CHAPTER 3

Probability

The study of probability grew out of the gambling dens of Europe over three centuries ago, when a few known mathematicians felt challenged to calculate optimum strategies for winning certain bets and, later, when patrons, often the aristocracy, further encouraged mathematicians with sums of money for providing such valued information. Much of what we study today is still based on this early work, especially in our analysis of proportion data and broadly in the use of basic terminology and underlying principles throughout the text.

In this chapter, we will study the concepts of probability mostly to help us understand and define what is meant by p, the proportion or percentage of a population that possesses a certain attribute, such as, the proportion or percentage of a population that possesses the attribute of red hair or diabetes. Later, we will discuss what we can expect when we sample from such a population.

In broad terms, we define **probability** as follows.

Probability
The proportion (or percentage) of times an event will occur in the long run, under similar circumstances. This probability is expressed as a number between 0 and 1 or as a percentage between 0 and 100.

Two methods we can use to obtain this probability are the empirical approximation and classical approach. We shall start with the empirical approximation. ▼

3.1 Probability Defined: Empirically

Empirical Approximation to Probability
The fraction of times an event *has actually occurred* over a great many experiments conducted over a long period of time under similar circumstances, expressed as

$$P(\text{an event}) \approx \frac{\text{Number of Times Event Has Actually Occurred}}{\text{Total Number of Experiments (or Attempts)}}$$

As the number of experiments increases, this empirical fraction gets closer and closer to the true probability.

It is important to note that use of the empirical fraction requires an experiment be performed a great many times over a long period and under similar circumstances.

Let's see how it works.

Example

Out of 100 consecutive tosses of a dart at a Thursday night tournament, a player strikes the bull's-eye 30 times. Calculate the probability of a bull's-eye for this player.

Solution

Although we might be tempted to say,

$$P(\text{an event}) \approx \frac{\text{Number of Times the Event Has Actually Occurred}}{\text{Total Number of Attempts}}$$

$$P(\text{bull's-eye}) \approx \frac{30}{100}\text{, this is incorrect.}$$

This does not meet our full definition of empirical probability because we have not performed the experiment a great many times over a long period, and therefore this may or may not be the true probability. So the answer to this question is: from this limited information, *we cannot determine the probability* of a bull's-eye for this player. ■

Well, then, you might ask, how *do* we obtain this probability? One way is to merely continue to perform the experiment (tossing the dart) a large number of times. As the total number of dart throws increase, this fraction (which we will call the **cumulative fraction**) gets closer and closer to the true probability, expressed in the following law.

▼ *Law of Large Numbers*

If we continue to repeat an experiment a great many times under similar circumstances, the cumulative fraction of successes will tend to draw closer and closer to the *true* probability.

Let's see how it works.

Say, for instance, for the first 100 tosses, we get 30 bull's-eyes. Then for the next 100 tosses, we get 12 bull's-eyes. Although the two fractions for each experiment would be 30/100 and 12/100, the *cumulative* fraction is 42 bull's-eyes (30 + 12) out of 200 total tosses, which would equal in percentage:

$$\frac{42 \text{ bull's-eyes}}{200 \text{ tosses}} = 21\%$$

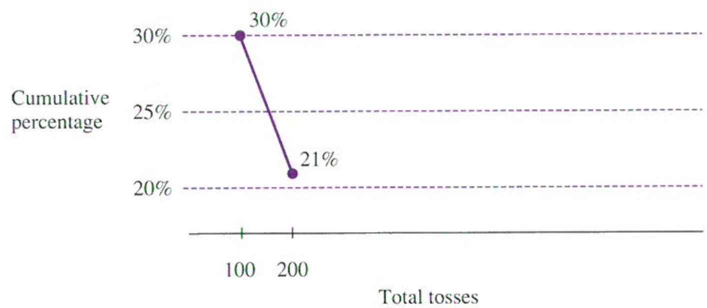
Note: 42/200 can be converted to a percentage as follows:

$$\text{Percent} = \text{Fraction} \times 100$$
$$= \frac{42}{\cancel{200}_2} \times \cancel{100}^1 = 21\%$$

We can present this cumulative fraction in chart form as follows:

Number of Tosses	Bull's-eyes	Cumulative Fraction	Cumulative Percentage
100	30	30/100	30%
100	12	42/200	21%

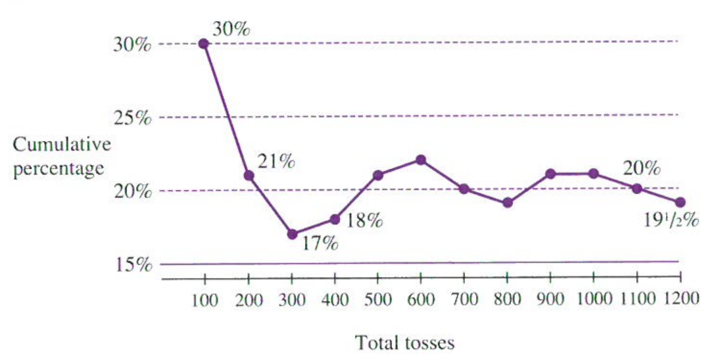
If we plotted these results on a graph, it would look as follows:



Now, what if we recorded a third set of 100 tosses where 9 bull's-eyes were achieved. The cumulative fraction would be 51 bull's-eyes (30 + 12 + 9) out of 300 total tosses, which equals 51/300 or 17%. And if we recorded a fourth set of 100 tosses, and so on for many sets of 100 tosses, we could represent this as follows:

Number of Tosses	Bull's-eyes	Cumulative Fraction	Cumulative Percentage
100	30	30/100	30%
100	12	42/200	21%
100	9	51/300	17%
100	21	72/400	18%
100	33	105/500	21%
100	27	132/600	22%
100	8	140/700	20%
100	12	152/800	19%
100	37	189/900	21%
100	21	210/1000	21%
100	10	220/1100	20%
100	14	234/1200	$19\frac{1}{2}$%

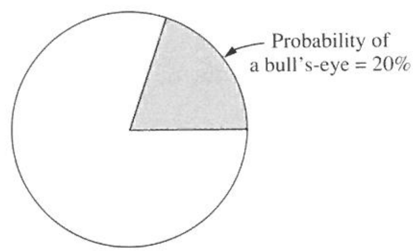
Note that the cumulative fraction adds up all the prior bull's-eyes and all the prior tosses and presents this as one fraction. If we plot each cumulative percentage, we get



Notice that even though the number of bull's-eyes for each 100 tosses drastically fluctuated in the chart, ranging from 8 to 37, creating a rather ragged pattern at the beginning, the long-term line begins to smooth out as the number of tosses increases.

Experience has shown, in the long run, provided no change in dart throwing ability of the player or other factors that might affect this ability, the line will eventually grow flat (horizontal), sticking very close to one particular value. When this happens over many tosses, we call this percentage the *true* probability of a bull's-eye for this player.

If we examine the preceding graph, we see the probability seems to be leveling at about 20%. If indeed 20% is the true probability, we might represent this probability on a circle graph as follows.



Probability, in essence, is a *population* value, the true percentage of times the player will hit the bull's-eye in the long run under similar circumstances. In effect, it represents what will happen in millions and millions of tosses. This probability may be presented as a percentage or decimal as follows:

$$P(\text{bull's-eye}) = 20\%$$
$$P(\text{bull's-eye}) = .20$$

Later, when we discuss sampling, this population value, 20%, will be referred to as the population proportion, p, expressed as

$$p = 20\% \quad \text{or}$$
$$p = .20$$

This empirical method for assigning a probability to an event is often used in the fields of psychology, education, biology, business, and medicine. For instance, if a surgeon says you have a 95% chance of surviving an operation, the surgeon is usually referring to an empirical probability. That is, in the long run, over many similar operations in the past, about 95 out of 100 people have survived.

Subjective Probability

One word of caution: too often, people will offer you probabilities off the top of their heads, which are no more than personal judgments. It is wise to request and examine the source of all probabilities offered when these probabilities are to be used in any subsequent decision-making process. Although some people are surprisingly astute in their ability to assess the probability of a situation, which is called **subjective probability,** others are not.

> ***Subjective Probability***
> Assigning a probability based on personal judgment.

In this text, we use only those probabilities derived from the empirical method, as described in the last few pages, or from the classical method, which is discussed in the next section.

