

## 3.2: Combinations

Having mastered permutations, we now consider combinations. Let  $U$  be a set with  $n$  elements; we want to count the number of distinct subsets of the set  $U$  that have exactly  $j$  elements. The empty set and the set  $U$  are considered to be subsets of  $U$ . The empty set is usually denoted by  $\phi$ .

### ✓ Example 3.2.1:

Let  $U = \{a, b, c\}$ . The subsets of  $U$  are

$$\phi, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}. \quad (3.2.1)$$

### Binomial Coefficients

The number of distinct subsets with  $j$  elements that can be chosen from a set with  $n$  elements is denoted by  $\binom{n}{j}$ , and is pronounced “ $n$  choose  $j$ .” The number  $\binom{n}{j}$  is called a *binomial coefficient*. This terminology comes from an application to algebra which will be discussed later in this section.

In the above example, there is one subset with no elements, three subsets with exactly 1 element, three subsets with exactly 2 elements, and one subset with exactly 3 elements. Thus,  $\binom{3}{0} = 1$ ,  $\binom{3}{1} = 3$ ,  $\binom{3}{2} = 3$ , and  $\binom{3}{3} = 1$ . Note that there are  $2^3 = 8$  subsets in all. (We have already seen that a set with  $n$  elements has  $2^n$  subsets; see Exercise 3.1.8) It follows that

$$\binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3} = 2^3 = 8, \quad (3.2.2)$$

$$\binom{n}{0} = \binom{n}{n} = 1. \quad (3.2.3)$$

Assume that  $n > 0$ . Then, since there is only one way to choose a set with no elements and only one way to choose a set with  $n$  elements, the remaining values of  $\binom{n}{j}$  are determined by the following :

### Theorem 3.2.1

For integers  $n$  and  $j$ , with  $0 < j < n$ , the binomial coefficients satisfy:

$$\binom{n}{j} = \binom{n-1}{j} + \binom{n-1}{j-1}. \quad (3.2.4)$$

#### Proof

We wish to choose a subset of  $j$  elements. Choose an element  $u$  of  $U$ . Assume first that we do not want  $u$  in the subset. Then we must choose the  $j$  elements from a set of  $n - 1$  elements; this can be done in  $\binom{n-1}{j}$  ways. On the other hand, assume that we do want  $u$  in the subset. Then we must choose the other  $j - 1$  elements from the remaining  $n - 1$  elements of  $U$ ; this can be done in  $\binom{n-1}{j-1}$  ways. Since  $u$  is either in our subset or not, the number of ways that we can choose a subset of  $j$  elements is the sum of the number of subsets of  $j$  elements which have  $u$  as a member and the number which do not—this is what Equation 3.1.1 states.

The binomial coefficient  $\binom{n}{j}$  is defined to be 0, if  $j < 0$  or if  $j > n$ . With this definition, the restrictions on  $j$  in Theorem 3.2.1 are unnecessary.

**Table 3.2.1:** Pascal's Triangle

	j=0	1	2	3	4	5	6	7	8	9	10
n=0	1										
1	1	1									
2	1	2	1								

	j=0	1	2	3	4	5	6	7	8	9	10
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

## Pascal's Triangle

The relation 3.1, together with the knowledge that

$$\binom{n}{0} = \binom{n}{n} = 1, \quad (3.2.5)$$

determines completely the numbers  $\binom{n}{j}$ . We can use these relations to determine the famous which exhibits all these numbers in matrix form (see **Table 3.2.1**).

The  $n$ th row of this triangle has the entries  $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ . We know that the first and last of these numbers are 1. The remaining numbers are determined by the recurrence relation Equation 3.1; that is, the entry  $\binom{n}{j}$  for  $0 < j < n$  in the  $n$ th row of Pascal's triangle is the of the entry immediately above and the one immediately to its left in the  $(n-1)$ st row. For example,  $\binom{5}{2} = 6 + 4 = 10$ .

This algorithm for constructing Pascal's triangle can be used to write a computer program to compute the binomial coefficients. You are asked to do this in Exercise 4.

While Pascal's triangle provides a way to construct recursively the binomial coefficients, it is also possible to give a formula for  $\binom{n}{j}$ .

### Theorem 3.2.1

The binomial coefficients are given by the formula

$$\binom{n}{j} = \frac{(n)_j}{j!}. \quad (3.2.6)$$

#### Proof

Each subset of size  $j$  of a set of size  $n$  can be ordered in  $j!$  ways. Each of these orderings is a  $j$ -permutation of the set of size  $n$ . The number of  $j$ -permutations is  $(n)_j$ , so the number of subsets of size  $j$  is

$$\frac{(n)_j}{j!}. \quad (3.2.7)$$

This completes the proof.

The above formula can be rewritten in the form

$$\binom{n}{j} = \frac{n!}{j!(n-j)!}. \quad (3.2.8)$$

This immediately shows that

$$\binom{n}{j} = \binom{n}{n-j}. \quad (3.2.9)$$

When using Equation 3.2 in the calculation of  $\binom{n}{j}$ , if one alternates the multiplications and divisions, then all of the intermediate values in the calculation are integers. Furthermore, none of these intermediate values exceed the final value. (See Exercise 3.2.40)

Another point that should be made concerning Equation [eq 3.4] is that if it is used to the binomial coefficients, then it is no longer necessary to require  $n$  to be a positive integer. The variable  $j$  must still be a non-negative integer under this definition. This idea is useful when extending the Binomial Theorem to general exponents. (The Binomial Theorem for non-negative integer exponents is given below as Theorem 3.2.5.)

## Poker Hands

### ✓ Example 3.2.2:

Poker players sometimes wonder why a beats a A poker hand is a random subset of 5 elements from a deck of 52 cards. A hand has four of a kind if it has four cards with the same value—for example, four sixes or four kings. It is a full house if it has three of one value and two of a second—for example, three twos and two queens. Let us see which hand is more likely. How many hands have four of a kind? There are 13 ways that we can specify the value for the four cards. For each of these, there are 48 possibilities for the fifth card. Thus, the number of four-of-a-kind hands is  $13 \cdot 48 = 624$ . Since the total number of possible hands is  $\binom{52}{5} = 2598960$ , the probability of a hand with four of a kind is  $624/2598960 = .00024$

Now consider the case of a full house; how many such hands are there? There are 13 choices for the value which occurs three times; for each of these there are  $\binom{4}{3} = 4$  choices for the particular three cards of this value that are in the hand. Having picked these three cards, there are 12 possibilities for the value which occurs twice; for each of these there are  $\binom{4}{2} = 6$  possibilities for the particular pair of this value. Thus, the number of full houses is  $13 \cdot 4 \cdot 12 \cdot 6 = 3744$ , and the probability of obtaining a hand with a full house is  $3744/2598960 = .0014$ . Thus, while both types of hands are unlikely, you are six times more likely to obtain a full house than four of a kind.

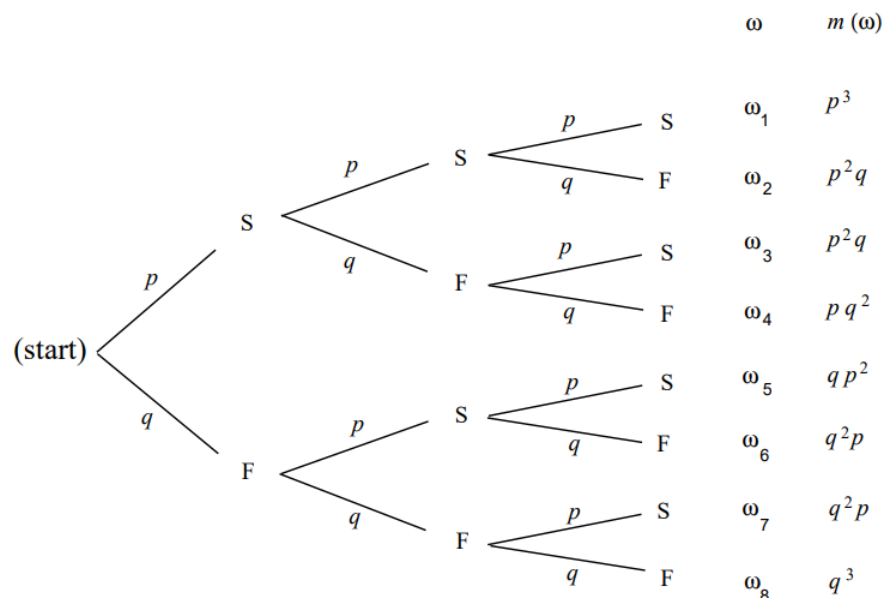


Figure 3.2.1: Tree diagram of three Bernoulli trials.

