3

Probability

3.2 Basic Terms of Probability

Objectives

- Learn the basic terminology of probability theory
- Be able to calculate simple probabilities
- Understand how probabilities are used in genetics

Basic Terms of Probability

Basic Probability Terms

experiment: a process by which an observation, or **outcome**, is obtained
sample space: the set S of all possible outcomes of an experiment
event: any subset E of the sample space S

(The term *sample space* really means the same thing as *universal set;* the only distinction between the two ideas is that *sample space* is used only in probability theory, whereas *universal set* is used in any situation in which sets are used.)

Basic Terms of Probability

If you roll a single die, then:

•

Basic Terms of Probability

If you roll a single die, then:

- the *experiment* is the rolling of the die

Basic Terms of Probability

If you roll a single die, then:

- the *experiment* is the rolling of the die
- the possible *outcomes* are 1, 2, 3, 4, 5, and 6

Basic Terms of Probability

If you roll a single die, then:

- the *experiment* is the rolling of the die
- the possible *outcomes* are 1, 2, 3, 4, 5, and 6
- the *sample space* $S = \{1, 2, 3, 4, 5, 6\}$

Basic Terms of Probability

If you roll a single die, then:

- the *experiment* is the rolling of the die
- the possible *outcomes* are 1, 2, 3, 4, 5, and 6
- the *sample space* $S = \{1, 2, 3, 4, 5, 6\}$
- there are several possible *events*, subsets of the sample space, including

$E_1 = \{3\}$ "you roll a 3"

$E_2 = \{2, 4, 6\}$ "you roll an even number"

$E_3 = \{1, 2, 3, 4, 5, 6\} = S$ "you roll a number between 1 and 6, inclusive"

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
“Rolling an even number”

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
"Rolling an even number" is the event $E_2 = \{2, 4, 6\}$, not an outcome;
it's a set of three separate outcomes.

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
"Rolling an even number" is the event $E_2 = \{2, 4, 6\}$, not an outcome;
it's a set of three separate outcomes.

Some events are distinguished from outcomes only in that set brackets are used with events and not outcomes.

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
"Rolling an even number" is the event $E_2 = \{2, 4, 6\}$, not an outcome;
it's a set of three separate outcomes.

Some events are distinguished from outcomes only in that set brackets are used with events and not outcomes.

Example: $E_1 = \{3\}$

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
"Rolling an even number" is the event $E_2 = \{2, 4, 6\}$, not an outcome;
it's a set of three separate outcomes.

Some events are distinguished from outcomes only in that set brackets are used with events and not outcomes.

Example: $E_1 = \{3\}$ is an event

Basic Terms of Probability

Note: an *event* is not the same as an *outcome*.

An *event* is a subset of the sample space;
An *outcome* is an element of the sample space.

Example:
"Rolling an even number" is the event $E_2 = \{2, 4, 6\}$, not an outcome;
it's a set of three separate outcomes.

Some events are distinguished from outcomes only in that set brackets are used with events and not outcomes.

Example: $E_1 = \{3\}$ is an event, and 3 is an outcome.
Each refers to "you roll a 3."

Basic Terms of Probability

More Probability Terms

A **certain event** is an event that is certain to happen. It is an event that's equal to the sample space S.

An **impossible event** is an event that cannot happen. It is an event that is equal to the empty set.

Basic Terms of Probability

Examples:

The event $E_3 = \{1, 2, 3, 4, 5, 6\} = S$ is a ***certain event***.

Basic Terms of Probability

Examples:

The event $E_3 = \{1, 2, 3, 4, 5, 6\} = S$ is a ***certain event***.

It's certain that one of the numbers from 1 to 6 will come up, and it's an event that's equal to the sample space.

Basic Terms of Probability

Examples:

The event $E_3 = \{1, 2, 3, 4, 5, 6\} = S$ is a ***certain event***.

It's certain that one of the numbers from 1 to 6 will come up, and it's an event that's equal to the sample space.

"A 17 comes up" is an ***impossible event***.

Basic Terms of Probability

Examples:

The event $E_3 = \{1, 2, 3, 4, 5, 6\} = S$ is a ***certain event***.

It's certain that one of the numbers from 1 to 6 will come up, and it's an event that's equal to the sample space.

"A 17 comes up" is an ***impossible event***.

17 is not a possible outcome.

Basic Terms of Probability

Examples:

The event $E_3 = \{1, 2, 3, 4, 5, 6\} = S$ is a ***certain event***.

It's certain that one of the numbers from 1 to 6 will come up, and it's an event that's equal to the sample space.

"A 17 comes up" is an ***impossible event***.

17 is not a possible outcome.

No outcome in the sample space S would result in a 17, so this event is the empty set.

Probabilities and Odds

Probabilities and Odds

The ***probability*** of an event is a measure of the likelihood that the event will occur.

