4 Probability

4.1 Probability of Events

Sweet Statistics

Sweet Statistics

Does this “sweet” picture suddenly make you hungry for some candy?

It is pretty hard to resist an M&M. Bet you have a favorite color. Let's see if your favorite color may change—depending on how hungry you are!

Suppose a large bag of M&M’s is opened and the resulting distribution of color counts is as shown in Table 4.1

Color	Count
Brown	91
Yellow	112
Red	102
Blue	151
Orange	137
Green	99
	692

M&M Colors by Count
Table 4.1

Sweet Statistics

If you were told that you could have all the M&M's of **one color** from this bag, which color would you choose? Remember you are very hungry!

Color	Percent
Brown	13.2
Yellow	16.2
Red	14.7
Blue	21.8
Orange	19.8
Green	14.3
	100.0

M&M Colors by Percentage

Table 4.2

Looks like "blue" is the way to go! It has the largest count for this bag of 692 M&M's. But, how does it compare to the rest of the colors?

Sweet Statistics

A convenient way to make a comparison is to use percentages. If we divide 151/692, we get 0.218 ≈ or 22%.

Thus, 22% of the M&M's in this bag are "blue."

Another way to consider this event is that if you were to select one M&M from a thoroughly mixed container without looking, there is a 22% chance of picking a blue M&M.

Sweet Statistics

You have just completed your first probability experiment! (Granted, actually doing the experiment, and eating the M&M’s, would have been more fun!)

We are now ready to define what is meant by probability. Specifically, we talk about “the probability that a certain event will occur.”

Probability of an event The relative frequency with which that event can be expected to occur.

Sweet Statistics

The probability of an event may be obtained in three different ways:

(1) *empirically*,

(2) *theoretically*, or

(3) *subjectively*.

The **empirical** method was just illustrated by the M&M's and their percentages and might be called **experimental** or **empirical probability**.

Sweet Statistics

This probability is the **observed relative frequency** with which an event occurs.

In the M&M example, we observed that 137 of the 692 M&M's were orange.

The observed empirical probability for the occurrence of orange was 137/692, or 0.198.

Sweet Statistics

The value assigned to the probability of event *A* as a result of experimentation can be found by means of the formula:

Empirical (Observed) Probability $P'(A)$

In words: $\text{empirical probability of A} = \dfrac{\text{number of times A occurred}}{\text{number of trials}}$

In algebra: $$P'(A) = \frac{n(A)}{n} \qquad \textbf{(4.1)}$$

Sweet Statistics

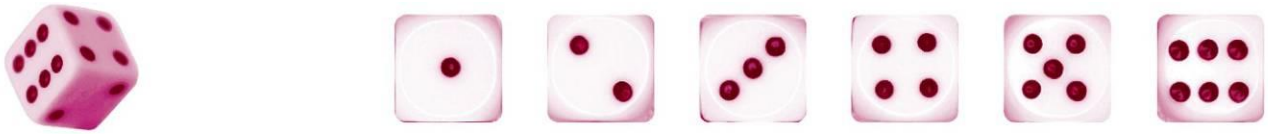
Notation for empirical probability: When the value assigned to the probability of an event results from experimental or empirical data, we will identify the probability of the event with the symbol P′( ).

The **theoretical** method for obtaining the probability of an event uses a *sample space*.

A **sample space** is a listing of all possible outcomes from the experiment being considered (denoted by the capital letter S).

Sweet Statistics

When this method is used, the sample space must contain **equally likely** sample points. For example, the sample space for the rolling of one die is S = {1, 2, 3, 4, 5, 6}.



The six possible outcomes from one roll

Each **outcome** (i.e., number) is equally likely.

An **event** is a subset of the sample space (denoted by a capital letter other than *S*; *A* is commonly used for the first event).

Sweet Statistics

Therefore, the *probability of an event A*, $P(A)$, is the ratio of the number of points that satisfy the definition of event A, $n(A)$, to the number of **sample points** in the entire sample space, $n(S)$.

Theoretical (Expected) Probability $P(A)$

In words:

$$\textit{theoretical probability of A} = \frac{\textit{number of times A occurs in sample space}}{\textit{number of elements in sample space}}$$

In algebra: $P(A) = \dfrac{n(A)}{n(S)}$, when the elements of S are equally likely **(4.2)**

Sweet Statistics

Notes

1. When the value assigned to the probability of an event results from a theoretical source, we will identify the probability of the event with the symbol *P*( ).