3.2 Probability Defined: Classically

The empirical definition of probability demands that we repeat an experiment a great many times before we can estimate the true probability. However, in practical situations this is not always possible. In a situation where we can be assured that every member of a set has an *equal chance of being selected*, then we have a second method for obtaining a probability. This method is called the classical approach.

> ***Classical Probability***
> In an experiment of n different possibilities, each having an *equal chance of occurring*, the probability a particular event will occur is equal to
>
> $$\frac{\text{Number of chances for success}}{\text{Total number of equally likely possibilities}}$$
>
> This can be expressed as
>
> $$P(\text{event}) = \frac{s}{n} \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely possibilities)}\end{array}$$
>
> Often the word, event, is replaced with the word, success, as follows:
>
> $$P(\text{success}) = \frac{s}{n}$$

Example

Suppose you attend a party of 20 people, of which 3 are famous TV celebrities. Now let's pretend a huge Green Giant were to walk up, lift the roof, reach down into the party and *randomly* pluck up 1 person by the collar. What is the probability the person selected is a famous TV celebrity?

Solution

This is a typical probability experiment. Out of $n = 20$ different possibilities (in this case, 20 different people), each having an *equal chance* of being selected, there are 3 chances for success. Since we have 3 chances for success out of 20 different equally likely possibilities, the probability of selecting a famous TV celebrity is given by the following formula:

$$P(\text{success}) = \frac{s \text{ (number of chances for success)}}{n \text{ (total equally likely possibilities)}}$$

$$P(\text{of selecting a famous TV celebrity}) = \frac{3}{20}$$ ■

Often these fractions are expressed as percentages or decimals. For instance, the fraction $\frac{3}{20}$ may be expressed as 15% or .15. To convert the fraction $\frac{3}{20}$ to its equivalent percentage or decimal, we perform the following operation:

$$\begin{aligned} \% &= \text{Fraction} \times 100 \\ &= \frac{3}{20} \times 100 \\ &= \frac{3}{\cancel{20}_1} \times \overset{5}{\cancel{100}} \\ &= 15\% \end{aligned}$$

$$\begin{aligned} \text{Decimal} &= \text{Numerator of fraction} \div \text{Denominator} \\ &= 3 \div 20 = .15 \end{aligned}$$

or

$$\begin{array}{r} .15 \\ 20\overline{)3.00} \\ -2\,0 \\ \hline 1\,00 \\ -1\,00 \\ \hline \end{array} = .15$$

So, instead of saying that the probability of selecting a famous TV celebrity is $\frac{3}{20}$, we might state it as 15% or .15. Whether we use $\frac{3}{20}$ or 15% or .15, these all indicate the same probability, which is 3 chances for success out of 20 possibilities.

Two Fundamental Properties

In dealing with probability fractions, there are two fundamental properties.

▼ *Property 1*

The probability of an event occurring will always be a number between 0 and 1, inclusive.

$$0 \leq P(\text{an event}) \leq 1$$

no chance to occur (0%)

certain to occur (100%)

A probability of $P = 0$ means the event has no chance of occurring. If at our party of 20 people we had *no* famous TV celebrities in attendance, then if the Giant reached in and selected one person, the probability of selecting a famous TV celebrity is 0 chances for success out of 20 possibilities or

$$P(\text{success}) = \frac{0}{20} = 0$$

$P = 0$ is the minimum probability. This means the event has no chance of occurring and can be expressed as $P = 0\%$.

A probability of $P = 1$ means the event is certain to occur. If at our party of 20 people we had 20 famous TV celebrities (in other words, all were famous TV celebrities), then if the Giant reached in and selected one person, the probability of selecting a famous TV celebrity is 20 chances for success out of 20 possibilities or

$$P(\text{success}) = \frac{20}{20} = 1$$

$P = 1$ is the maximum probability. This means the event is certain to occur and can be expressed as $P = 100\%$.

▼ *Property 2*

The probability of an event occurring *plus* the probability of the event *not* occurring = 1.

$$P(E) + P(\text{not } E) = 1$$

Basically this property says there is a 100% probability that the event will either occur or not occur.

Example

What if at our party of 20 people, of which 3 are famous TV celebrities, a Green Giant *randomly* plucks up one person by the collar, what is the probability the person selected is *not* a famous TV celebrity?

Solution

Since we now have 17 chances for success (people who are *not* TV celebrities) out of 20 equally likely possibilities,

$$P(\text{success}) = \frac{s \text{ (number of chances for success)}}{n \text{ (total equally likely possibilities)}}$$

$$P(\text{of } \textit{not} \text{ selecting a famous TV celebrity}) = \frac{17}{20}$$

This can also be expressed as 85% or .85. ■

Note that in our party example, the probability of *not* selecting a TV celebrity ($\frac{17}{20}$) plus the probability of selecting a TV celebrity ($\frac{3}{20}$) is equal to 1. This is a direct illustration of Property 2, as follows:

$$\text{Property 2: } P(E) + P(\text{not } E) = 1$$

$$\text{Expressed in fractions: } \frac{3}{20} + \frac{17}{20} = \frac{20}{20} = 1$$

$$\text{Expressed in percentages: } 15\% + 85\% = 100\% = 1$$

$$\text{Expressed in decimals: } .15 + .85 = 1.00 = 1$$

Example

For a 52-card deck, if we randomly select one card, what is the probability that the card will be

a. a club?
b. not a club?
c. a king or a queen?
d. a jack and a spade?

Solution

To apply our probability formula, we must be assured every card has an *equal chance* of being selected. *Random* selection guarantees this, so now we can proceed.

a. Each deck is divided into four suits: 13 clubs, 13 diamonds, 13 hearts, and 13 spades. Because we have 13 chances for success (13 clubs) out of $n =$ 52 equally likely possibilities,

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely possibilities)}\end{array}$$

$$P(\text{club}) = \frac{13}{52} \quad (\tfrac{13}{52} \text{ can also be expressed as } 25\% \text{ or } .25)$$

b. Since we have 13 clubs in a deck, we must have 39 (52 minus 13) *non*clubs, thus 39 chances for success.

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely possibilities)}\end{array}$$

$$P(\text{not club}) = \frac{39}{52} \quad (\tfrac{39}{52} \text{ can also be expressed as } 75\% \text{ or } .75)$$

c. Each deck has 4 kings and 4 queens. Since we have 8 chances for success ($4 + 4 = 8$) out of 52 equally likely possibilities,

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely possibilities)}\end{array}$$

$$P(\text{king or queen}) = \frac{8}{52} \quad \begin{array}{l}(\tfrac{8}{52} \text{ can also be expressed as}\\ \text{approximately } 15\% \text{ or } .15)\end{array}$$

d. Note that the conditions for success in this problem require that we have a jack *and* a spade. The word, *and,* in statistics means that *both* conditions must be met for success. Since in a deck of cards we have only one card that satisfies both conditions, namely the jack of spades, we have only 1 chance for success out of 52 equally likely possibilities.

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely possibilities)}\end{array}$$

$$P(\text{jack } \textit{and} \text{ spade}) = \frac{1}{52} \quad \begin{array}{l}(\tfrac{1}{52} \text{ can also be expressed as}\\ \text{approximately } 2\% \text{ or approximately } .02)\end{array}$$ ■

AND and OR Statements

Notice that the words **and** and **or** have very specific meanings in statistics.

And — The word *and* between two or more conditions or events implies *all* must be met for success.

Or — The word *or* between two or more conditions or events implies that *either* one or more may be met and that will give success.

Although many shortcut formulas are available to solve probability problems, each comes with restrictions that limit their use to a well-defined set of circumstances that can be quite confusing. It is best to first learn to solve these simple experiments (where we select *one* from a set of possibilities) by the methods described above. The following are offered for practice.

Practice Exercises

Practice 1

From a 52-card deck, if we *randomly* select one card, what is the probability the card will be

a. a heart *or* an ace?
b. a king *and* an ace?

Solution

a. Because we have 16 chances for success (13 hearts plus the aces of clubs, diamonds, and spades) out of 52 equally likely possibilities,

$$P(\text{heart } \textit{or} \text{ ace}) = \frac{16}{52}$$

Note that the ace of hearts was already counted in the 13 hearts.

b. An *and* statement means both conditions must be met for success. Since there are 0 chances for success,

$$P(\text{king } \textit{and} \text{ ace}) = \frac{0}{52} = 0$$

Note that in selecting *one* card, it is impossible to get both a king *and* an ace. ■

Practice 2

A chain of family video stores has their movies rated G, PG, R, X, or XX with the following probabilities:

$$P(\text{G}) = .31 \quad P(\text{R}) = .30 \quad P(\text{XX}) = .04$$
$$P(\text{PG}) = .25 \quad P(\text{X}) = .10$$

If you were to randomly select a video, what would be the probability the video would be rated

a. not G?
b. R or X or XX?

