = 1/6$ in (2.9) gives the answer in the previous example. The reasoning for the general formula is similar. There are $C_{n,k}$ ways of picking k of the n trials for successes to occur, and each pattern of k successes and $n-k$ failures has probability $p^k(1-p)^{n-k}$.

Theorem 2.2. *The binomial(n, p) distribution has mean np and variance $np(1-p)$.*

Proof using theory. Let $X_i = 1$ if the ith trial is a success and 0 otherwise. $S_n = X_1 + \cdots + X_n$ is the number of successes in n trials. Using (1.8) we see that $E\,S_n = nE\,X_i = np$; that is, the expected number of successes is the number of trials n times the success probability p on each trial.

Since $X_1, \ldots, X_n$ are independent, (6.11) implies that $\text{var}\,(S_n) = \text{var}\,(X_1) + \cdots + \text{var}\,(X_n) = n\,\text{var}\,(X_1)$. To compute $\text{var}\,(X_1)$, we note that $E\,X_1^2 = 1 \cdot p + 0 \cdot (1-p) = p$, so $\text{var}\,(X_1) = E\,X_1^2 - (E\,X_1)^2 = p - p^2 = p(1-p)$ and the desired result follows. □

Proof by computation. Using the definition of expected value,

$$E\,N = \sum_{m=0}^{n} m \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

The $m = 0$ term contributes nothing, so we can cancel m's and rearrange to get

$$np \sum_{m=1}^{n} m \frac{(n-1)!}{(m-1)!(n-m)!} p^{m-1}(1-p)^{n-m} = np$$

since the sum computes the total probability for the binomial$(n-1,\ p)$ distribution.

As in the case of the geometric, our next step is to compute

$$E(N(N-1) = \sum_{m=2}^{n} m(m-1) \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}$$

We have dropped the first two terms that contribute nothing, so we can cancel to get

$$n(n-1)p^2 \sum_{m=2}^{n} m(m-1) \frac{(n-2)!}{(m-2)!(n-m)!} p^{m-2}(1-p)^{n-m}$$

since the sum computes the total probability for the binomial$(n-2,\ p)$ distribution.

To finish up, we note that

$$\begin{aligned} \text{var}\,(N) &= EN^2 - (EN)^2 = E(N(N-1)) = EN - (EN)^2 \\ &= n(n-1)p^2 + np - n^2p^2 = n(p - p^2) = np(1-p) \end{aligned}$$

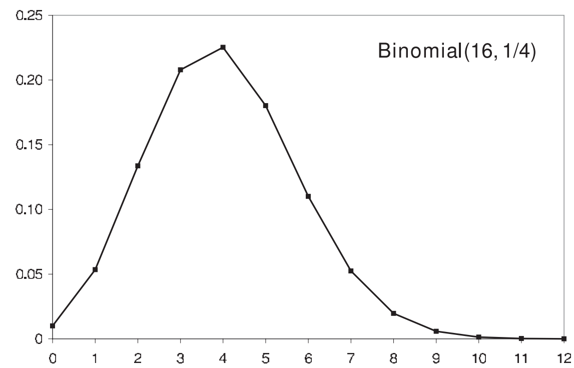
which completes the proof. □

Example 2.14 A student takes a test with 16 multiple-choice questions. Since she has never been to class she has to choose at random from the 4 possible answers. What is the probability that she will get exactly 3 right?

The number of trials is $n = 10$. Since she is guessing the probability of success $p = 1/4$, so using (2.9) the probability of $k = 3$ successes and $n - k = 7$ failures is

$$C_{16,3}(1/4)^3(3/4)^7 = \frac{16 \cdot 15 \cdot 14}{1 \cdot 2 \cdot 3} \cdot \frac{3^{13}}{4^{16}} = 560 \cdot \frac{1{,}594{,}323}{4{,}294{,}967{,}296} = 0.2079$$

In the same way we can compute the other probabilities. The results are given in next figure. With a TI-83 calculator these answers can be found by going to the DISTR menu and using binompdf$(16, 0.25, k)$. Here, pdf is short for probability density function.

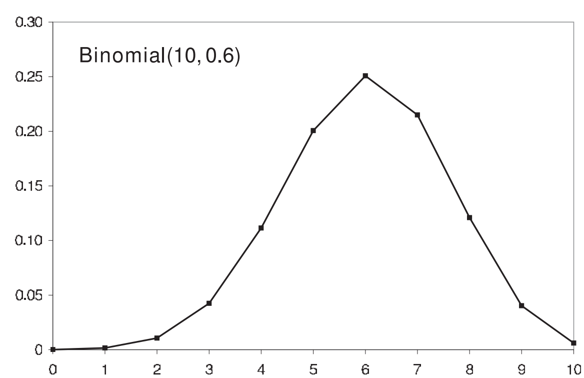


Example 2.15 A football team wins each week with probability 0.6 and loses with probability 0.4. If we suppose that the outcomes of their 10 games are independent, what is the probability that they will win exactly 8 games?

The number of trials is $n = 10$. We are told that the success probability $p = 0.6$, so by (2.9), the probability of $k = 8$ successes and $n - k = 2$ failures is

$$C_{10,8}(0.6)^8(0.4)^2 = \frac{10 \cdot 9}{1 \cdot 2}(0.6)^8(0.4)^2 = 0.1209$$

In the same way we can compute the other probabilities:



Example 2.16 **Aces at bridge.** When we draw 13 cards out of a deck of 52, each ace has a probability 1/4 of being chosen, but the four events are not independent. How does the probability of $k = 0, 1, 2, 3, 4$ aces compare with that of the binomial distribution with $n = 4$ and $p = 1/4$?

We first consider the probability of drawing two aces:

$$\frac{C_{4,2}C_{48,11}}{C_{52,13}} = \frac{6 \cdot \frac{48 \cdots 38}{11!}}{\frac{52 \cdots 40}{13!}} = 6 \cdot \frac{13 \cdot 12 \cdot 39 \cdot 38}{52 \cdot 51 \cdot 50 \cdot 49} = 0.2135$$

In contrast, the probability for the binomial is

$$C_{4,2}(1/4)^2(3/4)^2 = 0.2109$$

To compare the two formulas, note that $13/52 = 1/4$, $12/51 = 0.2352$, $39/50 = 0.78, 38/51 = 0.745$ versus $(1/4)^2(3/4)^2$ in the binomial formula. Similar computations show that if $D = 52 \cdot 51 \cdot 50 \cdot 49$, then the answers are

$$\begin{array}{ccc}
 & \text{Aces} & \text{Binomial} \\
0 & \dfrac{39 \cdot 38 \cdot 37 \cdot 36}{52 \cdot 51 \cdot 50 \cdot 49} & (3/4)^4 \\
1 & 4 \cdot \dfrac{13 \cdot 39 \cdot 38 \cdot 37}{52 \cdot 51 \cdot 50 \cdot 49} & 4 \cdot (1/4)(3/4)^3 \\
2 & 6 \cdot \dfrac{13 \cdot 12 \cdot 39 \cdot 38}{52 \cdot 51 \cdot 50 \cdot 49} & 6 \cdot (1/4)^2(3/4)^2 \\
3 & 4 \cdot \dfrac{13 \cdot 12 \cdot 11 \cdot 39}{52 \cdot 51 \cdot 50 \cdot 49} & 4 \cdot (1/4)^3(3/4) \\
4 & \dfrac{13 \cdot 12 \cdot 11 \cdot 10}{52 \cdot 51 \cdot 50 \cdot 49} & (1/4)^4
\end{array}$$