2. The prime symbol is *not used* with theoretical probabilities; it is used only for empirical probabilities.

When a probability experiment can be thought of as a sequence of events, a **tree diagram** often is a very helpful way to picture the sample space.

Example 4 – *Using Tree Diagrams*

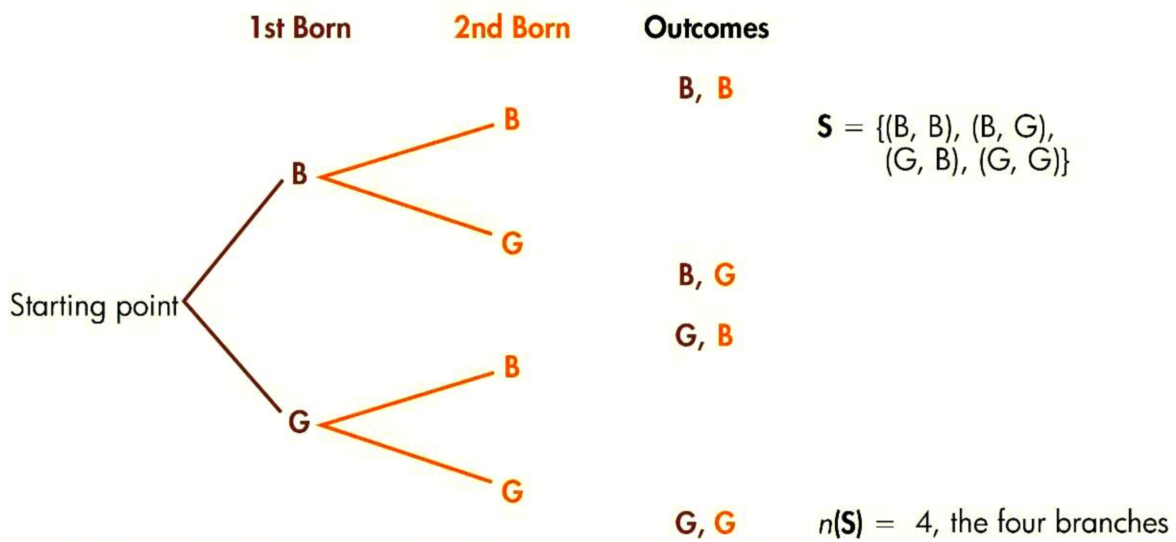
A family with two children is to be selected at random, and we want to find the probability that the family selected has one child of each gender.

Because there will always be a firstborn and a second-born child, we will use a tree diagram to show the possible arrangements of gender, thus making it possible for us to determine the probability.

Start by determining the sequence of events involved—firstborn and second-born in this case.

Example 4 – *Using Tree Diagrams* cont'd

Use the tree to show the possible outcomes of the first event (shown in brown on Figure 4.1) and then add branch segments to show the possible outcomes for the second event (shown in orange in Figure 4.1).



Tree Diagram Representation of Family with Two Children

Figure 4.1

Example 4 – *Using Tree Diagrams* cont'd

Notes

1. The two branch segments representing B and G for the second-born child must be drawn from each outcome for the firstborn child, thus creating the "tree" appearance.

2. There are four branches; each branch starts at the "tree root" and continues to an "end" (made up of two branch segments each), showing a possible outcome.

Because the branch segments are equally likely, assuming equal likeliness of gender, the four branches are then equally likely.

Example 4 – *Using Tree Diagrams* cont'd

This means we need only the count of branches to use formula (4.2) to find the probability of the family having one child of each gender.

The two middle branches, (B, G) and (G, B), represent the event of interest, so $n(A) = n(\text{one of each}) = 2$, whereas $n(S) = 4$ because there are a total of four branches.

