The Basic Practice of Statistics
Ninth Edition

David S. Moore William I. Notz

Chapter 12 and 13

Lecture Slides

In Chapter 12 and 13, we cover …

- The idea of probability
- Probability models and rules
- Finite and continuous probability models
- Random variables
- Personal probability
- The general addition rule
- Independence and the multiplication rule
- Conditional probability
- The general multiplication rule
- Showing events are independent

EXAMPLE 12.1 Does Anyone in Your Household Own a Gun? (1 of 5)

- What proportion of all U.S. adults say someone in their household owns a gun?
- A *Washington Post*/ABC News poll took a random sample of 1003 adults.
- The poll found that 461 of the people in the sample said that someone in their household owned a gun.

Sample proportion = $\frac{461}{1003}$ = 0.46 (that is, 46%)

EXAMPLE 12.1 Does Anyone in Your Household Own a Gun? (2 of 5)

Sample proportion = $\frac{461}{1003}$ = 0.46 (that is, 46%)

- If the sample was a simple random sample of all adults, then, as discussed in Ch. 8, all adults had the same chance to be among the chosen 1003.
- We don't know what percentage of all adults would say that someone in their household owned a gun, but we *estimate* that about 46% did at the time of the poll.
- This is a basic move in statistics: *use a result from a sample to estimate something about a population*.

EXAMPLE 12.1 Does Anyone in Your Household Own a Gun? (3 of 5)

- What if the *Washington Post*/ABC News poll took a second random sample of 1003 adults?
- It is almost certain that there would not be exactly 461 positive responses.

Random samples eliminate bias from the act of choosing a sample, but they can still fail to perfectly agree with the true population proportion because of the variability that results when we choose at random.

EXAMPLE 12.1 Does Anyone in Your Household Own a Gun? (4 of 5)

- If the variation when we take repeated samples from the same population is too great, we can’t trust the results of any one sample.

→ This is where we need facts about probability to make progress in statistics.

EXAMPLE 12.1 Does Anyone in Your Household Own a Gun? (5 of 5)

- When a poll uses chance to choose its samples, the laws of probability govern the behavior of the samples.
- The *Washington Post*/ABC News poll says that the probability is 0.95 that an estimate from one of its samples comes within ± 3.5 percentage points of the truth about the population of all adults.
- What does “probability is 0.95” mean?
- Our purpose in Ch. 12 and 13 is to understand the language of probability—but without going into the full mathematics of probability theory.

The idea of probability

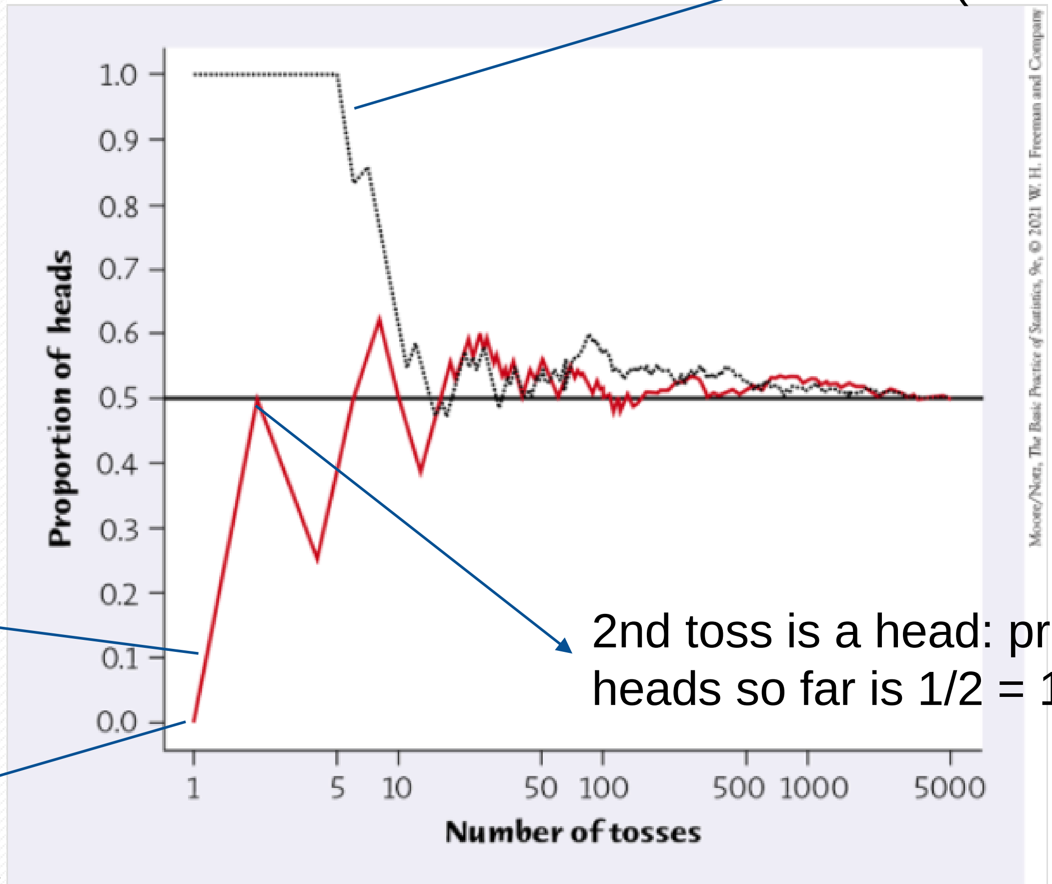
Chance behavior is unpredictable in the short run, but it has a regular and predictable pattern in the long run.

RANDOMNESS AND PROBABILITY

- We call a phenomenon random if individual outcomes are uncertain but there is, nonetheless, a regular distribution of outcomes in a large number of repetitions.
- The probability of any outcome of a random phenomenon is the proportion of times the outcome would occur in a very long series of repetitions.

EXAMPLE 12.2 Coin Tossing (two trials – A and B)

Trial B (black dotted line)



Trial A (red line)

2nd toss is a head: proportion of heads so far is 1/2 = 1/2

1st toss is a tail: proportion of heads so far is 0/1 =0

FIGURE 12.1

The proportion of tosses of a coin that give a head changes as we make more tosses. Eventually, however, the proportion approaches 0.5, the probability of a head. This figure shows the results of two trials of 5000 tosses each.

EXAMPLE 12.2 Coin Tossing (two trails – A and B)

