# Chapter 5 – Probability

## OUTLINE

1. Probability Rules
2. The Addition Rule and Complements
3. Independence and the Multiplication Rule
4. Conditional Probability and the General Multiplication Rule
5. Counting Techniques
6. Simulation
7. Putting It Together: Which Method Do I Use?

## Putting It Together

In Chapter 1, we learned the methods of collecting data. In Chapters 2 through 4, we learned how to summarize raw data using tables, graphs, and numbers. As far as the statistical process goes, we have discussed the collecting, organizing, and summarizing parts of the process.

Before we begin to analyze data, we introduce probability, which forms the basis of inferential statistics. Why? Well, we can think of the probability of an outcome as the likelihood of observing that outcome. If something has a high likelihood of happening, it has a high probability (close to 1). If something has a small chance of happening, it has a low probability (close to 0). For example, it is unlikely that we would roll five straight sixes when rolling a single die, so this result has a low probability. In fact, the probability of rolling five straight sixes is 0.0001286. If we were playing a game in which a player threw five sixes in a row with a single die, we would consider the player to be lucky (or a cheater) because it is such an unusual occurrence. Statisticians use probability in the same way. If something occurs that has a low probability, we investigate to find out “what’s up.”

## Section 5.1 Probability Rules

### Objectives

1. Understand Random Processes and the Law of Large Numbers
2. Apply the Rules of Probabilities
3. Compute and Interpret Probabilities Using the Empirical Method
4. Compute and Interpret Probabilities Using the Classical Method
5. Recognize and Interpret Subjective Probabilities
6. Objective 1: Understand Random Pro Understand Random Processes and the Law of Large Numbers

INSTRUCTOR: Do you recall the statistical process

STUDENT: Yes.

Step 1, identify the research objective.

Step 2, collect the data needed to answer

the questions posed in step 1.

Step 3, describe the data.

Step 4, perform inference.

INSTRUCTOR: Correct.

We covered steps 1and 2 in chapter 1.

Step 3 was covered in chapters 2 through 4.

In the next three chapters, we take a break

from the statistical process.

STUDENT: Why?

INSTRUCTOR: Recall in chapter 1, we

mentioned that inferential statistics uses methods

that generalize results obtained from a sample

to the population of interest and measures their reliability.

STUDENT: But how can we measure their reliability?

INSTRUCTOR: It turns out that the methods used

to generalize result from a sample to a population

are based on probability and probability models.

Probability forms the basis for inferential statistics

because it is used to measure the likelihood of observing

certain outcomes.

If an event has a high likelihood of occurring,

then it has a high probability, close to 1.

If an event has a low likelihood of occurring,

then it has a low probability, close to 0.

For example, it is unlikely that we

would roll five straight sixes when rolling a single die.

So, this result has a low probability.

In fact, the probability is 0.0001286.

If we were playing a game in which a player threw five sixes

in a row with a single die, we would consider the player

to be extremely lucky or a cheater

because it is such an unusual event assuming the die is fair.

Statisticians use probability in the same way.

If something occurs that has a low probability,

we investigate to find out what's up.

**The word "random" suggests an unpredictable result**

**or outcome.**

Predicting outcomes while facing uncertainty

is rather challenging.

For example, it would be difficult to predict

whether the outcome of flipping a fair coin

would be heads or tails for one particular flip.

However, if we flip a coin many times,

we may be able to determine the long-run proportion of times

a head is observed.

**The process of flipping a coin many times is a simulation.**

**Simulation is a technique used to recreate a random event.**

**Simulations can be tactile, as in actually physically flipping**

**a coin, or virtual, using a computer to pretend it's**

**flipping a coin.**

In both instances, the goal of the simulation

is to measure how often a certain outcome is observed,

such as a head in flipping coins.

To see this idea, we're going to simulate flipping a coin using

a statistical applet.

In this applet, the vertical axis

is going to represent the proportion of times

we observe a head, and the horizontal axis

is going to represent the number of coins that we flip.

**So in the applet, if I click One Run,**

**that's going to represent one coin flip.**

And you can see that we observe a tail.

The applet is going to keep a running total of the proportion

of heads observed, and so right now, we have 0 out of 1 heads.

I click One Run again, and this time, I observe a head.

And so now, the proportion of times that I observe a head

is 0.5, 1 out of 2.

Let me click One Run one more time, and now we get a tail.

And you can see in the Cartesian plane

that we're keeping a running total of the proportion

of times we observe a head.

So if I click Five Runs, that's going to be five coin flips.

I click Five Runs again, that's another five coin flips.

**And now you can see how we're continually**

**keeping track of the proportion of heads**

**in the Cartesian plane.**

So if I click 1,000 runs, that would

be like me flipping a coin 1,000 different times

in a random process.

Now, what you should notice is that the proportion

of times we observe a head settles down

to a specific value.

It settles down to 0.4936because 500 out of 1,013

flips of the coin resulted in heads.

If I hit Reset, I can do the same thing a second time.

In my first flip, I observe a tail.

In my second flip, I once again observe a tail,

so now the proportion of heads is 0 out of 2.

Three tails in a row--

and then I observe a head.

If I do this 1,000 times and another 1,000

and another 1,000, you can see the proportion of heads

again starts to settle down to a specific quantity.

In this case, you can see that the proportion of heads

is approaching 0.5.

So a random process represents scenarios

where the outcome of any particular trial

of an experiment is unknown, such as we don't know ahead

of time whether we're going to observe a head or a tail

when we flip a coin.

But the proportion or relative frequency

a particular outcome is observed approaches a specific value.

So if we go back to our coin-flipping applet,

you can see in the short run--

in other words, for a few flips of the coin,

we have a lot of variability in the proportion

of heads observed.

**But in the long run, the proportion of heads**

**settles down to a specific value.**

**In this case, it's approaching 0.5.**

Define simulation.

**is a technique used to recreate a random event**.**The process of flipping a coin many times is a simulation.**

**Simulation is a technique used to recreate a random event.**

**Simulations can be tactile, as in actually physically flipping**

**a coin, or virtual, using a computer to pretend it's**

**flipping a coin.**

1. Define a random process**. The word random suggests an unpredictable result or outcome. A random process represents scenarios where the outcome of any particular trial of an experiment is​ unknown, but the proportion a particular outcome is observed approaches a specific value.**

The **short run** is a few repetitions of the simulation, while the **long run** is many repetitions of the simulation. In this video, there is a lot of variability in the proportion of heads observed in the short run, while in the long run the proportion of heads approaches 0.5.

**The video you just watched illustrates a basic premise of probability.**

**Probability is the measure of the likelihood of a random phenomenon or chance behavior occurring. It deals with experiments that yield random short-term results or outcomes yet reveal long-term predictability.**

**The long-term proportion in which a certain outcome is observed is the probability of that outcome.**

**So we say that the probability of observing a head is 12 or 50% or 0.5 because as we flip the coin more times, the proportion of heads tends to 0.5. This phenomenon is referred to as the *Law of Large Numbers*.**

Objective 1, Page 6

1. Define probability. **Probability is the measure of the likelihood of a random phenomenon or chance behavior​ occurring, where the word random suggests an unpredictable result or outcome.**

State the Law of Large Numbers.

**The Law of Large Numbers**

As the number of repetitions of a probability experiment increases, the proportion with which a certain outcome is observed gets closer to the probability of the outcome.

Objective 1, Page 7

 *Answer the following after finishing Activity 1: The Law of Large Numbers.*

1. After rolling the die 1000 times, is the behavior in the short run (fewer rolls of the die) the same as it was with the first 1000 runs? Based on your results, what is the probability of rolling a 4 with a 10-sided die?

Objective 1, Page 9

 *Answer the following after watching the video.* 00:01>> The law of large numbers has intuitive feel,

that is, it seems to be common sense.

However, the law of large numbers

is often confused with a non-existent law

called the law of averages.

For example, consider a baseball announcer who says,

this player is due for a hit because he's

gone a number of at bats without getting a hit.

Or consider a mother of 4 girls, who says,

oh, I'm due for a boy on my next trial.

In both of these instances, there's

confusion between the law of large numbers

and what happens on the next trial

of a particular probability experiment.

For example, if you have four girls,

are you really more likely on the fifth child to have a boy?

Of course not, the likelihood of having a boy

is the same on the first trial, second trial,

up through the fifth trial.

Another way to think about this, the biology of having children

does not look at somebody and say, oh, you've had 4 girls,

now it's time for you to have a boy.

In other words, these trials, these random processes

have what's called a memory-less property.

Trials do not recall what has happened in the past

and use those trials to make changes in what's

going to happen in the future.

To help see that, we're going to use

StatCrunch to simulate having five children

10,000 different times.

What I'm going to do is go to Data Simulate Bernoulli.

I'm going to generate 10,000rows and five columns of data.

The Bernoulli parameter is going to be 0.5.

Basically what this means is that each column is

going to represent a child, each row

is going to represent a family.

We set the probability of success

at 0.5 because there's a 50% chance of having a girl

and a 50% chance of having a boy.

When I click Compute, all of our children are born.