Evaluating these expressions leads to the following probabilities:

	Aces	Binomial
0	0.3038	0.3164
1	0.4388	0.4218
2	0.2134	0.2109
3	0.0412	0.0468
4	0.00264	0.00390

Example 2.17 In 8 games of bridge, Harry had 6 hands without an ace. Should he doubt that the cards are being shuffled properly?

The number of hands with no ace has a binomial distribution with $n = 8$ and $p = 0.3038$. The probability of at least 6 hands without an ace is

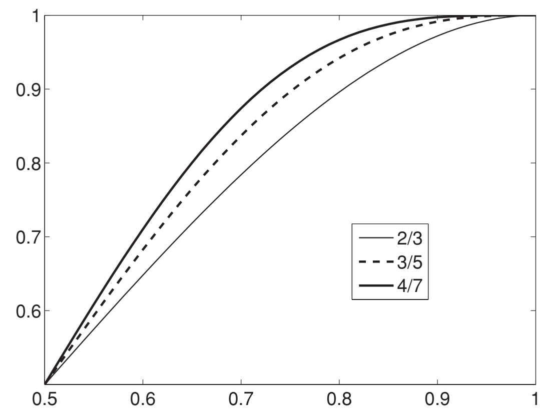
$$\sum_{k=6}^{8} C_{8,k}(0.3038)^k(0.6962)^{8-k} = 1 - \sum_{k=0}^{5} C_{8,k}(0.3038)^k(0.6962)^{8-k}$$

We have turned the probability around because on the TI-83 calculator the sum can be evaluated as binomcdf(8, 0.3038, 5) $= 0.9879$. Here, cdf is short for cumulative distribution function, that is, the probability of ≤ 5 hands without an ace. Thus the probability of luck this bad is 0.0121.

Example 2.18 **Tennis.** In men's tennis, the winner is the first to win 3 out of 5 sets. If Roger Federer wins a set against his opponent with probability 2/3, what is the probability w that he will win the match?

He can win in three sets, four or five, but he must win the last set, so

$$w = \left(\frac{2}{3}\right)^3 + C_{3,2}\left(\frac{2}{3}\right)^3 \frac{1}{3} + C_{4,2}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2$$
$$= \left(\frac{2}{3}\right)^3 (1 + 3(1/3) + 6(1/9)) = \frac{8}{27}\cdot\frac{8}{3} = 0.790$$



Replacing 2/3 by p and 1/3 by $(1 - p)$, we get the general solution for best 3 out of 5, which generalizes easily to other common formats:

2 out of 3: $p^2 + C_{2,1}p^2(1-p)$
3 out of 5: $p^3 + C_{3,2}p^3(1-p) + C_{4,2}p^3(1-p)^2$
4 out of 7: $p^4 + C_{4,3}p^4(1-p) + C_{5,3}p^4(1-p)^2 + C_{6,3}p^4(1-p)^3$

As we should expect, if $p > 1/2$, then the winning probability increases with the length of the series. When $p = 0.6$ the three values are 0.648, 0.68256, and 0.710208. The graph above compares the winning probabilities for a team that wins each game with probability p.

Multinomial distribution

The arguments above generalize easily to independent events with more than two possible outcomes. We begin with an example.

Example 2.19 Consider a die with 1 painted on three sides, 2 painted on two sides, and 3 painted on one side. If we roll this die 10 times, what is the probability that we get five 1's, three 2's and two 3's?

The answer is

$$\frac{10!}{5!\,3!\,2!}\left(\frac{1}{2}\right)^5\left(\frac{1}{3}\right)^3\left(\frac{1}{6}\right)^2$$

The first factor, by (2.8), gives the number of ways to pick five rolls for 1's, three rolls for 2's, and two rolls for 3's. The second factor gives the probability of any outcome with five 1's, three 2's, and two 3's. Generalizing from this example, we see that if we have k possible outcomes for our experiment with probabilities $p_1, \ldots, p_k$, then the probability of getting exactly n_i outcomes of type i in $n = n_1 + \cdots + n_k$ trials is

$$\frac{n!}{n_1! \cdots n_k!}\, p_1^{n_1} \cdots p_k^{n_k} \tag{2.10}$$

since the first factor gives the number of outcomes and the second the probability of each one.

Example 2.20 A baseball player gets a hit with probability 0.3, a walk with probability 0.1, and an out with probability 0.6. If he bats 4 times during a game and we assume that the outcomes are independent, what is the probability that he will get 1 hit, 1 walk, and 2 outs?

The total number of trials is $n = 4$. There are $k = 3$ categories: hit, walk, and out. $n_1 = 1$, $n_2 = 1$, and $n_3 = 2$. Plugging in to our formula the answer is

$$\frac{4!}{1!1!2!}(0.3)(0.1)(0.6)^2 = 0.1296$$

Example 2.21 The output of a machine is graded excellent 70% of the time, good 20% of the time, and defective 10% of the time. What is the probability that a sample of size 15 has 10 excellent, 3 good, and 2 defective items?

The total number of trials is $n = 15$. There are $k = 3$ categories: excellent, good, and defective. We are interested in outcomes with $n_1 = 10$, $n_2 = 3$, and $n_3 = 2$. Plugging in to our formula the answer is

$$\frac{15!}{10!\,3!\,2!} \cdot (0.7)^{10}(0.2)^3(0.1)^2$$

2.3 Poisson approximation to the binomial

X is said to have a **Poisson distribution** with parameter λ, or Poisson(λ) if

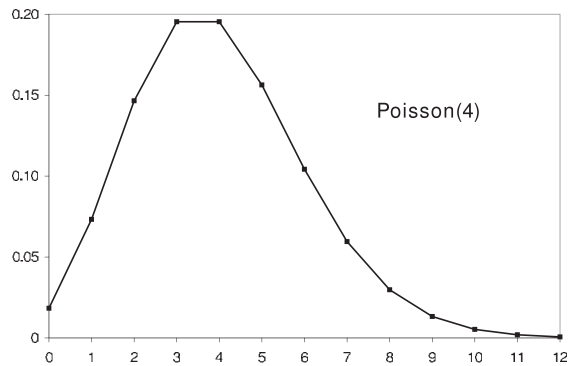
$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{for } k = 0, 1, 2, \ldots$$

Here, $\lambda > 0$ is a parameter. To see that this is a probability function we recall

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{2.11}$$

so the proposed probabilities are nonnegative and sum to 1.