Solution

a. $P(\text{G}) = .31$ means 31 out of 100 were rated G. Therefore, 69 $(100 - 31 = 69)$ must have been rated something other than G. So,

$$P(\text{not G}) = \frac{69}{100} \quad (\text{or } .69)$$

b. Out of 100, 30 were rated R, 10 were rated X, and 4 were rated XX. Thus, 44 $(30 + 10 + 4)$ were rated either R or X or XX out of 100. So,

$$P(\text{R or X or XX}) = \frac{44}{100} \quad (\text{or } .44)$$

■

Practice 3

In a regional survey of 1000 customers, a particular cable TV show was rated either favorable or unfavorable. Out of 760 favorable responses, 400 were female. And out of the 240 unfavorable responses, 80 were female.

If you were to randomly select one respondent from the survey, what would be the probability the respondent would be

a. female?
b. favorable and female?
c. favorable and male?
d. unfavorable *or* female?
e. male, given we already know the respondent voted favorable?

(Note: this is referred to as a **conditional probability**.)

Solution

The information above can be summarized as follows:

1000 Customers

Favorable	**Unfavorable**
400 F	80 F
360 M	160 M

a. Because 480 were female $(400 + 80)$,

$$P(\text{female}) = \frac{480}{1000} \quad (\text{or } .48)$$

b. Since it is stated that 400 of the favorable responses were female,

$$P(\text{favorable and female}) = \frac{400}{1000} \quad (\text{or } .40)$$

c. 760 rated cable TV favorably, of which 400 were female. This implies 360 were male. Thus,

$$P(\text{favorable and male}) = \frac{360}{1000} \quad (\text{or } .36)$$

d. An *or* statement implies *either* condition will give success. However, we must be careful not to count the same person twice. Because we have 240 unfavorable responses plus 400 *additional* females (from the favorable responses), we have 640 chances for success (240 + 400) out of 1000. Thus,

$$P(\text{unfavorable } or \text{ female}) = \frac{640}{1000} \quad (\text{or } .64)$$

Note: if we had reasoned there were 480 *total* females plus *240* total unfavorable responses, we would have counted 80 females *twice.*

e. The clause "given we already know the respondent voted favorable," is referred to as a *conditional* and the question referred to as a *conditional probability*. In effect, this conditional "given . . . favorable," limits the total set of possibilities to 760 favorable responses, out of which 360 males give us success. Thus, P(male, given we know the vote was favorable) = 360 chances for success out of 760 total possibilities.

$$P(\text{male, given favorable}) = \frac{360}{760} \quad (\text{or } .47).$$

Note: there were 400 females and 360 males in the 760 favorable responses.

Use of Mathematical Formulas in Simple Experiments

Although it is usually easier to solve simple experiments (where we select *one* from a set of possibilities) without using a formula, formulas are available and are most often preferred. We will demonstrate two common formulas with practice problem 3, parts d and e.

Practice 3(d)

Referring to practice problem 3, part d, if you were to randomly select one respondent from this survey, what would be the probability the respondent would be unfavorable *or* female?

Solution

The solution to practice problem 3, part d, can be solved by something known as the **addition rule:** Let E_1 = the first event and E_2 = the second event defined in a sample space, then

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

This subtracts out the elements counted twice.

Defining E_1 as unfavorable and E_2 as female,

$$P(\text{unfavorable } \textit{or} \text{ female}) = P(\text{unfavorable}) + P(\text{female}) - P(\text{unfavorable and female})$$

$$= \frac{240}{1000} + \frac{480}{1000} - \frac{80}{1000} = \frac{640}{1000}$$

Note: 80 females had to be subtracted out, expressed as the probability 80/1000, since they were counted twice, once in the unfavorable group and a second time in the female group.

If we have the situation where $P(E_1 \text{ and } E_2) = 0$, meaning the two events cannot occur together, then the events are referred to as **mutually exclusive** and the above formula reduces to $P(E_1 \text{ or } E_2) = P(E_1) + P(E_2)$.

Practice 3(e)

Referring to practice problem 3, part e, if you were to randomly select *one* respondent from this survey, what would be the probability the respondent would be male, given we already know the respondent voted favorable?

Solution

The solution to practice problem 3, part e, can also be solved by a form of multiplication rule (to be discussed more fully later in the chapter), defined as

$$P(E_1, \text{ given } E_2) = \frac{P(E_1 \text{ and } E_2)}{P(E_2)}$$

$$P(\text{male, given favorable}) = \frac{P(\text{male and favorable})}{P(\text{favorable})}$$

Again, this is referred to as a conditional probability.

$$= \frac{360/1000}{760/1000} = \frac{360}{760}$$

3.3 More Complex Experiments: Tree Diagram

In the preceding examples, we selected *one* from a set of equally likely possibilities. For instance, we selected *one* person from a party of 20 people, and we selected *one* card from a deck of 52. However, in more complex problems, where we select *two* or more from a set of possibilities, it is often helpful to list all the equally likely outcomes of the experiment using a technique known as a **tree diagram.**

Example

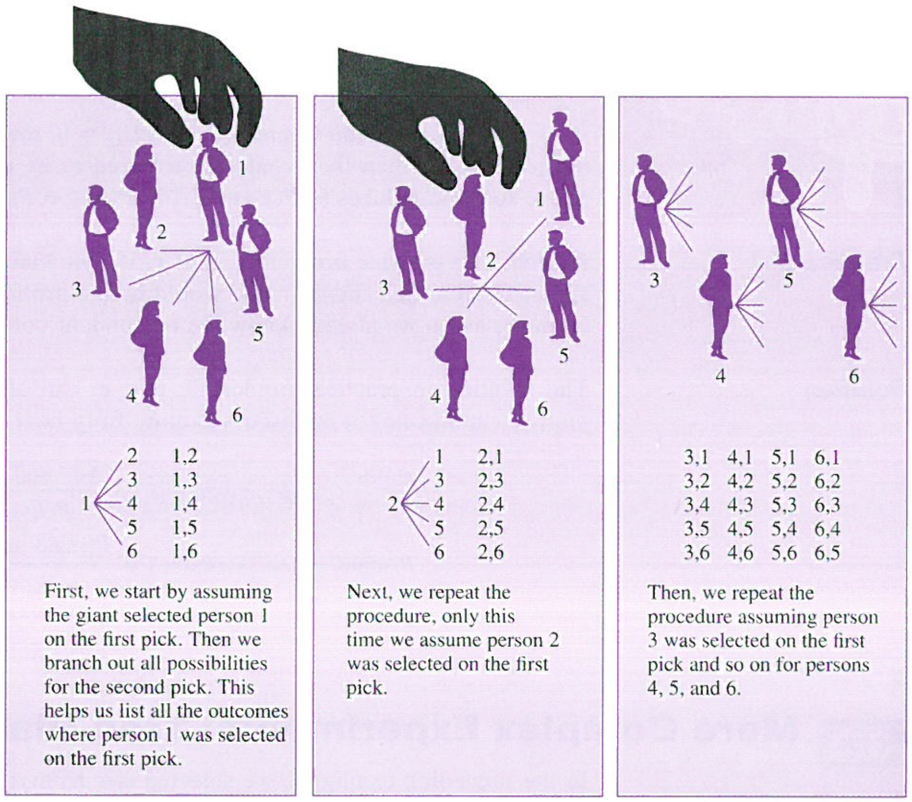
Suppose we return to our party example. Only this time we attend a party of six people, of which two are famous TV celebrities. We shall use the numbers, 1, 2, 3, and 4 to identify the *non*-TV celebrities and the numbers 5 and 6 to identify the two famous TV celebrities.

Now, what if a huge Green Giant were to walk up, lift the roof, and reach down into the party and randomly pluck up two people? What is the probability the two will both be famous TV celebrities?

Solution

If the Giant had selected one person from this party, this would have been a rather simple experiment. Selecting two is a little more complex as you will see.

To help us list all the equally likely ways we can select two people, we use a technique known as a *tree diagram*, as follows:



Our complete listing would be as follows.