So for example, in the first row you'll notice that I have all

0's.

Let's go ahead and say that 0is going to represent a boy,

and 1 is going to represent a girl.

So in this particular family, they had all boys.

As another example, let's look at row 2.

We have boy, girl, girl, girl, and girl.

Now, in our particular analysis, we're

only going to consider families where the first four

children were girls.

How can we identify them?

What I'm going to do is go to Stat, Summary Stats, Rows,

and I'm going to select the first four columns.

That's the first four children.

I'm going to find where the sum of the first four

children and then store this in the data table.

The sum is going to represent the number of girls that

are in my family because remember, 0 represents a boy,

1 represents a girl.

As I store the data in the data table,

I get a new column called row sum.

I'm going to re-title that girls for the number of girls

in the first four children.

So as we indicated, my first family

didn't have any girls, all boys.

In my second family, we had three girls, the second,

the third, and the fourth.

Again, we're only concerned with those families

where there are 4 girls on the first four children.

What we're then going to do is among those families

compute the proportion of times the fifth child was a boy

and compute the proportion of times

the fifth child was a girl.

How are we going to do that?

I'm going to go to Stat Tables Frequency.

I'm only going to look at Bernoulli 5, that's

my fifth child.

I'm only going to look at Bernoulli 5

where girls equals 4.

In other words, it's only going to compute

the proportions of the fifth child

where I had 4 girls prior.

I'm going to get a frequency and a relative frequency.

When I click Compute, what I notice

is that of these 601 times where I had 4 girls on the first four

children, 48.75% of the time the fifth child was a boy,

and 51.2% of the time the fifth child was a girl.

So it is certainly not the case that I'm

more likely to have a boy on my fifth child

if my first four children were all girls.

1. Explain the meaning of the sentence, “In a random process, the trials are memoryless**. It means essentially each trial has its own average dependent on the probability**
2. For a family whose first four children are girls, is the family more likely on the fifth child to have a boy? NO

Objective 1, Page 11

1. In probability, what is an experiment**? In probability, an experiment is any process with uncertain results that can be repeated**

Objective 1, Page 13

**DEFINITIONS CAREFUL**

The **sample space**, S, of a probability experiment is the collection of all possible outcomes for that experiment.

CAREFUL WITH THESE TWO

An **event** is any collection of outcomes from a probability experiment. An event consists of one or more outcomes. We denote events with one outcome, sometimes called *simple events,* as ei. In general, events are denoted using capital letters such as E.

Therefore, of 825,000 male 16-year-old drivers (825,000 repetitions of the experiment), the insurance company is fairly confident that about 30%, or 247,500, will have an accident. This prediction helps to establish insurance rates for any particular 16-year-old male driver.

**Example 1 *Identifying Events and the Sample Space of a Probability Experiment***

A probability experiment consists of rolling a single six-sided fair die. A fair die is one in which each possible outcome is equally likely. For example, rolling a two is just as likely as rolling a five.

1. Identify the outcomes of the probability experiment. The outcomes from rolling a single fair die are **e1=“rolling a one"={1}, e2=“rolling a two"={2}, e3=“rolling a three"={3}, e4=“rolling a four"={4}, e5=“rolling a five"={5}, and e6=“rolling a six"={6}.**
2. Define the sample space.
3. Define the event E = “roll an even number.”

#### Objective 2: Apply the Rules of Probabilities

Objective 2, Page 1

1. State Rules 1 and 2 of the rules of probabilities.
2. What is a probability model?

Objective 2, Page 2

**Example 2 *A Probability Model***

In a bag of peanut M&M milk chocolate candies, the colors of the candies can be brown, yellow, red, blue, orange, or green. Suppose that a candy is randomly selected from a bag. The table shows each color and the probability of drawing that color. Verify this is a probability model.

| **Color** | **Probability** |
| --- | --- |
| Brown | 0.12 |
| Yellow | 0.15 |
| Red | 0.12 |
| Blue | 0.23 |
| Orange | 0.23 |
| Green | 0.15 |

Objective 2, Page 4

* If an event is impossible, the probability of the event is 0.
* If an event is a certainty, the probability of the event is 1.
* The closer a probability is to 1, the more likely the event will occur.
* The closer a probability is to 0, the less likely the event will occur.
* For example, an event with probability 0.8 is more likely to occur than an event with probability 0.75.
* An event with probability 0.8 will occur about 80 times out of 100 repetitions of the experiment, whereas an event with probability 0.75 will occur about 75 times out of 100.

Objective 2, Page 6

1. What is an unusual event? What cutoff points do statisticians typically use for identifying unusual events?

Objective 2, Page 8

1. List the three methods in this section for determining the probability of an event.

#### Objective 3: Compute and Interpret Probabilities Using the Empirical Method

Objective 3, Page 1

1. Explain how to approximate probabilities using the empirical approach.

Objective 3, Page 2

**Example 3 *Using Relative Frequencies to Approximate Probabilities***

An insurance agent currently insures 182 teenage drivers (ages 16 to 19). Last year, 24 of the teenagers had to file a claim on their auto policy. Based on these results, the probability that a teenager will file a claim on his or her auto policy in a given year is

Objective 3, Page 4

**Example 4 *Building a Probability Model from a Random Process***

Pass the PigsTM is a Milton-Bradley game in which pigs are used as dice. Points are earned based on the way the pig lands. There are six possible outcomes when one pig is tossed. A class of 52 students rolled pigs 3939 times. The number of times each outcome occurred is recorded in the table.

(*Source:* [www.members.tripod.com/~passpigs/prob.html](http://www.members.tripod.com/~passpigs/prob.html))

| **Outcome** | **Frequency** |
| --- | --- |
| Side with no dot | 1344 |
| Side with dot | 1294 |
| Razorback | 767 |
| Trotter | 365 |
| Snouter | 137 |
| Leaning Jowler | 32 |

1. Use the results of the experiment to build a probability model for the way the pig lands.
2. Estimate the probability that a thrown pig lands on the “side with the dot”.
3. Would it be unusual to throw a “Leaning Jowler”?

Objective 3, Page 5

Surveys are probability experiments. Why? Each time a survey is conducted, a different random sample of individuals is selected. Therefore, the results of a survey are likely to be different each time the survey is conducted because different people are included.

#### Objective 4: Compute and Interpret Probabilities Using the Classical Method

Objective 4, Page 1

1. What requirement must be met in order to compute probabilities using the classical method?

Objective 4, Page 1(Continued)

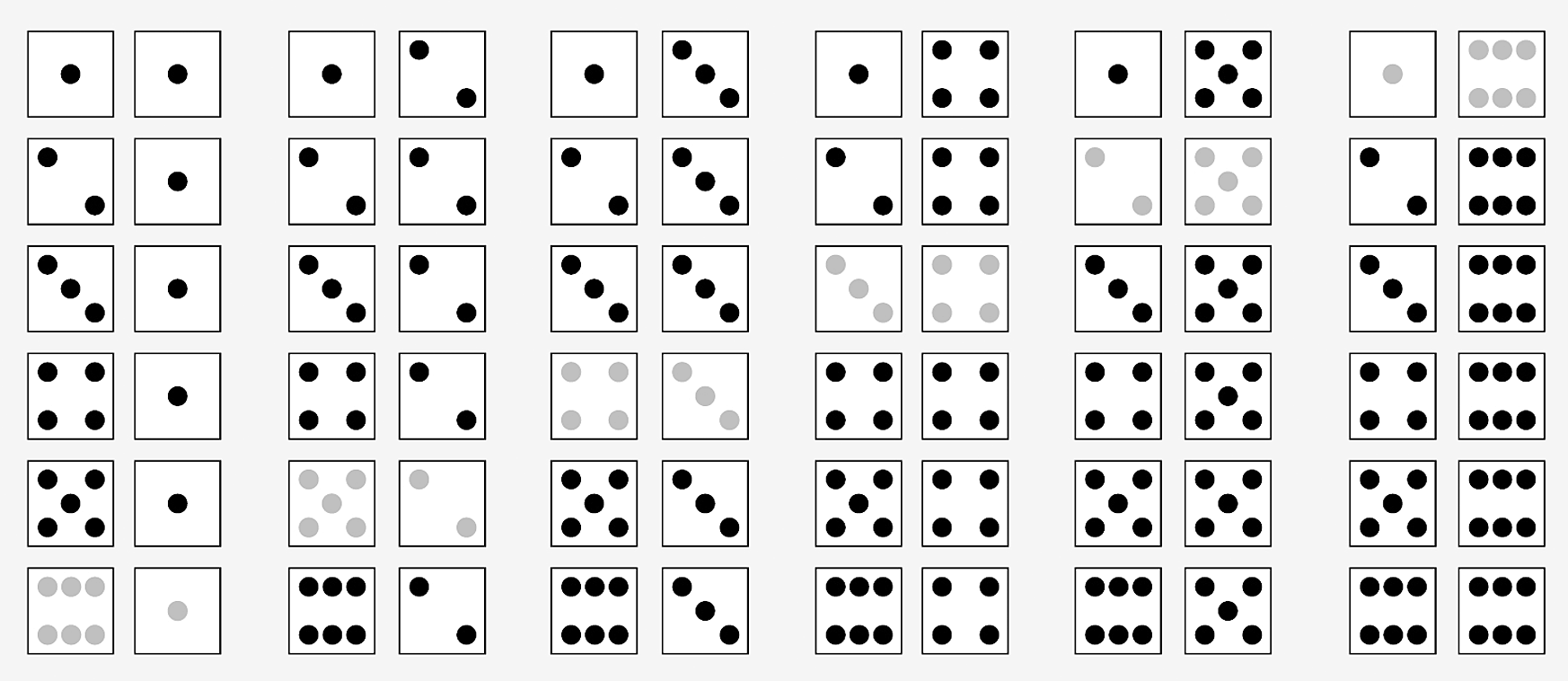
1. Explain how to compute probabilities using the classical method.