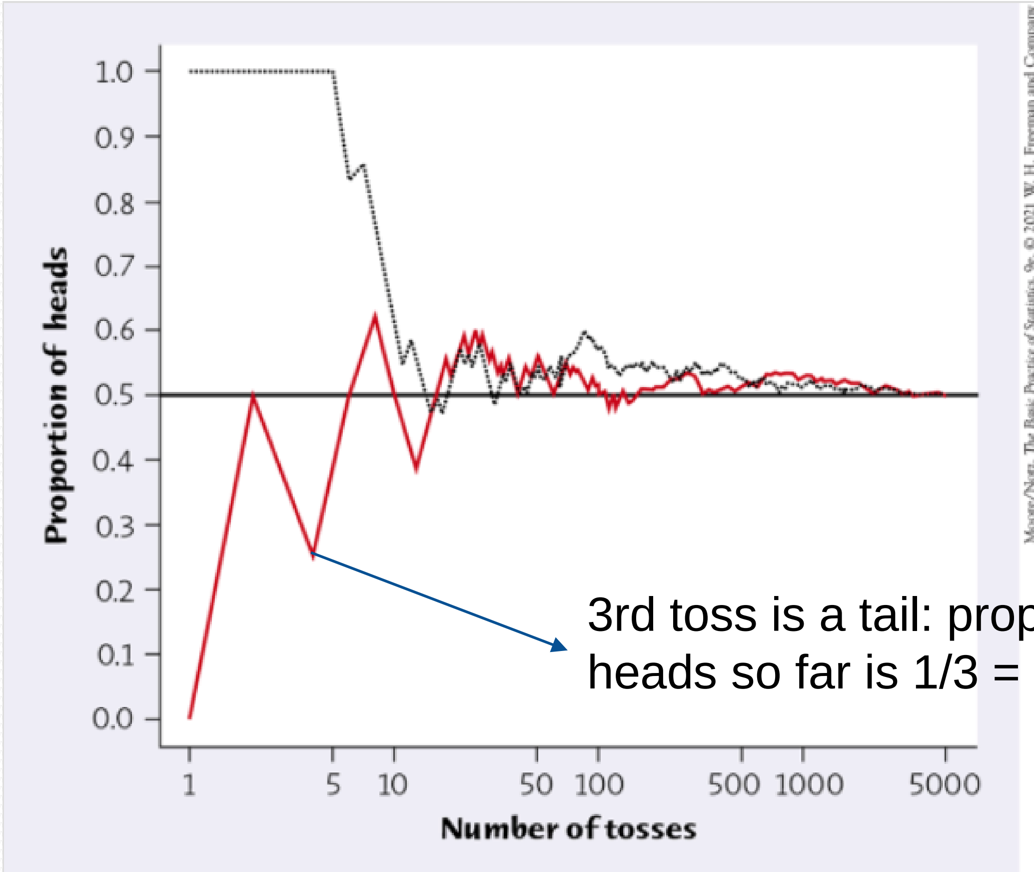
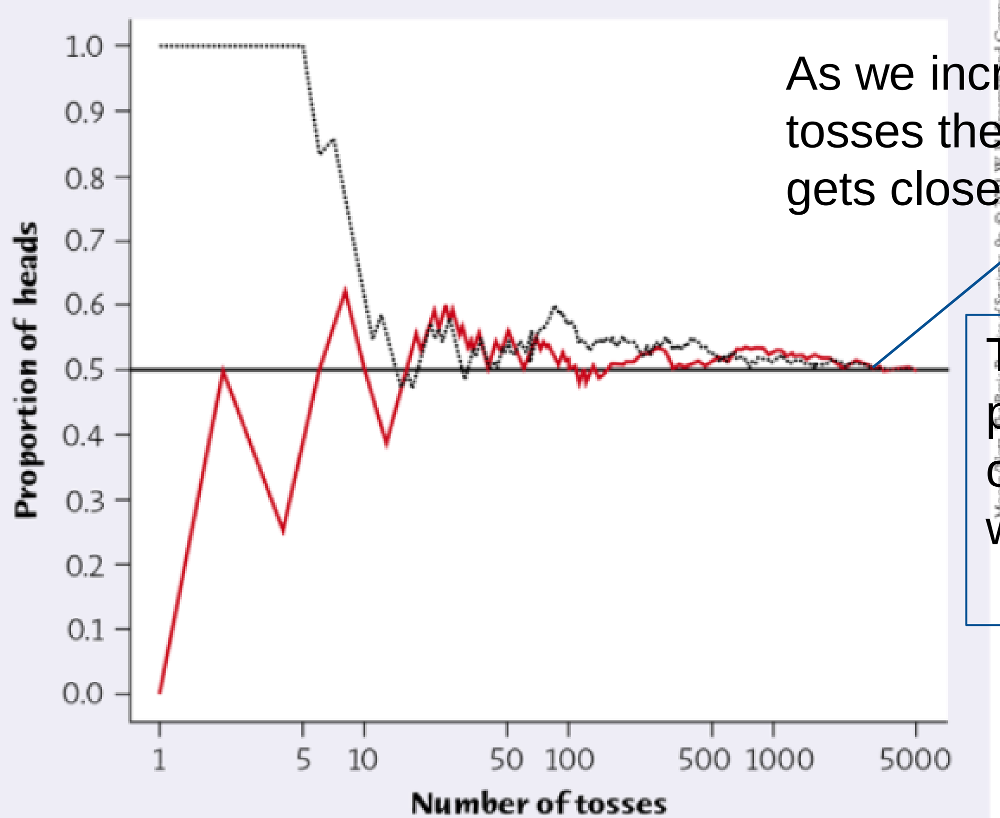


FIGURE 12.1

The proportion of tosses of a coin that give a head changes as we make more tosses. Eventually, however, the proportion approaches 0.5, the probability of a head. This figure shows the results of two trials of 5000 tosses each.

EXAMPLE 12.2 Coin Tossing (two trails – A and B)



As we increase the number of tosses the proportion of heads gets closer and closer to 0.5

That is the probability of obtaining a head when tossing a coin

FIGURE 12.1

The proportion of tosses of a coin that give a head changes as we make more tosses. Eventually, however, the proportion approaches 0.5, the probability of a head. This figure shows the results of two trials of 5000 tosses each.

Another example

- Visit https://www.statcrunch.com/applets/type3&dice

Probability models

Descriptions of chance behavior contain two parts: a list of possible outcomes and a probability for each outcome.

- The sample space *S* of a random phenomenon is the set of all possible outcomes.
- An event is an outcome or a set of outcomes of a random phenomenon. That is, an event is a subset of the sample space.
- A probability model is a mathematical description of a random phenomenon; it consists of two parts: a sample space *S* and a way of assigning probabilities to events.

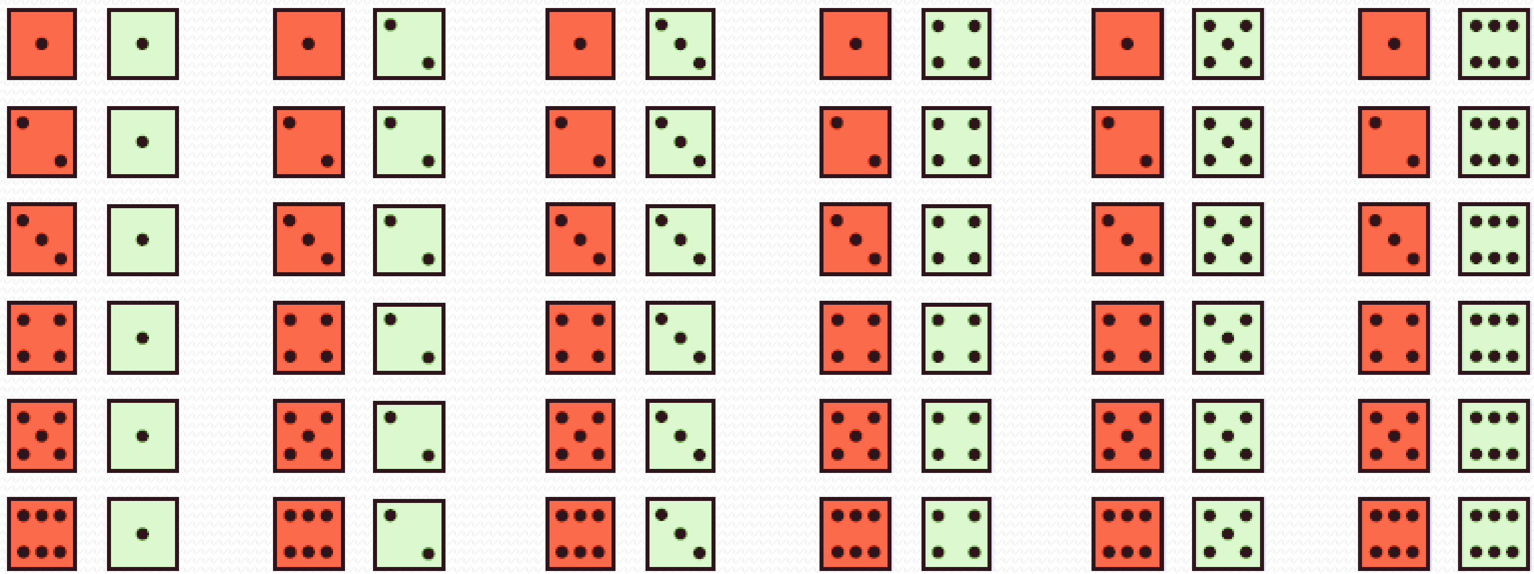
Probability models

Examples of sample spaces:

- A sample space *S* can be very simple or very complex.
- When we toss a coin once, there are only two outcomes: heads and tails. The sample space is S ={H,T}.
- When the *Washington Post*/ABC News poll draws a random sample of 1003 adults, the sample space contains all possible choices of 1003 of the 254 million adults in the USA. This *S* is extremely large.
- Each member of *S* is a possible sample, so *S* is the collection, or “space,” of all possible samples. This explains the term *sample space*.