The next figure shows the Poisson distribution with $\lambda = 4$.



Theorem 2.3. *The Poisson distribution has mean λ and variance λ.*

Proof. Since the $k = 0$ term makes no contribution to the sum,

$$EX = \sum_{k=1}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda}\frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

since $\sum_{k=1}^{\infty} P(X = (k-1)) = 1$. As in the case of the binomial and geometric our next step is to compute

$$E(X(X-1)) = \sum_{k=2}^{\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} = \lambda^2 \sum_{k=2}^{\infty} e^{-\lambda}\frac{\lambda^{k-2}}{(k-2)!} = \lambda^2$$

From this it follows that

$$\begin{aligned} \text{var}(X) &= EX^2 - (EX)^2 = E(X(X-1)) + EX - (EX)^2 \\ &= \lambda^2 + \lambda - \lambda^2 \end{aligned}$$

which completes the proof. □

Our next result explains why the Poisson distribution arises in a number of situations.

Theorem 2.4. *Suppose S_n has a binomial distribution with parameters n and p_n. If $p_n \to 0$ and $np_n \to \lambda$ as $n \to \infty$, then*

$$P(S_n = k) \to e^{-\lambda} \frac{\lambda^k}{k!} \tag{2.12}$$

In words, if we have a large number of independent events with small probability then the number that occurs has approximately a Poisson distribution. The key to the proof is the following fact.

Lemma. *If $p_n \to 0$ and $np_n \to \lambda$, then as $n \to \infty$*

$$(1 - p_n)^n \to e^{-\lambda} \tag{2.13}$$

Proof. Calculus tells us that if x is small then

$$\ln(1 - x) = -x - \frac{x^2}{2} - \cdots$$

Using this we have

$$\begin{aligned}(1 - p_n)^n &= \exp(n \ln(1 - p_n)) \\ &\approx \exp(-np_n - np_n^2/2) \approx \exp(-\lambda)\end{aligned}$$

In the last step we used the observation that $p_n \to 0$ to conclude that $np_n \cdot p_n/2$ is much smaller than np_n. □

Proof of (2.12). Since $P(S_n = 0) = (1 - p_n)^n$, (2.13) gives the result for $k = 0$. To prove the result for $k > 0$, we let $\lambda_n = np_n$ and observe that

$$\begin{aligned}P(S_n = k) &= C_{n,k} \left(\frac{\lambda_n}{n}\right)^k \left(1 - \frac{\lambda_n}{n}\right)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \frac{\lambda_n^k}{k!} \left(1 - \frac{\lambda_n}{n}\right)^n \left(1 - \frac{\lambda_n}{n}\right)^{-k} \\ &\to 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1\end{aligned}$$

Here, $n(n-1)\cdots(n-k+1)/n^k \to 1$ since there are k factors in the numerator, and for each fixed j, $(n-j)/n = 1 - (j/n) \to 1$. The last term $(1 - \{\lambda_n/n\})^{-k} \to 1$ since k is fixed and $1 - \{\lambda_n/n\} \to 1$. □

When we apply (2.12), we think, "If $S_n =$ binomial(n, p) and p is small then S_n is approximately Poisson(np)." The next example illustrates the use of this approximation and shows that the number of trials does not have to be very large for us to get accurate answers.

Example 2.22 Suppose we roll two dice 12 times and we let D be the number of times a double 6 appears. Here, $n = 12$ and $p = 1/36$, so $np = 1/3$. We now compare $P(D = k)$ with the Poisson approximation for $k = 0, 1, 2$.

$$k = 0 \text{ exact answer:} \quad P(D = 0) = \left(1 - \frac{1}{36}\right)^{12} = 0.7132$$

$$\text{Poisson approximation:} \quad P(D = 0) = e^{-1/3} = 0.7165$$

$$k = 1 \text{ exact answer:} \quad P(D = 1) = C_{12,1} \frac{1}{36} \left(1 - \frac{1}{36}\right)^{11}$$

$$= \left(1 - \frac{1}{36}\right)^{11} \cdot \frac{1}{3} = 0.2445$$

$$\text{Poisson approximation:} \quad P(D = 1) = e^{-1/3} \frac{1}{3} = 0.2388$$

$$k = 2 \text{ exact answer:} \quad P(D = 2) = C_{12,2} \left(\frac{1}{36}\right)^2 \left(1 - \frac{1}{36}\right)^{10}$$

$$= \left(1 - \frac{1}{36}\right)^{10} \cdot \frac{12 \cdot 11}{36^2} \cdot \frac{1}{2!} = 0.0384$$

$$\text{Poisson approximation:} \quad P(D = 2) = e^{-1/3} \left(\frac{1}{3}\right)^2 \frac{1}{2!} = 0.0398$$

Example 2.23 **Death by horse kick.** Ladislaus Bortkiewicz published a book about the Poisson distribution titled *The Law of Small Numbers* in 1898. In this book he analyzed the number of German soldiers kicked to death by cavalry horses between 1875 and 1894 in each of 14 cavalry corps, arguing that it fits the Poisson distribution. I. J. Good and others have argued that the Poisson distribution should be called the Bortkiewicz distribution, but then it would be very difficult to say or write.

Example 2.24 **V-2 rocket hits in London during World War II.** The area under study was divided into 576 areas of equal size. There were a total of 537 hits or an average of 0.9323 per subdivision. Using the Poisson distribution the probability a subdivision is not hit is $e^{-0.9323} = 0.3936$. Multiplying by 576 we see that the expected number not hit was 226.71, which agrees well with the 229 that were observed not to be hit.

Example 2.25 **Shark attacks.** In the summer of 2001 there were 6 shark attacks in Florida, while the yearly average is 2. Is this unusual?

In an article in the September 7, 2001, *National Post*, Professor David Kelton of Penn State University argued that this was a random event. “Just because you see events happening in a rash this does not imply that there is some physical driver causing them to happen. It is characteristic of random processes that they have bursty behavior." He did not seem to realize that the probability of six shark attacks under the Poisson distribution is

$$e^{-2}\frac{2^6}{6!} = 0.01203$$

This probability can be found with the TI-83 by using Poissonpdf(2, 6) on DISTR menu. If we want the probability of at least 6, we would use 1 − Poissoncdf(2, 5).