## Bernoulli Trials

Our principal use of the binomial coefficients will occur in the study of one of the important chance processes called

### Definition 3.2.1

A *Bernoulli* is a sequence of  $n$  chance experiments such that

1. Each experiment has two possible outcomes, which we may call and
2. The probability  $p$  of success on each experiment is the same for each experiment, and this probability is not affected by any knowledge of previous outcomes. The probability  $q$  of failure is given by  $q = 1 - p$ .

### Example 3.2.3

The following are Bernoulli trials processes:

1. A coin is tossed ten times. The two possible outcomes are heads and tails. The probability of heads on any one toss is  $1/2$ .
2. An opinion poll is carried out by asking 1000 people, randomly chosen from the population, if they favor the Equal Rights Amendment—the two outcomes being yes and no. The probability  $p$  of a yes answer (i.e., a success) indicates the proportion of people in the entire population that favor this amendment.
3. A gambler makes a sequence of 1-dollar bets, betting each time on black at roulette at Las Vegas. Here a success is winning 1 dollar and a failure is losing 1 dollar. Since in American roulette the gambler wins if the ball stops on one of 18 out of 38 positions and loses otherwise, the probability of winning is  $p = 18/38 = .474$ .

To analyze a Bernoulli trials process, we choose as our sample space a binary tree and assign a probability distribution to the paths in this tree. Suppose, for example, that we have three Bernoulli trials. The possible outcomes are indicated in the tree diagram shown in Figure 3.4. We define  $X$  to be the random variable which represents the outcome of the process, i.e., an ordered triple of S's and F's. The probabilities assigned to the branches of the tree represent the probability for each individual trial. Let the outcome of the  $i$ th trial be denoted by the random variable  $X_i$ , with distribution function  $m_i$ . Since we have assumed that outcomes on any one trial do not affect those on another, we assign the same probabilities at each level of the tree. An outcome  $\omega$  for the entire experiment will be a path through the tree. For example,  $\omega_3$  represents the outcomes SFS. Our frequency interpretation of probability would lead us to expect a fraction  $p$  of successes on the first experiment; of these, a fraction  $q$  of failures on the second; and, of these, a fraction  $p$  of successes on the third experiment. This suggests assigning probability  $pqp$  to the outcome  $\omega_3$ . More generally, we assign a distribution function  $m(\omega)$  for paths  $\omega$  by defining  $m(\omega)$  to be the product of the branch probabilities along the path  $\omega$ . Thus, the probability that the three events S on the first trial, F on the second trial, and S on the third trial occur is the product of the probabilities for the individual events. We shall see in the next chapter that this means that the events involved are in the sense that the knowledge of one event does not affect our prediction for the occurrences of the other events.

## Binomial Probabilities

We shall be particularly interested in the probability that in  $n$  Bernoulli trials there are exactly  $j$  successes. We denote this probability by  $b(n, p, j)$ . Let us calculate the particular value  $b(3, p, 2)$  from our tree measure. We see that there are three paths which have exactly two successes and one failure, namely  $\omega_2$ ,  $\omega_3$ , and  $\omega_5$ . Each of these paths has the same probability  $p^2q$ . Thus  $b(3, p, 2) = 3p^2q$ . Considering all possible numbers of successes we have

$$\begin{aligned} b(3, p, 0) &= q^3, \\ b(3, p, 1) &= 3pq^2, \\ b(3, p, 2) &= 3p^2q, \\ b(3, p, 3) &= p^3. \end{aligned}$$

We can, in the same manner, carry out a tree measure for  $n$  experiments and determine  $b(n, p, j)$  for the general case of  $n$  Bernoulli trials.

### Theorem 3.2.2

Given  $n$  Bernoulli trials with probability  $p$  of success on each experiment, the probability of exactly  $j$  successes is

$$b(n, p, j) = \binom{n}{j} p^j q^{n-j} \quad (3.2.10)$$

where  $q = 1 - p$ .

### Proof

We construct a tree measure as described above. We want to find the sum of the probabilities for all paths which have exactly  $j$  successes and  $n - j$  failures. Each such path is assigned a probability  $p^j q^{n-j}$ . How many such paths are there? To specify a path, we have to pick, from the  $n$  possible trials, a subset of  $j$  to be successes, with the remaining  $n - j$  outcomes being failures. We can do this in  $\binom{n}{j}$  ways. Thus the sum of the probabilities is

$$b(n, p, j) = \binom{n}{j} p^j q^{n-j} . \quad (3.2.11)$$

### ✓ Example 3.2.4:

A fair coin is tossed six times. What is the probability that exactly three heads turn up? The answer is

$$b(6, .5, 3) = \binom{6}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^3 = 20 \cdot \frac{1}{64} = .3125 . \quad (3.2.12)$$

### ✓ Example 3.2.5:

A die is rolled four times. What is the probability that we obtain exactly one 6? We treat this as Bernoulli trials with *success* = "rolling a 6" and *failure* = "rolling some number other than a 6." Then  $p = 1/6$ , and the probability of exactly one success in four trials is

$$b(4, 1/6, 1) = \binom{4}{1} \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^3 = .386 . \quad (3.2.13)$$

To compute binomial probabilities using the computer, multiply the function `choose( $n, k$ )` by  $p^k q^{n-k}$ . The program **BinomialProbabilities** prints out the binomial probabilities  $b(n, p, k)$  for  $k$  between  $kmin$  and  $kmax$ , and the sum of these probabilities. We have run this program for  $n = 100$ ,  $p = 1/2$ ,  $kmin = 45$ , and  $kmax = 55$ ; the output is shown in Table 3.8. Note that the individual probabilities are quite small. The probability of exactly 50 heads in 100 tosses of a coin is about 0.08. Our intuition tells us that this is the most likely outcome, which is correct; but, all the same, it is not a very likely outcome.

**Table 3.2.2** Binomial probabilities for  $n = 100$ ,  $p = 1/2$ .

$k$	$b(n, p, k)$
45	.0485
46	.0580
47	.0666
48	.0735
49	.0780
50	.0796
51	.0780
52	.0735
53	.0666
54	.0580
55	.0485

## Binomial Distributions

### Definition 3.2.6

Let  $n$  be a positive integer, and let  $p$  be a real number between 0 and 1. Let  $B$  be the random variable which counts the number of successes in a Bernoulli trials process with parameters  $n$  and  $p$ . Then the distribution  $b(n, p, k)$  of  $B$  is called the *binomial distribution*.

We can get a better idea about the binomial distribution by graphing this distribution for different values of  $n$  and  $p$  (see Figure 3.2.2). The plots in this figure were generated using the program **BinomialPlot**.

We have run this program for  $p = .5$  and  $p = .3$ . Note that even for  $p = .3$  the graphs are quite symmetric. We shall have an explanation for this in Chapter 9. We also note that the highest probability occurs around the value  $np$ , but that these highest probabilities get smaller as  $n$  increases. We shall see in Chapter 6 that  $np$  is the *mean* or *expected* value of the binomial distribution  $b(n, p, k)$ .