Probability models

Example: Give a probability model for the chance process of rolling two fair, six-sided dice—one that is red and one that is green.



Sample space
36 outcomes

Since the dice are fair, each outcome is equally likely.
Each outcome has probability 1/36.

Probability rules

Rule 1. **Any probability is a number between 0 and 1.** Any proportion is a number between 0 and 1, so any probability is also a number between 0 and 1.

Rule 2. **All possible outcomes together must have probability 1.** Because some outcome must occur on every trial, the sum of the probabilities for all possible outcomes must be exactly 1.

Rule 3. ***If two events have no outcomes in common*, the probability that one or the other occurs is the sum of their individual probabilities.**

Rule 4. **The probability that an event does not occur is 1 minus the probability that the event does occur.** The probability that an event occurs and the probability that it does not occur always add to 1, or 100%.

Probability rules

The probability rules in formal language:

Rule 1. The probability $P(A)$ of any event A satisfies $0 \leq P(A) \leq 1$.

Rule 2. If S is the sample space in a probability model, $P(S) = 1$.

Rule 3. Two events A and B are disjoint if they have no outcomes in common and, thus, can never occur together. If A and B are disjoint, then

$$P(A \text{ or } B) = P(A) + P(B)$$

This is the addition rule for disjoint events.

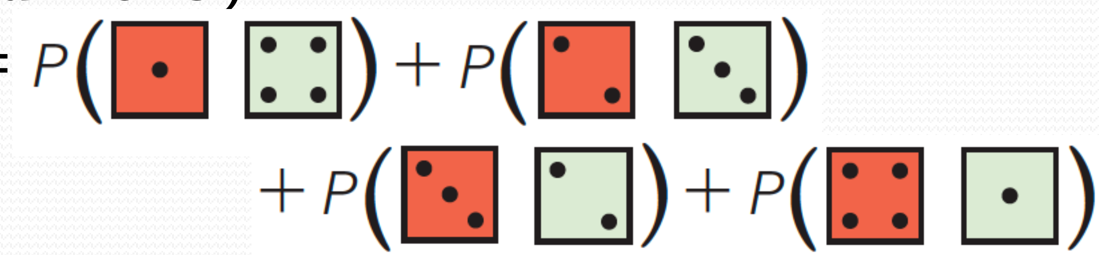
Rule 4. For any event A,

$$P(A \text{ does not occur}) = 1 - P(A)$$

Probability rules (example)

We will use the probability model from the dice example to illustrate some of the probability rules.

- Finding $P(\text{roll a sum of 5})$:

$P(\text{roll a sum of 5}) =$ 

That is, the sum of the outcomes of the two dice is of 5

$$= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{4}{36}$$

because these four disjoint outcomes make up entirely the event "roll a sum of 5."

- Finding $P(\text{roll } \mathit{not} \text{ a sum of 5})$:

$$P(\text{roll } \mathit{not} \text{ a sum of 5}) = 1 - P(\text{roll a sum of 5}) = {}^{32}/_{36}$$

Finite probability models

- One way to find the probability of an event is to assign a probability to each of the individual outcomes that make up the event and then add these probabilities. This idea works well when there are only a finite (fixed and limited) number of outcomes.

- A probability model with a finite sample space is called a finite probability model.

- To assign probabilities in a finite model, list the probabilities of all the individual outcomes. These probabilities must be numbers between 0 and 1 that add to exactly 1. The probability of any event is the sum of the probabilities of the outcomes making up the event.

- Finite probability models are sometimes called discrete probability models. Statisticians often refer to finite probability models as discrete.

Finite probability model (example)

The first digits of numbers in legitimate financial records often follow a model known as Benford's law. Call the first digit of a randomly chosen record X. Benford's law gives the following probability model for X.

First digit (X)	1	2	3	4	5	6	7	8	9
Probability	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046

(a) Show that this is a legitimate finite (or discrete) probability model.

Each probability is between 0 and 1, and

$$\mathbf{0.301 + 0.176 + \cdots + 0.046 = 1}$$

(b) Find the probability that the first digit for the chosen number is not a 1.

$$\begin{aligned} \boldsymbol{P}(\textbf{not } \mathbf{1}) &= \mathbf{1} - \boldsymbol{P}(\mathbf{1}) \\ &= \mathbf{1} - \mathbf{0.301} = \mathbf{0.699} \end{aligned}$$

Continuous probability models

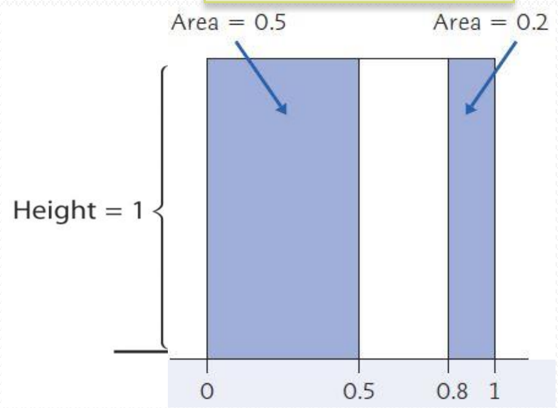
- Suppose we want to choose a number at random between 0 and 1, allowing any number between 0 and 1 as the outcome.
- We cannot assign probabilities to each individual value because there is an infinite continuum of possible values.

- A continuous probability model assigns probability as an area under a density curve. The area under the curve and above any range of values is the probability of an outcome in that range.

Uniform Distribution

Example: Find the probability of getting a random number that is less than or equal to 0.5 or greater than 0.8.

$$P(X \leq 0.5 \text{ or } X > 0.8)$$
$$= P(X \leq 0.5) + P(X > 0.8)$$
$$= 0.5 + 0.2 = 0.7$$



Normal probability models

- We can use any density curve to assign probabilities. The density curves that are most familiar to us are the Normal curves.
- <u>Normal distributions are continuous probability models</u>.
- Like all continuous probability models, the Normal assigns probability 0 to every individual outcome.

Normal probability models (Example 12.10)

- If we look at the heights of all young women, we find that they closely follow the Normal distribution, with mean $\mu = 64.1$ inches and standard deviation $\sigma = 3.7$ inches.

What is the probability that a randomly chosen young women has a height (X) between 68 and 70 inches?

$$P(68 \leq X \leq 70) = \mathrm{P}\left(\frac{68 - 64.1}{3.7} \leq \frac{X - 64.1}{3.7} \leq \frac{70 - 64.1}{3.7}\right)$$
$$= \mathrm{P}(1.05 \leq z \leq 1.59)$$
$$= \mathrm{P}(z \leq 1.59) - P(z \leq 1.05)$$
$$= 0.9441 - 0.8531 = 0.0910$$

- This is the same calculation we learned in Ch. 3, but now instead of asking about proportion/percentage we ask about the probability. The probability equals to the proportion.

Random variables

- A random variable is a variable whose value is a numerical outcome of a random phenomenon.
- The probability distribution of a random variable X tells us what values X can take and how to assign probabilities to those values.

- A finite random variable has a finite list of possible outcomes.
- Random variables that can take on any value in an interval, with probabilities given as areas under a density curve, are called continuous.

Some exercises from Ch. 12

12.6 Role-Playing Games. Computer games in which the players take the roles of characters are very popular. They go back to earlier tabletop games such as Dungeons & Dragons. These games use many different types of dice. A four-sided die has faces with one of the numbers 1, 2, 3, or 4 appearing at the bottom of each visible face.