Probabilities and Odds

The ***probability*** of an event is a measure of the likelihood that the event will occur.

If a single die is rolled, the outcomes are *equally likely*; a 3 is just as likely to come up as any other number.

Probabilities and Odds

The ***probability*** of an event is a measure of the likelihood that the event will occur.

If a single die is rolled, the outcomes are *equally likely*; a 3 is just as likely to come up as any other number.

There are 6 possible outcomes, so a 3 should come up about one out of every six rolls.

Probabilities and Odds

The ***probability*** of an event is a measure of the likelihood that the event will occur.

If a single die is rolled, the outcomes are *equally likely*; a 3 is just as likely to come up as any other number.

There are 6 possible outcomes, so a 3 should come up about one out of every six rolls.

The probability of event E_1 ("a 3 comes up") is $\frac{1}{6}$.

number of elements in $E_1 = \{3\}$

number of elements in $S = \{1, 2, 3, 4, 5, 6\}$

Probabilities and Odds

If an experiment's outcomes are equally likely, then the probability of an event E is the number of outcomes in the event divided by the number of outcomes in the sample space, or $n(E) / n(S)$.

Probability can be thought of as "success over a total."

Probability of an Event

The **probability** of an event E, denoted by $p(E)$, is

$$p(E) = \frac{n(E)}{n(S)}$$

if the experiment's outcomes are equally likely.
(*Think: Success over total.*)

Probabilities and Odds

Many people use the words *probability* and *odds* interchangeably. However, the words have different meanings.

The *odds* in favor of an event are

the number of ways the event can occur compared to the number of ways the event *can fail to occur,* or "success compared to *failure*"
(if the experiment's outcomes are equally likely).

Probabilities and Odds

The odds of event E_1 (“a 3 comes up”) are 1 to 5 (or 1:5), since a three can come up in one way and can fail to come up in five ways.

Probabilities and Odds

The odds of event E_1 (“a 3 comes up”) are 1 to 5 (or 1:5), since a three can come up in one way and can fail to come up in five ways.

Similarly, the odds of event E_3 (“a number between 1 and 6 inclusive comes up”) are 6 to 0 (or 6:0), since a number between 1 and 6 inclusive can come up in six ways and can fail to come up in zero ways.

Probabilities and Odds

Odds of an Event

The **odds** of an event E with equally likely outcomes, denoted by $o(E)$, are given by

$$o(E) = n(E) : n(E')$$

(Think: Success compared with failure.)

Probabilities and Odds

Odds of an Event

The **odds** of an event E with equally likely outcomes, denoted by $o(E)$, are given by

$$o(E) = n(E) : n(E')$$

(Think: Success compared with failure.)

In addition to the above meaning, the word *odds* can also refer to "house odds," which has to do with how much you will be paid if you win a bet at a casino. The odds of an event are sometimes called the true odds to distinguish them from the house odds.

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

b. the probability of event E_1, “getting heads”

c. the odds of event E_1, “getting heads”

d. the probability of event E_2, “getting heads or tails”

e. the odds of event E_2, “getting heads or tails”

f. an impossible event, and its probability

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

b. the probability of event E_1, "getting heads"

c. the odds of event E_1, "getting heads"

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space S = {H, T}

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

c. the odds of event E_1, "getting heads"

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)}$$

c. the odds of event E_1, "getting heads"

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

c. the odds of event E_1, "getting heads"

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$E'_1 = \{T\}$

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$E'_1 = \{T\}$

$o(E_1) = n(E_1) : n(E_1')$

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$E'_1 = \{T\}$

$o(E_1) = n(E_1) : n(E_1') = 1:1$

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H, T\}$

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$$p(E_2) = \frac{n(E_2)}{n(S)}$$

$E_2 = \{H,T\}$

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H, T\}$

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2}$$

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1') = 1:1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

$E'_2 = \{\} = \varnothing$

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1{:}1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

$E'_2 = \{\} = \varnothing$

$o(E_2) = n(E_2) : n(E_2\,')$

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1:1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

$E'_2 = \{\} = \varnothing$

$o(E_2) = n(E_2) : n(E_2\,') = 2:0$

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1:1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

$E'_2 = \{\} = \varnothing$

$o(E_2) = n(E_2) : n(E_2\,') = 2:0 = 1:0$

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Flipping a Coin*

A coin is flipped. Find the following.

a. the sample space

$E_1 = \{H\}$

b. the probability of event E_1, "getting heads"

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{1}{2}$$

one out of every two possible outcomes is a success

c. the odds of event E_1, "getting heads"

$o(E_1) = n(E_1) : n(E_1\,') = 1:1$

for every one possible success, there is one possible failure

d. the probability of event E_2, "getting heads or tails"

$E_2 = \{H,T\}$ E_2 is a certain event

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{2}{2} = \frac{1}{1}$$

e. the odds of event E_2, "getting heads or tails"

$E'_2 = \{\} = \varnothing$ there are no possible failures $o(E_2) = n(E_2) : n(E_2\,') = 2:0 = 1:0$

f. an impossible event, and its probability

sample space $S = \{H, T\}$

Example 1 – *Solution*

cont'd

f. *Finding an impossible event:*

"Getting a 5" is an impossible event, because we're flipping a coin.