The following example gives a nice way to see the binomial distribution, when  $p = 1/2$ .

### Example 3.2.6:

A *Galton board* is a board in which a large number of BB-shots are dropped from a chute at the top of the board and deflected off a number of pins on their way down to the bottom of the board. The final position of each slot is the result of a number of random deflections either to the left or the right. We have written a program **GaltonBoard** to simulate this experiment.

We have run the program for the case of 20 rows of pins and 10,000 shots being dropped. We show the result of this simulation in Figure 3.6

Note that if we write 0 every time the shot is deflected to the left, and 1 every time it is deflected to the right, then the path of the shot can be described by a sequence of 0's and 1's of length  $n$ , just as for the  $n$ -fold coin toss.

The distribution shown in Figure 3.6 is an example of an empirical distribution, in the sense that it comes about by means of a sequence of experiments. As expected, this empirical distribution resembles the corresponding binomial distribution with parameters  $n = 20$  and  $p = 1/2$ .

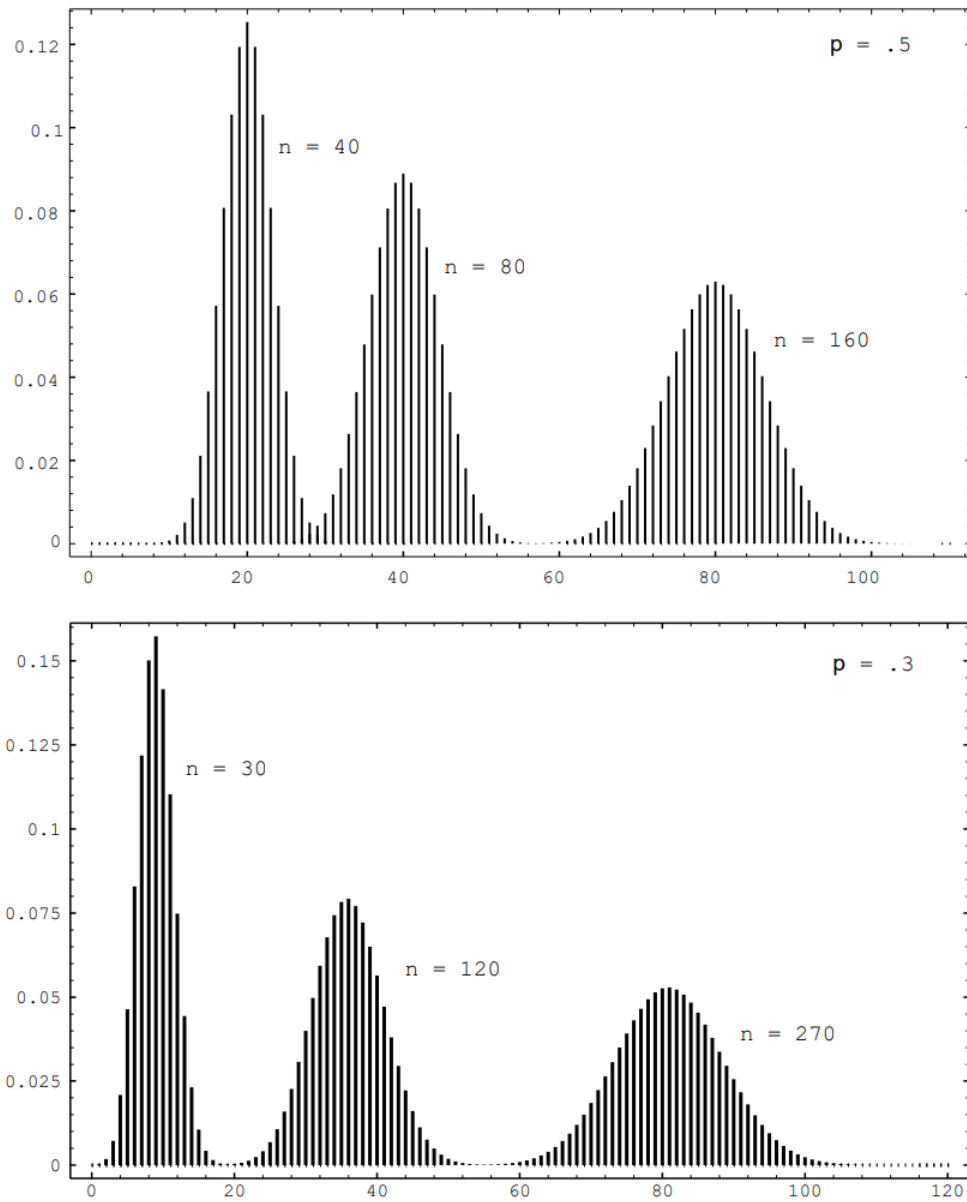


Figure 3.2.2: Binomial distributions

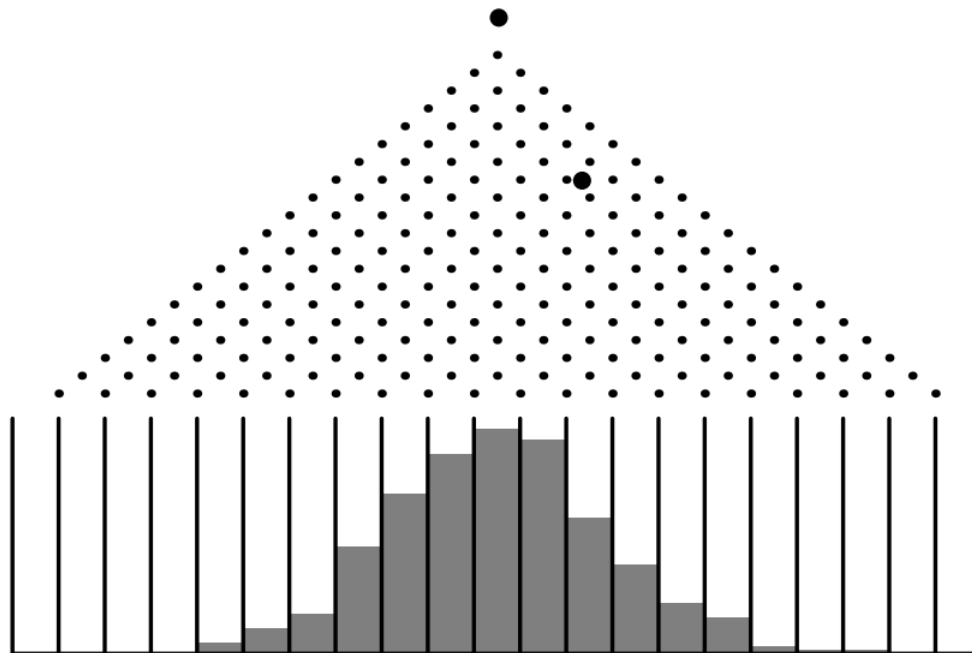


Figure \(\PageIndex{3}\): Simulation of the Galton board.

## Hypothesis Testing

### ✓ Example 3.2.7:

Suppose that ordinary aspirin has been found effective against headaches 60 percent of the time, and that a drug company claims that its new aspirin with a special headache additive is more effective. We can test this claim as follows: we call their claim the and its negation, that the additive has no appreciable effect, the Thus the null hypothesis is that  $p = .6$ , and the alternate hypothesis is that  $p > .6$ , where  $p$  is the probability that the new aspirin is effective.

We give the aspirin to  $n$  people to take when they have a headache. We want to find a number  $m$ , called the for our experiment, such that we reject the null hypothesis if at least  $m$  people are cured, and otherwise we accept it. How should we determine this critical value?