Example 2.26 **Alliteration in Shakespeare.** Did Shakespeare consciously choose words with the same sounds or did lines like "full fathom five thy father lies" just occur by chance. Psychologist B. F. Skinner addressed this question in two papers (one in 1939 in *The Psychological Record*, Vol. 3, pp. 186–192, and one in 1941 in *The American Journal of Psychology*, Vol. 54, pp. 64–79). He looked at 100 sonnets (for a total of 1,400 lines) and counted the number of times *s* sounds appeared in a line. The next table compares the counts to those expected under a Poisson distribution with the same mean.

s sounds	0	1	2	3	4
Observed	702	501	161	29	7
Expected	685	523	162	26	2

It turns out that most of the discrepancy in the last two cells goes away if the same word on a line is not counted more than once. However, even without this manipulation the similarity to the Poisson is remarkable. This example comes from a 1989 article by Diaconis and Mosteller on coincidences in the *Journal of the American Statistical Association*, Vol. 84, pp. 853–861.

Example 2.27 **Birthday problem, II.** If we are in a group of $n = 183$ individuals, what is the probability that no one else has our birthday?

By (2.13), the probability is

$$\left(1 - \frac{1}{365}\right)^{182} \approx e^{-182/365} = 0.6073$$

From this we see that in order to have a probability of about 0.5, we need $365 \ln 2 = 253$ people as we calculated before.

Example 2.28 **Birthday problem, I.** Consider now a group of $n = 25$ and ask our original question: What is the probability that two people have the same birthday?

The events $A_{i,j}$ that persons i and j have the same birthday are only pairwise independent, so strictly speaking (2.12) does not apply. However, it gives a reasonable approximation. The number of pairs of people is $C_{25,2} = 300$, while the probability of a match for a given pair is 1/365, so by (2.13) the probability of no match is

$$\approx \exp(-300/365) = 0.4395$$

versus the exact probability of 0.4313 from the table in Section 1.1.

General Poisson approximation result. The Poisson distribution is often used as a model for the number of people who go to a fast-food restaurant between 12 and 1, the number of people who make a cell phone call between 1:45 and 1:50, or the number of traffic accidents in a day. To explain the reasoning in the last case we note that any one person has a small probability of having an accident on a given day, and it is reasonable to assume that the events $A_i =$ "the ith person has an accident" are independent. However, it is not reasonable to assume that the probabilities of having an accident $p_i = P(A_i)$ are all the same, nor is it reasonable to assume that all women have the same probability of giving birth, but fortunately the Poisson approximation does not require this.

Theorem 2.5. *Consider independent events A_i, $i = 1, 2, \ldots, n$, with probabilities $p_i = P(A_i)$. Let N be the number of events that occur, let $\lambda = p_1 + \cdots + p_n$, and let Z have a Poisson distribution with parameter λ. Then, for any set of integers B,*

$$|P(N \in B) - P(Z \in B)| \le \sum_{i=1}^{n} p_i^2 \tag{2.14}$$

We can simplify the right-hand side by noting

$$\sum_{i=1}^{n} p_i^2 \le \max_i p_i \sum_{i=1}^{n} p_i = \lambda \max_i p_i$$

This says that if all the p_i are small then the distribution of N is close to a Poisson with parameter λ. Taking $B = \{k\}$, we see that the individual probabilities $P(N = k)$ are close to $P(Z = k)$, but this result says more. The probabilities of events such as $P(3 \le N \le 8)$ are close to $P(3 \le Z \le 8)$ and we have an explicit bound on the error.

For a concrete situation, consider Example 2.22, where $n = 12$ and all the $p_i = 1/36$. In this case the error bound is

$$\sum_{i=1}^{12} p_i^2 = 12\left(\frac{1}{36}\right)^2 = \frac{1}{108} = 0.00926$$

while the error for the approximation for $k = 1$ is 0.0057.

Example 2.29 The previous example justifies the use of the Poisson distribution in modeling the number of visits to a Web site in a minute. Suppose that the average number of visitors per minute is $\lambda = 5$, but that the site will crash if there are 12 visitors or more. What is the probability that the site will crash?

The probability of 12 or more visitors is

$$1 - \sum_{k=0}^{11} e^{-5} \frac{5^k}{k!}$$

This would be tedious to do by hand, but is easy if we use the TI-83 calculator. Using the distributions menu,

$$\sum_{k=0}^{11} e^{-5} \frac{5^k}{k!} = \text{Poissoncdf}(5, 11) = 0.994547$$

Subtracting this from 1, we have the answer 0.005453.

Example 2.30 **Births in Ithaca.** The Poisson distribution can be used for births as well as for deaths. There were 63 births in Ithaca, NY, between March 1 and April 8, 2005, a total of 39 days, or 1.615 per day. The next table gives the observed number of births per day and compares with the prediction from the Poisson distribution.

	0	1	2	3	4	5	6
Observed	9	12	9	5	3	0	1
Poisson	7.75	12.52	10.11	5.44	2.19	0.71	0.19

Example 2.31 **Wayne Gretsky.** He scored a remarkable 1,669 points in 696 games as an Edmonton Oiler, for a rate of $1{,}669/696 = 2.39$ points per game. From the Poisson formula with $k = 0$, the probability of Gretzky having a pointless game is $e^{-2.39} = 0.090$. The next table compares that actual number of games with the numbers predicted by the Poisson approximation.

Points	Games	Poisson
0	69	63.27
1	155	151.71
2	171	181.90
3	143	145.40
4	79	87.17
5	57	41.81
6	14	16.71
7	6	5.72
8	2	1.72
9	0	0.46

Coincidences. If we see an event with a chance of one in a hundred million then we are amazed, even though each day in the United States such events will happen to three people. This observation helps explain why some incredible things occur. Some times, as in our first example, it is just a miscalculation.

Example 2.32 **Pick 4 coincidence.** To quote a United Press story on September 10, 1981, reported by James Hanley in his article in the *American Statistician*, Vol. 46, pp. 197–202: "Lottery officials say that there is 1 chance in 100 million that the same four digit lottery number would be drawn in Massachusetts and New York on the same night. That's just what happened Tuesday. The number 8902 came up paying $5842 in Massachusetts and $4,500 New York."

These lotteries pick four-digit numbers, so each number has a 1 in 10^4 chance. The probability that 8902 is chosen in both states is 1 in 10^8 but then again some number will be chosen in Massachusetts and after that has been done the probability that the same number is chosen in New York is 1 in 10^4. When you take into account that there are a half dozen states that have similar games and drawings occur twice a day in New York, one should expect this to happen every few years.

Example 2.33 **Lottery double winner.** The following item was reported in the February 14, 1986, edition of the *New York Times*: A New Jersey woman, Evelyn Adams, won the lottery twice within a span of 4 months raking in a total of 5.4 million dollars. She won the jackpot for the first time on October 23, 1985, in the Lotto 6/39, in which you pick 6 numbers out of 39. Then she won the jackpot in the new Lotto 6/42 on February 13, 1986. Lottery officials calculated the probability of this as roughly one in 17.1 trillion. What do you think of this statement?