Objective 4, Page 2

**Example 5 *Computing Probabilities Using the Classical Approach***

A pair of fair dice is rolled. Fair die are die where each outcome is equally likely. The possible outcomes of this experiment are shown in Figure 1.

**Figure 1**



1. Compute the probability of rolling a seven.
2. Compute the probability of rolling "snake eyes"; that is, compute the probability of rolling a two.
3. Comment on the likelihood of rolling a seven versus rolling a two.

Objective 4, Page 4

1. As the number of trials of an experiment increase, how does the empirical probability of an event occurring compare to the classical probability of that event occurring?

Objective 4, Page 5

**Example 6 *Computing Probabilities Using Equally Likely Outcomes***

Sophia has three tickets to a concert, but Yolanda, Michael, Kevin, and Marissa all want to go to the concert with her. To be fair, Sophia wants to randomly select the two people who will go with her.

1. Determine the sample space of the experiment. In other words, list all possible simple random samples of size *n* = 2.
2. Compute the probability of the event "Michael and Kevin attend the concert."
3. Compute and interpret the probability of the event "Marissa attends the concert."

Objective 4, Page 7

**Example 7 *Comparing the Classical Method and Empirical Method***

Suppose that a survey asked 500 families with three children to disclose the gender of their children and found that 180 of the families had two boys and one girl.

1. Estimate the probability of having two boys and one girl in a three-child family, using the empirical method.
2. Compute and interpret the probability of having two boys and one girl in a three-child family, using the classical method and assuming boys and girls are equally likely.

Objective 4, Page 8

Empirical probabilities and classical probabilities often differ in value, but as the number of repetitions of a probability experiment increases, the empirical probability should get closer to the classical probability according to the Law of Large Numbers.

#### Objective 5: Recognize and Interpret Subjective Probabilities

Objective 5, Page 1

1. What is a subjective probability? Explain why subjective probabilities are used.

## Section 5.2 The Addition Rule and Complements

### Objectives

1. Use the Addition Rule for Disjoint Events
2. Use the General Addition Rule
3. Compute the Probability of an Event Using the Complement Rule

#### Objective 1: Use the Addition Rule for Disjoint Events

Objective 1, Page 1

 *Answer the following as you watch the video.*

1. What does it mean for two events to be disjoint?
2. In a Venn diagram, what does the rectangle represent? What does a circle represent?
3. How can you tell from a Venn diagram that two events are not disjoint?
4. For disjoint events *E* and *F*, how is *P*(*E* or *F*) related to *P*(*E*) and *P*(*E*)?
5. State the Addition Rule for Disjoint Events.

Objective 1, Page 3

**Example 1 *Benford’s Law and the Addition Rule for Disjoint Events***

Our number system consists of the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Because we do not write numbers such as 12 as 012, the first significant digit in any number must be 1, 2, 3, 4, 5, 6, 7, 8, or 9. Although we may think that each digit appears with equal frequency so that each digit has a  probability of being the first significant digit, this is not true. In 1881, Simon Newcomb discovered that first-digits do not occur with equal frequency. The physicist Frank Benford discovered the same result in 1938. After studying a great deal of data, he assigned probabilities of occurrence for each of the first digits, as shown in Table 2. The probability model is now known as Benford's Law and plays a major role in identifying fraudulent data on tax returns and accounting books.

**Table 2**

| Digit | 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 |

Data from The First Digit Phenomenon, T. P. Hill, American Scientist, July–August, 1998

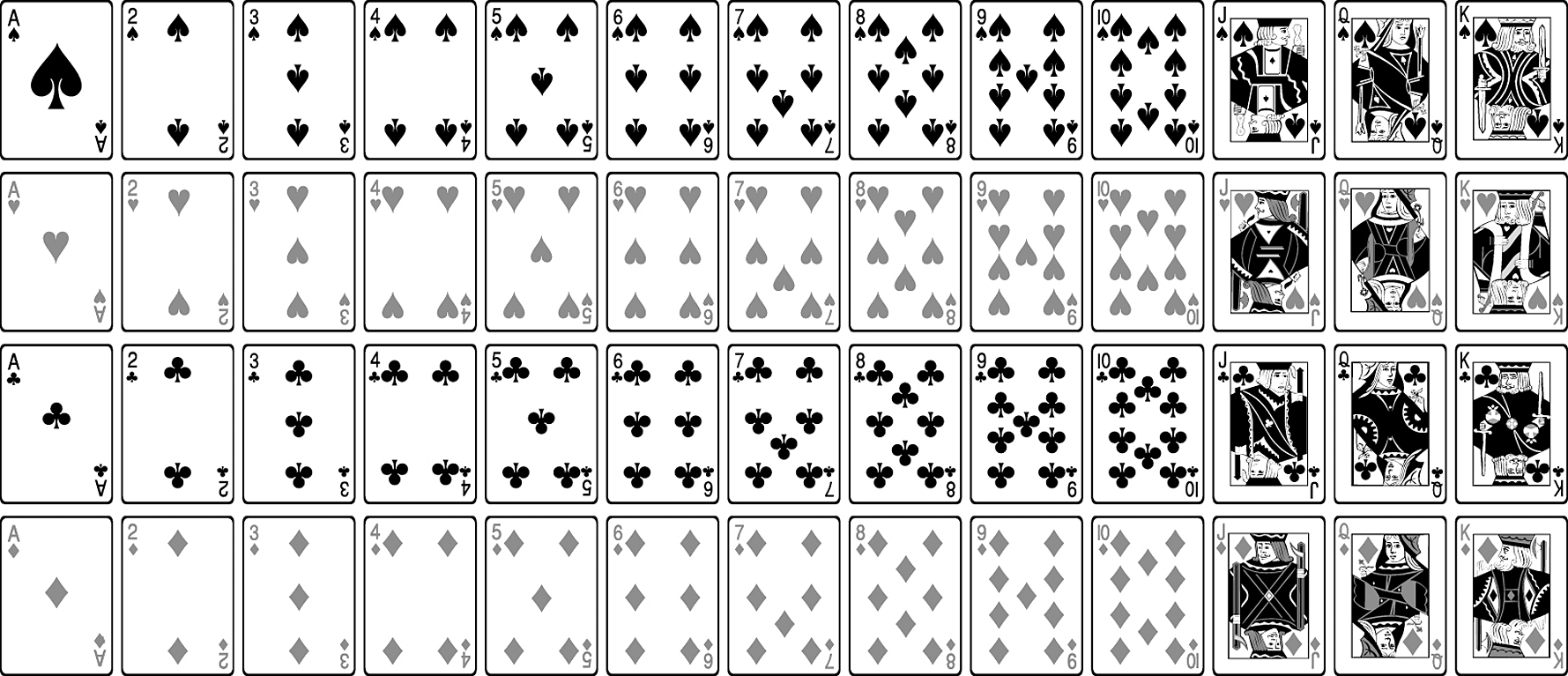
1. Verify that Benford's Law is a probability model.
2. Use Benford's Law to determine the probability that a randomly selected first digit is 1 or 2.
3. Use Benford's Law to determine the probability that a randomly selected first digit is at least 6.

Objective 1, Page 4

**Example 2 *A Deck of Cards and the Addition Rule for Disjoint Events***

Suppose that a single card is selected from a standard 52-card deck, such as the one shown in Figure 3.

**Figure 3**



1. Compute the probability of the event *E* = “drawing a king.”
2. Compute the probability of the event *E* = “drawing a king” or *F* = “drawing a queen” or *G* = “drawing a jack.”

#### Objective 2: Use the General Addition Rule

Objective 2, Page 1

1. State the General Addition Rule.
2. Explain why we subtract *P*(*E* and *F*) when using the General Addition Rule.

Objective 2, Page 3

**Example 3 *Computing Probabilities for Events That Are Not Disjoint***

Suppose a single card is selected from a standard 52-card deck. Compute the probability of the event *E* = “drawing a king” or *F* = “drawing a diamond.”

Objective 2, Page 5

A table that relates two categories of data is called a **contingency table** (or **two-way table**).

The **row variable** is the variable that describes each row in the contingency table.

The **column variable** is the variable that describes each column in the contingency table.