1,2	2,1	3,1	4,1	5,1	6,1
1,3	2,3	3,2	4,2	5,2	6,2
1,4	2,4	3,4	4,3	5,3	6,3
1,5	2,5	3,5	4,5	5,4	6,4
1,6	2,6	3,6	4,6	5,6	6,5

$n = 30$ equally likely outcomes
$s = 2$ chances for success (circled)

This is called a **sample space** of equally likely outcomes. It represents all the ways the giant could have reached in and first selected one person, then reached in and selected a second person.

Note that outcome 5,6 is considered different from outcome 6,5. It is important to maintain the order of selection. This ensures that we have a complete listing of *all* equally likely outcomes.

Now we are ready to answer our question.

Since we have 2 chances for success (circled above) out of 30 equally likely outcomes,

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely outcomes*)}\end{array}$$

$$P\left(\begin{smallmatrix}\text{celebrity}\\ \text{1st pick}\end{smallmatrix} \textit{ and } \begin{smallmatrix}\text{celebrity}\\ \text{2nd pick}\end{smallmatrix}\right) = \frac{2}{30} \quad \text{(or 6.7\%)}$$

■

Now let's suppose we asked the same question but selected in a different manner.

Example Suppose at our party of six, of which two are famous TV celebrities, a Green Giant plucked up one person by the collar, *replaced* that one person into the party, then later returned and plucked up one person again.

What is the probability both picks would be famous TV celebrities?

Solution If we list our sample space of equally likely outcomes, we get

1,1	2,1	3,1	4,1	5,1	6,1
1,2	2,2	3,2	4,2	5,2	6,2
1,3	2,3	3,3	4,3	5,3	6,3
1,4	2,4	3,4	4,4	5,4	6,4
1,5	2,5	3,5	4,5	5,5	6,5
1,6	2,6	3,6	4,6	5,6	6,6

n = 36 equally likely outcomes
s = 4 chances for success (circled)

Note the addition of outcomes 1,1 and 2,2 and 3,3 and 4,4 and 5,5 and 6,6. Since we replaced the first person, we must include the possibility that the *same* person might be chosen twice.

Because we now have 4 chances for success (circled above) out of 36 equally likely outcomes,

$$P(\text{success}) = \frac{s}{n} \quad \begin{array}{l}\text{(number of chances for success)}\\ \text{(total equally likely outcomes)}\end{array}$$

$$P\left(\begin{smallmatrix}\text{celebrity}\\ \text{1st pick}\end{smallmatrix} \textit{ and } \begin{smallmatrix}\text{celebrity}\\ \text{2nd pick}\end{smallmatrix}\right) = \frac{4}{36} \quad \text{(or 11.1\%)}$$

■

We can also list sample spaces for three or more selections, as follows.

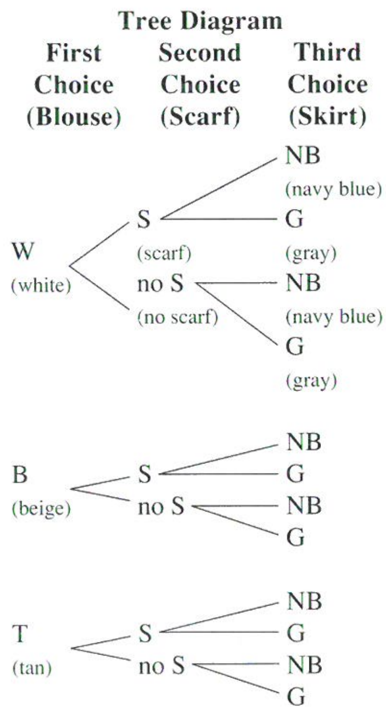
*Note: In more complex experiments, the word outcomes is used in place of possibilities.

Example

In a woman's wardrobe of 3 blouses (white, beige, and tan), 1 scarf, and 2 skirts (navy blue and gray), a blouse and a skirt must be worn but a *scarf* is optional.

If we assume the woman *randomly* selected from each group, what is the probability the woman will be wearing a scarf *and* gray skirt?

Solution



Equally Likely Outcomes

W, S, NB
W, S, G
W, no S, NB
W, no S, G

B, S, NB
B, S, G
B, no S, NB
B, no S, G

T, S, NB
T, S, G
T, no S, NB
T, no S, G

n = 12 equally likely outcomes
s = 3 chances for success

Since we have 3 chances for success (circled) out of 12 equally likely ways an outfit can be put together,

$$P(\text{success}) = \frac{s}{n} \quad \frac{\text{(number of chances for success)}}{\text{(total equally likely outcomes)}}$$

$$P(\text{scarf and gray skirt}) = \frac{3}{12} \quad \text{(or 25\%)}$$

Note that we randomly sampled from each category. If color matching or personal taste were involved, the solution would be more complex.

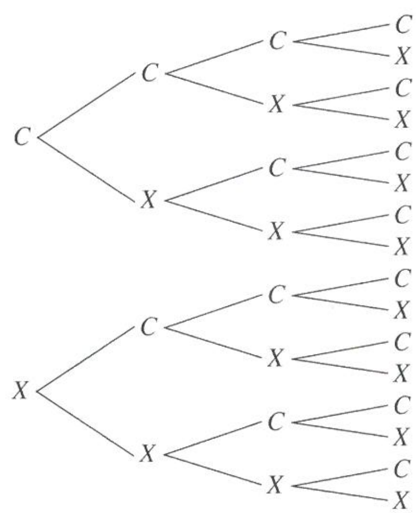
Example

Suppose we take a true–false quiz of four questions and we *randomly* guess on each answer, what is the probability of getting every question correct?

Solution

Let C = correct answer and X = incorrect answer, our sample space of equally likely outcomes would be as follows.

First Question **Second Question** **Third Question** **Fourth Question** **Equally Likely Outcomes**



C, C, C, C
C, C, C, X
C, C, X, C
C, C, X, X
C, X, C, C
C, X, C, X
C, X, X, C
C, X, X, X
X, C, C, C
X, C, C, X
X, C, X, C
X, C, X, X
X, X, C, C
X, X, C, X
X, X, X, C
X, X, X, X

n = 16 equally likely outcomes
s = 1 chance for success (circled)

Because we have 1 chance for success (circled) out of 16 equally likely outcomes,

$$P(\text{success}) = \frac{s \text{ (number of chances for success)}}{n \text{ (total equally likely possibilities)}}$$

$$P(C \text{ and } C \text{ and } C \text{ and } C) = \frac{1}{16} \quad \text{(or approximately 6.3\%)}$$

■

3.4 More Complex Experiments: Multiplication Rules

Complex experiments (in which we select two or more from a set of possibilities) may also be solved by formula. However, we must be careful since each formula comes with restrictions that limits its use to a well-defined set of circumstances.

Dependent and Independent Events

In the case where two events are **dependent** (that is, the occurrence or nonoccurrence of one event affects the probability associated with the other event), we use the **general multiplication rule,** as follows.

▼ *General Multiplication Rule*

For two dependent events

If E_1 and E_2 are two defined events in a sample space, then

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2, \text{ given } E_1 \text{ has already occurred})$$

$P(E_2$, given E_1 has already occurred) is often symbolically expressed as $P(E_2 \mid E_1)$. $P(E_1$ and $E_2)$ also equals $P(E_2)P(E_1$, given E_2 has already occurred).

In special cases, where the two events have no effect on each other's probability, we call the events **independent** (that is, the occurrence or nonoccurrence of one event has no effect on the probability of the other event). In which case, the multiplication rule is greatly simplified to:

▼ *Special Multiplication Rule*

For two independent events

If E_1 and E_2 are two independent events in a sample space, then,

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2)$$

Note that in $P(E_2)$ we eliminated the condition "given E_1 has already occurred."

For three or more independent events

If we can guarantee all events are independent, then we can expand our formula to three or more events as follows:

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \ldots) = P(E_1)P(E_2)P(E_3) \ldots$$

Let's see how these rules apply to the examples we studied in the prior section.

Example

We attend a party of 6 people, of which 2 are famous TV celebrities. If some Green Giant *randomly* plucked up 2 people by the collar, what is the probability both will be famous TV celebrities? (Note: this is the first example solved in section 3.3 using a tree diagram.)

Solution

First we define "selecting a celebrity on the 1st pick" as *event one* (E_1) and "selecting a celebrity on the 2nd pick" as *event two* (E_2).