It is easy to see where they get this from. The probability of a person picked in advance of the lottery getting all six numbers right both times is

$$\frac{1}{C_{39,6}} \cdot \frac{1}{C_{42,6}} = \frac{1}{17.1 \times 10^{12}}$$

One can immediately reduce the odds against this event by noting that the first lottery had some winner, who if they played only one ticket in the second lottery had a $1/C_{42,6}$ chance.

The odds drop even further when you consider that there are a large number of people who submit more than one entry for the twice weekly drawing and that wins on October 23, 1985, and February 13, 1986, is not the only combination. Suppose for concreteness that each week 1 million people play the lottery and each buys exactly five tickets. The probability of one person winning on a given week is

$$p_1 = \frac{5}{C_{42,6}} = 9.531 \times 10^{-7}$$

The number of times one person will win a jackpot in the next year (100 twice-weekly drawings) is roughly Poisson with mean

$$\lambda_1 = 100\, p_1 = 9.531 \times 10^{-5}$$

The probability that a given player wins the jackpot two or more times is

$$p_0 = 1 - e^{-\lambda_1} - e^{-\lambda_1}\lambda_1 = 4.54 \times 10^{-9}$$

The number of double winners in a population of 1 million players is Poisson with mean

$$\lambda_0 = (1{,}000{,}000)\, p_0 = 4.54 \times 10^{-3}$$

so the probability of at least one double winner is $1 - e^{-0.00454} \approx 0.00454$. If you take into account that many states have lotteries and we were just looking at 1 year, we see that a double winner is not unusual at all.

My favorite double-winner story is Maureen Wilcox. In June 1980, she bought tickets for both the Massachusetts Lottery and the Rhode Island Lottery. She picked the winning numbers for both lotteries. Unfortunately for her, her Massachusetts numbers won in Rhode Island and vice versa.

Example 2.34 **Scratch-off triple winner.** 81-year-old Keith Selix won three lottery prizes totaling \$81,000 from scratch-off games in the 12 months preceding May 3, 2006. He won \$30,00 twice in "Wild Crossword" games and \$21,000 playing "Double Blackjack." Again we want to calculate the probability of this.

The odds of winning in these games are 89,775 to 1 and 119,700 to 1 respectively. One of the reasons Selix won so many times in 2006 is that he spent about \$200 a week or more than \$10,000 a year on scratch-off games. If the games cost \$1 then this would be 10,000 plays with an approximate 1/100,000 chance of winning. Thus his expected number of wins would be 0.1 and the probability of exactly three wins would be

$$e^{-0.1}\frac{(0.1)^3}{3!} \quad \text{or} \quad < \frac{1}{60{,}000}$$

Example 2.35 **Sally Clark.** Sometimes coincidences are not happy events like lottery wins. In 1999, a British jury convicted Sally Clark of murdering her two children who had died suddenly at the ages of 11 and 8 weeks respectively of sudden infant death syndrome or "cot deaths." There was no physical or other evidence of a murder, nor was there a motive. Most likely, the jury was convinced by a pediatrician who said that a baby had a probability of roughly 1/8,500 of dying a cot death, so having two children die this way had probability roughly 1/73,000,000.

There are two problems with the computation: (i) Many families have children who die this way, so the first factor of 1/8,500 should be dropped. (ii) Two cot deaths in the same family are not independent events; once one occurs the second child faces an increased risk of about 1/100 of dying this way. Thus the probability that it would happen again without foul play is 1/100. If this number had been presented to the jury, Sally probably would not have had to spend 3 years in jail before the verdict was overturned.

2.4 Card games and other urn problems

A number of problems in probability have the following form.

Example 2.36 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 $P(B) = 1{,}848/4{,}845 = 0.3814$.

We now consider two gambling games in which numbered balls are picked out of urns.

Example 2.37 **New York State lottery.** As mentioned in Section 2.1, if there are 59 numbered balls and 6 are picked, the number of outcomes is

$$C_{59,6} = 45{,}057{,}474$$

In the lottery you do win some money if at least three of your six numbers are chosen. The problem is to compute the probability of winning these other prizes.

Five out of six has probability

$$\frac{C_{6,5}C_{53,1}}{C_{59,6}} = \frac{6 \cdot 53}{C_{59,6}} = \frac{1}{141{,}690}$$

Four out of six has probability

$$\frac{C_{6,4}C_{53,2}}{C_{59,6}} = \frac{15 \cdot 1{,}378}{C_{59,6}} = \frac{1}{2{,}180}$$

Three out of six has probability

$$\frac{C_{6,3}C_{53,3}}{C_{59,6}} = \frac{20 \cdot 23{,}426}{C_{59,6}} = \frac{1}{96}$$

To add more prizes, a bonus number has been added to the card. You win if you match 5 out of 6 and also the bonus number that has probability

$$\frac{C_{6,5}C_{52,1}}{C_{59,6}}\frac{1}{52} = \frac{6}{C_{59,6}} = \frac{1}{7{,}509{,}579}$$

It is hard to compute the expected value because the rewards for 6 out of 6 and 5 out of 6 plus the bonus number depend on the number of weeks that there has been no winner, and all prizes with the exception of 3 out of 6 which always pays \$1 depend on both the number of people who play and the number of people who win. However, one can get some idea of the expected value by noting that 54.7% of the money bet is returned in prizes, 32.9% to education, and 12.4% to various operating expenses.

The next table gives data for the number of winners and winning amounts for the month of January 2008.

Date	6	5+	5	4	3
1/2	10M	503,347	2,284	23	1
	0	0	15	1,601	36,719
1/5	11M	557,243	1,635	29	1
	0	1	25	1,502	36,203
1/9	12M	46,833	2,200	34	1
	1	0	16	1,128	26,319
1/12	3M	96,488	2,354	32	1
	0	1	16	1,299	30,689
1/16	4M	41,839	2,645	36	1
	0	0	12	951	24,147
1/19	5M	91,440	984	21	1
	0	0	37	1,909	42,927
1/23	6M	133,696	1,885	26	1
	0	0	17	1,319	31,131
1/26	7M	184,664	1,104	26	1
	0	1	34	1,582	37,648
1/30	8M	38,956	461	11	1
	2	1	60	2,700	45,386

As you can see from the table, the big prize starts at 3 million and increases by 1 million each week when there are no winners. The 5 out of 6 plus bonus number is about 40,000 times the number of weeks since the previous winner. One can get a pretty good idea of the number of people who played each week by multiplying the number of 3 out of 6 winners by 96 (or 100 which is easier).

Example 2.38 **Keno.** In this game the casino picks 20 balls out of 80 numbered balls. Before the draw you may, for example, pick 10 numbers and bet \$1. In this case you win \$1 if 4 of your numbers are chosen; \$2 for 5; \$20 for 6; \$105 for 7; \$500 for 8; \$5,000 for 9; and \$12,000 if all 10 are chosen. We want to compute the expected value of the bet.