Objective 2, Page 6

**Example 4 *Using the Addition Rule with Contingency Tables***

Use the data in Table 3 to answer parts (A) through (D).

**Table 3**

|  | **Gender** | **Gender** |
| --- | --- | --- |
| **Marital Status** | **Males (in millions)** | **Females (in millions)** |
| **Never married** | 44.1 | 39 |
| **Married** | 66.7 | 67.5 |
| **Widowed** | 3.5 | 11.4 |
| **Divorced** | 10.7 | 14.8 |

Data from U.S. Census Bureau, Current Population Reports

1. Determine the probability that a randomly selected U.S. resident 15 years and older is male.
2. Determine the probability that a randomly selected U.S. resident 15 years and older is widowed.
3. Determine the probability that a randomly selected U.S. resident 15 years and older is widowed or divorced.
4. Determine the probability that a randomly selected U.S. resident 15 years and older is male or widowed.

#### Objective 3: Compute the Probability of an Event Using the Complement Rule

Objective 3, Page 1

1. Define the complement of an event *E*.
2. State the Complement Rule.

Objective 3, Page 2

**Example 5 *Illustrating the Complement Rule***

According to the American Veterinary Medical Association, 31.6% of American households own a dog. What is the probability that a randomly selected household does not own a dog?

Objective 3, Page 4

**Example 6 *Computing Probabilities Using Complements***

The data in Table 4 represent the travel time to work for residents of Hartford County, Connecticut.

**Table 4**

| **Travel Time** | **Frequency** |
| --- | --- |
| Less than 5 minutes | 24,358 |
| 5 to 9 minutes | 39,112 |
| 10 to 14 minutes | 62,124 |
| 15 to 19 minutes | 72,854 |
| 20 to 24 minutes | 74,386 |
| 25 to 29 minutes | 30,099 |
| 30 to 34 minutes | 45,043 |
| 35 to 39 minutes | 11,169 |
| 40 to 44 minutes | 8045 |
| 45 to 59 minutes | 15,650 |
| 60 to 89 minutes | 5451 |
| 90 or more minutes | 4895 |

Data from United States Census Bureau

1. What is the probability that a randomly selected resident has a travel time of 90 or more minutes?
2. What is the probability that a randomly selected resident of Hartford County, Connecticut will have a travel time less than 90 minutes?

## Section 5.3 Independence and the Multiplication Rule

### Objectives

1. Identify Independent Events
2. Use the Multiplication Rule for Independent Events
3. Compute At-least Probabilities

#### Objective 1: Identify Independent Events

Objective 1, Page 1

 *Answer the following as you watch the video.*

1. Define independent events and dependent events.
2. Explain why the events “draw a heart” and “roll an even number” are independent.
3. Explain why the events “woman 1 survives the year” and “woman 2 survives the year” are dependent if the two women live in the same complex.
4. When we take a very small sample from a very large finite population, we make the assumption of independence even though the events are technically dependent. State the general rule of thumb for assuming independence.

Objective 1, Page 4

1. Are disjoint events independent?

#### Objective 2: Use the Multiplication Rule for Independent Events

Objective 2, Page 1

1. State the Multiplication Rule for Independent Events.

Objective 2, Page 2

**Example 1 *Computing Probabilities of Independent Events***

In the game of roulette, the wheel has slots numbered 0, 00, and 1 through 36. A metal ball rolls around a wheel until it falls into one of the numbered slots. What is the probability that the ball will land in the slot numbered 17 two times in a row?

Objective 2, Page 3

1. State the Multiplication Rule for *n* Independent Events.

Objective 2, Page 4

**Example 2 *Life Expectancy***

The probability that a randomly selected 24-year-old male will survive the year is 0.9986 according to the National Vital Statistics Report, Vol. 56, No. 9.

1. What is the probability that three randomly selected 24-year-old males will survive the year?
2. What is the probability that twenty randomly selected 24-year-old males will survive the year?

#### Objective 3: Compute At-least Probabilities

Objective 3, Page 1

Usually, when computing probabilities involving the phrase *at least*, use the Complement Rule.

The phrase *at least* means “greater than or equal to.”

Objective 3, Page 2

**Example 3 *Computing At-least Probabilities***

The probability that a randomly selected female aged 60 years will survive the year is 0.99186 according to the National Vital Statistics Report. What is the probability that at least one of 500 randomly selected 60-year-old females will die during the course of the year?

Objective 3, Page 4

### Summary: Rules of Probability

**Rule 1** The probability of any event must be between 0 and 1, inclusive. If we let *E* denote any event, then .

**Rule 2** The sum of the probabilities of all outcomes in the sample space must equal 1. That is, if the sample space , then



**Rule 3** If *E* and *F* are disjoint events, then *P(E or F) = P(E) + P(F).* If *E* and *F* are not disjoint events, then .

**Rule 4** If *E* represents any event and  represents the complement of *E*, then .

**Rule 5** If *E* and *F* are independent events, then

**

Notice that *or* probabilities use the Addition Rule, whereas *and* probabilities use the Multiplication Rule. Accordingly, *or* probabilities imply addition, whereas *and* probabilities imply multiplication.

## Section 5.4 Conditional Probability and the General Multiplication Rule

### Objectives

1. Compute Conditional Probabilities
2. Compute Probabilities Using the General Multiplication Rule

#### Objective 1: Compute Conditional Probabilities

Objective 1, Page 1

1. What does the notation  represent?

Objective 1, Page 3

**Example 1 *An Introduction to Conditional Probability***

Suppose a single die is rolled. What is the probability that the die comes up three? Now suppose that the die is rolled a second time, but we are told the outcome will be an odd number. What is the probability that the die comes up three?

Objective 1, Page 5

1. State the Conditional Probability Rule.

Objective 1, Page 6

**Example 2 *Conditional Probabilities on Marital Status and Gender***

The data in Table 5 represent the marital status and gender of U.S. residents aged 15 years and older in 2016.

**Table 5**

|  | Males (in millions) | Females (in millions) | Totals (in millions) |
| --- | --- | --- | --- |
| Never Married | 44.1 | 39.0 | 83.1 |
| Married | 66.7 | 67.5 | 134.2 |
| Widowed | 3.5 | 11.4 | 14.9 |
| Divorced | 10.7 | 14.8 | 25.5 |
| Totals (in millions) | 125.0 | 132.7 | 257.7 |

1. Compute the probability that a randomly selected individual never married, given that the individual is male.
2. Compute the probability that a randomly selected individual is male, given that the individual never married.

Objective 1, Page 8

**Example 3 *Birth Weights of Preterm Babies***

Suppose that 12.2% of all births are preterm. (Preterm means that the gestation period of the pregnancy is less than 37 weeks.) Also, 0.2% of all births result in a preterm baby who weighs 8 pounds, 13 ounces or more. What is the probability that a randomly selected baby weighs 8 pounds, 13 ounces or more, given that the baby is preterm? Is this unusual? Data based on the Vital Statistics Reports.

#### Objective 2: Compute Probabilities Using the General Multiplication Rule

Objective 2, Page 1

1. State the General Multiplication Rule.

Objective 2, Page 2

**Example 4 *Using the General Multiplication Rule***

The probability that a driver who is speeding gets pulled over is 0.8. The probability that a driver gets a ticket, given that he or she is pulled over, is 0.9. What is the probability that a randomly selected driver who is speeding gets pulled over and gets a ticket?

Objective 2, Page 4

**Example 5 *Acceptance Sampling***

Suppose that of 100 circuits sent to a manufacturing plant, 5 are defective. The plant manager receiving the circuits randomly selects two and tests them. If both circuits work, she will accept the shipment. Otherwise, the shipment is rejected. What is the probability that the plant manager discovers at least one defective circuit and rejects the shipment?

Objective 2, Page 6

**Example 6 *Favorite Other***

In a study to determine whether preferences for self are more or less prevalent than preferences for others, researchers first asked individuals to identify the person who is most valuable and likeable to you, or favorite other.

Of the 1519 individuals surveyed, 42 had chosen themselves as their favorite other.

Source: Gebauer JE, et al. Self-Love or Other-Love? Explicit Other-Preference but Implicit Self-Preference. PLoS ONE 7(7):e41789. doi:10.1371/journal.prone.0041789

1. Suppose we randomly select 1 of the 1519 individuals surveyed. What is the probability that he or she chose themselves as their favorite other?
2. If two individuals from this group are randomly selected, what is the probability that both chose themselves as their favorite other?
3. Compute the probability of randomly selecting two individuals from this group who selected themselves as their favorite other assuming independence.

Objective 2, Page 7

If small random samples are taken from large populations without replacement, it is reasonable to assume independence of the events. As a rule of thumb, if the sample size, *n*, is less than 5% of the population size, *N*, we treat the events as independent. That is, if *n* < 0.05*N*, treat the events as independent.