The only possible outcomes are H and T.

Example 1 – *Solution*

cont'd

f. *Finding an impossible event:*

"Getting a 5" is an impossible event,

because we're flipping a coin.

The only possible outcomes are H and T.

"Getting a 5" is not a possible outcome, so $5 \notin S$.

This event is $\varnothing = \{\ \}$.

Example 1 – *Solution*

cont'd

f. *Finding an impossible event:*

"Getting a 5" is an impossible event,

because we're flipping a coin.

The only possible outcomes are H and T.

"Getting a 5" is not a possible outcome, so $5 \notin S$.

This event is $\varnothing = \{\ \}$.

The probability of this event is $p(\varnothing) = \dfrac{n(\varnothing)}{n(S)} = \dfrac{0}{2} = 0$

Relative Frequency versus Probability

Relative Frequency versus Probability

When we found that the probability of heads was $\frac{1}{2}$, we never actually tossed a coin.

It does not always make sense to calculate probabilities theoretically; sometimes they must be found empirically, the way a batting average is calculated.

Relative Frequency versus Probability

When we found that the probability of heads was $\frac{1}{2}$, we never actually tossed a coin.

It does not always make sense to calculate probabilities theoretically; sometimes they must be found empirically, the way a batting average is calculated.

For example, in 8,389 times at bat, Babe Ruth had 2,875 hits. His batting average was $\frac{2{,}875}{8{,}389} \approx 0.343.$

In other words, his probability of getting a hit was 0.343.

Relative Frequency versus Probability

Sometimes a probability can be found either *theoretically* or *empirically*. We have already found that the theoretical probability of heads is $\frac{1}{2}$.

Relative Frequency versus Probability

Sometimes a probability can be found either *theoretically* or *empirically*. We have already found that the theoretical probability of heads is $\frac{1}{2}$.

We could also flip a coin eight times, have heads come up one time (this is the frequency of heads), and calculate

$$\frac{\text{number of heads}}{\text{total number of flips}} = \frac{1}{8}$$

Relative Frequency versus Probability

Sometimes a probability can be found either *theoretically* or *empirically*. We have already found that the theoretical probability of heads is $\frac{1}{2}$.

We could also flip a coin eight times, have heads come up one time (this is the frequency of heads), and calculate

$$\frac{\text{number of heads}}{\text{total number of flips}} = \frac{1}{8}$$

This is called the ***relative frequency*** of heads.

Relative Frequency versus Probability

Sometimes a probability can be found either *theoretically* or *empirically*. We have already found that the theoretical probability of heads is $\frac{1}{2}$.

We could also flip a coin eight times, have heads come up one time (this is the frequency of heads), and calculate

$$\frac{\text{number of heads}}{\text{total number of flips}} = \frac{1}{8}$$

This is called the ***relative frequency*** of heads.

The probability of heads is $\frac{1}{2}$, but the relative frequency of heads, in our particular experiment, is $\frac{1}{8}$.

Relative Frequency versus Probability

Relative Frequency

If you repeat an experiment a number of times and count how frequently a particular outcome occurs, then the **relative frequency** of that outcome is

$$\text{relative frequency} = \frac{\text{number of times the outcome occurs}}{\text{total number of repetitions}}$$

Relative Frequency versus Probability

Relative Frequency

If you repeat an experiment a number of times and count how frequently a particular outcome occurs, then the **relative frequency** of that outcome is

$$\text{relative frequency} = \frac{\text{number of times the outcome occurs}}{\text{total number of repetitions}}$$

If you repeat an experiment a small number of times, anything can happen, and a *relative frequency* may or may not equal the corresponding *probability*.

Relative Frequency versus Probability

Relative Frequency

If you repeat an experiment a number of times and count how frequently a particular outcome occurs, then the **relative frequency** of that outcome is

$$\text{relative frequency} = \frac{\text{number of times the outcome occurs}}{\text{total number of repetitions}}$$

If you repeat an experiment a small number of times, anything can happen, and a *relative frequency* may or may not equal the corresponding *probability*.

In the above discussion:

$$\left.\begin{array}{l}\text{relative frequency of heads} = 1/8 \\ \text{probability of heads} = 1/2\end{array}\right\}$$

these aren't equal or close

Relative Frequency versus Probability

However, if you repeat an experiment a large number of times, a relative frequency will *tend to be close* to the corresponding probability, even though anything can happen.