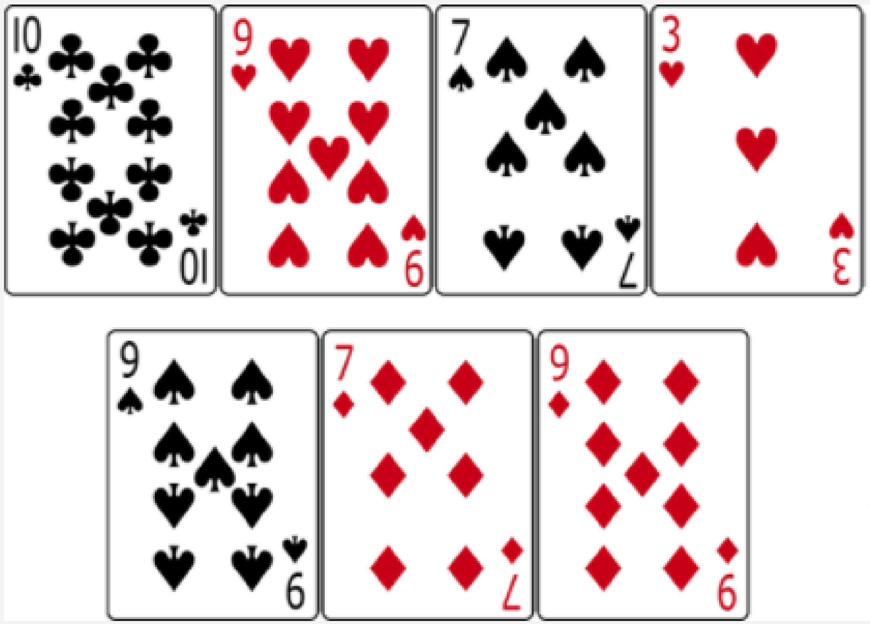
a. What is the sample space for rolling a four-sided die twice (numbers on first and second rolls)? Follow the example of Figure 12.2.
b. What is the assignment of probabilities to outcomes in this sample space? Assume that the die is perfectly balanced and follow the method of Example 12.4.

12.6 (a) The accompanying table illustrates the 16 possible pair combinations in the sample space.

	1	2	3	4
1	1 1	1 2	1 3	1 4
2	2 1	2 2	2 3	2 4
3	3 1	3 2	3 3	3 4
4	4 1	4 2	4 3	4 4

(b) Each of the 16 outcomes has probability 1/16.

12.38 Drawing Cards. You are about to draw a card at random (that is, all choices have the same probability) from a set of seven cards. Although you can't see the cards, here they are:

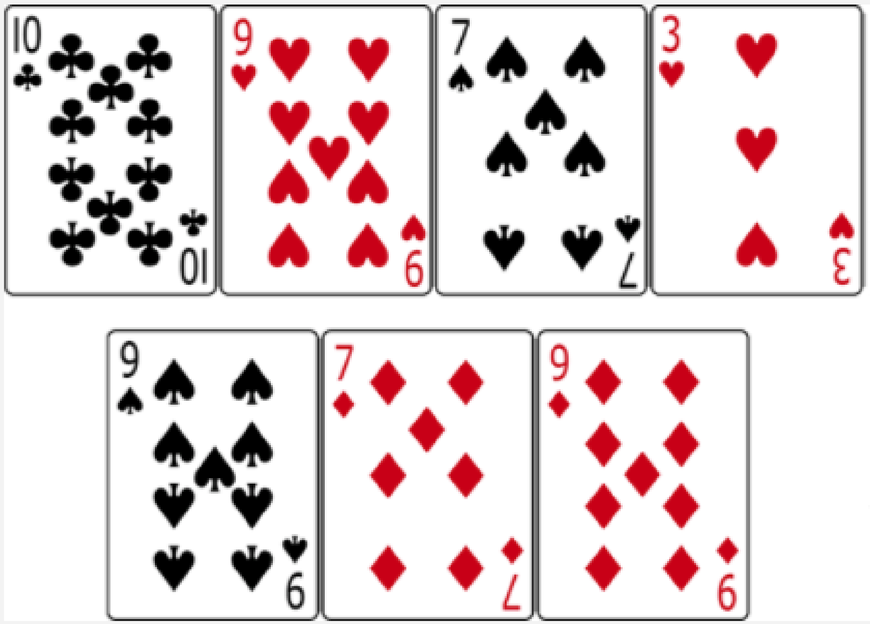


a. What is the probability that you draw a 9?

b. What is the probability that you draw a red 9?

c. What is the probability that you do not draw a 7?

12.38 Drawing Cards. You are about to draw a card at random (that is, all choices have the same probability) from a set of seven cards. Although you can't see the cards, here they are:



a. What is the probability that you draw a 9?

b. What is the probability that you draw a red 9?

c. What is the probability that you do not draw a 7?

12.38 (Of the seven cards, there are three 9s, two red 9s, and two 7s.)

(a) P(draw a 9) = 3/7.

(b) P(draw a red 9) = 2/7.

(c) P(don't draw a 7) = 1 – P(draw a 7) = 1 – 2/7 = 5/7.

12.39 **Loaded Dice**. There are many ways to produce crooked dice. To *load* a die so that 6 comes up too often and 1 (which is opposite 6) comes up too seldom, add a bit of lead to the filling of the spot on the 1 face. If a die is loaded so that 6 comes up with probability 0.2 and the probabilities of the 2, 3, 4, and 5 faces are not affected, what is the assignment of probabilities to the six faces?

12.39 Loaded Dice. There are many ways to produce crooked dice. To *load* a die so that 6 comes up too often and 1 (which is opposite 6) comes up too seldom, add a bit of lead to the filling of the spot on the 1 face. If a die is loaded so that 6 comes up with probability 0.2 and the probabilities of the 2, 3, 4, and 5 faces are not affected, what is the assignment of probabilities to the six faces?

12.39 The probabilities of 2, 3, 4, and 5 are unchanged (1/6), so P(1 or 6) must still be 1/3. If $P(6) = 0.2$, then $P(1) = 1/3 - 0.2 = 2/15$.

Face	⚀	⚁	⚂	⚃	⚄	⚅
Probability	$0.1\overline{3}$	1/6	1/6	1/6	1/6	0.2

In Chapter 13, we cover …

- The general addition rule
- Independence and the multiplication rule
- Conditional probability
- The general multiplication rule
- Showing events are independent

Probability rules

Everything in this chapter follows from the four rules we learned in Chapter 12:

Rule 1. For any event A, $0 \leq P(A) \leq 1$.

Rule 2. If S is the sample space, $P(S) = 1$.

Rule 3. The addition rule for disjoint events says that if A and B are disjoint events, then