The number of possible draws is astronomically large:

$$C_{80,20} = 3.5353 \times 10^{18}$$

The probability that k of your numbers are chosen is

$$p_k = \frac{C_{10,k}C_{70,20-k}}{C_{80,20}}$$

When $k = 0$, this is

$$\frac{C_{70,20}}{C_{80,20}} = \frac{70!60!}{80!50!} = \frac{60 \cdot 59 \cdot \cdots \cdot 51}{80 \cdot 79 \cdot \cdots \cdot 71} = 0.045791$$

To compute the other probabilities it is useful to note that for $1 \le m \le n$,

$$\begin{aligned} C_{n,m} = \frac{n!}{m!(n-m)!} &= \frac{n+1-m}{m} \cdot \frac{n!}{(m-1)!(n+1-m)!} \\ &= \frac{n+1-m}{m} \cdot C_{n,m-1} \end{aligned}$$

so we have

$$p_k = p_{k-1} \cdot \frac{11-k}{k} \cdot \frac{21-k}{50+k}$$

Writing w_k for the winning when k of our numbers are drawn, using this recursion and the result for p_0 gives

k	p_k	w_k	$w_k p_k$
0	0.045791		0
1	0.179571		0
2	0.295257		0
3	0.267402		0
4	0.147319	1	0.147319
5	0.051428	2	0.102855
6	0.011479	20	0.229588
7	0.001611	105	0.169701
8	0.000135	500	0.067710
9	6.12×10^{-6}	5,000	0.030603
10	1.12×10^{-7}	12,000	0.001347
4–10	0.2120		0.7486

Thus, we win something about 21.2% of the time and our average winning is a little less than 75 cents, a typical expected value for Keno bets. The last column shows the contribution of the different payoffs to the expected value.

Example 2.39 **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 8 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 nonhearts, so in the 26 that remain there are $13 - 8 = 5$ hearts and $39 - 18 = 21$ nonhearts. To construct a hand for West with 3 hearts and 10 nonhearts we must pick 3 of the 5 hearts, which can be done in $C_{5,3}$ ways, and 10 of the 21 nonhearts 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; that is, West might have 3 or 1.

Example 2.40 **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 (a court case *De Martini v. Power*), 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

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

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. In election considered in *Ipolito v. Power*, the winning margin was 1,422 to 1,405 but 101 votes had to be thrown out. The judge rules that "it does not strain the probabilities to assume a likelihood that the questioned votes produced or could produce a change in the result." Do you agree with this assessment? We return to this question in Example 6.22.

Example 2.41

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 that 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

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

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

Example 2.42

Capture–recapture experiments. An ecology graduate student goes to a pond and captures $k = 60$ 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 beetles?

To turn this into a precisely formulated problem, we suppose that no beetles enter or leave the population between the two visits. With this assumption, if there were N 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 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 $N \leq kr/m$. 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, equals the proportion of marked beetles in the sample, m/r. Plugging in the numbers from our example, $kr/m = (60 \cdot 50)/12 = 250$, so the probability is maximized when $N = 250$.

2.5 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 extend this formula to $n > 2$ events. We begin with $n = 3$ events:

$$\begin{aligned} P(A \cup B \cup C) = P(A) &+ P(B) + P(C) \\ &- P(A \cap B) - P(A \cap C) - P(B \cap C) \\ &+ P(A \cap B \cap C) \end{aligned} \tag{2.15}$$

Proof. As in the proof of the formula for two events, 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 2.43

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,

$$P(A_1) = P(A_2) = P(A_3) = 1/6$$
$$P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_2 \cap A_3) = 1/36$$
$$P(A_1 \cap A_2 \cap A_3) = 1/216$$

So plugging into (2.15) 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.

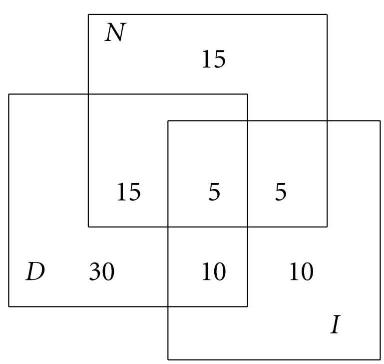
Example 2.44

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 *Cornell Daily Sun* and the *New York Times*, 15 read the *Cornell Daily Sun* and the *Ithaca Journal*, 10 read the *New York 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 who read only D is $60 - 30 = 30$. In a similar way, we compute that there are $40 - 25 = 15$ students who read only N and $30 - 20 = 10$ students who read only I. Adding up the numbers in the seven regions gives a total of 90, as we found before.

2.5.1 Inclusion–exclusion formula

Formula (2.15) generalizes to n events:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n) \tag{2.16}$$

In words, we take all possible intersections of $1, 2, \ldots, n$ events and the signs of the sums alternate.

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

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

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

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

which proves the desired result. □

Example 2.45 You pick 7 cards out of deck of 52. What is the probability that you have a three of a kind, that is, exactly three cards of some denomination (for example, 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

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

A_1 can occur in $C_{4,3} C_{48,4} = 778{,}320$ ways and $A_1 \cap A_2$ 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 having two sets of three of a kind.

Example 2.46 Suppose we roll a die 15 times. What is the probability that we do not see each of the 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

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

Turning to the second term, 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

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

For 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

$$\sum_{i<j<k} P(A_i \cap A_j \cap A_k) = 20(3.05 \times 10^{-5}) = 0.00061$$

At this point the pattern should be clear:

$$\begin{aligned}
&C_{6,1}(5/6)^{15} - C_{6,2}(4/6)^{15} + C_{6,3}(3/6)^{15} - C_{6,4}(2/6)^{15} + C_{6,5}(1/6)^{15} \\
&\quad = 0.389433 - 0.03426 + 6.1 \times 10^{-4} - 1.045 \times 10^{-6} + 1.276 \times 10^{-11} \\
&\quad = 0.355787
\end{aligned}$$

2.5.2 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.

$$P\left(\bigcup_{i=1}^{n} A_i\right) \le \sum_{i=1}^{n} P(A_i) \tag{2.17}$$

$$\ge \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) \tag{2.18}$$

$$\le \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \tag{2.19}$$

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

$$\begin{aligned}
&\leq 0.389433\\
&\geq 0.389433 - 0.03426 = 0.355177\\
&\leq 0.389433 - 0.03426 + 6.1 \times 10^{-4} = 0.355738
\end{aligned}$$

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 exactly 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 twice and the second once, with a net total of 1. If $k \geq 3$, the first term counts it k times and the second $C_{k,2} = k(k-1)/2 > k$ times so the net number of countings is < 0.

The third formula is similar.

In k sets	Counted
1	$1 - 0 + 0 = 1$
2	$2 - 1 + 0 = 1$
3	$3 - 3 + 1 = 1$

When ≥ 4, the number of countings is

$$C_{k,1} - C_{k,2} + C_{k,3} > k - \frac{k(k-1)}{2} + (k-1)(k-2) \geq 0 \qquad \square$$

Example 2.47 **The streak.** In the summer of 1941, Joe DiMaggio achieved 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?