Relative Frequency versus Probability

However, if you repeat an experiment a large number of times, a relative frequency will *tend to be close* to the corresponding probability, even though anything can happen.

If we flip a coin a hundred times, we would probably find that the relative frequency of heads was close to $\frac{1}{2}$. Perhaps the frequency would be 47.

relative frequency of heads = $47/100 = 0.47$
probability of heads = $1/2 = 0.5$

these are close

Relative Frequency versus Probability

If you flipped a coin a thousand times, you would probably find that the relative frequency of heads to be even closer to $\frac{1}{2}$.

This relationship between probabilities and relative frequencies is called the **Law of Large Numbers.**

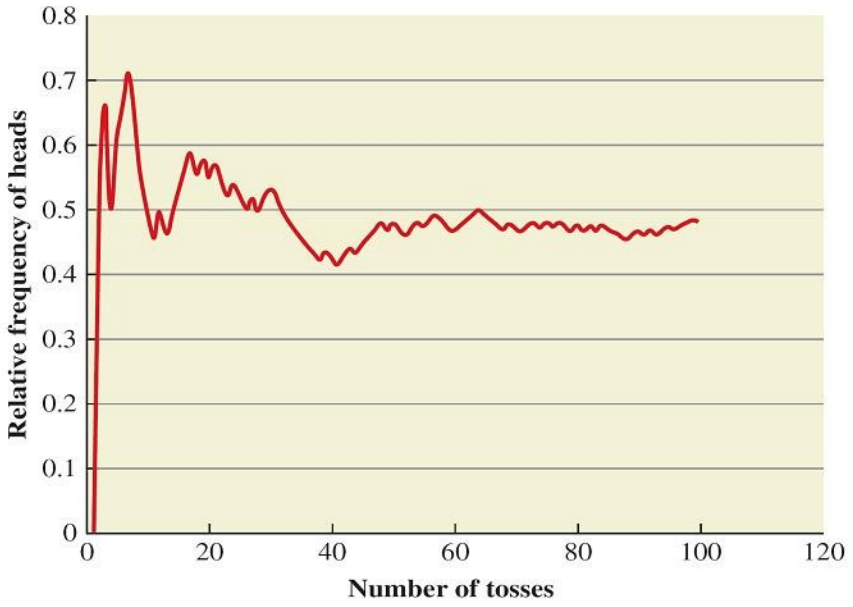
The Law of Large Numbers

Law of Large Numbers

If an experiment is repeated a large number of times, the relative frequency of an outcome will tend to be close to the probability of that outcome.

The Law of Large Numbers

The graph in Figure 3.2 shows the result of a simulated coin toss, using a computer and a random number generator rather than an actual coin.



The relative frequency of heads after 100 simulated coin tosses.

Figure 3.2

The Law of Large Numbers

What if we used a real coin, rather than a computer, and we tossed the coin a lot more?

The Law of Large Numbers

What if we used a real coin, rather than a computer, and we tossed the coin a lot more? Three different mathematicians have performed such an experiment:

- In the eighteenth century, Count Buffon tossed a coin 4,040 times. He obtained 2,048 heads, for a relative frequency of $2{,}048/4{,}040 \approx 0.5069$.

- During World War II, the South African mathematician John Kerrich tossed a coin 10,000 times while he was imprisoned in a German concentration camp. He obtained 5,067 heads, for a relative frequency of $5{,}067/10{,}000 = 0.5067$.

The Law of Large Numbers

- In the early twentieth century, the English mathematician Karl Pearson tossed a coin 24,000 times! He obtained 12,012 heads, for a relative frequency of 12,012/24,000 = 0.5005.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

b. Find the event *E* "getting exactly one heads."

c. Find the probability of event *E*.

d. Use the Law of Large Numbers to interpret the probability of event *E*.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

$S = \{(H, H), (H, T), (T, T), (T, H)\}$

b. Find the event E "getting exactly one heads."

c. Find the probability of event E.

d. Use the Law of Large Numbers to interpret the probability of event E.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

$S = \{(H, H), (H, T), (T, T), (T, H)\}$

b. Find the event E "getting exactly one heads."

$E = \{(H, T), (T, H)\}$

c. Find the probability of event E.

d. Use the Law of Large Numbers to interpret the probability of event E.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

S = {(H, H), (H, T), (T, T), (T, H)}

b. Find the event E "getting exactly one heads."

E = {(H, T), (T, H)}

c. Find the probability of event E. $P(E) = \frac{n(E)}{n(S)}$

d. Use the Law of Large Numbers to interpret the probability of event E.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

S = {(H, H), (H, T), (T, T), (T, H)}

b. Find the event E "getting exactly one heads."