Since the two events are *dependent* (that is, whether or not we select a celebrity on the 1st pick affects the probability of selecting a celebrity on the 2nd pick), thus we use the general multiplication rule:

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2, \text{ given } E_1 \text{ has already occurred})$$

$$P\left(\substack{\text{celebrity}\\\text{1st pick}} \text{ and } \substack{\text{celebrity}\\\text{2nd pick}}\right) = P\left(\substack{\text{celebrity}\\\text{1st pick}}\right) P\left(\substack{\text{celebrity}\\\text{2nd pick}}, \text{ given we chose a celebrity on 1st pick}\right)$$

$$= \frac{2}{6} \cdot \frac{1}{5} = \frac{2}{30} \quad (6.7\%)$$

Notice that getting a celebrity on the first pick was $\frac{2}{6}$ as we would expect. However, now, we assume the first pick occurred, that is, we picked a celebrity out of the party. Thus, we have only 5 people left at the party with only 1 celebrity left, so the probability of selecting a celebrity on the second pick is 1 chance of success out of 5 possibilities or $\frac{1}{5}$. When we multiply $\frac{2}{6}$ and $\frac{1}{5}$, we get the same answer as we did in the prior section $\frac{2}{30}$ (or 6.7%). ■

One additional comment about *dependence*. Note in the above case, whether or not we selected a celebrity on the 1st pick affects the probability associated with selecting a celebrity on the 2nd pick. In other words, if a celebrity was chosen on the 1st pick, then P(celebrity on 2nd pick) equals $\frac{1}{5}$ (as stated above). However, if a celebrity was *not* chosen on the 1st pick, the P(celebrity on 2nd pick) equals $\frac{2}{5}$, which is quite different. This is what is meant when we say the occurrence or nonoccurrence of one event affects the probability associated with another event—that is, the two events are dependent.

The multiplication rule can also be expanded to include three or more *dependent* events, demonstrated in the following example.

Example

Three cards are randomly selected from a 52-card deck. Calculate the probability all will be kings.

Solution

We can expand the multiplication rule as follows: $P(E_1 \text{ and } E_2 \text{ and } E_3) = P(E_1) P(E_2, \text{ given } E_1 \text{ occurred}) P(E_3, \text{ given } E_1 \text{ and } E_2 \text{ occurred})$

Let
K_1 = selecting a king on 1st pick
K_2 = selecting a king on 2nd pick
K_3 = selecting a king on 3rd pick

$P(K_1 \text{ and } K_2 \text{ and } K_3) = P(K_1)P(K_2, \text{ given } K_1 \text{ occurred})\, P(K_3, \text{ given } K_1 \text{ and } K_2 \text{ occurred})$

$$= \frac{4}{52} \cdot \frac{3}{51} \cdot \frac{2}{50} = \frac{24}{132{,}600} \quad \text{(near 0\%)}$$

Note that getting a king on the 1st pick was $\frac{4}{52}$, as we would expect. However, let's assume that this occurred—that is, we picked a king out of the deck. Thus, we have only 3 kings left out of 51 cards, so the probability of getting a king on the 2nd pick is $\frac{3}{51}$. Now, we assume the first two events occurred, that is, both kings were picked from the deck, so now we have 2 kings left out of 50 cards. Thus, the probability of getting a king on the 3rd pick is $\frac{2}{50}$. ■

As you can see, these formulas do save us time. In other words, we do not have to construct sample spaces of equally likely outcomes to calculate probabilities. For instance, in the example given above, we would have had to construct a sample space of 132,600 equally likely outcomes to solve this problem. A formidable task, indeed. But as easy as formulas are to use, they do come with restrictions that limit their use to a well-defined set of circumstances. And we must be careful to apply them exactly as presented.

Let's consider the second example from section 3.3 to demonstrate the conditions and restrictions for use of the *special* multiplication rule (for independent events).

Example Suppose at our party of 6, of which 2 are famous TV celebrities, a Green Giant plucked up one person by the collar, *replaced that one person* into the party, then later returned and plucked up one person again. (Note: This is the second example solved in section 3.3 using a tree diagram.)

What is the probability both picks would be famous TV celebrities?

Solution The two events are now *independent* since we replaced the first person. In other words, whether we select or do not select a celebrity on the 1st pick in no way influences the probability of selecting a celebrity on the 2nd pick.

$$P(E_1 \text{ and } E_2) = P(E_1)P(E_2)$$

$$P\left({\text{celebrity} \atop \text{1st pick}} \text{ and } {\text{celebrity} \atop \text{2nd pick}}\right) = \frac{2}{6} \cdot \frac{2}{6} = \frac{4}{36} \quad (11.1\%)$$

Notice that the probability on the 1st pick was $\frac{2}{6}$. However, because we replaced the person back into the party, we still had six people at the party with two famous TV celebrities. So the probability of success on the 2nd pick was also $\frac{2}{6}$. When multiplied, the answer is $\frac{4}{36}$ (or 11.1%), which is the same answer we achieved when listing our sample space of equally likely outcomes in section 3.3. ■

One comment about *independence* in the preceding example: note the probability on the 2nd pick would have been $\frac{2}{6}$ no matter what the outcome of the 1st pick (whether we picked a celebrity or not on the 1st pick). This is what is meant by independence.

If we can guarantee all events are independent, that is, the outcome of one event in no way influences the probability of any other event, we can use the expanded form of the special multiplication rule, as follows:

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \ldots) = P(E_1)P(E_2)P(E_3) \ldots$$

Example Suppose at our party of 6, of which 2 are famous TV celebrities, the Green Giant chooses 3 people, one at a time, *but replaces* each after the person was chosen. What is the probability all 3 will be famous TV celebrities?

Solution Because replacement in this case assures independence

$$P\left({\text{celebrity} \atop \text{1st pick}} \text{ and } {\text{celebrity} \atop \text{2nd pick}} \text{ and } {\text{celebrity} \atop \text{3rd pick}}\right) = \frac{2}{6} \cdot \frac{2}{6} \cdot \frac{2}{6} = \frac{8}{216} \quad (\text{or } 3.7\%)$$ ■

Example

In a woman's wardrobe of 3 blouses (white, beige, and tan), 1 scarf, and 2 skirts (navy blue and gray), a blouse and a skirt must be worn but a *scarf* is optional. (Note: This is the third example solved in section 3.3 using a tree diagram.)

If we assume the woman randomly selected from each group, what is the probability the woman will be wearing a scarf *and* gray skirt?

Solution

If we *randomly* select from each group, we essentially have three independent choices, since a choice from one group does not affect the probability of a choice from any other group. Thus

$$\begin{aligned} P(E_1 \text{ and } E_2 \text{ and } E_3) &= P(E_1)P(E_2)P(E_3) \\ P(\text{any blouse and scarf and gray skirt}) &= P(\text{any blouse})\,P(\text{scarf})\,P(\text{gray skirt}) \\ &= \frac{3}{3}\cdot\frac{1}{2}\cdot\frac{1}{2} \\ &= \frac{3}{12} \quad (\text{or } 25\%) \end{aligned}$$

Notice we must choose a blouse and all choices give us success. Because a scarf is optional, we have two choices (scarf or no scarf) of which one gives us success (scarf). We must choose a skirt and there is one chance for success (a gray skirt) out of two possibilities. ■

Note that we *randomly* selected from each category. If color matching or personal taste were involved, we would have lost our independence. In other words, say, if the woman selected a tan blouse, she may have personal preference for wearing the scarf and navy blue skirt, in which case the probabilities associated with the 2nd and 3rd choices would be greatly affected. These type of problems (where the choices are dependent) are much more difficult to solve, since they require knowledge of the extent one choice affects another.

Violation of independence is often the reason why statistical studies go awry since this requirement of independence is necessary for many of the formulas we use later in the text when we sample. If for some reason independence is violated, we may be obliged to use the general multiplication rule (which assesses the effect each event has on the probability of subsequent events). Unfortunately, in many experiments, especially those involving people such as in psychological or educational studies, these effects are quite difficult to assess and sometimes impossible. So it is important when we design a study to do our best to ensure independence when we can.

Let's examine this property of independence with one more example.

Example

Suppose we take a true–false quiz of 4 questions and we *randomly* guess on each question, what is the probability of getting every question correct? (Note: This is the fourth example solved in section 3.3 using a tree diagram.)