First note that we can make two kinds of errors. The first, often called a *type 1 error* in statistics, is to reject the null hypothesis when in fact it is true. The second, called a *type 2 error* is to accept the null hypothesis when it is false. To determine the probability of both these types of errors we introduce a function  $\alpha(p)$ , defined to be the probability that we reject the null hypothesis, where this probability is calculated under the assumption that the null hypothesis is true. In the present case, we have

$$\alpha(p) = \sum_{m \leq k \leq n} b(n, p, k). \quad (3.2.14)$$

Note that  $\alpha(.6)$  is the probability of a type 1 error, since this is the probability of a high number of successes for an ineffective additive. So for a given  $n$  we want to choose  $m$  so as to make  $\alpha(.6)$  quite small, to reduce the likelihood of a type 1 error. But as  $m$  increases above the most probable value  $np = .6n$ ,  $\alpha(.6)$ , being the upper tail of a binomial distribution, approaches 0. Thus  $m$  makes a type 1 error less likely.

Now suppose that the additive really is effective, so that  $p$  is appreciably greater than .6; say  $p = .8$ . (This alternative value of  $p$  is chosen arbitrarily; the following calculations depend on this choice.) Then choosing  $m$  well below  $np = .8n$  will increase  $\alpha(.8)$ , since now  $\alpha(.8)$  is all but the lower tail of a binomial distribution. Indeed, if we put  $\beta(.8) = 1 - \alpha(.8)$ , then  $\beta(.8)$  gives us the probability of a type 2 error, and so *decreasing*  $m$  makes a type 2 error less likely.



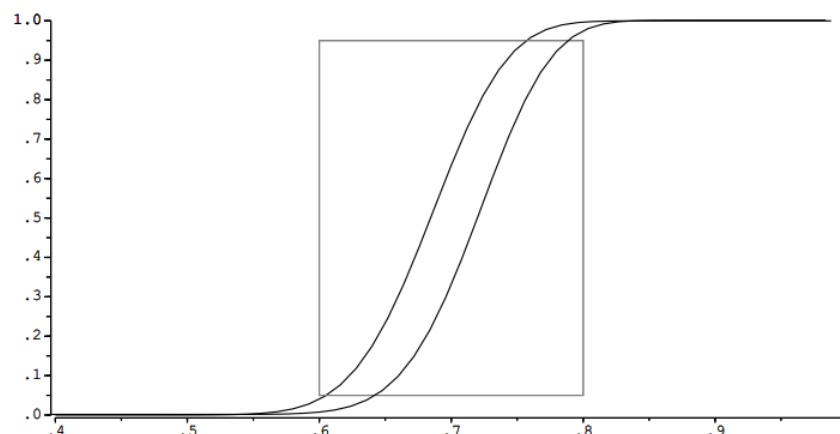
The manufacturer would like to guard against a type 2 error, since if such an error is made, then the test does not show that the new drug is better, when in fact it is. If the alternative value of  $p$  is chosen closer to the value of  $p$  given in the null hypothesis (in this case  $p = .6$ ), then for a given test population, the value of  $\beta$  will increase. So, if the manufacturer's statistician chooses an alternative value for  $p$  which is close to the value in the null hypothesis, then it will be an expensive proposition (i.e., the test population will have to be large) to reject the null hypothesis with a small value of  $\beta$ .

What we hope to do then, for a given test population  $n$ , is to choose a value of  $m$ , if possible, which makes both these probabilities small. If we make a type 1 error we end up buying a lot of essentially ordinary aspirin at an inflated price; a type 2 error means we miss a bargain on a superior medication. Let us say that we want our critical number  $m$  to make each of these undesirable cases less than 5 percent probable.

We write a program **PowerCurve** to plot, for  $n = 100$  and selected values of  $m$ , the function  $\alpha(p)$ , for  $p$  ranging from .4 to 1. The result is shown in Figure [fig 3.9]. We include in our graph a box (in dotted lines) from .6 to .8, with bottom and top at heights .05 and .95. Then a value for  $m$  satisfies our requirements if and only if the graph of  $\alpha$  enters the box from the bottom, and leaves from the top (why?—which is the type 1 and which is the type 2 criterion?). As  $m$  increases, the graph of  $\alpha$  moves to the right. A few experiments have shown us that  $m = 69$  is the smallest value for  $m$  that thwarts a type 1 error, while  $m = 73$  is the largest which thwarts a type 2. So we may choose our critical value between 69 and 73. If we're more intent on avoiding a type 1 error we favor 73, and similarly we favor 69 if we regard a type 2 error as worse. Of course, the drug company may not be happy with having as much as a 5 percent chance of an error. They might insist on having a 1 percent chance of an error. For this we would have to increase the number  $n$  of trials (see Exercise 3.2.28).

## Binomial Expansion

We next remind the reader of an application of the binomial coefficients to algebra. This is the *binomial expansion* from which we get the term binomial coefficient.



5

Figure 3.2.4: The power curve

### Theorem 3.2.7: (Binomial Theorem)

The quantity  $(a + b)^n$  can be expressed in the form

$$(a + b)^n = \sum_{j=0}^n \binom{n}{j} a^j b^{n-j}. \quad (3.2.15)$$

To see that this expansion is correct, write

$$(a + b)^n = (a + b)(a + b) \cdots (a + b). \quad (3.2.16)$$

### Proof

When we multiply this out we will have a sum of terms each of which results from a choice of an  $a$  or  $b$  for each of  $n$  factors. When we choose  $j$   $a$ 's and  $(n-j)$   $b$ 's, we obtain a term of the form  $a^j b^{n-j}$ . To determine such a term, we have to specify  $j$  of the  $n$  terms in the product from which we choose the  $a$ . This can be done in  $\binom{n}{j}$  ways. Thus, collecting these terms in the sum contributes a term  $\binom{n}{j} a^j b^{n-j}$ .

For example, we have

$$\begin{aligned}(a+b)^0 &= 1 \\ (a+b)^1 &= a+b \\ (a+b)^2 &= a^2 + 2ab + b^2 \\ (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3.\end{aligned}$$

We see here that the coefficients of successive powers do indeed yield Pascal's triangle.

### Corollary 3.2.1

The sum of the elements in the  $n$ th row of Pascal's triangle is  $2^n$ . If the elements in the  $n$ th row of Pascal's triangle are added with alternating signs, the sum is 0.

### Proof

The first statement in the corollary follows from the fact that

$$2^n = (1+1)^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}, \quad (3.2.17)$$

and the second from the fact that

$$0 = (1-1)^n = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \cdots + (-1)^n \binom{n}{n}. \quad (3.2.18)$$

The first statement of the corollary tells us that the number of subsets of a set of  $n$  elements is  $2^n$ . We shall use the second statement in our next application of the binomial theorem.

We have seen that, when  $A$  and  $B$  are any two events (cf. Section [sec 1.2]),

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (3.2.19)$$

We now extend this theorem to a more general version, which will enable us to find the probability that at least one of a number of events occurs.

## Inclusion-Exclusion Principle

### Theorem 3.2.1

Let  $P$  be a probability distribution on a sample space  $\Omega$ , and let  $\{A_1, A_2, \dots, A_n\}$  be a finite set of events. Then

$$\begin{aligned}P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) \\ &\quad + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \cdots\end{aligned}$$

That is, to find the probability that at least one of  $n$  events  $A_i$  occurs, first add the probability of each event, then subtract the probabilities of all possible two-way intersections, add the probability of all three-way intersections, and so forth.