Thus,

$$P\,(\text{one of each gender in family of two children}) = \frac{2}{4}$$

$$= \frac{1}{2}$$

$$= 0.5$$

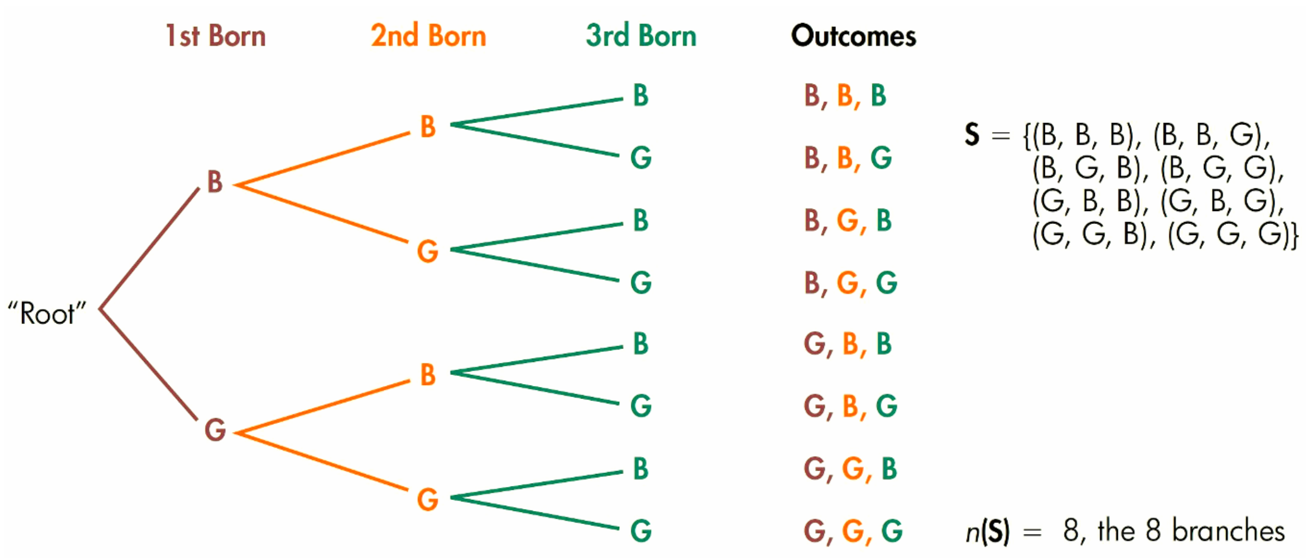
Example 4 – *Using Tree Diagrams* cont’d

Now let’s consider selecting a family of three children and finding the probability of “at least one boy” in that family. Again the family can be thought of as a sequence of three events—firstborn, second-born, and third born.

To create a tree diagram of this family, we need to add a third set of branch segments to our two-child family tree diagram.

Example 4 – *Using Tree Diagrams* cont'd

The green branch segments represent the third child (see Figure 4.2).



Tree Diagram Representation of Family with Three Children

Figure 4.2

Example 4 – *Using Tree Diagrams* cont'd

Again, because the branch segments are equally likely, assuming equal likeliness of gender, the eight branches are then equally likely.

This means we need only the count of branches to use formula (4.2) to find the probability of the family having at least one boy.

The top seven branches all have one or more boys, the equivalent of "at least one."