E = {(H, T), (T, H)}

c. Find the probability of event E. $P(E) = \frac{n(E)}{n(S)} = \frac{2}{4}$

d. Use the Law of Large Numbers to interpret the probability of event E.

Example 3 – *Flipping a Pair of Coins*

A pair of coins is flipped.

a. Find the sample space.

S = {(H, H), (H, T), (T, T), (T, H)}

b. Find the event E "getting exactly one heads."

E = {(H, T), (T, H)}

c. Find the probability of event E. $P(E) = \frac{n(E)}{n(S)} = \frac{2}{4} = \frac{1}{2} = 50\%$

d. Use the Law of Large Numbers to interpret the probability of event E.

Example 3 – *Solution*

cont'd

d. According to the Law of Large Numbers, if an experiment is repeated a large number of times, the relative frequency of that outcome will tend to be close to the probability of the outcome.

Example 3 – *Solution*

cont'd

d. According to the Law of Large Numbers, if an experiment is repeated a large number of times, the relative frequency of that outcome will tend to be close to the probability of the outcome.

Here, that means that if we were to toss a pair of coins many times, we should expect to get exactly one heads about half (or 50%) of the time. Realize that this is only a prediction, and we might never get exactly one heads.

Mendel's Use of Probabilities

Mendel's Use of Probabilities

In his experiments with plants, Gregor Mendel pollinated peas until he produced pure-red plants (that is, plants that would produce only red-flowered offspring) and pure-white plants.

He then cross-fertilized these pure reds and pure whites and obtained offspring that had only red flowers. This amazed him, because the accepted theory of the day incorrectly predicted that these offspring would all have pink flowers.

Mendel's Use of Probabilities

He explained this result by postulating that there is a "determiner" that is responsible for flower color. These determiners are now called **genes**.

Each plant has two flower color genes, one from each parent.

Mendel reasoned that these offspring had to have inherited a red gene from one parent and a white gene from the other. These plants had one red flower gene and one white flower gene, but they had red flowers.
That is, the red flower gene is **dominant**, and the white flower gene is **recessive**.

Mendel's Use of Probabilities

Basic Genetic Terms

A **gene** is a biological determiner of some inherited quality in a living being. Genes come in pairs, with one gene coming from one parent and the other gene from the other parent.

A **dominant** gene is one that determines a particular inherited quality regardless of what gene it is paired with.

A **recessive** gene is one that determines a particular inherited quality only if it is paired with another gene that's exactly the same.

A pair of genes can be described by using ordered pairs, just as we did with a pair of coins in Example 3, where the first component refers to the first parent's contribution and the second component refers to the second parent's contribution.

Mendel's Use of Probabilities

By tradition, we use capital letters for dominant genes and lowercase letters for recessive genes.

If, in Mendel's plant experiment, we use R to stand for the dominant red gene and w to stand for the recessive white gene, then

Mendel's Use of Probabilities

By tradition, we use capital letters for dominant genes and lowercase letters for recessive genes.

If, in Mendel's plant experiment, we use R to stand for the dominant red gene and w to stand for the recessive white gene, then

•(R, R) refers to a pair of two dominant red genes—this pure-red plant can produce only red-flowered offspring because it can pass on only red genes

Mendel's Use of Probabilities

By tradition, we use capital letters for dominant genes and lowercase letters for recessive genes.

If, in Mendel's plant experiment, we use R to stand for the dominant red gene and w to stand for the recessive white gene, then

- (R, R) refers to a pair of two dominant red genes—this pure-red plant can produce only red-flowered offspring because it can pass on only red genes

- (w, w) refers to a pair of two recessive white genes—this pure-white plant can produce only white-flowered offspring because it can pass on only white genes

Mendel's Use of Probabilities

When Mendel cross-fertilized pure-red plants with pure-white plants, the offspring all had red flowers.

Mendel's Use of Probabilities

When Mendel cross-fertilized pure-red plants with pure-white plants, the offspring all had red flowers. This can be explained with the **punnett square in** Figure 3.3.

	R	R
w	(R, w)	(R, w)
w	(R, w)	(R, w)

← first parent's genes

← possible offspring

← possible offspring

↑ second parent's genes

A Punnett square for the first generation.

Figure 3.3

Mendel's Use of Probabilities

The offspring of this experiment are all (R, w).

R is a dominant gene, so these plants all had red flowers. Then these first-generation offspring were cross-fertilized, and Mendel found that approximately three-fourths of the resulting second-generaton offspring had red flowers and one-fourth had white flowers.

Mendel's Use of Probabilities

The offspring of this experiment are all (R, w).

R is a dominant gene, so these plants all had red flowers. Then these first-generation offspring were cross-fertilized, and Mendel found that approximately three-fourths of the resulting second-generaton offspring had red flowers and one-fourth had white flowers. This can be explained with the Punnett square in Figure 3.4.