$$P(A \text{ or } B) = P(A) + P(B)$$

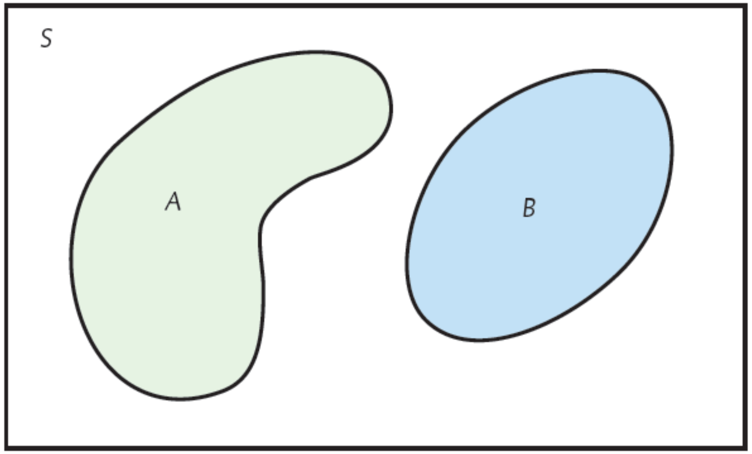
Rule 4. For any event A,

$$P(A \text{ does not occur}) = 1 - P(A)$$

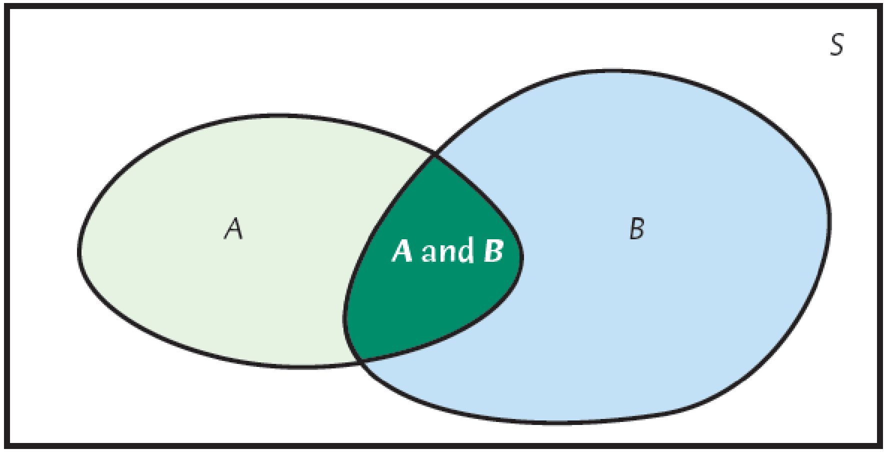
Venn diagrams

Sometimes it is helpful to draw a picture to display relations among several events. A picture that shows the sample space S as a rectangular area and shows events as areas within S is called a Venn diagram.

Two events that are disjoint.



Two events that are not disjoint. The event {A and B} consists of the outcomes that they have in common.



The general addition rule

We know that if *A* and *B* are disjoint events:

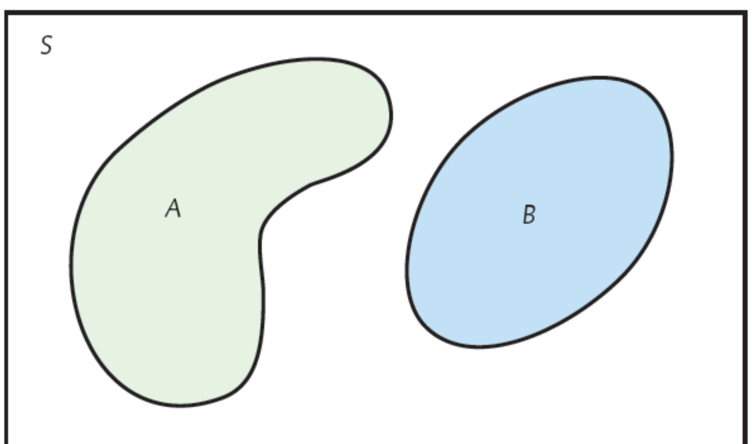
$$P(A \text{ or } B) = P(A) + P(B)$$

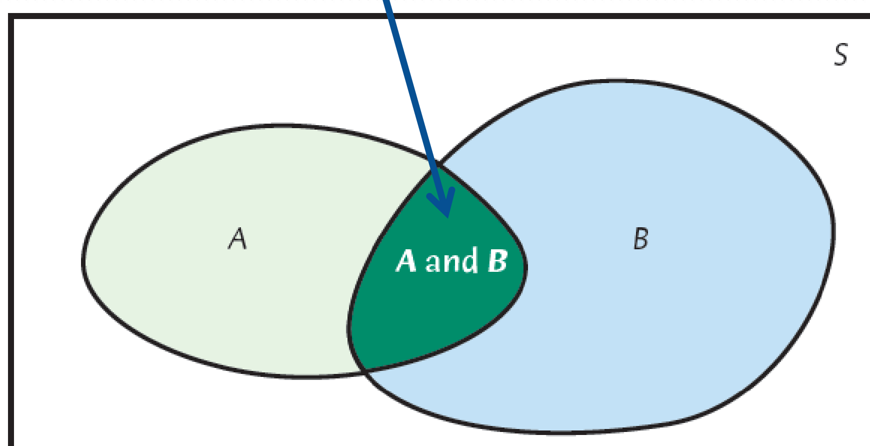
<u>Addition rule for *any* two events</u>

For any two events *A* and *B*:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

Outcomes here are double-counted by $P(A) + P(B)$.





The general addition rule (Example 13.1)

- A 2019 survey found that 65% of adults had read a print book (B) in the preceding 12 months, 25% had read a book in digital format (D), and 18% had read both a print book and a book in digital format.
- Choose an adult at random. Then

$$\begin{aligned} P(B \text{ or } D) &= P(B) + P(D) - P(B \text{ and } D) \\ &= 0.65 + 0.25 - 0.18 = 0.72 \end{aligned}$$

- That is, 72% of adults had read a book in either print, digital, or both forms in the preceding 12 months.

The general addition rule (Example 13.1)

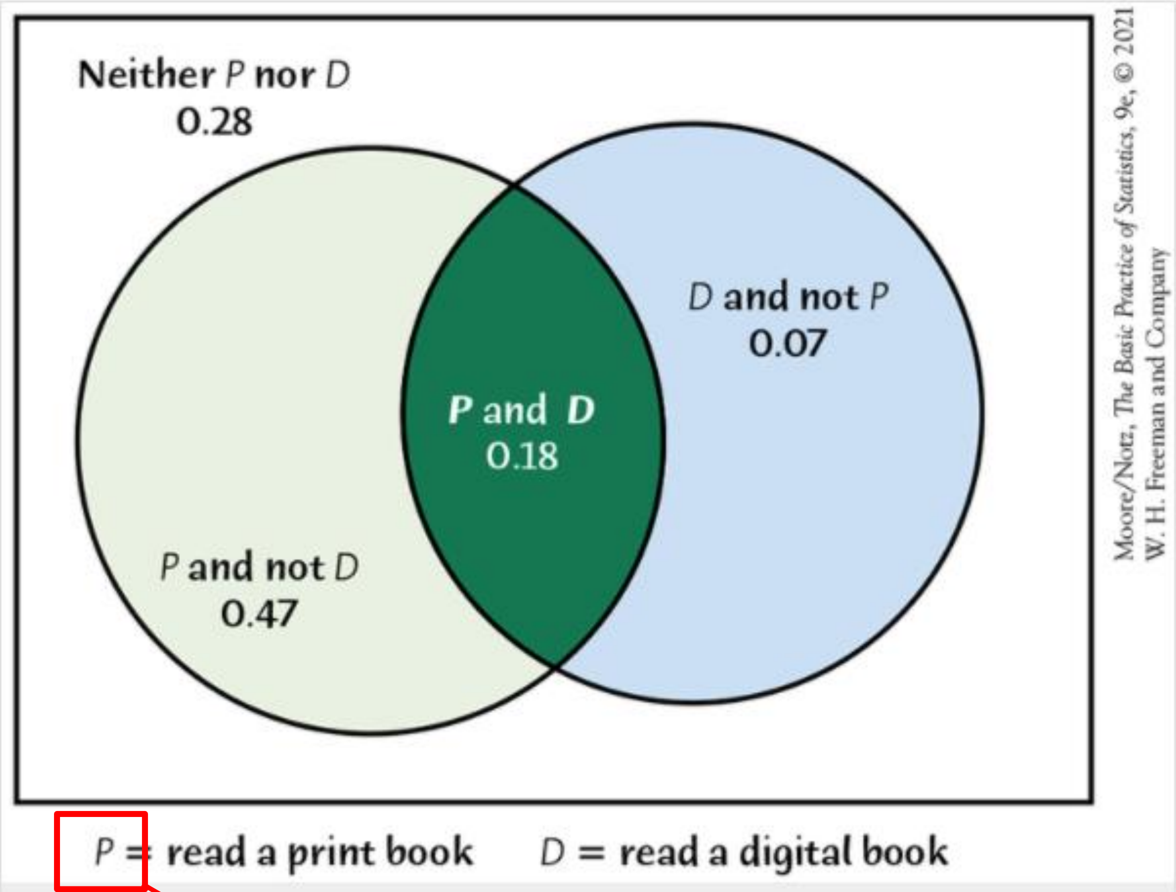


FIGURE 13.4

Venn diagram and probabilities for print versus digital format, for Example 13.1.

$$P(B \text{ or } D) = P(B) + P(D) - P(B \text{ and } D) = 0.65 + 0.25 - 0.18 = 0.72$$

Or

P(B or D) = P(B and not D)
+ P(B and D)
+ P(D and not B)
= 0.47 + 0.18 + 0.07 = 0.72

In the previous slide we called this B

Independence and the multiplication rule

- If two events A and B do not influence each other (if knowledge about one does not change the probability we assign to the other), then the events are said to be independent of each other.