Example 4 – *Using Tree Diagrams* cont'd

P(at least one boy in a family of three children) = $\frac{7}{8}$

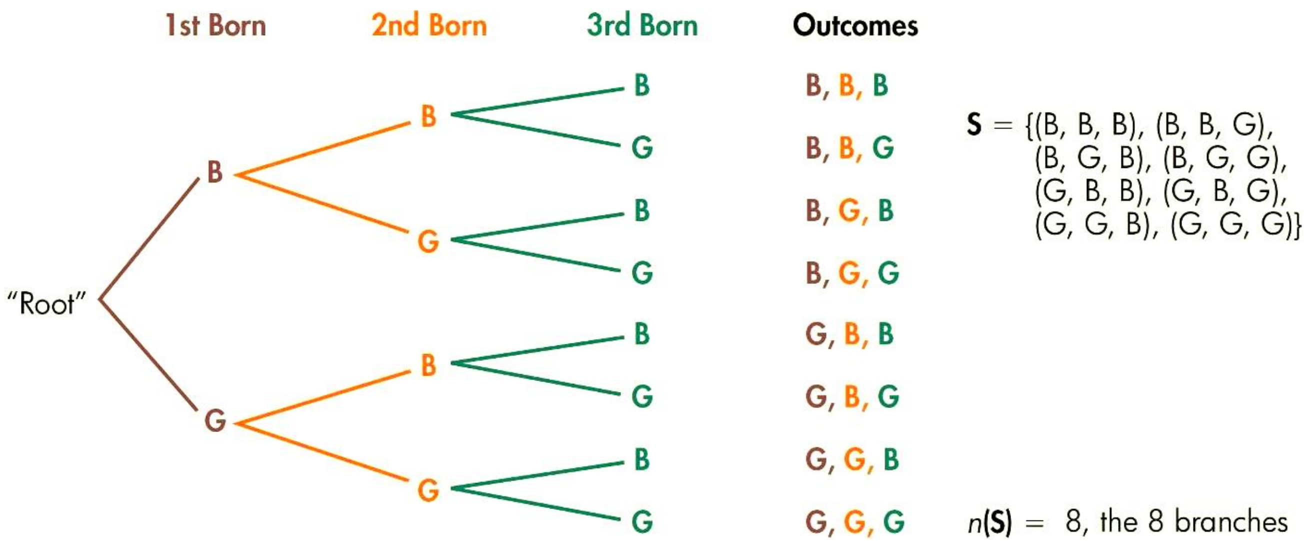
= 0.875

Let's consider one other question before we leave this example. What is the probability that the third child in this family of three children is a girl?

The question is actually an easy one; the answer is 0.5, because we have assumed equal likelihood of either gender.

Example 4 – *Using Tree Diagrams* cont'd

However, if we look at the tree diagram in Figure 4.2, there are two ways to view the answer.



Tree Diagram Representation of Family with Three Children

Figure 4.2

Example 4 – *Using Tree Diagrams* cont'd

First, if you look at only the branch segments for the third-born child, you see one of two is for a girl in each set, thus $\frac{1}{2}$, or 0.5.

Also, if you look at the entire tree diagram, the last child is a girl on four of the eight branches; thus $\frac{4}{8}$, or 0.5.

Sweet Statistics

When a probability question provides information about the events in the form of the probability of the various events, the number of items per set, or the percentage of each set, a **Venn diagram** often is a very helpful way to display the sample space or the information.

Venn diagrams can be used to find both theoretical and empirical probabilities.

Sweet Statistics

Special attention should always be given to the sample space. Like the statistical population, the sample space must be well defined. Once the sample space is defined, you will find the remaining work much easier.

A **subjective probability** generally results from personal judgment. Your local weather forecaster often assigns a probability to the event “precipitation.”

For example, “There is a 20% chance of rain today,” or “There is a 70% chance of snow tomorrow.”

Sweet Statistics

In such cases, the only method available for assigning probabilities is personal judgment. These probability assignments are called *subjective probabilities*.

The accuracy of subjective probabilities depends on an individual’s ability to correctly assess the situation.

Properties of Probability Numbers

Properties of Probability Numbers

Whether the probability is *empirical*, *theoretical*, or *subjective*, the following properties must hold.

Property 1

In words: "A probability is always a numerical value between zero and one."

In algebra: $0 \leq \text{each } P(A) \leq 1$ or $0 \leq \text{each } P'(A) \leq 1$

Notes about Property 1:

1. The probability is 0 if the event cannot occur.

2. The probability is 1 if the event occurs every time.

Properties of Probability Numbers

3. Otherwise, the probability is a fractional number between 0 and 1.

Property 2

In words: "The sum of the probabilities for all outcomes of an experiment is equal to exactly one."

In algebra: $\sum_{\text{all outcomes}} P(A) = 1$ or $\sum_{\text{all outcomes}} P'(A) = 1$

Note about Property 2: The list of "all outcomes" must be a nonoverlapping set of events that includes all the possibilities (**all-inclusive**).

Properties of Probability Numbers

Notes about probability numbers:

1. Probability represents a relative frequency, whether from a sample space or a sample.

2. $P(A)$ is the ratio of the number of times an event can be expected to occur divided by the number of possibilities. $P'(A)$ is the ratio of the number of times an event did occur divided by the number of data.

3. The numerator of the probability ratio must be a positive number or zero.

Properties of Probability Numbers

4. The denominator of the probability ratio must be a positive number (greater than zero).

5. As a result of Notes 1 through 4 above, the probability of an event, whether it be empirical, theoretical, or subjective, will always be a numerical value between zero and one, inclusive.