Objective 2, Page 9

1. State the definition for independence using conditional probabilities.

## Section 5.5 Counting Techniques

### Objectives

1. Solve Counting Problems Using the Multiplication Rule
2. Solve Counting Problems Using Permutations
3. Solve Counting Problems Using Combinations
4. Solve Counting Problems Involving Permutations with Nondistinct Items
5. Compute Probabilities Involving Permutations and Combinations

#### Objective 1: Solve Counting Problems Using the Multiplication Rule

Objective 1, Page 2

**Example 1 *Counting the Number of Possible Meals***

The fixed-price dinner at Mabenka Restaurant provides the following choices:

Appetizer: soup or salad

Entrée: baked chicken, broiled beef patty, baby beef liver, or roast beef au jus

Dessert: ice cream or cheesecake

How many different meals can be ordered?

Objective 1, Page 3

1. State the Multiplication Rule of Counting.

Objective 1, Page 4

**Example 2 *Counting Airport Codes (Repetition Allowed)***

The International Airline Transportation Association (IATA) assigns three-letter codes to represent airport locations. For example, the code for Fort Lauderdale International Airport is FLL. How many different airport codes are possible?

Objective 1, Page 5

**Example 3 *Counting (Without Repetition)***

Three members from a 14-member committee are to be randomly selected to serve as chair, vice-chair, and secretary. The first person selected is the chair, the second is the vice-chair, and the third is the secretary. How many different committee structures are possible?

Objective 1, Page 7

1. Give the definition of *n* factorial.

Objective 1, Page 10

**Example 4 *The Traveling Salesperson***

You have just been hired as a book representative for Pearson Education.

On your first day, you must travel to seven schools to introduce yourself.

How many different routes are possible?

#### Objective 2: Solve Counting Problems Using Permutations

Objective 2, Page 1

1. State the definition of a permutation.

Objective 2, Page 2

1. State the formula for the number of permutations of *n* distinct objects taken *r* at a time.

Objective 2, Page 3

**Example 5 *Computing Permutations***

Evaluate:

1. 
2. 

Objective 2, Page 5

**Example 6 *Betting the Trifecta***

In how many ways can horses in a ten-horse race finish first, second, and third?

#### Objective 3: Solve Counting Problems Using Combinations

Objective 3, Page 1

1. State the definition of a combination.

Objective 3, Page 2

**Example 7 *Listing* Combinations**

Roger, Ken, Tom, and Jay are going to play golf. They will randomly select teams of two players each. List all possible team combinations. That is, list all the combinations of the four people Roger, Ken, Tom, and Jay taken two at a time. What is ?

Objective 3, Page 4

1. State the formula for the number of combinations of *n* distinct objects taken *r* at a time.

Objective 3, Page 5

**Example 8 *Computing Combinations***

Evaluate:

1. 
2. 
3. 

Objective 3, Page 7

**Example 9 *Simple Random Samples***

How many different simple random samples of size 4 can be obtained from a population whose size is 20?

#### Objective 4: Solve Counting Problems Involving Permutations with Nondistinct Items

Objective 4, Page 1

**Example 10 *DNA Sequence***

A DNA sequence consists of a series of letters representing a DNA strand that spells out the genetic code. There are four possible letters (A, C, G, and T), each representing a specific nucleotide base in the DNA strand (adenine, cytosine, guanine, and thymine, respectively).

How many distinguishable sequences can be formed using two As, two Cs, three Gs, and one T?

Objective 4, Page 2

1. State the formula for permutations with nondistinct items.

Objective 4, Page 3

**Example 11 *Arranging Flags***

How many different vertical arrangements are there of ten flags if five are white, three are blue, and two are red?

Objective 4, Page 5

**Summary: Combinations and Permutations**

|  | **Description** | **Formula** |
| --- | --- | --- |
| **Combination** | The selection of *r* objects from a set of *n* different objects when the order in which the objects are selected does not matter (so *AB* is the same as *BA*) and an object cannot be selected more than once (repetition is not allowed) |  |
| **Permutation of Distinct Items with Replacement** | The selection of *r* objects from a set of *n* different objects when the order in which the objects are selected matters (so *AB* is different from *BA*) and an object may be selected more than once (repetition is allowed) |  |
| **Permutation of Distinct Items without Replacement** | The selection of *r* objects from a set of *n* different objects when the order in which the objects are selected matters so (*AB* is different from *BA*) and an object cannot be selected more than once (repetition is not allowed) |  |
| **Permutation of Nondistinct Items without Replacement** | The number of ways *n* objects can be arranged (order matters) in which there are  of one kind, of a second kind, …, and  of a *k*th kind, where |  |

#### Objective 5: Compute Probabilities Involving Permutations and Combinations

Objective 5, Page 2

**Example 12 *Winning the Lottery***

In the Illinois Lottery, an urn contains balls numbered 1 to 52. From this urn, six balls are randomly chosen without replacement. For a $1 bet, a player chooses two sets of six numbers. To win, all six numbers must match those chosen from the urn. The order in which the balls are picked does not matter. What is the probability of winning the lottery?

Objective 5, Page 3

**Example 13 *Acceptance Sampling***

A shipment of 120 fasteners that contains 4 defective fasteners was sent to a manufacturing plant. The plant's quality control manager randomly selects and inspects 5 fasteners. What is the probability that exactly 1 of the inspected fasteners is defective?

## Section 5.6 Simulation

### Objective

1. Use Simulation to Obtain Probabilities

#### Objective 1: Use Simulation to Obtain Probabilities

Objective 1, Page 1

 *Answer the following while watching the video.*

1. List two historical uses of simulation.

Objective 1, Page 2

**Example 1 *Getting Out of Jail in Monopoly***

In the board game Monopoly, a player can get out of jail in one of three ways.

1. The player pays a $50 fine.
2. The player uses a “Get Out of Jail” card.
3. The player rolls doubles.

If the player does not roll doubles after three rolls, the player must pay the $50 fine. Use simulation to determine the probability that a player will not roll doubles after three consecutive rolls.

Objective 1, Page 5

When collecting data for an observational study, it is important that individuals are randomly selected to be in the study. This allows the results of the study to be extended to the population from which the individuals were randomly selected.

When collecting data for a designed experiment, it is important that the individuals are randomly assigned to the various treatment groups in the study. This allows us to make statements of causation between the levels of treatment and the response variable in the study.

Objective 1, Page 6

**Example 2 *Random Selection–Qualitative Response***

Unplugging refers to eliminating the use of social media, cell phones, and other technology. According to Harris Interactive, the proportion of adult Americans (aged 18 years or older) who attempt to “unplug” at least once a week is 0.45. There are approximately 241,000,000 Americans aged 18 years or older in the United States.

1. Simulate obtaining a simple random sample of size 500 from the population. How many of the individuals sampled unplug? How many do not unplug? What proportion unplug at least once a week?
2. Simulate obtaining a second simple random sample of size 500 from the population. How many of the individuals sampled unplug? How many do not unplug? What proportion unplug at least once a week? Why will the results of the first sample differ from those in the second sample?
3. Now simulate obtaining at least 2000 more simple random samples of size 500 from the population. Based on the simulation, what is the probability of obtaining a random sample where the proportion who unplug at least once a week is greater than 0.50? Would it be unusual to obtain a sample proportion greater than 0.5 from this population? Explain.