MULTIPLICATION RULE FOR INDEPENDENT EVENTS

- Two events A and B are independent if knowing that one occurs does not change the probability we assign to the other occurring.
- If A and B are independent, then

$$\boldsymbol{P(A \text{ and } B) = P(A)P(B)}$$

Notes about this multiplication rule

- The multiplication rule extends to collections of more than two events, provided that all events are independent.
- *Caution! Be careful not to confuse* disjointness *and* independence*. If A and B are disjoint, then the fact that A occurs tell us that B* cannot *occur*—very *dependent!*
- *Caution! The special multiplication rule* $P(A \text{ and } B) = P(A)P(B)$ *holds if A and B are independent, but* not *otherwise.*

Conditional probability

- The probability we assign to an event can change if we know that some other event has occurred. This idea is the key to many applications of probability.
- When we are trying to find the probability that one event will happen, *given the information that the other event is already known to have occurred*, we are trying to determine a conditional probability.

- When $P(A) > 0$, the **conditional probability** of B given A is

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

Conditional probability (example)

Consider imported motor vehicle sales in the United States:

	NAFTA	Other	Total
Light truck/car	4,337,091	3,881,650	8,218,741
Medium/heavy truck	189,722	40,995	230,717
Total	4,526,813	3,922,645	8,449,458

If we consider *only the medium and heavy trucks*, what is the likelihood of a randomly selected sale being the sale of a "NAFTA" vehicle?

$$P(\text{NAFTA}|\text{Medium/heavy truck}) = \frac{189{,}722}{230{,}717} = \frac{\frac{189{,}722}{8{,}449{,}458}}{\frac{230{,}717}{8{,}449{,}458}}$$

$$= \frac{P(\text{NAFTA and Medium/heavy truck})}{P(\text{Medium/heavy truck})}$$

- This is an example of a conditional probability.

The general multiplication rule (part I)

- The definition of conditional probability reminds us that, in principle, all probabilities, including conditional probabilities, can be found from the assignment of probabilities to events that describe a random phenomenon.
- More often, however, conditional probabilities are part of the information given to us in a probability model. The definition of conditional probability then turns into a rule for finding the probability that both of two events occur.

The general multiplication rule (part II)

- The definition of conditional probability leads to a more general version of the multiplication rule.

MULTIPLICATION RULE FOR ANY TWO EVENTS

- The probability that two events *A* and *B* happen together can be found by

$$P(A \text{ and } B) = P(A)P(B|A)$$

- Here $P(B|A)$ is the conditional probability that *B* occurs, given the information that *A* occurs.
- This may be extended to any number of events:

$$P(A \text{ and } B \text{ and } C) = P(A)P(B|A)P(C|\text{both } A \text{ and } B)$$

The general multiplication rule (example)

- The Pew Internet and Technology Project finds that 91% of Gen X-ers use the Internet and that 89% of online Gen X-ers say the Internet is a good thing for them personally. What percent of Gen X-ers are online and say the Internet is a good thing for them personally?
- Use the product rule:

$P(\text{online}) = 0.91$, and $P(\text{Internet is good for them personally}|\text{online}) = 0.89$ (why conditional?), so

$$P(\text{online and Internet is good for them personally})$$

$$= P(\text{online}) \times P(\text{Internet is good for them personally}|\text{online})$$

$$= 0.91 \times 0.89 = 0.8099$$

Showing events are independent

- The conditional probability $P(B \mid A)$ is generally not equal to the unconditional probability $P(B)$.
- If knowing that A occurs gives no additional information about B, then A and B are independent events.

INDEPENDENT EVENTS

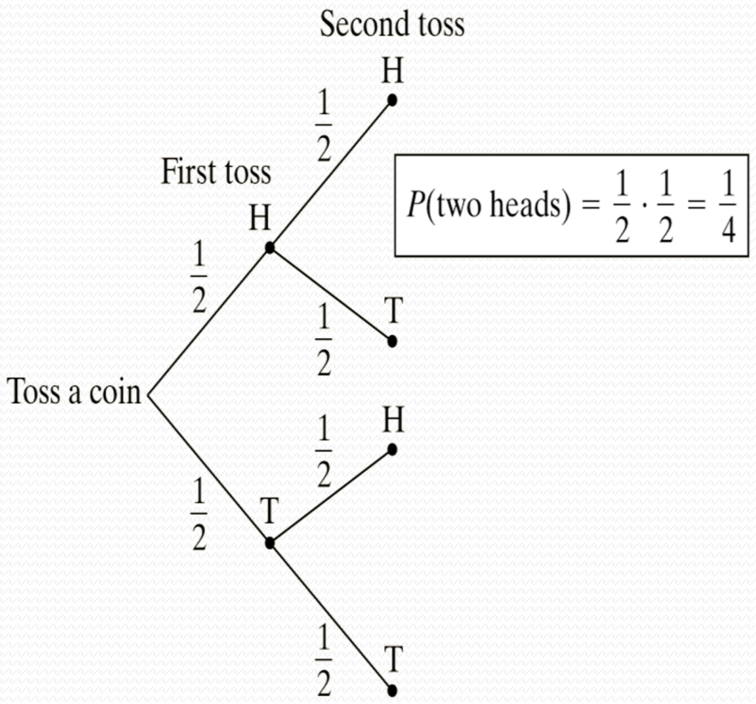
- Two events A and B that both have positive probability are **independent** if

$$P(B \mid A) = P(B)$$

Tree diagrams

- Another way to model chance behavior that involves a sequence of outcomes is to construct a tree diagram.

Consider flipping a coin twice. What is the probability of getting two heads?