6. The rules for probability are the same for all three types of probability: empirical, theoretical, and subjective.

How Are Empirical and Theoretical Probabilities Related?

How Are Empirical and Theoretical Probabilities Related?

Consider the rolling of one die and define event A as the occurrence of a “1.”

An ordinary die has six equally likely sides, so the theoretical probability of event A is $P(A) = \frac{1}{6}$

What does this mean?

Do you expect to see one “1” in each trial of six rolls? Explain. If not, what results do you expect?

How Are Empirical and Theoretical Probabilities Related?

If we were to roll the die several times and keep track of the proportion of the time event A occurs, we would observe an empirical probability for event A.

What value would you expect to observe for $P'(A)$? Explain.

How are the two probabilities $P(A)$ and $P'(A)$ related? Explain.

To gain some insight into this relationship, let's perform an experiment.

Example 6 – *Demonstration-Law of Large Numbers*

The experiment will consist of 20 trials. Each trial of the experiment will consist of rolling a die six times and recording the number of times the “1” occurs. Perform 20 trials.

Each row of Table 4.3 shows the results of one trial; we conduct 20 trials, so there are 20 rows.

Trial	Column 1: Number of 1s Observed	Column 2: Relative Frequency	Column 3: Cumulative Relative Frequency	Trial	Column 1: Number of 1s Observed	Column 2: Relative Frequency	Column 3: Cumulative Relative Frequency
1	1	1/6	1/6 = 0.17	11	1	1/6	10/66 = 0.15
2	2	2/6	3/12 = 0.25	12	0	0/6	10/72 = 0.14
3	0	0/6	3/18 = 0.17	13	2	2/6	12/78 = 0.15
4	1	1/6	4/24 = 0.17	14	1	1/6	13/84 = 0.15
5	0	0/6	4/30 = 0.13	15	1	1/6	14/90 = 0.16
6	1	1/6	5/36 = 0.14	16	3	3/6	17/96 = 0.18
7	2	2/6	7/42 = 0.17	17	0	0/6	17/102 = 0.17
8	2	2/6	9/48 = 0.19	18	1	1/6	18/108 = 0.17
9	0	0/6	9/54 = 0.17	19	0	0/6	18/114 = 0.16
10	0	0/6	9/60 = 0.15	20	1	1/6	19/120 = 0.16

Experimental Results of Rolling a Die Six Times in Each Trial

Table 4.3

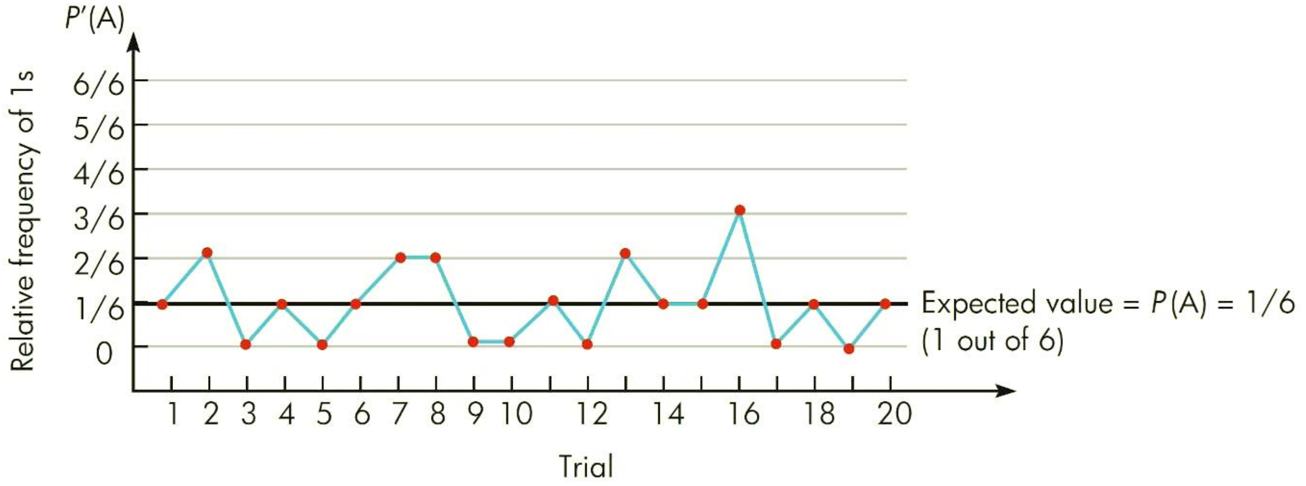
Example 6 – *Demonstration-Law of Large Numbers* cont'd