### Proof

If the outcome  $\omega$  occurs in at least one of the events  $A_i$ , its probability is added exactly once by the left side of Equation [eq 3.5]. We must show that it is added exactly once by the right side of Equation [eq 3.5]. Assume that  $\omega$  is in exactly  $k$  of the sets. Then its probability is added  $k$  times in the first term, subtracted  $\binom{k}{2}$  times in the second, added  $\binom{k}{3}$  times in the third term, and so forth. Thus, the total number of times that it is added is

$$\binom{k}{1} - \binom{k}{2} + \binom{k}{3} - \cdots (-1)^{k-1} \binom{k}{k}. \quad (3.2.20)$$

But

$$0 = (1 - 1)^k = \sum_{j=0}^k \binom{k}{j} (-1)^j = \binom{k}{0} - \sum_{j=1}^k \binom{k}{j} (-1)^{j-1}. \quad (3.2.21)$$

Hence,

$$1 = \binom{k}{0} = \sum_{j=1}^k \binom{k}{j} (-1)^{j-1}. \quad (3.2.22)$$

If the outcome  $\omega$  is not in any of the events  $A_i$ , then it is not counted on either side of the equation.

## Hat Check Problem

### ✓ Example 3.2.8:

We return to the hat check problem discussed in Section 1.1, that is, the problem of finding the probability that a random permutation contains at least one fixed point. Recall that a permutation is a one-to-one map of a set  $A = \{a_1, a_2, \dots, a_n\}$  onto itself. Let  $A_i$  be the event that the  $i$ th element  $a_i$  remains fixed under this map. If we require that  $a_i$  is fixed, then the map of the remaining  $n - 1$  elements provides an arbitrary permutation of  $(n - 1)$  objects. Since there are  $(n - 1)!$  such permutations,  $P(A_i) = (n - 1)!/n! = 1/n$ . Since there are  $n$  choices for  $a_i$ , the first term of Equation [eq 3.5] is 1. In the same way, to have a particular pair  $(a_i, a_j)$  fixed, we can choose any permutation of the remaining  $n - 2$  elements; there are  $(n - 2)!$  such choices and thus

$$P(A_i \cap A_j) = \frac{(n - 2)!}{n!} = \frac{1}{n(n - 1)}. \quad (3.2.23)$$

The number of terms of this form in the right side of Equation [eq 3.5] is

$$\binom{n}{2} = \frac{n(n - 1)}{2!}. \quad (3.2.24)$$

Hence, the second term of Equation [eq 3.5] is

$$-\frac{n(n - 1)}{2!} \cdot \frac{1}{n(n - 1)} = -\frac{1}{2!}. \quad (3.2.25)$$

Similarly, for any specific three events  $A_i, A_j, A_k$ ,

$$P(A_i \cap A_j \cap A_k) = \frac{(n - 3)!}{n!} = \frac{1}{n(n - 1)(n - 2)}, \quad (3.2.26)$$

and the number of such terms is

$$\binom{n}{3} = \frac{n(n - 1)(n - 2)}{3!}, \quad (3.2.27)$$

making the third term of Equation [eq 3.5] equal to  $1/3!$ . Continuing in this way, we obtain

$$P(\text{at least one fixed point}) = 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots (-1)^{n-1} \frac{1}{n!} \quad (3.2.28)$$

and

$$P(\text{no fixed point}) = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!} . \quad (3.2.29)$$

From calculus we learn that

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots + \frac{1}{n!}x^n + \cdots . \quad (3.2.30)$$

Thus, if  $x = -1$ , we have

$$\begin{aligned} e^{-1} &= \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} + \cdots \\ &= .3678794 . \end{aligned}$$

Therefore, the probability that there is no fixed point, i.e., that none of the  $n$  people gets his own hat back, is equal to the sum of the first  $n$  terms in the expression for  $e^{-1}$ . This series converges very fast. Calculating the partial sums for  $n = 3$  to 10 gives the data in Table [table 3.7].

Table ✓ 3.2.3 Hat check problem.

$n$	Probability that no one gets his own hat back
3	.333333
4	.375
5	.366667
6	.368056
7	.367857
8	.367882
9	.367879
10	.367879

After  $n = 9$  the probabilities are essentially the same to six significant figures. Interestingly, the probability of no fixed point alternately increases and decreases as  $n$  increases. Finally, we note that our exact results are in good agreement with our simulations reported in the previous section.

## Choosing a Sample Space

We now have some of the tools needed to accurately describe sample spaces and to assign probability functions to those sample spaces. Nevertheless, in some cases, the description and assignment process is somewhat arbitrary. Of course, it is to be hoped that the description of the sample space and the subsequent assignment of a probability function will yield a model which accurately predicts what would happen if the experiment were actually carried out. As the following examples show, there are situations in which “reasonable” descriptions of the sample space do not produce a model which fits the data.

In Feller’s book,<sup>14</sup> a pair of models is given which describe arrangements of certain kinds of elementary particles, such as photons and protons. It turns out that experiments have shown that certain types of elementary particles exhibit behavior which is accurately described by one model, called while other types of elementary particles can be modelled using Feller:

*We have here an instructive example of the impossibility of selecting or justifying probability models by arguments. In fact, no pure reasoning could tell that photons and protons would not obey the same probability laws.*

We now give some examples of this description and assignment process.

### ✓ Example 3.2.9:

In the quantum mechanical model of the helium atom, various parameters can be used to classify the energy states of the atom. In the triplet spin state ( $S = 1$ ) with orbital angular momentum 1 ( $L = 1$ ), there are three possibilities, 0, 1, or 2, for the total angular momentum ( $J$ ). (It is not assumed that the reader knows what any of this means; in fact, the example is more illustrative if the reader does know anything about quantum mechanics.) We would like to assign probabilities to the three possibilities for  $J$ . The reader is undoubtedly resisting the idea of assigning the probability of  $1/3$  to each of these outcomes. She should now ask herself why she is resisting this assignment. The answer is probably because she does not have any “intuition” (i.e., experience) about the way in which helium atoms behave. In fact, in this example, the probabilities  $1/9$ ,  $3/9$ , and  $5/9$  are assigned by the theory. The theory gives these assignments because these frequencies were observed and further parameters were developed in the theory to allow these frequencies to be predicted.

### ✓ Example 3.2.10:

Suppose two pennies are flipped once each. There are several “reasonable” ways to describe the sample space. One way is to count the number of heads in the outcome; in this case, the sample space can be written  $\{0, 1, 2\}$ . Another description of the sample space is the set of all ordered pairs of  $H$ ’s and  $T$ ’s, i.e.,

$$\{(H, H), (H, T), (T, H), (T, T)\}. \quad (3.2.31)$$

Both of these descriptions are accurate ones, but it is easy to see that (at most) one of these, if assigned a constant probability function, can claim to accurately model reality. In this case, as opposed to the preceding example, the reader will probably say that the second description, with each outcome being assigned a probability of  $1/4$ , is the “right” description. This conviction is due to experience; there is no proof that this is the way reality works.

The reader is also referred to Exercise 3.2.26 for another example of this process.

## Historical Remarks

The binomial coefficients have a long and colorful history leading up to Pascal’s *Treatise on the Arithmetical Triangle*<sup>15</sup> where Pascal developed many important properties of these numbers. This history is set forth in the book by *Pascal’s Arithmetical Triangle* by A. W. F. Edwards.<sup>16</sup> Pascal wrote his triangle in the form shown in Table 3.10.

**Table 3.2.4**

1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	
1	3	6	10	15	21	28	36		
1	4	10	20	35	56	84			
1	5	15	35	70	126				
1	6	21	56	126					
1	7	28	84						
1	8	36							
1	9								
1									

(3.2.32)

**Table 3.2.5** Figurate numbers

natural numbers	1	2	3	4	5	6	7	8	9
triangular numbers	1	3	6	10	15	21	28	36	45
tetrahedral numbers	1	4	10	20	35	56	84	120	165

(3.2.33)

Edwards traces three different ways that the binomial coefficients arose. He refers to these as the *figurate numbers*, the *combinatorial numbers* and the *binomial numbers*. They are all names for the same thing (which we have called binomial coefficients) but that they are all the same was not appreciated until the sixteenth century.