	R	w
R	(R, R)	(w, R)
w	(R, w)	(w, w)

← first parent's genes

← possible offspring

← possible offspring

↑ second parent's genes

A Punnett square for the second generation.

Figure 3.4

Mendel's Use of Probabilities

We can see in the Punnett square that:

- the possible outcomes are (R, R), (R, w), (w, R), and (w, w)
- the sample space for these second-generation offspring is

$$S = \{(R, R), (R, w), (w, R), (w, w)\}$$

- the event "having red flowers" is $E_1 = \{(R, R), (R, w), (w, R)\}$, because R dominates w
- the event "having white flowers" is $E_2 = \{(w, w)\}$

Mendel's Use of Probabilities

The probability of having red flowers is

$$p(E_1) = \frac{n(E_1)}{n(S)} = \frac{3}{4}$$

The probability of having white flowers is

$$p(E_2) = \frac{n(E_2)}{n(S)} = \frac{1}{4}$$

The law of large numbers tells us that approximately three-fourths of the second generation offspring should have red flowers and approximately one-fourth should have white. This is in agreement with Mendel's observations.

Mendel's Use of Probabilities

Outcomes (R, w) and (w, R) are genetically identical; it does not matter which gene is inherited from which parent. For this reason, geneticists do not use the ordered pair notation but instead refer to each of these two outcomes as "Rw."

The only difficulty with this convention is that it makes the sample space appear to be S = {RR, Rw, ww}, which consists of only three elements, when in fact it consists of four elements.

Mendel's Use of Probabilities

This distinction is important; if the sample space consisted of three equally likely elements, then the probability of a red-flowered offspring would be $\frac{2}{3}$ rather than $\frac{3}{4}$.

Mendel knew that the sample space had to have four elements, because his cross fertilization experiments resulted in a relative frequency very close to $\frac{3}{4}$, not $\frac{2}{3}$.

Mendel's Use of Probabilities

Ronald Fisher, a noted British statistician, used statistics to deduce that Mendel fudged his data. Mendel's relative frequencies were unusually close to the theoretical probabilities, and Fisher found that there was only about a 0.00007 chance of such close agreement.

Others have suggested that perhaps Mendel did not willfully change his results but rather continued collecting data until the numbers were in close agreement with his expectations.

Mendel's Use of Probabilities

We stop here

The rest is for self-development!

Probabilities in Genetics: Inherited Diseases

Probabilities in Genetics: Inherited Diseases

Cystic Fibrosis:

Cystic fibrosis is an inherited disease characterized by abnormally functioning exocrine glands that secrete a thick mucus, clogging the pancreatic ducts and lung passages. Most patients with cystic fibrosis die of chronic lung disease; until recently, most died in early childhood.

This early death made it extremely unlikely that an afflicted person would ever parent a child. Only after the advent of Mendelian genetics did it become clear how a child could inherit the disease from two healthy parents.

In the 1960s, patients with cystic fibrosis died at around six months of age. Now, patients tend to die in their late 30s or
early 40s.

Probabilities in Genetics: Inherited Diseases

In 1989, a team of Canadian and American doctors announced the discovery of the gene that is responsible for most cases of cystic fibrosis.

As a result of that discovery, a new therapy for cystic fibrosis is being developed.

Researchers splice a therapeutic gene into a cold virus and administer it through an affected person's nose. When the virus infects the lungs, the gene becomes active.

It is hoped that this will result in normally functioning cells, without the damaging mucus.

Probabilities in Genetics: Inherited Diseases

Since the discovery of the responsible gene, there have been many studies of this new gene therapy, most with a small number of participants.

In March 2012, a British team of scientists announced a new study with 130 participants.

They use an inhaler to breathe in a copy of the “good” cystic fibrosis gene once a month for a year. The results of this study will be known in spring 2014.

Probabilities in Genetics: Inherited Diseases

Cystic fibrosis occurs in about 1 out of every 2,000 births in the Caucasian population and only in about 1 in 250,000 births in the non-Caucasian population. It is one of the most common inherited diseases in North America.

One in 25 Americans carries a single gene for cystic fibrosis. Children who inherit two such genes develop the disease; that is, cystic fibrosis is recessive.

Probabilities in Genetics: Inherited Diseases

There are tests that can be used to determine whether a person carries the gene. However, they are not accurate enough to use for the general population.

They are much more accurate with people who have a family history of cystic fibrosis, so The American College of Obstetricians and Gynecologists recommends testing only for couples with a personal or close family history of cystic fibrosis.

Example 4 – *Probabilities and Cystic Fibrosis*

Each of two prospective parents carries one cystic fibrosis gene. Cystic fibrosis is recessive, so neither parent has the disease.

a. Find the probability that their child would have cystic fibrosis.

b. Find the probability that their child would be free of symptoms.

c. Find the probability that their child would be free of symptoms but could pass the cystic fibrosis gene on to his or her own child.