Column 1 lists the number of 1s observed in each trial (set of six rolls); column 2 lists the observed relative frequency for each trial; and column 3 lists the cumulative relative frequency as each trial was completed.

Example 6 – *Demonstration-Law of Large Numbers*

cont'd

Figure 4.4a shows the fluctuation (above and below) of the observed probability, $P'(A)$ (Table 4.3, column 2), about the theoretical probability, $P(A) = \frac{1}{6}$,

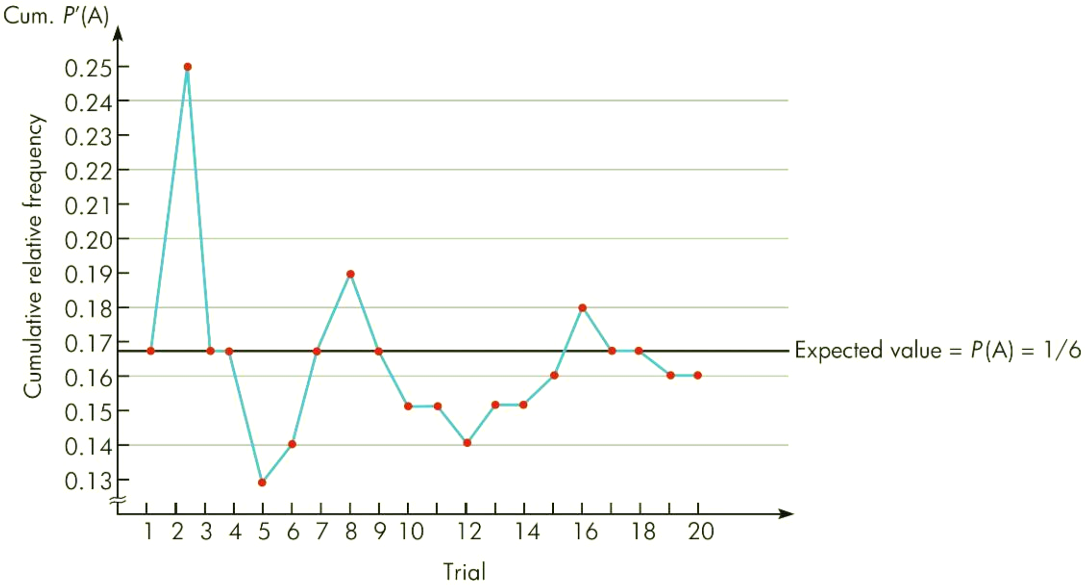


Fluctuations Found in the Die-Tossing Experiment - (a) Relative Frequency

Figure 4.4

Example 6 – *Demonstration-Law of Large Numbers* cont'd

whereas Figure 4.4b shows the fluctuation of the cumulative relative frequency (Table 4.3, column 3) and how it becomes more stable.



Fluctuations Found in the Die-Tossing Experiment - (b) Cumulative Relative Frequency