The *figurate numbers* date back to the Pythagorean interest in number patterns around 540 . The Pythagoreans considered, for example, triangular patterns shown in Figure 3.2.5 The sequence of numbers

$$1, 3, 6, 10, \dots \quad (3.2.34)$$

obtained as the number of points in each triangle are called *triangular numbers*. From the triangles it is clear that the  $n$ th triangular number is simply the sum of the first  $n$  integers. The *tetrahedral numbers* are the sums of the triangular numbers and were obtained by the Greek mathematicians Theon and Nicomachus at the beginning of the second century . The tetrahedral number 10, for example, has the geometric representation shown in Figure 3.2.6. The first three types of figurate numbers can be represented in tabular form as shown in Table 3.2.5

natural \ numbers	1	2	3	4	5	6	7	8	9	
triangular \ numbers	1	3	6	10	15	21	28	36	45	(3.2.35)
tetrahedral \ numbers	1	4	10	20	35	56	84	120	165	

These numbers provide the first four rows of Pascal's triangle, but the table was not to be completed in the West until the sixteenth century.

In the East, Hindu mathematicians began to encounter the binomial coefficients in combinatorial problems. Bhaskara in his of 1150 gave a rule to find the number of medicinal preparations using 1, 2, 3, 4, 5, or 6 possible ingredients.<sup>17</sup> His rule is equivalent to our formula

$$\binom{n}{r} = \frac{(n)_r}{r!} . \quad (3.2.36)$$

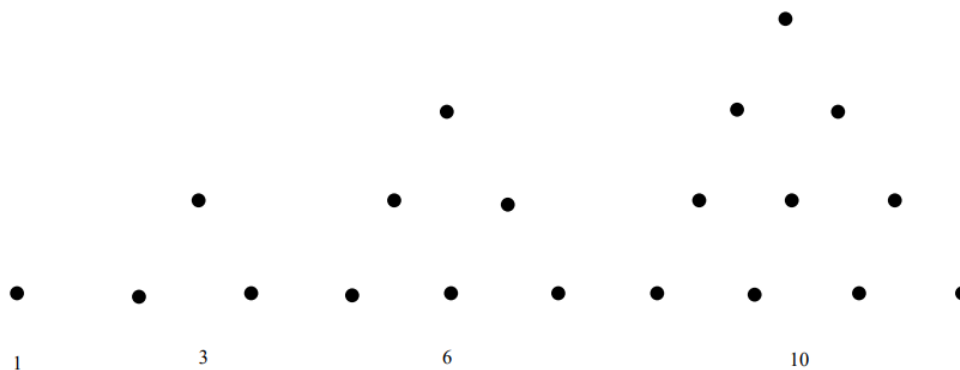


Figure 3.2.5: Pythagorean triangular patterns

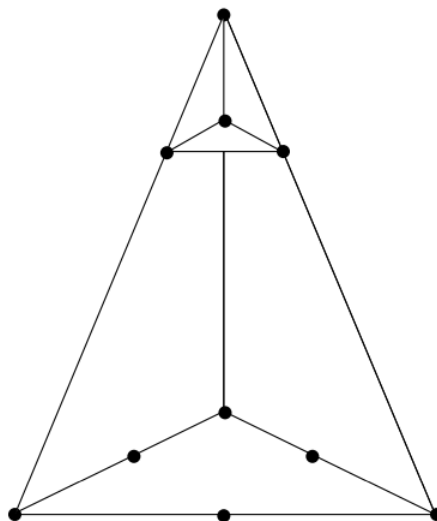


Figure 3.2.6: Geometric representation of the tetrahedral number 10.

Table 3.2.6: Outcomes for the roll of two dice

11					
12	22				
13	23	33			
14	24	34	44		
15	25	35	45	55	
16	26	36	46	56	66

(3.2.37)

The binomial numbers as coefficients of  $(a + b)^n$  appeared in the works of mathematicians in China around 1100. There are references about this time to “the tabulation system for unlocking binomial coefficients.” The triangle to provide the coefficients up to the eighth power is given by Chu Shih-chieh in a book written around 1303 (see Figure [fig 3.12]).<sup>18</sup> The original manuscript of Chu’s book has been lost, but copies have survived. Edwards notes that there is an error in this copy of Chu’s triangle. Can you find it? (: Two numbers which should be equal are not.) Other copies do not show this error.

The first appearance of Pascal’s triangle in the West seems to have come from calculations of Tartaglia in calculating the number of possible ways that  $n$  dice might turn up.<sup>19</sup> For one die the answer is clearly 6. For two dice the possibilities may be displayed as shown in Table [table 3.2.7].

Table 3.2.7 Outcomes for the roll of two dice

11					
12	22				
13	23	33			
14	24	34	44		
15	25	35	45	55	
16	26	36	46	56	66

(3.2.38)

Displaying them this way suggests the sixth triangular number  $1 + 2 + 3 + 4 + 5 + 6 = 21$  for the throw of 2 dice. Tartaglia “on the first day of Lent, 1523, in Verona, having thought about the problem all night,”<sup>20</sup> realized that the extension of the figurate table gave the answers for  $n$  dice. The problem had suggested itself to Tartaglia from watching people casting their own horoscopes by means of a *Book of Fortune* selecting verses by a process which included noting the numbers on the faces of three dice. The 56 ways that three dice can fall were set out on each page. The way the numbers were written in the book did not suggest the





wins the match. It is desired to find the probability that A wins the match at a time when A has won  $a$  games and B has won  $b$  games. (See Exercises 4.1.40-4.1.42.)

Pascal solved the problem by backward induction, much the way we would do today in writing a computer program for its solution. He referred to the combinatorial method of Fermat which proceeds as follows: If A needs  $c$  games and B needs  $d$  games to win, we require that the players continue to play until they have played  $c + d - 1$  games. The winner in this extended series will be the same as the winner in the original series. The probability that A wins in the extended series and hence in the original series is

$$\sum_{r=c}^{c+d-1} \frac{1}{2^{c+d-1}} \binom{c+d-1}{r} . \quad (3.2.39)$$

Even at the time of the letters Pascal seemed to understand this formula.

Suppose that the first player to win  $n$  games wins the match, and suppose that each player has put up a stake of  $x$ . Pascal studied the value of winning a particular game. By this he meant the increase in the expected winnings of the winner of the particular game under consideration. He showed that the value of the first game is

$$\frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} x . \quad (3.2.40)$$

His proof of this seems to use Fermat's formula and the fact that the above ratio of products of odd to products of even numbers is equal to the probability of exactly  $n$  heads in  $2n$  tosses of a coin. (See Exercise 39.)

Pascal presented Fermat with the table shown in Table 3.2.8

Table 3.2.8 Pascal's solution for the problem of points.

From my opponent's 256	6	5	4	3	2	1
positions I get, for the	games	games	games	games	games	games
1st game	63	70	80	96	128	256
2nd game	63	70	80	96	128	
3rd game	56	60	64	64		
4th game	42	40	32			
5th game	24	16				
6th game						

He states:

*You will see as always, that the value of the first game is equal to that of the second which is easily shown by combinations. You will see, in the same way, that the numbers in the first line are always increasing; so also are those in the second; and those in the third. But those in the fourth line are decreasing, and those in the fifth, etc. This seems odd.*<sup>23</sup>

The student can pursue this question further using the computer and Pascal's backward iteration method for computing the expected payoff at any point in the series.