Solution

If we guess randomly, this ensures *independence*, that is, whether we are correct or not correct on one guess in no way affects the probability of being correct on any other guess. Thus,

$$P(E_1 \text{ and } E_2 \text{ and } E_3 \text{ and } E_4) = P(E_1)P(E_2)P(E_3)P(E_4)$$

Let

$$\begin{aligned} C_1 &= \text{correct on 1st question} \\ C_2 &= \text{correct on 2nd question} \\ C_3 &= \text{correct on 3rd question} \\ C_4 &= \text{correct on 4th question} \end{aligned}$$

$$\begin{aligned} P(C_1 \text{ and } C_2 \text{ and } C_3 \text{ and } C_4) &= P(C_1)P(C_2)P(C_3)P(C_4) \\ &= \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{16} \quad (6.3\%) \end{aligned}$$

Remember, for independence, the outcome on one event must in no way affect the probabilities associated with any other event.

Counting Principle

Although the counting principle can be used in many circumstances, one of its uses is to quickly count the total number of possible outcomes for an experiment.

▼ *Counting Principle*

For a sequence of events, in which the first event can occur in *a* ways, the second in *b* ways, and the third in *c* ways, and so on, the total number of ways the events can occur together is

$$a \cdot b \cdot c \ldots$$

Example

For our woman's wardrobe experiment of 3 blouses, scarf or no scarf, and 2 skirts, how many outcomes are possible?

Solution

Because the first event, picking a blouse, can occur in 3 ways; the second event, choosing a scarf, in 2 ways; and the third event, selecting a skirt, in 2 ways,

$$\begin{aligned} \text{Total number of possible outcomes} &= 3 \cdot 2 \cdot 2 \\ &= 12 \end{aligned}$$

■

Example

At our party of 6 people with 2 famous TV celebrities, how many ways can we select 2 people, given that

a. we replace the first person before we select the second?
b. we do not replace the first person?

Solution

a. Since the first event, selecting the first person, can occur in 6 ways, and the second event, selecting the second person, can occur in 6 ways, then

$$\begin{aligned}\text{Total number of possible outcomes} &= 6 \cdot 6\\ &= 36\end{aligned}$$

b. Since the first event can occur in 6 ways and the second event can now only occur in 5 ways,

$$\begin{aligned}\text{Total number of possible outcomes} &= 6 \cdot 5\\ &= 30\end{aligned}$$ ■

Other circumstances in which we may use the counting principle are as follows.

Example In the 714 telephone area, how many different telephone numbers are possible?

Solution For a telephone number we have seven events:

___ ___ ___ - ___ ___ ___ ___

For the first event, selection of the first digit, we have 8 choices (2, 3, 4, 5, 6, 7, 8, and 9). Note that we cannot use the digit 0 or 1 as the first digit of a telephone number. For each other digit, we have 10 choices (0, 1, 2, 3, 4, 5, 6, 7, 8, and 9).

$$\begin{aligned}\begin{matrix}\text{Total number of ways these}\\ \text{seven events can occur}\end{matrix} &= 8 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10 \cdot 10\\ &= 8{,}000{,}000 \text{ telephone numbers}\end{aligned}$$ ■

Example Suppose a particular state wishes to use 3 letters followed by 3 digits for an automobile license plate. How many different license plates are possible?

Solution For this license plate, we have six selections,

___ ___ ___ - ___ ___ ___

As long as there are no restrictions on which letters or digits can be used, we have 26 choices each for the first three selections and 10 choices each for the last three selections.

$$\begin{aligned}\text{Total number of ways these six events can occur} &= 26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10\\ &= 17{,}576{,}000 \text{ license plates}\end{aligned}$$ ■

In probability experiments, the counting formula is often used to count the total number of equally likely outcomes for an experiment, denoted by n in the probability formula,

$$P = \frac{s}{n} \quad \begin{matrix}\text{(number of chances for success)}\\ \text{(total equally likely outcomes)}\end{matrix}$$

The counting formula can give you the total number of possible outcomes rather quickly, however the formula generally does not lend itself for use in calculating s, the number of chances for success.

3.5 Early Gambling Experiments Leading to Discovery of the Normal Curve

Early gambling experiments usually involving the tossing of coins and dice form the theoretical underpinning for many of our formulas and the statistical procedures we use today in statistics (such as, chi-square analysis and tests of proportions, which are discussed at length in chapters 10 and 11). However, these early experiments also paved the way for the original discovery of one of our most fundamental statistical tools, the normal curve, which is the subject of chapter 4. It is in these regards that we present the following.

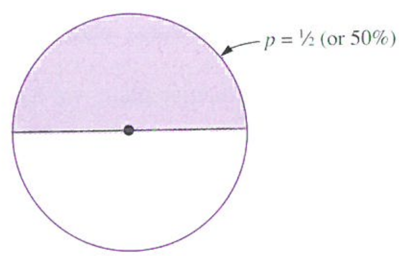
Example

A fair coin is tossed 4 times. Bets are taken as to the number of heads that would turn up.

a. Use a tree diagram to list all equally likely outcomes.
b. Construct a histogram demonstrating how many times we would expect 0 heads, 1 head, 2 heads, 3 heads, and 4 heads to occur.
c. Find the probability of achieving 2 heads.
d. Would you bet even money on 2 heads? Explain your reason.

Rationale

Essentially we are sampling from a huge population of coin flips, where 50% possess the attribute of heads.



In other words, if we were to flip this coin millions and millions of times, 50% (or extremely close to 50%) would be heads.

Now, if we were to sample from this huge population of coin flips, in this case we are sampling 4 tosses (that is, we are selecting a sample size of $n = 4$), what may we expect to happen? We know from theory and a long history of experience, that if we were to randomly sample from such a population,

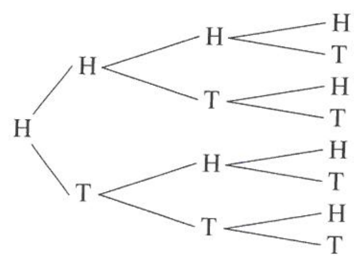
$p_s \approx p$ The sample proportion, p_s, will be approximately equal to the population proportion, p.

That is, since the population consists of 50% heads, the sample should consist of *approximately* 50% heads. In the case of $n = 4$ tosses, we should get *approximately* 2 heads. However, we can also get 1 head or 4 heads. How can we determine the percentage of times we can expect each of these outcomes to occur?

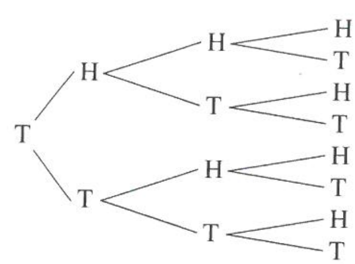
One way is to use a tree diagram to list all the equally likely outcomes that can occur when 4 coins are tossed and simply count the number of times we achieved zero heads, one head, two heads, three heads, and four heads, as follows:

Solution

a. We use a tree diagram to help us list the sample space of equally likely outcomes.



H, H, H, H
H, H, H, T
H, H, T, H
H, H, T, T
H, T, H, H
H, T, H, T
H, T, T, H
H, T, T, T



T, H, H, H
T, H, H, T
T, H, T, H
T, H, T, T
T, T, H, H
T, T, H, T
T, T, T, H
T, T, T, T

Notice we have 16 equally likely outcomes.

b. To construct the histogram,

H, H, H, H
H, H, H, T
H, H, T, H
H, H, T, T

H, T, H, H
H, T, H, T
H, T, T, H
H, T, T, T

T, H, H, H
T, H, H, T
T, H, T, H
T, H, T, T

T, T, H, H
T, T, H, T
T, T, T, H
T, T, T, T

First, we count how many times we obtained 0 heads. This occurred on one occasion, circled at left, so we indicate this with one slash mark in our tally.