Figure 4.4

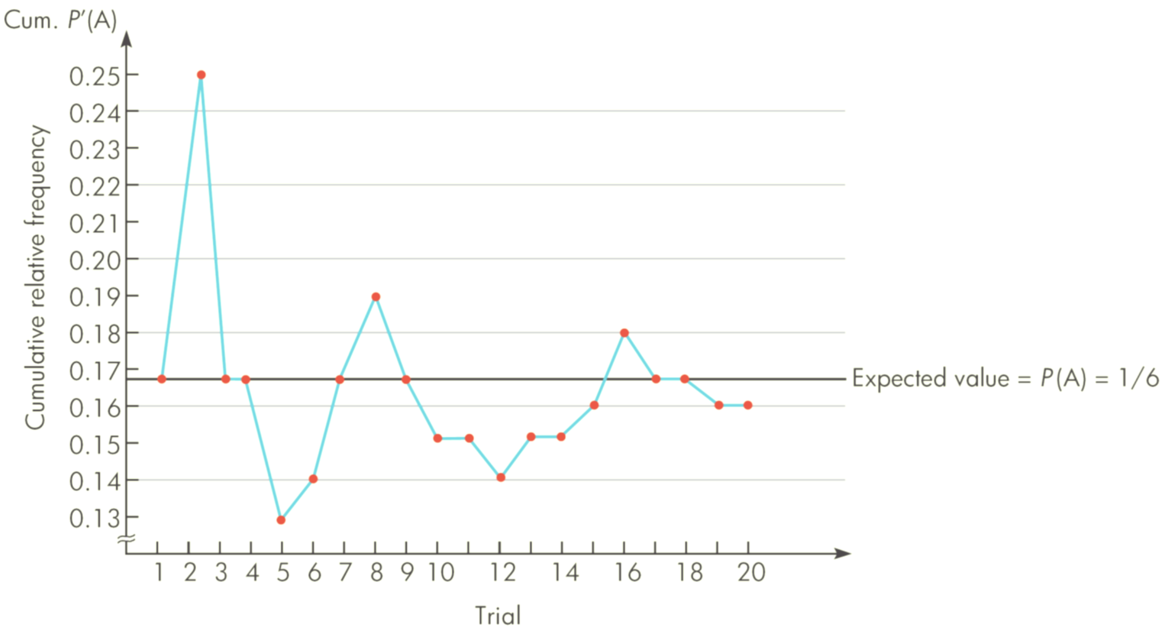
Example 6 – *Demonstration-Law of Large Numbers*

cont'd

In fact, the cumulative relative frequency becomes relatively close to the theoretical or expected probability $\frac{1}{6}$, or $0.166\overline{6}$ = 0.167.

How Are Empirical and Theoretical Probabilities Related?

A cumulative graph such as that shown in Figure 4.4b demonstrates the idea of a **long-term average** and is often referred to as the *law of large numbers*.



(b) Cumulative Relative Frequency

Figure 4.4

How Are Empirical and Theoretical Probabilities Related?

Law of large numbers As the number of times an experiment is repeated increases, the ratio of the number of successful occurrences to the number of trials will tend to approach the theoretical probability of the outcome for an individual trial.

The law of large numbers is telling us that the larger the number of experimental trials, n, the closer the empirical probability, $P'(A)$. This concept has many applications.

How Are Empirical and Theoretical Probabilities Related?

The preceding die-tossing experiment is an example in which we can easily compare actual results against what we expected to happen; it gave us a chance to verify the claim of the law of large numbers.

Probabilities as Odds

Probabilities as Odds

Probabilities can be and are expressed in many ways; we see and hear many of them in the news nearly every day (most of the time, they are subjective probabilities).

Odds are a way of expressing probabilities by expressing the number of ways an event can happen compared to the number of ways it can't happen.

The statement "It is four times more likely to rain tomorrow (R) than not rain (NR)" is a probability statement that can be expressed as odds: "The odds are 4 to 1 in favor of rain tomorrow" (also written 4: 1).

Probabilities as Odds

The relationship between odds and probability is shown here.