In his treatise, Pascal gave a formal proof of Fermat's combinatorial formula as well as proofs of many other basic properties of binomial numbers. Many of his proofs involved induction and represent some of the first proofs by this method. His book brought together all the different aspects of the numbers in the Pascal triangle as known in 1654, and, as Edwards states, "That the Arithmetical Triangle should bear Pascal's name cannot be disputed."<sup>24</sup>

The first serious study of the binomial distribution was undertaken by James Bernoulli in his published in 1713.<sup>25</sup> We shall return to this work in the historical remarks in Chapter 8

## Exercises

**3.2.1**

Compute the following:

- a.  $\binom{6}{3}$
- b.  $b(5, .2, 4)$
- c.  $\binom{7}{2}$
- d.  $\binom{26}{26}$
- e.  $b(4, .2, 3)$
- f.  $\binom{6}{2}$
- g.  $\binom{10}{9}$
- h.  $b(8, .3, 5)$

**3.2.2**

In how many ways can we choose five people from a group of ten to form a committee?

**3.2.3**

How many seven-element subsets are there in a set of nine elements?

**3.2.4**

Using the relation Equation 3.1 write a program to compute Pascal's triangle, putting the results in a matrix. Have your program print the triangle for  $n = 10$ .

**3.2.5**

Use the program **BinomialProbabilities** to find the probability that, in 100 tosses of a fair coin, the number of heads that turns up lies between 35 and 65, between 40 and 60, and between 45 and 55.

**3.2.6**

Charles claims that he can distinguish between beer and ale 75 percent of the time. Ruth bets that he cannot and, in fact, just guesses. To settle this, a bet is made: Charles is to be given ten small glasses, each having been filled with beer or ale, chosen by tossing a fair coin. He wins the bet if he gets seven or more correct. Find the probability that Charles wins if he has the ability that he claims. Find the probability that Ruth wins if Charles is guessing.

**3.2.7**

Show that

$$b(n, p, j) = \frac{p}{q} \binom{n-j+1}{j} b(n, p, j-1), \quad (3.2.41)$$

for  $j \geq 1$ . Use this fact to determine the value or values of  $j$  which give  $b(n, p, j)$  its greatest value. : Consider the successive ratios as  $j$  increases.

**3.2.8**

A die is rolled 30 times. What is the probability that a 6 turns up exactly 5 times? What is the most probable number of times that a 6 will turn up?

**3.2.9**

Find integers  $n$  and  $r$  such that the following equation is true:

$$\binom{13}{5} + 2\binom{13}{6} + \binom{13}{7} = \binom{n}{r}. \quad (3.2.42)$$

**3.2.10**

In a ten-question true-false exam, find the probability that a student gets a grade of 70 percent or better by guessing. Answer the same question if the test has 30 questions, and if the test has 50 questions.

### 3.2.11

A restaurant offers apple and blueberry pies and stocks an equal number of each kind of pie. Each day ten customers request pie. They choose, with equal probabilities, one of the two kinds of pie. How many pieces of each kind of pie should the owner provide so that the probability is about .95 that each customer gets the pie of his or her own choice?

### 3.2.12

A poker hand is a set of 5 cards randomly chosen from a deck of 52 cards. Find the probability of a

- royal flush (ten, jack, queen, king, ace in a single suit).
- straight flush (five in a sequence in a single suit, but not a royal flush).
- four of a kind (four cards of the same face value).
- full house (one pair and one triple, each of the same face value).
- flush (five cards in a single suit but not a straight or royal flush).
- straight (five cards in a sequence, not all the same suit). (Note that in straights, an ace counts high or low.)

### 3.2.13

If a set has  $2n$  elements, show that it has more subsets with  $n$  elements than with any other number of elements.

### 3.2.14

Let  $b(2n, .5, n)$  be the probability that in  $2n$  tosses of a fair coin exactly  $n$  heads turn up. Using Stirling's formula (Theorem 3.1), show that  $b(2n, .5, n) \sim 1/\sqrt{\pi n}$ . Use the program **BinomialProbabilities** to compare this with the exact value for  $n = 10$  to 25.

### 3.2.15

A baseball player, Smith, has a batting average of .300 and in a typical game comes to bat three times. Assume that Smith's hits in a game can be considered to be a Bernoulli trials process with probability .3 for Find the probability that Smith gets 0, 1, 2, and 3 hits.

### 3.2.16

The Siwash University football team plays eight games in a season, winning three, losing three, and ending two in a tie. Show that the number of ways that this can happen is

$$\binom{8}{3} \binom{5}{3} = \frac{8!}{3! 3! 2!} . \quad (3.2.43)$$

### 3.2.17

Using the technique of Exercise 16, show that the number of ways that one can put  $n$  different objects into three boxes with  $a$  in the first,  $b$  in the second, and  $c$  in the third is  $n!/(a! b! c!)$ .

### 3.2.18

Baumgartner, Prosser, and Crowell are grading a calculus exam. There is a true-false question with ten parts. Baumgartner notices that one student has only two out of the ten correct and remarks, "The student was not even bright enough to have flipped a coin to determine his answers." "Not so clear," says Prosser. "With 340 students I bet that if they all flipped coins to determine their answers there would be at least one exam with two or fewer answers correct." Crowell says, "I'm with Prosser. In fact, I bet that we should expect at least one exam in which no answer is correct if everyone is just guessing." Who is right in all of this?

### 3.2.19

A gin hand consists of 10 cards from a deck of 52 cards. Find the probability that a gin hand has

- all 10 cards of the same suit.
- exactly 4 cards in one suit and 3 in two other suits.
- a 4, 3, 2, 1, distribution of suits.

### 3.2.20

A six-card hand is dealt from an ordinary deck of cards. Find the probability that:

1. All six cards are hearts.
2. There are three aces, two kings, and one queen.
3. There are three cards of one suit and three of another suit.

### 3.2.21

A lady wishes to color her fingernails on one hand using at most two of the colors red, yellow, and blue. How many ways can she do this?

### 3.2.22

How many ways can six indistinguishable letters be put in three mail boxes? : One representation of this is given by a sequence  $|LL|L|LLL|$  where the  $|$ 's represent the partitions for the boxes and the L's the letters. Any possible way can be so described. Note that we need two bars at the ends and the remaining two bars and the six L's can be put in any order.

### 3.2.23

Using the method for the hint in Exercise 22, show that  $r$  indistinguishable objects can be put in  $n$  boxes in

$$\binom{n+r-1}{n-1} = \binom{n+r-1}{r}$$

different ways.

### 3.2.24

A travel bureau estimates that when 20 tourists go to a resort with ten hotels they distribute themselves as if the bureau were putting 20 indistinguishable objects into ten distinguishable boxes. Assuming this model is correct, find the probability that no hotel is left vacant when the first group of 20 tourists arrives.

### 3.2.25

An elevator takes on six passengers and stops at ten floors. We can assign two different equiprobable measures for the ways that the passengers are discharged: (a) we consider the passengers to be distinguishable or (b) we consider them to be indistinguishable (see Exercise 23 for this case). For each case, calculate the probability that all the passengers get off at different floors.

### 3.2.26

You are playing with Prosser but you suspect that his coin is unfair. Von Neumann suggested that you proceed as follows: Toss Prosser's coin twice. If the outcome is HT call the result if it is TH call the result If it is TT or HH ignore the outcome and toss Prosser's coin twice again. Keep going until you get either an HT or a TH and call the result win or lose in a single play. Repeat this procedure for each play. Assume that Prosser's coin turns up heads with probability  $p$ .

1. Find the probability of HT, TH, HH, TT with two tosses of Prosser's coin.
2. Using part (a), show that the probability of a win on any one play is  $1/2$ , no matter what  $p$  is.

### 3.2.27

John claims that he has extrasensory powers and can tell which of two symbols is on a card turned face down (see Example 12). To test his ability he is asked to do this for a sequence of trials. Let the null hypothesis be that he is just guessing, so that the probability is  $1/2$  of his getting it right each time, and let the alternative hypothesis be that he can name the symbol correctly more than half the time. Devise a test with the property that the probability of a type 1 error is less than .05 and the probability of a type 2 error is less than .05 if John can name the symbol correctly 75 percent of the time.