A useful trick. Suppose for the moment that we know the probability p that Joe DiMaggio gets a hit in one game and that successive games are independent. Assuming a 154-game season, we could let A_i be the probability that a player got hits in games $i+1, \ldots, i+56$ for $0 \leq i \leq 98$. Using (2.17) it follows that the probability of the streak is

$$\leq 99\, p^{56}$$

As we will see in a minute this overestimates the actual answer by a factor of $1/(1-p)$. 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 no hit in game i but has hits in games $i+1, i+2, \ldots, i+56$, where $1 \leq i \leq 98$. Ignoring the probability of having hits in games 1, 2, ..., 56, the event of interest $S = \cup_{i=1}^{98} B_i$, so

$$P(S) \le q \equiv 98\, p^{56}(1-p)$$

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 hit in game j, while B_j requires no hit. If $56 + i < j \le 98$, then $P(B_i \cap B_j) = P(B_i)P(B_j)$. To simplify the arithmetic we note that in either case $P(B_i \cap B_j) \le P(B_i)P(B_j)$, so

$$\sum_{1 \le i < j \le 98} P(B_i \cap B_j) \le C_{98,2}\, p^{112}(1-p)^2 \le \frac{q^2}{2}$$

This is the number we have to subtract from the upper bound to get the lower bound, so we have

$$q \ge P(S) \ge q - \frac{q^2}{2} \tag{2.20}$$

Since q will end up being very small, the ratio of the two bounds is $1 - (q/2) \approx 1$.

To compute the probability p that Joe DiMaggio gets a hit in one game, we will introduce two somewhat questionable assumptions: (i) A player gets exactly four at bats per game (during the streak, DiMaggio averaged 3.98 at bats per game) and (ii) the outcomes of different at bats are independent with the probability of a hit being 0.325, Joe DiMaggio's lifetime batting average. From assumptions (i) and (ii) it follows that the probability

$$p = 1 - (0.675)^4 = 0.7924$$

and using (2.20) we have

$$P(S) \approx q = 98(0.7924)^{56}(0.2076) = 4.46 \times 10^{-5}$$

To interpret our result, note that the probability in (2.20) is roughly 1/22,000, so even if there were 220 players with 0.325 batting averages, it would take 100 years for this to occur again.

Example 2.48 **A less famous streak.** *Sports Illustrated* reports that a high school football team in Bloomington, Indiana, lost 21 straight pregame 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 won the coin flip in game i but lost it in games $i+1, \ldots, i+21$. Using the reasoning that led to (2.20),

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

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

$$(1 - 79(0.5)^{22})^{15,000} = 0.7539$$

that is, with probability 0.2461 some team will have this happen to them. As a check on the last calculation, note that (2.17) gives an upper bound of

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

2.6 Blackjack

In this book we analyze craps and roulette, casino games where the player has a substantial disadvantage. In the case of blackjack, a little strategy, which we explain in this section, can make the game almost even. To begin we 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 (for example, 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 that 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 that the dealer's ending total is k when he has a hard total of j, that is, 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 to work 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	0.13981	0.13491	0.12966	0.12403	0.11799	**0.35361**
3	0.13503	0.13048	0.12558	0.12033	0.11470	**0.37387**
4	0.13049	0.12594	0.12139	0.11648	0.11123	**0.39447**
5	0.12225	0.12225	0.11770	0.11315	0.10825	**0.41640**
6	0.16544	0.10627	0.10627	0.10171	0.09716	**0.42315**
7	**0.36857**	0.13780	0.07863	0.07863	0.07407	0.26231
8	0.12857	**0.35934**	0.12857	0.06939	0.06939	0.24474
9	0.12000	0.12000	**0.35076**	0.12000	0.06082	0.22843
10	0.11142	0.11142	0.11142	**0.34219**	0.11142	0.21211
Ace	0.13079	0.13079	0.13079	0.13079	**0.36156**	0.11529

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, so we will not attempt it here. The first analysis was performed in the mid-1950s (see Baldwin et al. in *Journal of the American Statistical Association*, Vol. 51, pp. 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 downcard was a 10. If her upcard 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 that 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 upcard is 9 or less
- If your total is 9 and the dealer's upcard is 3 through 6
- If you have A and 2 through 7 and the dealer's upcard is 4, 5, or 6

Again the doubling down rules can be explained by assuming that we are going to get a 10. Some Nevada casinos allow doubling down only 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 split 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 upcard is 3 through 7.
- Split 9's when the dealer's upcard is 2 through 9, but not 7.

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" (that is, 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 hardwork, 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. The book *Bringing Down the House* gives an entertaining account of MIT students using the strategy to win money at blackjack.

2.7 Exercises

Permutations and combinations

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

2. 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?

3. 16 horses race in the Kentucky Derby. How many possible results are there for win, place, and show (first, second, and third)?

4. 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?

5. 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 or (b) not important?

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

7. A commercial for Glade Plug-ins says that by inserting 2 of a choice of 11 scents into the device, you can make more than 50 combinations. If we exclude the boring choice of two of the same scent, how many possibilities are there?

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

9. (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 that the three letters and the three numbers are different?

10. 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)?

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

12. 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?

13. A person has 12 friends and will invite 7 to a party. (a) How many choices are possible if Al and Bob are feuding 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?

14. A basketball team has 5 players more than 6 feet tall and 6 who are less than 6 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?

15. The Duke basketball team has 10 women who can play guard and 12 tall women who can play the other three positions. At the start of the game, the coach gives the referee a starting lineup that lists who will play left guard, right guard, left forward, center, and right forward. In how many ways can this be done?

16. 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?

17. 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 and (b) must not sit next to each other? (Two seating patterns that differ only by a rotation of the table are considered the same.)

18. 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 8 markers be placed on an 8×8 grid of squares so that there is at most one in each row or column?

19. A BINGO card is a 5×5 grid. The center square is a free space and has no number. The first column is filled with five distinct numbers from 1 to 15, the second with five numbers from 16 to 30, the middle column with four numbers from 31 to 45, the fourth with five numbers from 46 to 60, and the fifth with five numbers from 61 to 75. Since the object of the game is to get five in a row horizontally, vertically, or diagonally, the order is important. How many BINGO cards are there?

20. Continuing with the setup from the previous problem, in BINGO numbers are drawn from 1 to 75 without replacement. When a number is called you put a marker on that square. If you have five in a row horizontally, vertically, or diagonally, you have a BINGO. What is the probability you will have a BINGO after (a) four numbers are called? (b) After five?

Multinomial counting problems

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

22. 12 different toys are to be divided among 3 children so that each one gets 4 toys. How many ways can this be done?

23. 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?

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

25. 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?

26. 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?