	Tally
0 Heads	|
1 Head	
2 Heads	
3 Heads	
4 Heads	

H, H, H, H
H, H, H, T
H, H, T, H
H, H, T, T

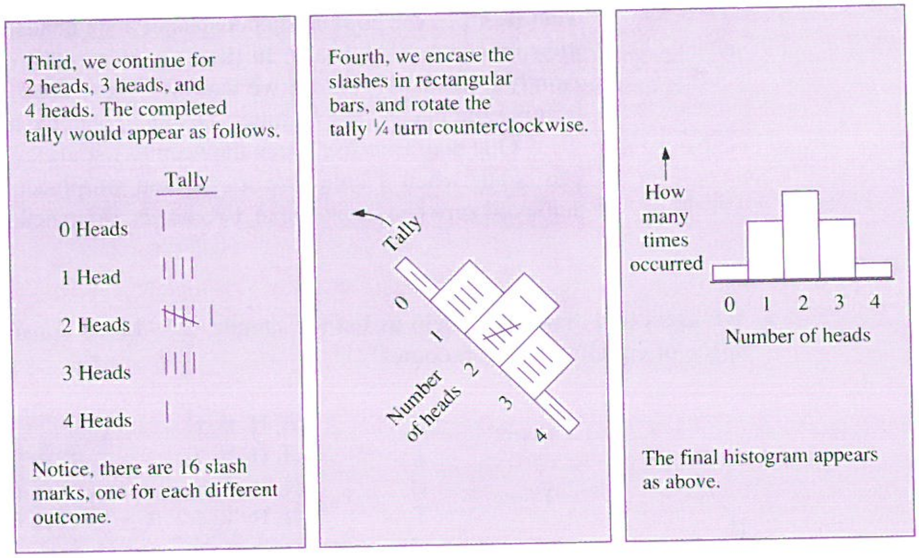
H, T, H, H
H, T, H, T
H, T, T, H
H, T, T, T

T, H, H, H
T, H, H, T
T, H, T, H
T, H, T, T

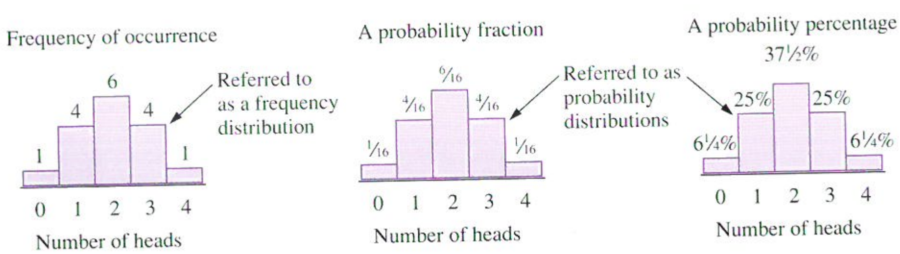
T, T, H, H
T, T, H, T
T, T, T, H
T, T, T, T

Second, we count how many times we obtained 1 head. This occurred on four occasions, circled at left, so we indicate this with four slash marks in our tally.

	Tally
0 Heads	|
1 Head	||||
2 Heads	
3 Heads	
4 Heads	



The histogram bars are usually labeled in one of three different ways in terms of



No matter which way the bars are labeled, however, they all give us the same information, as follows:

$$
\begin{aligned}
P(0 \text{ heads}) &= 1 \text{ chance in } 16 = 1/16 = 6\tfrac{1}{4}\% \\
P(1 \text{ head}) &= 4 \text{ chances in } 16 = 4/16 = 25\% \\
P(2 \text{ heads}) &= 6 \text{ chances in } 16 = 6/16 = 37\tfrac{1}{2}\% \\
P(3 \text{ heads}) &= 4 \text{ chances in } 16 = 4/16 = 25\% \\
P(4 \text{ heads}) &= 1 \text{ chance in } 16 = 1/16 = 6\tfrac{1}{4}\%
\end{aligned}
$$

c. So, to answer the question of the probability of 2 heads, we merely look at the results. You have *6 chances out of 16* = $37\frac{1}{2}\%$.

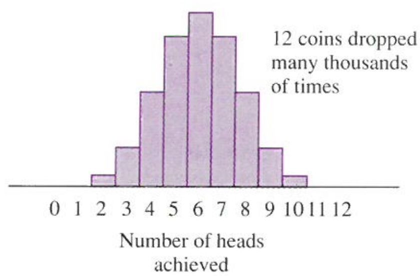
d. If you did bet even money on 2 heads, unfortunately in the long run you would lose your money. Out of every 16 times you played the game, you can expect to win on 6 occasions and lose on 10. Sometimes this is expressed as odds, 6:10 (meaning, in the long run, you would average 6 wins and 10 losses out of each 16 plays). ■

Although the above probabilities were based on mathematical analysis (that is, by constructing sample spaces), experience has shown these to give reasonably accurate estimates of what we would expect to occur in the long run in actual practice. In other words, if we were to drop 4 coins on a table thousands and thousands of times and record the number of heads achieved on each drop, we would find we would get

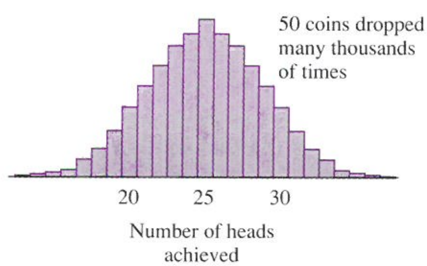
0 heads approximately	$6\frac{1}{4}$% of the time
1 head approximately	25% of the time
2 heads approximately	$37\frac{1}{2}$% of the time
3 heads approximately	25% of the time
4 heads approximately	$6\frac{1}{4}$% of the time

Now, what if we decided to drop $n = 12$ coins or $n = 50$ coins on a table, what would we expect to get?

Well, we can determine this either mathematically, by constructing a sample space, or we can actually drop 12 coins (or 50 coins) thousands and thousands of times and tally the result.*



Suppose we drop $n = 12$ coins on a table thousands and thousands of times and record the number of heads achieved on each drop, we would get a distribution something like this figure.



Suppose we drop $n = 50$ coins on a table thousands and thousands of times and recorded the number of heads achieved on each drop, we would get a distribution something like this figure.

Look at the two histograms above. Both are symmetrical around the value we would most likely expect to occur. In the case of dropping $n = 12$ coins, we would most likely expect approximately 6 heads (50% of 12 = 6). And indeed we do most often get 6 heads. However on a great many occasions we get somewhat more than 6 heads, and on a great many occasions somewhat less, with the heights of the histogram bars seeming to fall off in the shape of a bell.

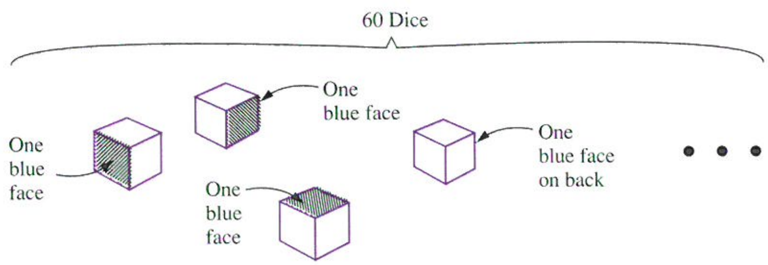
*Actually, simpler techniques and formulas are available to calculate these probabilities, which are discussed at the end of chapter 4 and in chapter 11.

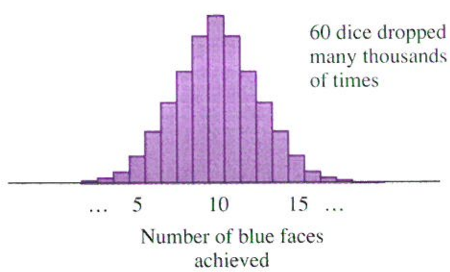
Notice we have a similar situation in dropping $n = 50$ coins. The distribution is symmetrical around the value we would most likely expect, in this case 25 heads (50% of 50 = 25). However on a great many occasions we get somewhat more than 25 heads, and on a great many occasions somewhat less, with the heights of the histogram bars again falling off in the shape of a bell.

This bell-shaped pattern appears repeatedly with these coin experiments.

Now, you might ask, this may happen with coin tosses, where the probability of a head for a coin toss is $\frac{1}{2}$ (50%), but what if we sampled from a different population, say die tosses, where the probability of a particular face turning up is $\frac{1}{6}$ ($16\frac{2}{3}$%)? What happens then?

Okay, let's take 60 dice and paint one face on each blue (for identification purposes).





Say we drop $n = 60$ dice on a table thousands and thousands of times, and each time recorded the number of blue faces that turned up. If we tally the results into a histogram, we would get a distribution something like this figure.

Notice the shape of the distribution. It is symmetrical about the value we would most likely expect, in this case 10 blue faces. In other words, we would expect about 1 die in 6 to turn up blue, or 10 in 60. And indeed 10 blue faces would be our most frequently occurring value if we had constructed a sample space or in practice if we actually dropped 60 dice many thousands of times. However, on many occasions we get somewhat more than 10 blue faces and on many occasions somewhat less, again with the heights of the histogram bars falling off in the shape of a bell.