## Section 5.7 Putting It Together: Which Method Do I Use?

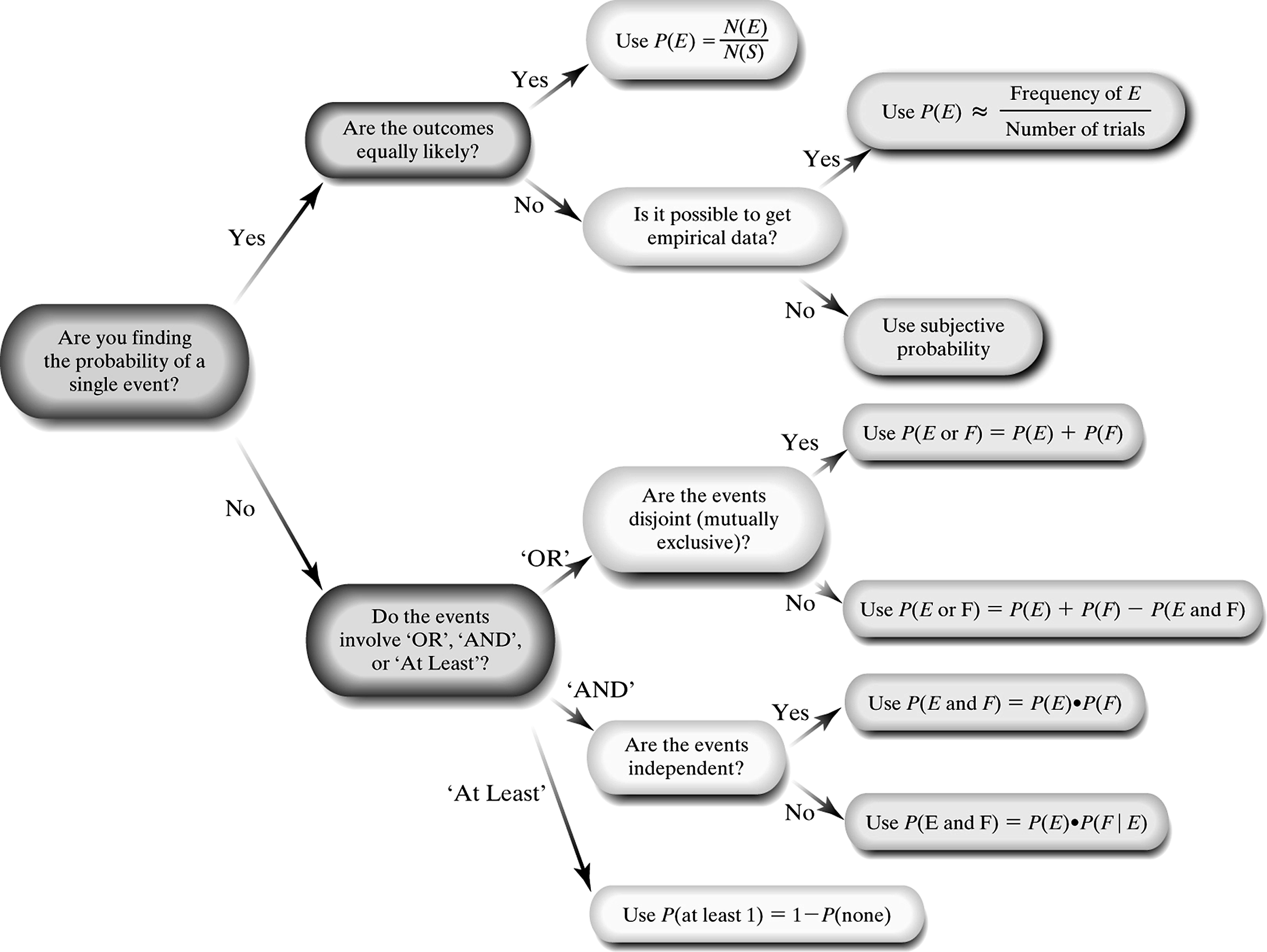
### Objectives

1. Determine the Appropriate Probability Rule to Use
2. Determine the Appropriate Counting Technique to Use

#### Objective 1: Determine the Appropriate Probability Rule to Use

Objective 1, Page 1

**Flowchart for Probability Rules**



1. What are three options when finding the probability of a single event?

OBJECTIVE 1, PAGE 1 (CONTINUED)

1. What must you determine when working with events involving the word “AND”?
2. What must you determine when working with events involving the word “OR”?

Objective 1, Page 2

**Example 1 *Probability: Which Rule Do I Use?***

In the game show Deal or No Deal?, a contestant is presented with 26 suitcases that contain amounts ranging from $0.01 to $1,000,000. The contestant must pick an initial case that is set aside as the game progresses. The amounts are randomly distributed among the suitcases prior to the game as shown in Table 7. What is the probability that the contestant picks a case worth at least $100,000?

**Table 7**

| **Prize** | **Number of Suitcases** |
| --- | --- |
| $0.01–$100 | 8 |
| $200–$1000 | 6 |
| $5000–$50,000 | 5 |
| $100,000–$1,000,000 | 7 |

Objective 1, Page 3

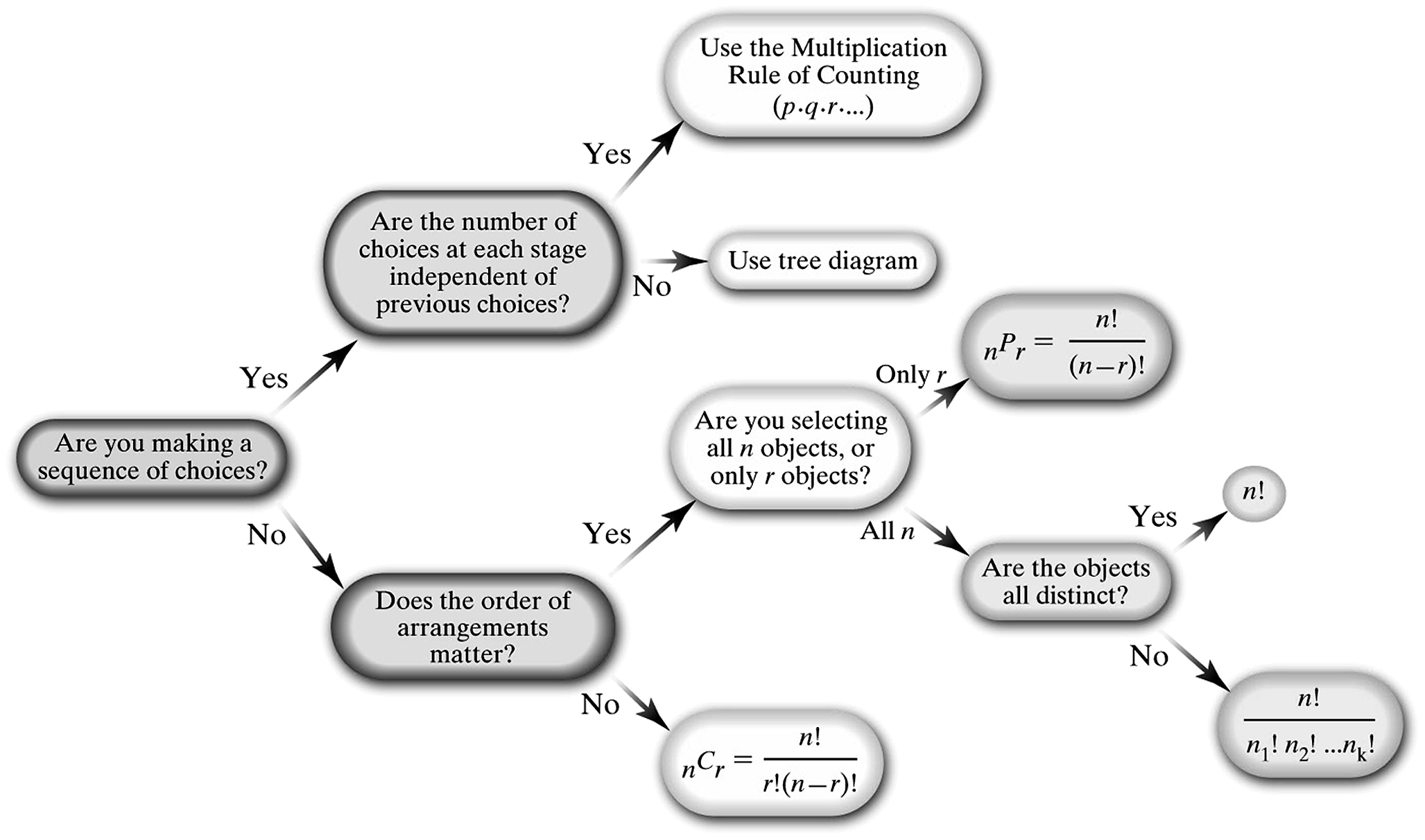
**Example 2 *Probability: Which Rule Do I Use?***

According to a Harris poll, 14% of adult Americans have one or more tattoos, 50% have pierced ears, and 65% of those with one or more tattoos also have pierced ears. What is the probability that a randomly selected adult American has one or more tattoos and pierced ears?

#### Objective 2: Determine the Appropriate Counting Technique to Use

Objective 2, Page 1

**Flowchart for Counting Techniques**



1. What counting techniques can be used when working with a sequence of choices? Explain when to use each strategy.
2. What counting techniques can be used when working with the number of arrangements of items? Explain when to use each strategy.

Objective 2, Page 2

**Example 3 *Counting: Which Technique Do I Use?***

The Hazelwood city council consists of 5 men and 4 women. How many different subcommittees can be formed that consist of 3 men and 2 women?

Objective 2, Page 3

**Example 4 *Counting: Which Technique Do I Use?***

The Daytona 500, the season opening NASCAR event, has 43 drivers in the race. In how many different ways could the top four finishers (first, second, third, and fourth place) occur?

**Chapter 6 – Discrete Probability Distributions**

**OUTLINE**

1. Discrete Random Variables
2. The Binomial Probability Distribution
3. The Poisson Probability Distribution

**Putting It Together**

Recall, the probability of an event is the long-term proportion with which the event is observed. That is, if we conduct an experiment 1000 times and observe an outcome 300 times, we estimate that the probability of the outcome is . The more times we conduct the experiment, the more accurate this empirical probability will be. This is the Law of Large Numbers. We can also use counting techniques to obtain theoretical probabilities if the outcomes in the experiment are equally likely. This is called classical probability.

A probability model lists the possible outcomes of a probability experiment and each outcome’s probability. A probability model must satisfy the rules of probability. In particular, all probabilities must be between 0 and 1, inclusive, and the sum of the probabilities must equal 1.

Now we introduce probability models for random variables. A random variable is a numerical measure of the outcome to a probability experiment. So, rather than listing specific outcomes of a probability experiment, such as heads or tails, we might list the number of heads obtained in three flips of a coin. We begin by discussing random variables and describe the distribution of discrete random variables (shape, center, and spread). We then discuss two specific discrete probability distributions: the binomial probability distribution and the Poisson probability distribution.