27. 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?

Binomial and multinomial distributions

28. A die is rolled 8 times. What is the probability that we will get exactly two 3's?

29. Mary knows the answers to 20 of the 25 multiple-choice questions on the Psychology 101 exam, but she has skipped several of the lectures; she must take random guesses for the other five. Assuming each question has four answers, what is the probability that she will get exactly 3 of the last 5 questions right?

30. In 1997, 10.8% of female smokers smoked cigars. In a sample of size 10 female smokers, what is the probability that (a) exactly 2 of the women smoke cigars? (b) At most 1 smokes cigars?

31. A 1994 report revealed that 32.6% of U.S. births were to unmarried women. A parenting magazine selected 30 women who gave birth in 1994 at random. (a) What is the probability that exactly 10 of the women were unmarried? (b) Using your calculator determine the probability that in the sample at most 10 are unmarried.

32. 20% of all students are left-handed. A class of size 20 meets in a room with 5 left-handed and 18 right-handed chairs. Use your calculator to find the probability that each student will have a chair to match their needs.

33. David claims to be able to distinguish brand B beer from brand H, but Alice claims that he just guesses. They set up a taste test with 10 small glasses of beer. David wins if he gets 8 or more right. What is the probability that he will win (a) if he is just guessing? (b) If he gets the right answer with probability 0.9?

34. The following situation comes up the game of Yahtzee. We have three rolls of five dice and want to get three sixes or more. On each turn we reroll any dice that are not 6's. What is the probability that we succeed?

35. A baseball pitcher throws a strike with probability 0.5 and a ball with probability 0.5. He is facing a batter who never swings at a pitch. What is the probability that he strikes out, that is, gets three strikes before four balls?

36. A baseball player is said to "hit for the cycle" if he has a single, a double, a triple, and a home run all in one game. Suppose these four types of hits have probabilities 1/6, 1/20, 1/120, and 1/24. What is the probability of hitting for the cycle if he gets to bat (a) four times and (b) five times? (c) Using $P(\cup_i A_i) \leq \sum_i P(A_i)$ shows that the answer to (b) is at most 5 times the answer to (a). What is the ratio of the two answers?

Poisson approximation

37. Compare the Poisson approximation with the exact binomial probabilities when (a) $n = 10,\ p = 0.1$, (b) $n = 20,\ p = 0.05$, and (c) $n = 40,\ p = 0.025$.

38. Use the Poisson approximation to compute the probability that you will roll at least one double 6 in 24 trials. How does this compare with the exact answer?

39. The probability of a three of a kind in poker is approximately 1/50. Use the Poisson approximation to compute the probability that you will get at least one three of a kind if you play 20 hands of poker.

40. Calls to a toll-free hotline service are made randomly at rate 2 per minute. The service has five operators, none of whom is currently busy. Use the Poisson distribution to estimate the probability that in the next minute there are < 5 calls.

41. In one of the New York state lottery games, a number is chosen at random between 0 and 999. Suppose you play this game 250 times. Use the Poisson approximation to estimate the probability that you will never win and compare this with the exact answer.

42. If you bet \$1 on number 13 at roulette (or on any other number) then you win \$35 if that number comes up, an event of probability $1/38$, and you lose your dollar otherwise. Suppose you play 70 times. Use the Poisson approximation to estimate the probability that (a) you have won 0 times and lost \$70, and (b) you have won 1 time and lost \$34. (c) If you win 2 times you have won \$2. Combine the results of (a) and (b) to conclude that the probability that you will have won more money than you have lost is larger than 1/2.

43. In a particular Powerball drawing 210,850,582 tickets were sold. The chance of winning the lottery is 1 in 80,000,000. Use the Poisson approximation to estimate the probability that there is exactly one winner.

44. Suppose that the probability of a defect in a foot of magnetic tape is 0.002. Use the Poisson approximation to compute the probability that a 1,500-foot roll will have no defects.

45. Suppose 1% of a certain brand of Christmas lights is defective. Use the Poisson approximation to compute the probability that in a box of 25 there will be at most one defective bulb.

46. In February 2000, 2.8% of Colorado's labor force was unemployed. Calculate the probability that in a group of 50 workers exactly one is unemployed.

47. An insurance company insures 3,000 people, each of whom has a 1/1,000 chance of an accident in 1 year. Use the Poisson approximation to compute the probability that there will be at most 2 accidents.

48. Suppose that 1% of people in the population are more than 6 feet 3 inches tall. What is the chance that in a group of 200 people picked at random from the population at least four people will be more than 6 feet 3 inches tall.

49. In an average year in Mythica there are 8 fires. Last year there were 12 fires. How likely is it to have 12 or more fires just by chance?

50. An airline company sells 160 tickets for a plane with 150 seats, knowing that the probability that a passenger will not show up for the flight is 0.1. Use the Poisson approximation to compute the probability that they will have enough seats for all the passengers who show up.

51. Books from a certain publisher contain an average of 1 misprint per page. What is the probability that on at least one page in a 300-page book there are five misprints?

Urn problems

52. 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?

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

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

55. 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?

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

57. An electronics store receives a shipment of 30 calculators of which 4 are defective. Six of these calculators are selected to be sent to a local high school. What is the probability that exactly one is defective?

58. 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.

59. (a) How many ways 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 that the team will have 2 boys and 2 girls.

60. The following probability problem arose in a court case concerning possible discrimination 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 that we would get this outcome if the five nurses who failed were chosen at random?

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

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

63. 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?

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

65. 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 1,000 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.

66. 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 that 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 blackjack?

67. 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?

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

69. 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 1 week five different children are chosen?

70. An investor picks 3 stocks out of 10 recommended by his broker. Of these, 6 will show a profit in the next year. What is the probability that the investor will pick (a) 3, (b) 2, (c) 1, (d) 0 profitable stocks?

71. Four red cards (that is, 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 that he would do this if he were guessing?

72. A town council considers the question of closing down an "adult" theater. The five men on the council all vote against this and the three women vote in favor. What is the probability that 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?

73. 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?

74. A town has four TV repairmen. 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.

75. 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

76. 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.

Probabilities of unions

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

78. Suppose you draw 7 cards out of a deck of 52. What is the probability that you will have (a) exactly 5 cards of one suit? (b) At least 5 cards of one suit?

79. 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 and (b) exactly one newspaper?

80. 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?

81. 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 that there is no 1.

82. 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"?

83. We pick a number between 0 and 999, then a computer picks one at random from that range. Use (2.15) to compute the probability that 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.)

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

85. You pick 13 cards out of a deck of 52. Let $A =$ "you have exactly 6 cards in at least one suit" and $B =$ "you have exactly 6 spades." The first Bonferroni inequality says that $P(A) \leq 4P(B)$. Compute $P(A)$ and $P(A)/P(B)$.

86. 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.

87. 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 $1 - (35/36)^6$.

88. 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.