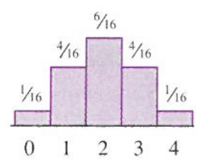
Bell-shaped distributions kept occurring with amazing regularity in these coin experiments and in the dice experiments when the number of dice dropped was large. This repetitive bell-shaped pattern in gambling experiments led De Moivre (in 1733) to the initial discovery of one of the most powerful predictive tools we have in all of statistics and the topic of our next chapter, the normal distribution.

3.6 Additional Probability Topics

Although the material presented thus far provides the main thrust of the techniques needed for future chapters, several additional topics are often encountered and are presented here. Specifically we introduce the mean and standard deviation of a discrete probability distribution, expected value, and permutations and combinations (combinations are further discussed in chapter 11, section 11.1).

Mean and Standard Deviation of a Discrete Probability Distribution

In section 3.5, we constructed a distribution that offered probabilities associated with achieving a given number of heads when a fair coin is tossed four times, as follows:



Number of heads achieved in four tosses

This is called a discrete probability distribution. In such a distribution, each outcome for an experiment is associated with a specific probability of occurrence and the sum of all probabilities equals 1.00. (Note in the above case, $\frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = \frac{16}{16} = 1.00$.)

To calculate the mean and standard deviation of such a probability distribution, we use the following two formulas.

Mean

$$\mu = \Sigma\, xp\,(x)$$

x: one outcome
p(*x*): probability of achieving this outcome

For this example,
$\mu = \Sigma xp(x)$
$\mu = 0(1/16) + 1(4/16) + 2(6/16) + 3(4/16) + 4(1/16)$
$\mu = 2$

Standard Deviation

$$\sigma = \sqrt{\Sigma\,(x - \mu)^2 p\,(x)}$$

For this example,
$\sigma = \sqrt{\Sigma(x - \mu)^2 p\,(x)}$
$\sigma = \sqrt{(0 - 2)^2(1/16) + (1 - 2)^2(4/16) + (2 - 2)^2(6/16) + (3 - 2)^2(4/16) + (4 - 2)^2(1/16)}$
$\sigma = \sqrt{1} = 1$

> ***Probability Distribution***
> A distribution that offers the probabilities associated with each possible outcome of an experiment, such that the sum of these probabilities always equals 1.00.

> ***Discrete Probability Distribution***
> A probability distribution where each possible outcome of an experiment can only be one of a limited number of discrete values, that is, a value that when presented on a number line occupies only a distinct isolated point.

In other words, in the above experiment, we could only achieve 0, 1, 2, 3, or 4 heads as outcomes, a limited number of distinct isolated points on a number line. Note we could never achieve $1\frac{1}{4}$ heads or $3\frac{1}{2}$ heads or any values other than these limited isolated values of 0, 1, 2, 3, or 4.

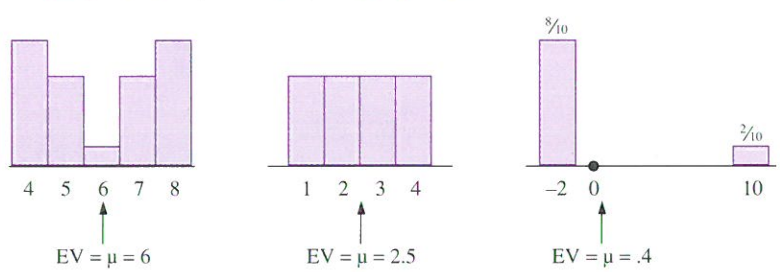
The term **discrete** is discussed again in section 4.4 when we introduce one of the most frequently encountered discrete probability distributions, called the binomial sampling distribution.

Expected Value

In the example just presented, we calculated the mean and standard deviation (μ and σ) of a discrete probability distribution. This mean (μ) is often referred to as the **expected value.**

> ***Expected Value***
> The long-range average of a repeated experiment, essentially the population mean, μ.
>
> $$\text{Expected value (EV)} = \mu = \Sigma\, xp(x)$$

In our experiment of tossing four coins, since $\mu = 2$ heads, this is the expected value. Note in this particular experiment, the expected value is also the most frequently occurring outcome (with probability $\frac{6}{16}$; refer to the histogram at the beginning of the section). However this is not always the case, as shown in the following discrete probability distributions:



Note in the first example, the expected value, μ, has the *least* probability of occurring, and in the other two examples the expected value, μ, has *no* probability of occurring. Keep in mind, expected value is merely an average, the average value you would likely calculate if you repeated an experiment many many thousands of times, added up all the outcomes, and divided by n, the total number of values.

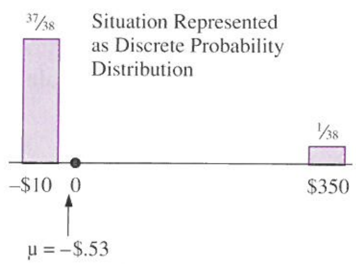
This concept of expected value (average value for a repeated experiment) is used extensively in business and scientific investigations but in honor of its origins in gambling, we introduce the following example:

Example

Roulette is a game played by placing a bet that a ball will tumble into a particular pocket of a spinning wheel. There are 38 available pockets (numbered 1 to 36, 0, and 00), so your chance of winning is $\frac{1}{38}$, while your chance of losing is $\frac{37}{38}$. On a \$10 bet, say, a gambling house will return \$360 if you win, however the house will keep your \$10 if you lose. What is the EV, the expected value, for this experiment? In other words, if you played the game many thousands of times, added up all the winnings and losses, and divided by n, the number of times you played, what would be your expected average (the expected value)?

Solution

This is solved as follows:



$$EV = \Sigma x\, p(x)$$
$$EV = (-10)(^{37}\!/_{38}) + (350)(^{1}\!/_{38})$$
$$\boxed{EV = -\$.53}$$

Important
Note the true gain for a win is \$350, not \$360 (since \$10 was originally yours)

Thus the expected value, μ, for this experiment is $-\$.53$, meaning if you repeated this experiment (betting \$10) many thousands of times, added up all your losses ($-\$10$'s) and wins ($+\$350$), and divided by n, the number of times you placed a bet, the average would be about $-\$.53$. In other words, in the long run, over many thousands of bets, you will average losing 53¢ for each bet you place. Of course, from the point of view of the gambling house, they gain, *in the long run,* on average, 53¢ for each bet you place. ■

Permutations and Combinations

In section 3.4, we introduced a useful method for quickly counting the total number of possible outcomes for an experiment, called the *counting principle.* In addition to the counting principle, other counting techniques are available, such as the **permutation** and **combination,** denoted $P_{n,s}$ and $C_{n,s}$ respectively.

A **permutation** counts the number of ways *n different* objects can be arranged, *s* at a time, where order of arrangement is important.

$$P_{n,s} = \frac{n!}{(n-s)!}$$

Factorial Symbol, !, is defined as

$$n! = n(n-1)(n-2)\ldots 1$$

For example,

$$5! = 5(4)(3)(2)(1) = 120$$
$$3! = 3(2)(1) = 6$$
$$1! = 1$$
$$0! = 1 \text{ by definition}$$

Let's look at permutations.

Example Suppose there are 5 books (A, B, C, D, and E) that are to be placed in 3 available positions on a shelf. How many different arrangements are possible?

Solution If order of arrangement is important, meaning book arrangement A, B, C is considered different from book arrangement A, C, B (even though the same books are used), we use the permutation formula, as follows:

$$P_{n,r} = \frac{n!}{(n-r)!}$$

$$P_{5,3} = \frac{5!}{(5-3)!} = \frac{5!}{2!}$$

$$P_{5,3} = \frac{5(4)(3)(2)(1)}{(2)(1)}$$

$$P_{5,3} = 60 \text{ ways}$$

60 Ways Listed

ABC	BAC	CAB	DAB	EAB
ABD	BAD	CAD	DAC	EAC
ABE	BAE	CAE	DAE	EAD
ACB	BCA	CBA	DBA	EBA
ACD	BCD	CBD	DBC	EBC
ACE	BCE	CBE	DBE	EBD
ADB	BDA	CDA	DCA	ECA
ADC	BDC	CDB	DCB	ECB
ADE	BDE	CDE	DCE	ECD
AEB	BEA	CEA	DEA	EDA
AEC	BEC	CEB	DEB	EDB
AED	BED	CED	DEC	EDC

■

In situations where order of arrangement is *not* important, say for instance, if from 5 books (A, B, C, D, and E) we select reading lists of 3 (demonstrated in the following example), we use one of the most popular counting devices in mathematics called the combination.