Example 4 – *Solution*

We will denote the recessive cystic fibrosis gene with the letter c and the normal disease-free gene with an N.

Each parent is Nc and therefore does not have the disease. Figure 3.5 shows the Punnett square for the child.

	N	c
N	(N, N)	(c, N)
c	(N, c)	(c, c)

A Punnett square for Example 4

Figure 3.5

Example 4 – *Solution*

cont'd

a. Cystic fibrosis is recessive, so only the (c, c) child will have the disease. The probability of such an event is 1/4.

b. The (N, N), (c, N), and (N, c) children will be free of symptoms. The probability of this event is

$$p(\text{healthy}) = p((\text{N, N})) + p((\text{c, N})) + p((\text{N, c}))$$

$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4}$$

$$= \frac{3}{4}$$

Example 4 – *Solution*

cont'd

c. The (c, N) and (N, c) children would never suffer from any symptoms but could pass the cystic fibrosis gene on to their own children.

The probability of this event is

$$p((\mathrm{c}, \mathrm{N})) + p((\mathrm{N}, \mathrm{c})) = \frac{1}{4} + \frac{1}{4}$$

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

Probabilities in Genetics: Inherited Diseases

In Example 4, the Nc child is called a **carrier** because that child would never suffer from any symptoms but could pass the cystic fibrosis gene on to his or her own child. Both of the parents were carriers.

Sickle-Cell Anemia

Sickle-cell anemia is an inherited disease characterized by a tendency of the red blood cells to become distorted and deprived of oxygen. Although it varies in severity, the disease can be fatal in early childhood. More often, patients have a shortened life span and chronic organ damage.

Probabilities in Genetics: Inherited Diseases

Newborns are now routinely screened for sickle-cell disease. The only true cure is a bone marrow transplant from a sibling without sickle-cell anemia; however, this can cause the patient's death, so it is done only under certain circumstances.

Until recently, about 10% of the children with sickle-cell anemia had a stroke before they were twenty-one.

But in 2009, it was announced that the rate of these strokes has been cut in half thanks to a new specialized ultrasound scan that identifies the individuals who have a high stroke risk. There are also medications that can decrease the episodes of pain.

Probabilities in Genetics: Inherited Diseases

Approximately 1 in every 500 black babies is born with sickle-cell anemia, but only 1 in 160,000 nonblack babies has the disease.

This disease is **codominant:** A person with two sickle-cell genes will have the disease, while a person with one sickle-cell gene will have a mild, nonfatal anemia called sickle-cell trait. Approximately 8%–10% of the black population has sickle-cell trait.

Huntington's Disease

Huntington's disease, caused by a dominant gene, is characterized by nerve degeneration causing spasmodic movements and progressive mental deterioration.

Probabilities in Genetics: Inherited Diseases

The symptoms do not usually appear until well after reproductive age has been reached; the disease usually hits people in their forties.

Death typically follows 20 years after the onset of the symptoms.

No effective treatment is available, but physicians can now assess with certainty whether someone will develop the disease, and they can estimate when the disease will strike.

Many of those who are at risk choose not to undergo the test, especially if they have already had children.

Probabilities in Genetics: Inherited Diseases

Singer-songwriter Woody Guthrie started to show signs of Huntington's disease in his late 30s. He died from the disease when he was 55.

Bob Dylan, Bruce Springsteen, John Mellencamp, Joe Strummer, Jay Farrar, and Wilco have all acknowledged Guthrie as a major influence and have all recorded CDs of Woody's songs.

Woody's wife Marjorie formed the Committee to Combat Huntington's Disease, which has stimulated research, increased public awareness, and provided support for families in many countries.

Probabilities in Genetics: Inherited Diseases

Woody's singer-songwriter son Arlo Guthrie chose to not be tested for Huntington's disease. Fortunately, the disease never manifested itself.

In August 1999, researchers in Britain, Germany, and the United States discovered what causes brain cells to die in people with Huntington's disease.

This discovery may eventually lead to a treatment. In 2008, a new drug that reduces the uncontrollable spasmodic movements was approved.

Probabilities in Genetics: Inherited Diseases

Tay-Sachs Disease

Tay-Sachs disease is a recessive disease characterized by an abnormal accumulation of certain fat compounds in the spinal cord and brain.

Most typically, a child with Tay-Sachs disease starts to deteriorate mentally and physically at six months of age. After becoming blind, deaf, and unable to swallow, the child becomes paralyzed and dies before the age of four years.

Probabilities in Genetics: Inherited Diseases

There is no effective treatment. The disease occurs once in 3,600 births among Ashkenazi Jews (Jews from central and eastern Europe), Cajuns, and French Canadians but only once in 600,000 births in other populations.