**Section 6.1 Discrete Random Variable**

**Objectives**

1. Distinguish between Discrete and Continuous Random Variables
2. Identify Discrete Probability Distributions
3. Graph Discrete Probability Distributions
4. Compute and Interpret the Mean of a Discrete Random Variable
5. Interpret the Mean of a Discrete Random Variable as an Expected Value
6. Compute the Standard Deviation of a Discrete Random Variable

***Objective 1: Distinguish between Discrete and Continuous Random Variables***

Objective 1, Page 1

1. Give the definition of a random variable.

Objective 1, Page 2

1. There are two types of random variables, discrete and continuous. Explain the difference between them.

Objective 1, Page 4

**Example 1 *Distinguishing between Discrete and Continuous Random Variables***

Determine whether the random variable is a discrete random variable or a continuous random variable.

1. The number of as earned in a section of statistics with 15 students enrolled
2. The number of cars that travel through a McDonald’s drive-through in the next hour
3. The speed of the next car that passes a state trooper

***Objective 2: Identify Discrete Probability Distributions***

Objective 2, Page 1

1. Give the definition of a probability distribution.

Objective 2, Page 2

**Example 2 *A Discrete Probability Distribution***

Suppose we ask a basketball player to shoot three free throws. Let the random variable *X* represent the number of shots made; so *x* = 0, 1, 2, or 3. Table 1 shows a probability distribution for the random variable *X*.

**Table 1**

| ***x*** | ***P(x)*** |
| --- | --- |
| 0 | 0.01 |
| 1 | 0.1 |
| 2 | 0.38 |
| 3 | 0.51 |

1. What does the notation *P*(*x*) represent?
2. Explain what *P*(3) = 0.51 represents.

Objective 2, Page 3

1. State the rules for a discrete probability distribution.

Objective 2, Page 5

**Example 3 *Identifying Discrete Probability Distributions***

Which of the following is a discrete probability distribution?

| A) | A) | B) | B) | C) | C) |
| --- | --- | --- | --- | --- | --- |
| *x* | *P(x)* | *x* | *P(x)* | *x* | *P(x)* |
| 0 | 0.16 | 0 | 0.16 | 0 | 0.16 |
| 1 | 0.18 | 1 | 0.18 | 1 | 0.18 |
| 2 | 0.22 | 2 | 0.22 | 2 | 0.22 |
| 3 | 0.10 | 3 | 0.10 | 3 | 0.10 |
| 4 | 0.30 | 4 | 0.30 | 4 | 0.30 |
| 5 | 0.01 | 5 |  | 5 | 0.04 |

***Objective 3: Graph Discrete Probability Distributions***

Objective 3, Page 1

1. In the graph of a discrete probability distribution, what do the horizontal axis and the vertical axis represent?
2. When graphing a discrete probability distribution, how do we emphasize that the data is discrete?

Objective 3, Page 2

**Example 4 *Graph a Discrete Probability Distribution***

Graph the discrete probability distribution given in Table 1.

**Table 1**

| ***x*** | ***P(x)*** |
| --- | --- |
| 0 | 0.01 |
| 1 | 0.10 |
| 2 | 0.38 |
| 3 | 0.51 |

Objective 3, Page 3

Graphs of discrete probability distributions help determine the shape of the distribution.

Recall that we describe distributions as skewed left, skewed right, or symmetric.

***Objective 4: Compute and Interpret the Mean of a Discrete Random Variable***

Objective 4, Page 1

 *Watch the video to learn about the derivation of the formula for finding the mean of a discrete random variable.*

Objective 4, Page 2

1. State the formula for the mean of a discrete random variable.

Objective 4, Page 3

**Example 5 *Computing the Mean of a Discrete Random Variable***

Compute the mean of the discrete probability distribution given in Table 1.

**Table 1**

| ***x*** | ***P(x)*** |
| --- | --- |
| 0 | 0.01 |
| 1 | 0.10 |
| 2 | 0.38 |
| 3 | 0.51 |

Objective 4, Page 4

 *Answer the following after watching the video.*

1. As the number of repetitions of the experiments increases, what does the mean value of the *n* trials approach?
2. As the number of repetitions of the experiments increases, what happens to the difference between the mean outcome and the mean of the probability distribution?

Objective 4, Page 5

**Example 6 *Interpretation of the Mean of a Discrete Random Variable***

The basketball player from Example 2 is asked to shoot three free throws 100 times. Compute the mean number of free throws made.

In each simulation, what value is the graph (that shows the mean number of free throws made) drawn towards?

***Objective 5: Interpret the Mean of a Discrete Random Variable as an Expected Value***

Objective 5, Page 1

Because the mean of a random variable represents what we would expect to happen in the long run, it is also called the expected value, *E*(*X*). The interpretation of the expected value is the same as the interpretation of the mean of a discrete random variable.

Objective 5, Page 2

**Example 7 *Computing the Expected Value of a Discrete Random Variable***

A term life insurance policy will pay a beneficiary a certain sum of money upon the death of the policy holder. These policies have premiums that must be paid annually. Suppose a life insurance company sells a $250,000 one-year term life insurance policy to a 49-year-old female for $530. According to the National Vital Statistics Report, Vol. 47, No. 28, the probability that the female will survive the year is 0.99791. Compute the expected value of this policy to the insurance company.

***Objective 6: Compute the Standard Deviation of a Discrete Random Variable***

Objective 6, Page 1

1. State the formula for computing the standard deviation of a discrete random variable.

Objective 6, Page 2

**Example 8 *Computing the Standard Deviation of a Discrete Random Variable***

Compute the standard deviation of the discrete random variable given in Table 1.

**Table 1**

| ***x*** | ***P(x)*** |
| --- | --- |
| 0 | 0.01 |
| 1 | 0.10 |
| 2 | 0.38 |
| 3 | 0.51 |

Objective 6, Page 4

The variance of the discrete random variable, , is the value under the square root in the computation of the standard deviation.

**Section 6.2 The Binomial Probability Distribution**

**Objectives**

1. Determine Whether a Probability Experiment is a Binomial Experiment
2. Compute Probabilities of Binomial Experiments
3. Compute the Mean and Standard Deviation of a Binomial Random Variable
4. Graph a Binomial Probability Distribution

***Objective 1: Determine Whether a Probability Experiment is a Binomial Experiment***

Objective 1, Page 1

The binomial probability distribution is a discrete probability distribution that describes probabilities for experiments in which there are two mutually exclusive (disjoint) outcomes. These two outcomes are generally referred to as success (such as making a free throw) and failure (such as missing a free throw). Experiments in which only two outcomes are possible are referred to as binomial experiments, provided that certain criteria are met.

Objective 1, Page 2

 *Answer the following as you watch the video.*

1. What are the four criteria for a binomial experiment?
2. What do *n*, *p*, and  represent when working with a binomial probability distribution?

Objective 1, Page 2 (continued)

1. If *X* is a binomial random variable that denotes the number of successes in *n* independent trials of an experiment, what are the possible values of *X*?

Objective 1, Page 3

**Example 1 *Identifying Binomial Experiments***

Determine which of the following probability experiments qualify as binomial experiments. For those that are binomial experiments, identify the number of trials, probability of success, probability of failure, and possible values of the random variable X.

1. An experiment in which a basketball player who historically makes 80% of his free throws is asked to shoot three free throws and the number of free throws made is recorded.
2. According to a recent Harris Poll, 28% of Americans state that chocolate is their favorite flavor of ice cream. Suppose a simple random sample of size 10 is obtained and the number of Americans who choose chocolate as their favorite ice cream flavor is recorded.
3. A probability experiment in which three cards are drawn from a deck without replacement and the number of aces is recorded.

***Objective 2: Compute Probabilities of Binomial Experiments***

Objective 2, Page 1

 *Watch the video to learn how the binomial probability distribution function is developed.*

Objective 2, Page 2

1. In the formula, what does  represent?
2. In the formula, what do 0.07 and 1 represent?
3. In the formula, what do 0.93 and 3 represent?

Objective 2, Page 3

1. State the Binomial Probability Distribution Function (pdf).

Objective 2, Page 4

1. Fill in the math symbol that is associated with the given phrase.

Phrase math symbol

at least or no less than or greater than or equal to

more than or greater than

fewer than or less than

no more than or at most or less than or equal to

exactly or equals or is

Objective 2, Page 5

**Example 2 *Using the Binomial Probability Distribution Function***

According to CTIA, 55% of all U.S. households are wireless-only households (no landline).

1. What is the probability of obtaining exactly ten wireless-only households based on a random sample of fifteen households?
2. What is the probability of obtaining fewer than three wireless-only households based on a random sample of fifteen households?
3. What is the probability of obtaining at least three wireless-only households based on a random sample of fifteen households?
4. What is the probability of obtaining between five and seven, inclusive, wireless-only households based on a random sample of twenty households?