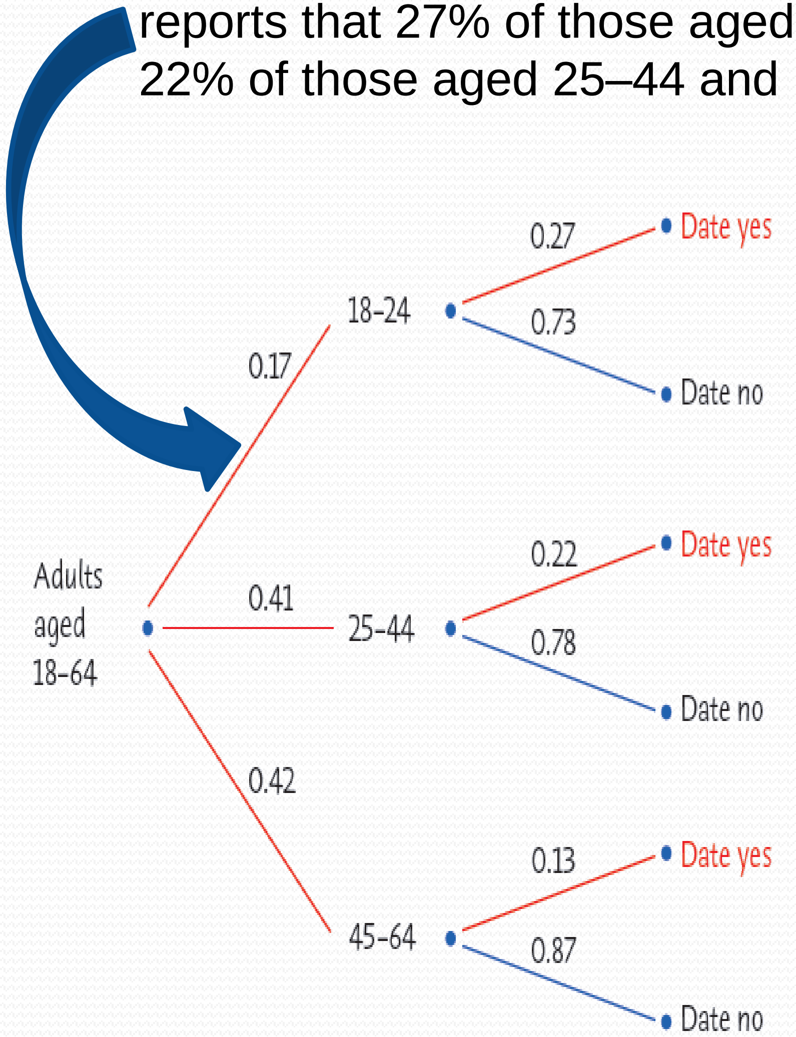
Sample Space

HH HT TH TT

Thus, P(two heads) = P(HH) = 1/4.

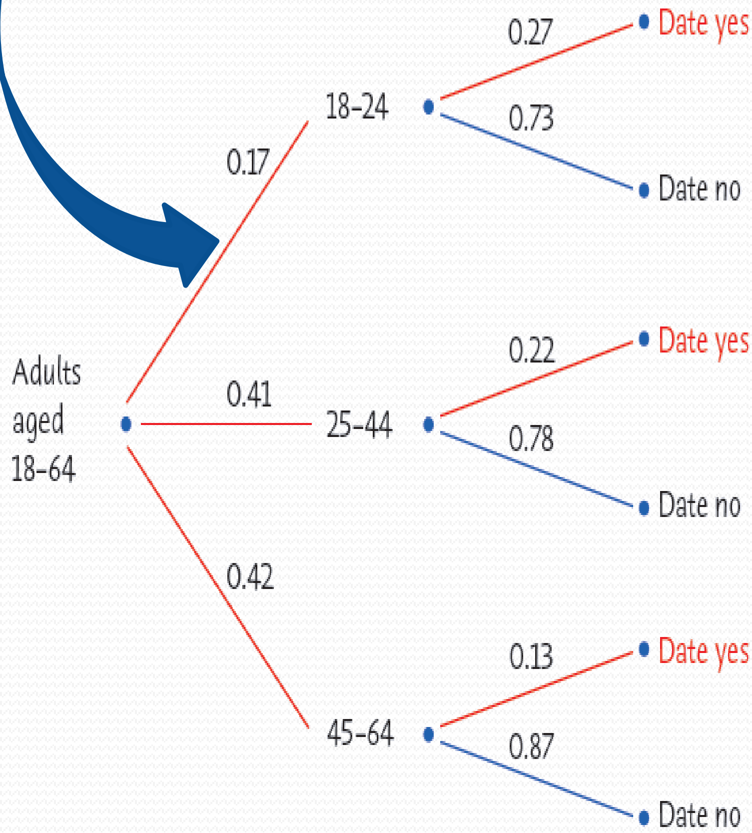
Tree diagrams (example)

Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.



Tree diagrams (example)

Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.



$\rightarrow P(18 \text{ to } 24) = 0.17$

$P(\text{online date yes} \mid 18 \text{ to } 24) = 0.27$

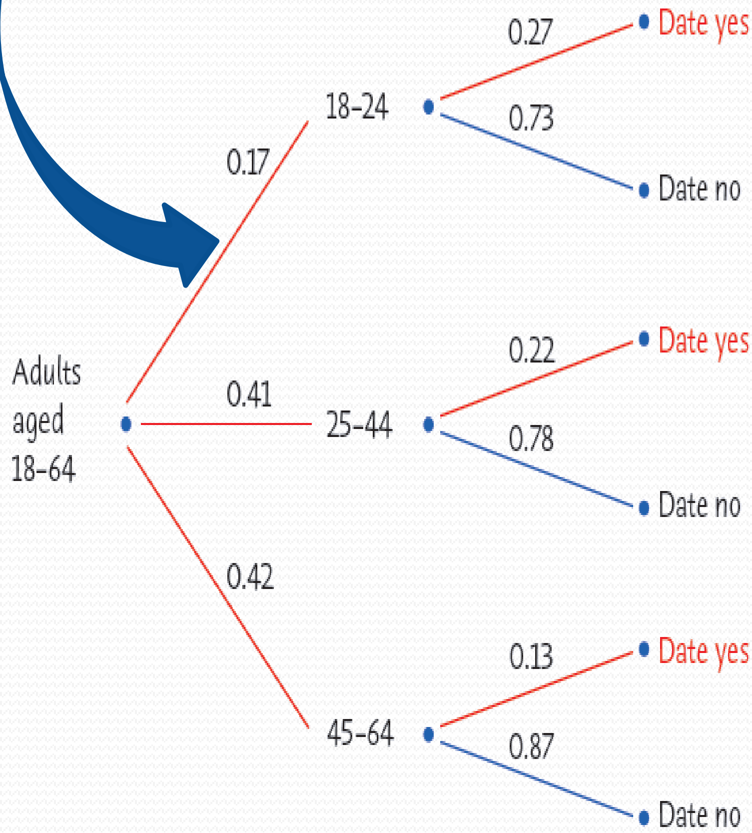
$$\boxed{\begin{aligned} &P(18 \text{ to } 24 \text{ and online date yes}) = \\ &\quad = P(18 \text{ to } 24) * P(\text{online date yes} \mid 18 \text{ to } 24) \end{aligned}}$$
$$= (0.17)(0.27)$$
$$= 0.0459$$

Multiplication rule for any two events, using conditional probability

What is *P*(25 to 44 and online date yes)?

Tree diagrams (example)

Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.



$\rightarrow P(\text{18 to 24}) = 0.17$

$P(\text{online dating} \mid \text{18 to 24}) = 0.27$

$$\boxed{\begin{aligned} &P(\text{18 to 24 and online dating}) \\ &\quad = P(\text{18 to 24}) * P(\text{online dating} \mid \text{18 to 24}) \end{aligned}}$$
$$= (0.17)(0.27)$$
$$= 0.0459$$

Multiplication rule for any two events, using conditional probability

What is P(25 to 44 and online dating)?

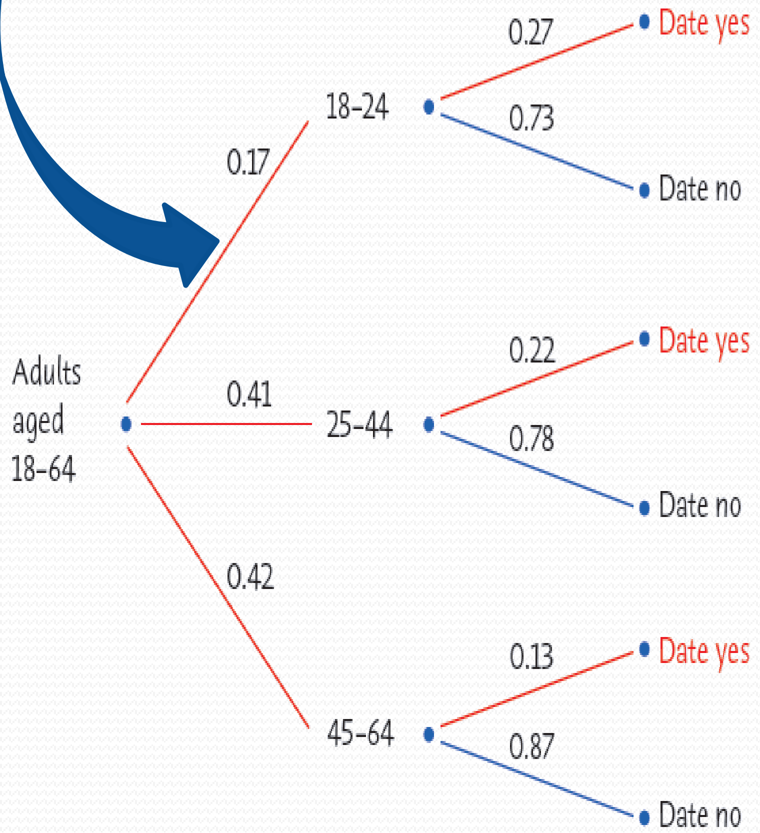
$P(\text{ 25 to 44 and online dating}) =$