If the odds in favor of an event A are ***a* to *b*** (or ***a*:*b***), then

1. The odds against event A are ***b* to *a*** (or ***b*:*a***).

2. The probability of event A is $\boldsymbol{P(A)} = \dfrac{a}{a+b}$.

3. The probability that event A will not occur is

$$\boldsymbol{P(\text{not } A)} = \frac{b}{a+b}.$$

Probabilities as Odds

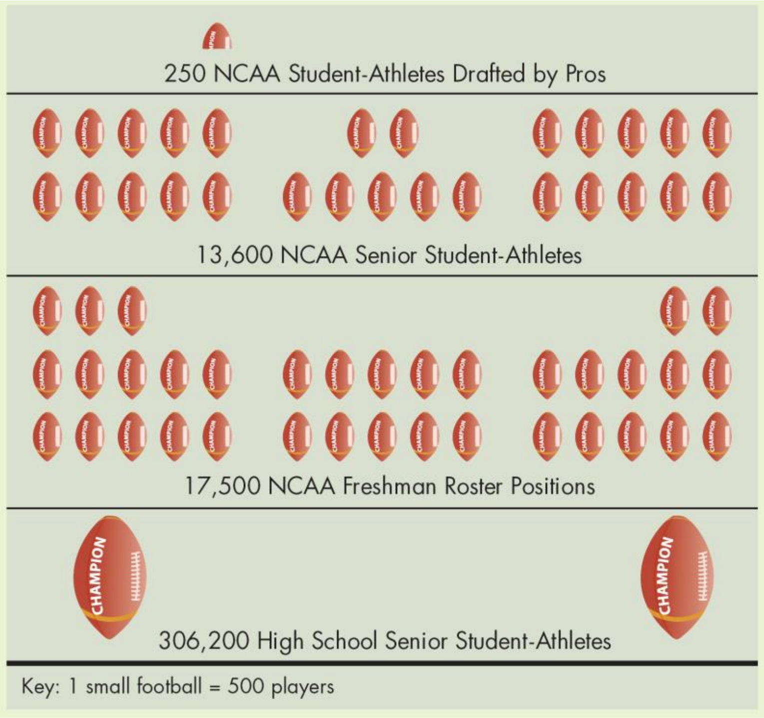
To illustrate this relationship, consider the statement "The odds favoring rain tomorrow are 4 to 1." Using the preceding notation, *a* = 4 and *b* = 1.

Therefore, the probability of rain tomorrow is $\frac{4}{4 + 1}$, or $\frac{4}{5}$ = 0.8.

The odds against rain tomorrow are 1 to 4 (or 1:4), and the probability that there will be no rain tomorrow is $\frac{1}{4 + 1}$, or $\frac{1}{5}$ = 0.2.

Applied Example 8 – *Making it to the Next Level*

Many young men aspire to become professional athletes. Only a few make it to the big time, as indicated in the following graph.



Applied Example 8 – *Making it to the Next Level* cont'd

For every 13,600 college senior football players, only 250 are drafted by a professional team; that translates to a probability of only 0.018 (250/13,600).

Student-Athletes	**Football**
High School Senior Student-Athletes	306,200
NCAA Freshman Roster Positions	17,500
NCAA Senior Student-Athletes	13,600
NCAA Student-Athletes Drafted	250

Source: http://www.ncaa.org/

Applied Example 8 – *Making it to the Next Level* cont'd

There are many other interesting specifics hidden in this information. For example, many high school boys dream of being a pro football player, but according to these numbers, the probability of a high school senior even being drafted by the pros is only 0.000816 (250/306,200).

Once a player has made a college football team, he might be very interested in the odds that he will play as a senior.

Of the 17,500 players making a college freshman team, 13,600 play as seniors, while 3,900 do not.

Applied Example 8 – *Making it to the Next Level* cont’d

Thus, if a player has made a college team, the odds he will play as a senior are 13,600 to 3,900, which reduces to 136 to 39. The college senior who is playing is interested in his chances of making the next level.

We see that of the 13,600 college seniors, only 250 are drafted by the pros, while 13,350 are not; thus the odds against him making the next level are 13,350 to 250, which reduces to 267 to 5.

Odds are strongly against him being drafted, and the odds against him making the team are somewhat stronger.

Comparison of Probability and Statistics

Comparison of Probability and Statistics

Probability and **statistics** are two separate but related fields of mathematics. It has been said that “probability is the vehicle of statistics.”

That is, if it were not for the laws of probability, the theory of statistics would not be possible.

Let’s illustrate the relationship and the difference between these two branches of mathematics by looking at two boxes.

Comparison of Probability and Statistics

We know that the probability box contains five blue, five red, and five white poker chips.

Probability tries to answer questions such as, “If one chip is randomly drawn from this box, what is the chance that it will be blue?”

On the other hand, in the statistics box we don’t know what the combination of chips is.

We draw a sample and, based on the findings in the sample, make conjectures about what we believe to be in the box.

Comparison of Probability and Statistics

Note the difference: Probability asks you about the chance that something specific, like drawing a blue chip, will happen when you know the possibilities (that is, you know the population).

Statistics, on the other hand, asks you to draw a sample, describe the sample (descriptive statistics), and then make inferences about the population based on the information found in the sample (inferential statistics).