### 3.2.28

In Example 3.2.12 assume the alternative hypothesis is that  $p = .8$  and that it is desired to have the probability of each type of error less than .01. Use the program **PowerCurve** to determine values of  $n$  and  $m$  that will achieve this. Choose  $n$  as small as possible.

### 3.2.29

A drug is assumed to be effective with an unknown probability  $p$ . To estimate  $p$  the drug is given to  $n$  patients. It is found to be effective for  $m$  patients. The for estimating  $p$  states that we should choose the value for  $p$  that gives the highest probability of getting what we got on the experiment. Assuming that the experiment can be considered as a Bernoulli trials process with probability  $p$  for success, show that the maximum likelihood estimate for  $p$  is the proportion  $m/n$  of successes.

### 3.2.30

Recall that in the World Series the first team to win four games wins the series. The series can go at most seven games. Assume that the Red Sox and the Mets are playing the series. Assume that the Mets win each game with probability  $p$ . Fermat observed that even though the series might not go seven games, the probability that the Mets win the series is the same as the probability that they win four or more game in a series that was forced to go seven games no matter who wins the individual games.

1. Using the program **PowerCurve** of Example 3.11 find the probability that the Mets win the series for the cases  $p = .5$ ,  $p = .6$ ,  $p = .7$ .
2. Assume that the Mets have probability .6 of winning each game. Use the program **PowerCurve** to find a value of  $n$  so that, if the series goes to the first team to win more than half the games, the Mets will have a 95 percent chance of winning the series. Choose  $n$  as small as possible.

### 3.2.31

Each of the four engines on an airplane functions correctly on a given flight with probability .99, and the engines function independently of each other. Assume that the plane can make a safe landing if at least two of its engines are functioning correctly. What is the probability that the engines will allow for a safe landing?

### 3.2.32

A small boy is lost coming down Mount Washington. The leader of the search team estimates that there is a probability  $p$  that he came down on the east side and a probability  $1 - p$  that he came down on the west side. He has  $n$  people in his search team who will search independently and, if the boy is on the side being searched, each member will find the boy with probability  $u$ . Determine how he should divide the  $n$  people into two groups to search the two sides of the mountain so that he will have the highest probability of finding the boy. How does this depend on  $u$ ?

### \*3.2.33

$2n$  balls are chosen at random from a total of  $2n$  red balls and  $2n$  blue balls. Find a combinatorial expression for the probability that the chosen balls are equally divided in color. Use Stirling's formula to estimate this probability. Using **BinomialProbabilities**, compare the exact value with Stirling's approximation for  $n = 20$ .

### 3.2.34

Assume that every time you buy a box of Wheaties, you receive one of the pictures of the  $n$  players on the New York Yankees. Over a period of time, you buy  $m \geq n$  boxes of Wheaties.

1. Use Theorem 3.8 to show that the probability that you get all  $n$  pictures is

$$1 - \binom{n}{1} \left( \frac{n-1}{n} \right)^m + \binom{n}{2} \left( \frac{n-2}{n} \right)^m - \dots + (-1)^{n-1} \binom{n}{n-1} \left( \frac{1}{n} \right)^m.$$

: Let  $E_k$  be the event that you do not get the  $k$ th player's picture.

2. Write a computer program to compute this probability. Use this program to find, for given  $n$ , the smallest value of  $m$  which will give probability  $\geq .5$  of getting all  $n$  pictures. Consider  $n = 50$ , 100, and 150 and show that  $m = n \log n + n \log 2$  is a good estimate for the number of boxes needed. (For a derivation of this estimate, see Feller.<sup>26</sup>)

### 3.2.35

Prove the following *binomial identity*

$$\binom{2n}{n} = \sum_{j=0}^n \binom{n}{j}^2. \quad (3.2.44)$$

: Consider an urn with  $n$  red balls and  $n$  blue balls inside. Show that each side of the equation equals the number of ways to choose  $n$  balls from the urn.

**3.2.36** Let  $j$  and  $n$  be positive integers, with  $j \leq n$ . An experiment consists of choosing, at random, a  $j$ -tuple of integers whose sum is at most  $n$ .

1. Find the size of the sample space. : Consider  $n$  indistinguishable balls placed in a row. Place  $j$  markers between consecutive pairs of balls, with no two markers between the same pair of balls. (We also allow one of the  $n$  markers to be placed at the end of the row of balls.) Show that there is a 1-1 correspondence between the set of possible positions for the markers and the set of  $j$ -tuples whose size we are trying to count.
2. Find the probability that the  $j$ -tuple selected contains at least one 1.

### 3.2.37

Let  $n \pmod m$  denote the remainder when the integer  $n$  is divided by the integer  $m$ . Write a computer program to compute the numbers  $\binom{n}{j} \pmod m$  where  $\binom{n}{j}$  is a binomial coefficient and  $m$  is an integer. You can do this by using the recursion relations for generating binomial coefficients, doing all the arithmetic using the basic function  $\text{mod}(n, m)$ . Try to write your program to make as large a table as possible. Run your program for the cases  $m = 2$  to 7. Do you see any patterns? In particular, for the case  $m = 2$  and  $n$  a power of 2, verify that all the entries in the  $(n - 1)$  st row are 1. (The corresponding binomial numbers are odd.) Use your pictures to explain why this is true.

### 3.2.38

Lucas<sup>27</sup> proved the following general result relating to Exercise 37. If  $p$  is any prime number, then  $\binom{n}{j} \pmod p$  can be found as follows: Expand  $n$  and  $j$  in base  $p$  as  $n = s_0 + s_1p + s_2p^2 + \cdots + s_kp^k$  and  $j = r_0 + r_1p + r_2p^2 + \cdots + r_kp^k$ , respectively. (Here  $k$  is chosen large enough to represent all numbers from 0 to  $n$  in base  $p$  using  $k$  digits.) Let  $s = (s_0, s_1, s_2, \dots, s_k)$  and  $r = (r_0, r_1, r_2, \dots, r_k)$ . Then

$$\binom{n}{j} \pmod p = \prod_{i=0}^k \binom{s_i}{r_i} \pmod p \quad (3.2.45)$$

For example, if  $p = 7$ ,  $n = 12$ , and  $j = 9$ , then

$$\begin{aligned} 12 &= 5 \cdot 7^0 + 1 \cdot 7^1, \\ 9 &= 2 \cdot 7^0 + 1 \cdot 7^1, \end{aligned}$$

so that

$$\begin{aligned} s &= (5, 1), \\ r &= (2, 1), \end{aligned}$$

and this result states that

$$\binom{12}{9} \pmod p = \binom{5}{2} \binom{1}{1} \pmod 7. \quad (3.2.46)$$

Since  $\binom{12}{9} = 220 = 3 \pmod 7$ , and  $\binom{5}{2} = 10 = 3 \pmod 7$ , we see that the result is correct for this example.

Show that this result implies that, for  $p = 2$ , the  $(p^k - 1)$  st row of your triangle in Exercise 37 has no zeros.

### 3.2.39

Prove that the probability of exactly  $n$  heads in  $2n$  tosses of a fair coin is given by the product of the odd numbers up to  $2n - 1$  divided by the product of the even numbers up to  $2n$ .

### 3.2.40

Let  $n$  be a positive integer, and assume that  $j$  is a positive integer not exceeding  $n/2$ . Show that in Theorem 3.5, if one alternates the multiplications and divisions, then all of the intermediate values in the calculation are integers. Show also that none of these intermediate values exceed the final value.

---

This page titled [3.2: Combinations](#) is shared under a [GNU Free Documentation License 1.3](#) license and was authored, remixed, and/or curated by [Charles M. Grinstead & J. Laurie Snell](#) ([American Mathematical Society](#)) via [source content](#) that was edited to the style and standards of the LibreTexts platform.