$P(\text{25 to 44}) * P(\text{online dating} \mid \text{25 to 44})$

$= 0.41 * 0.22 = 0.0902$

Tree diagrams (example)

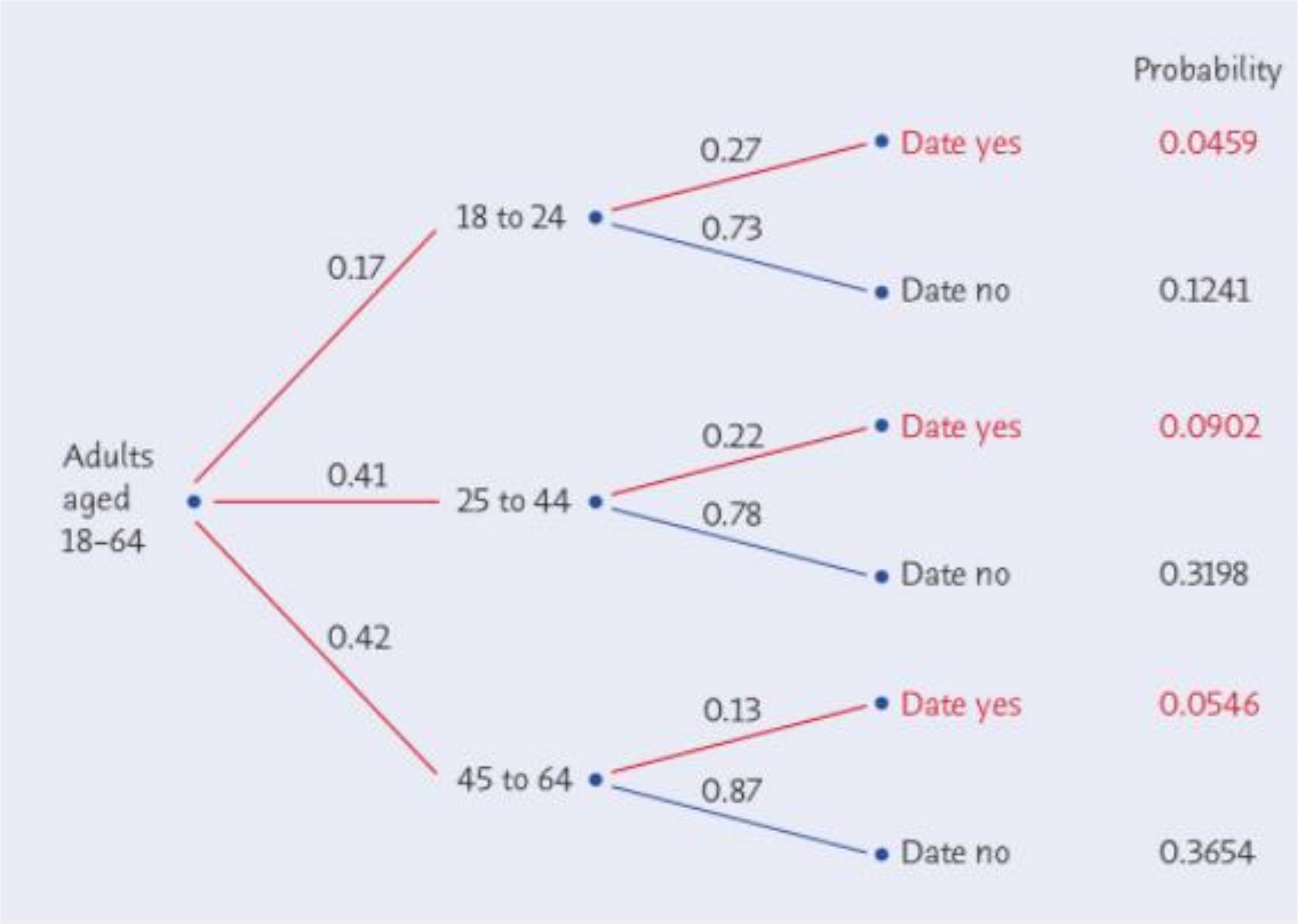
Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.



$$P(\text{18 to 24 and online dating}) = 0.0459$$
$$P(\text{25 to 44 and online dating}) = 0.0902$$
$$P(\text{45 to 64 and online dating}) = 0.0546$$

Tree diagrams (example)

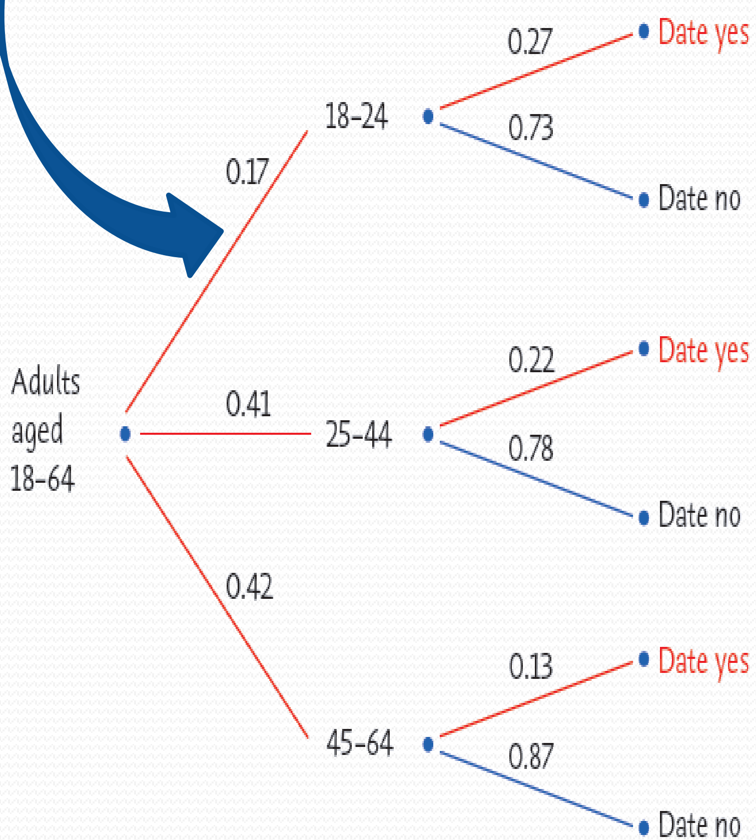
Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.



Tree diagrams (example)

Looking only at adults under 65, about 17% are 18–24 years old, another 41% are 25–44 years old, and the remaining 42% are 45–64 years old. Pew Research reports that 27% of those aged 18–24 have used online dating sites, along with 22% of those aged 25–44 and 13% of those aged 45–64.

What is the probability that an adult under 65 has used an online dating site?



Three disjoint paths in the diagram

$$
\begin{aligned}
&P(\text{online dating yes}) \\
&= P\left(\begin{array}{c}\text{18 to 24 and online dating or} \\ \text{25 to 44 and online dating or} \\ \text{45 to 64 and online dating}\end{array}\right) \\
&= (0.27)(0.17) + (0.41)(0.22) + (0.42)(0.13) \\
&= 0.0459 + 0.0902 + 0.0546 \\
&= 0.1907
\end{aligned}
$$