Carrier-detection tests and fetal-monitoring tests are available. The successful use of these tests, combined with an aggressive counseling program, has resulted in a decrease of 90% of the incidence of this disease.

Probabilities in Genetics: Inherited Diseases

Hemophilia

When you break a blood vessel, the bleeding stops because your blood clots. Hemophilia is an inherited disease that impairs blood clotting. As a result, hemophiliacs experience prolonged bleeding. Without treatment, the disease is crippling, and often fatal. About 6 in 50,000 people are born with hemophilia.

The gene that causes hemophilia is located on the X chromosome, one of the two chromosomes that determine gender. Because of this, the disease is called an *X-linked disorder.*

Probabilities in Genetics: Inherited Diseases

We'll call an X chromosome that has the hemophilia gene "Xh," and we'll call an X chromosome that doesn't have the hemophilia gene "X."

The X gene dominates the Xh gene. A male has one Y and one X chromosome (either X or Xh), and a female has two X chromosomes (either X or Xh).

Example 5 – *Probabilities and Hemophilia*

A prospective mother has an X chromosome and an Xh chromosome. A prospective father has one X chromosome (X, not Xh) and one Y chromosome.

a. Find the probability that their child would be a hemophiliac male.

b. Find the probability that their child would be a hemophiliac female.

c. Find the probability that their child would be carrier.

d. Find the probability that their child would be symptom free.

Example 5 – *Solution*

There are four possible outcomes, as shown in Figure 3.6:

mother → father ↓	X	Xh
X	(X, X)	(Xh, X)
Y	(X, Y)	(Xh, Y)

A Punnett square for Example 5.

Figure 3.6

- The (X, X) child is female because she has two X chromosomes. She has no hemophilia gene and is symptom free.

Example 5 – *Solution*

cont'd

- The (Xh, X) child is female because she has two X chromosomes. X dominates Xh, so she does not have hemophilia. She is a carrier, because she would not suffer from any symptoms but could pass her Xh gene on to a child.
- The (X, Y) child is male because he has an X chromosome and a Y chromosome. He has no hemophilia gene and is symptom free.
- The (Xh, Y) child is male because he has an X chromosome and a Y chromosome. He has hemophilia, because he has an Xh gene, but no X gene to dominate it.

Example 5 – *Solution*

cont'd

a. The (Xh, Y) child is a boy with hemophilia. The probability of this event is ¼.

b. The (X, X) child and the (Xh, X) child are girls, but neither has hemophilia. The probability of a hemophiliac female is 0/4 = 0.

c. The (Xh, X) child is a carrier. The probability of this event is ¼.

d. The (X, X) child, the (Xh, X) child, and the (X, Y) child are all symptom free. The probability of this event is ¾.

Genetic Screening

Genetic Screening

There are no conclusive tests that will tell a parent whether he or she is a cystic fibrosis carrier, nor are there conclusive tests that will tell whether a fetus has the disease.

A new test resulted from the 1989 discovery of the location of most cystic fibrosis genes, but that test will detect only 85% to 95% of the cystic fibrosis genes, depending on the individual's ethnic background.

The extent to which this test will be used has created quite a controversy.

Genetic Screening

Individuals who have relatives with cystic fibrosis are routinely informed about the availability of the new test.

The controversial question is whether a massive genetic screening program should be instituted to identify cystic Fibrosis carriers in the general population, regardless of family history.

This is an important question, considering that four in five babies with cystic fibrosis are born to couples with no previous family history of the condition.

Genetic Screening

Opponents of routine screening cite a number of important concerns. The existing test is rather inaccurate; 5% to 15% of the cystic fibrosis carriers would be missed.

It is not known how health insurers would use this genetic information—insurance firms could raise rates or refuse to carry people if a screening test indicated a presence of cystic fibrosis.

Also, some experts question the adequacy of quality assurance for the diagnostic facilities and for the tests themselves.

Genetic Screening

Supporters of routine testing say that the individual should be allowed to decide whether to be screened.

Failing to inform people denies them the opportunity to make a personal choice about their reproductive future.

An individual who is found to be a carrier could choose to avoid conception, to adopt, to use artificial insemination by a donor, or to use prenatal testing to determine whether a fetus is affected—at which point the additional controversy regarding abortion could enter the picture.

The Failures of Genetic Screening

The Failures of Genetic Screening

The history of genetic screening programs is not an impressive one. In the 1970s, mass screening of blacks for sickle-cell anemia was instituted.

This program caused unwarranted panic; those who were told they had the sickle-cell trait feared that they would develop symptoms of the disease and often did not understand the probability that their children would inherit the disease.

Some people with sickle-cell trait were denied health insurance and life insurance and were treated as if they had, or could later develop, sickle-cell anemia. This discrimination was short-lived due to the passage of the National Sickle-Cell Anemia Control Act in 1972.