***Objective 3: Compute the Mean and Standard Deviation of a Binomial Random Variable***

Objective 3, Page 1

1. State the formulas for the mean (or expected value) and standard deviation of a binomial random variable.

Objective 3, Page 2

**Example 3 *Finding the Mean and Standard Deviation of a Binomial Random Variable***

According to CTIA, 55% of all U.S. households are wireless-only households. In a simple random sample of 500 households, determine the mean and standard deviation number of wireless-only households.

***Objective 4: Graph a Binomial Probability Distribution***

Objective 4, Page 1

To graph a binomial probability distribution, first find the probabilities for each possible value of the random variable. Then follow the same approach as was used to graph discrete probability distributions.

Objective 4, Page 2

**Example 4 *Graph a Binomial Probability Distribution***

1. Graph a binomial probability distribution with n = 10 and p = 0.2. Comment on the shape of the distribution.
2. Graph a binomial probability distribution with n = 10 and p = 0.5. Comment on the shape of the distribution.
3. Graph a binomial probability distribution with n = 10 and p = 0.8. Comment on the shape of the distribution.

Objective 4, Page 3

1. What is the shape of the binomial probability distribution if *p* < 0.5, if *p* = 0.5, and if *p* > 0.5?

Objective 4, Page 5

 *Answer the following after Activity 1: The Role of n, the Number of Trials of a Binomial Experiment, on Distribution Shape*

1. As *n* increases, describe what happens to the shape of a binomial probability distribution.

Objective 4, Page 6

1. Under what conditions will a binomial probability distribution be approximately bell-shaped?
2. Explain how to determine if an observation in a binomial experiment is unusual.

Objective 4, Page 7

**Example 5 *Using the Mean, Standard Deviation, and Empirical Rule to Check for Unusual Results in a Binomial Experiment***

According to CTIA, 55% of all U.S. households are wireless-only households. In a simple random sample of 500 households, 301 were wireless-only. Is this result unusual?

**Section 6.3 The Poisson Probability Distribution**

**Objectives**

1. Determine Whether a Probability Experiment Follows a Poisson Process
2. Compute Probabilities of a Poisson Random Variable
3. Find the Mean and Standard Deviation of a Poisson Random Variable

***Objective 1: Determine Whether a Probability Experiment Follows a Poisson Process***

Objective 1, Page 1

1. For what situations in the Poisson probability distribution used?

Objective 1, Page 2

**Example 1 *Illustrating a Poisson Process***

A McDonald's® manager knows from experience that cars arrive at the drive-through at an average rate of two cars per minute between the hours of 12:00 noon and 1:00 PM. The random variable X, the number of cars that arrive between 12:20 and 12:40, follows a Poisson process.

Objective 1, Page 3

1. Under what conditions does a random variable *X* follow a Poisson process?

***Objective 2: Compute Probabilities of a Poisson Random Variable***

Objective 2, Page 1

1. State the Poisson Probability Distribution Function.

Objective 2, Page 2

**Example 2 *Computing Probabilities of a Poisson Process***

A McDonald's manager knows that cars arrive at the drive-through at the average rate of two cars per minute between the hours of 12 noon and 1:00 PM. Find the following probabilities.

1. Find the probability that exactly six cars arrive between 12 noon and 12:05 PM.
2. Find the probability that fewer than six cars arrive between 12 noon and 12:05 PM.
3. Find the probability that at least six cars arrive between 12 noon and 12:05 PM.

***Objective 3: Find the Mean and Standard Deviation of a Poisson Random Variable***

Objective 3, Page 1

1. State the formula for the mean and standard deviation of a Poisson random variable.

Objective 3, Page 2

1. Restate the Poisson probability distribution function in terms of its mean.

Objective 3, Page 3

**Example 3 *Beetles and the Poisson Distribution***

A biologist performs an experiment in which 2000 Asian beetles are allowed to roam in an enclosed area of 1000 square feet. The area is divided into 200 subsections of 5 square feet each.

1. If the beetles spread themselves evenly throughout the enclosed area, how many beetles would you expect in each subsection?
2. What is the standard deviation of X, the number of beetles in a particular subsection?
3. What is the probability of finding exactly eight beetles in a particular subsection?
4. Would it be unusual to find more than 16 beetles in a particular subsection?

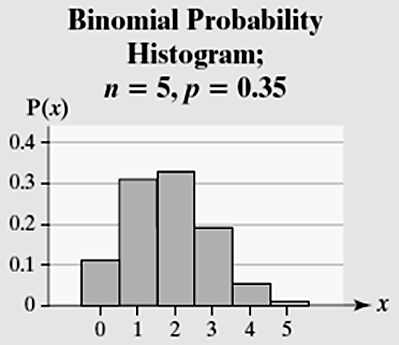
**Chapter 7 – The Normal Probability Distribution**

**OUTLINE**

1. Properties of the Normal Distribution
2. Applications of the Normal Distribution
3. Assessing Normality
4. The Normal Approximation to the Binomial Probability Distribution

**Putting It Together**

In Chapter 6, we introduced discrete probability distributions. We computed probabilities using probability distribution functions. However, we could also determine the probability of any discrete random variable from its probability histogram. For example, the figure below shows the probability histogram for the binomial random variable *X* with *n* = 5 and *p* = 0.35.



From this probability histogram, we see that  Notice that the width of each rectangle in the probability histogram is 1. Since the area of a rectangle equals height times width, we can think of *P*(1) as the area of the rectangle corresponding to *x* = 1. Thinking of probability in this way makes the transition from computing discrete probabilities to finding continuous probabilities much easier.

In this chapter, we discuss two continuous distributions: the *uniform distribution* and the *normal distribution*. Most of the discussion will focus on the normal distribution, which has many applications.

**Section 7.1 Properties of the Normal Distribution**

**Objectives**

1. Use the Uniform Probability Distribution
2. Graph a Normal Curve
3. State the Properties of the Normal Curve
4. Explain the Role of Area in the Normal Density Function

***Objective 1: Use the Uniform Probability Distribution***

Objective 1, Page 1

We discuss a uniform distribution to see the relation between area and probability.

Objective 1, Page 2

**Example 1 *The Uniform Distribution***

Assume that United Parcel Service is supposed to deliver a package to your front door and the arrival time is somewhere between 10 AM and 11 AM. Let the random variable *X* represent the time from 10 AM when the delivery is supposed to take place.

The delivery could be at 10 AM (*x* = 0) or at 11 AM (*x* = 60), with all one-minute intervals of time between *x* = 0 and *x* = 60 equally likely. That is to say, your package is just as likely to arrive between 10:15 and 10:16 as it is to arrive between 10:40 and 10:41.

The random variable *X* can be any value in the interval from 0 to 60, that is,  Because any two intervals of equal length between 0 and 60, inclusive, are equally likely, the random variable *X* is said to follow a uniform probability distribution.

Objective 1, Page 3

 *Answer the following after watching the video.*

1. What two properties must a probability density function (pdf) satisfy?

Objective 1, Page.3(continued)

1. If the possible values of a uniform density function go from 0 through *n*, what is the height of the rectangle?
2. What does the area under the graph of a probability density function over an interval represent?

***Objective 2: Graph a Normal Curve***

Objective 2, Page 1

Not all continuous random variables follow a uniform distribution.

In Figure 1, as the class width of the histogram decreases, the histogram becomes closely approximated by the smooth red curve. For this reason, we can use the curve to *model* the probability distribution of this continuous random variable.

Objective 2, Page 2

1. What does it mean to say that a continuous random variable is normally distributed?

Objective 2, Page 3

*Answer the following while watching the video.*

1. What value of *x* is associated with the peak of a normal curve?
2. What values of *x* are associated with the inflection points of a normal curve?

Objective 2, Page 4

1. Sketch and label the graph from Figure 2.

Objective 2, Page 5

 *Answer the following after Activity 1: The Role of  and  in a Normal Curve.*

1. What happens to the graph as the mean increases? What happens to the graph as the mean decreases?
2. What happens to the graph as the standard deviation increases? What happens to the graph as the standard deviation decreases?

***Objective 3: State the Properties of the Normal Curve***

Objective 3, Page 1

1. State the seven properties of the normal curve.

***Objective 4: Explain the Role of Area in the Normal Density Function***

Objective 4, Page 1

 *Watch the video to see an example of a normally distributed random variable*.

The area under the normal curve can be used to model the probability histogram and the actual proportion in a given interval.

Objective 4, Page 2

1. Suppose that a random variable *X* is normally distributed with mean  and standard deviation . Give two representations for the area under the normal curve for any interval of values of the random variable *X*.

**Example 2 *Interpreting the Area Under a Normal Curve***

The serum total cholesterol for males 20 to 29 years old is approximately normally distributed with mean  and standard deviation , based on data obtained from the National Health and Nutrition Examination Survey.

1. Draw a normal curve with the parameters labeled.
2. An individual with total cholesterol greater than 200 is considered to have high cholesterol. Shade the region under the normal curve to the right of *x* = 200.
3. Suppose that the area under the normal curve to the right of *x* = 200 is 0.2903. (You will learn how to find this area in the next section.) Provide two interpretations of this result.

**Section 7.2 Applications of the Normal Distribution**

**Objectives**

1. Find and Interpret the Area under a Normal Curve
2. Find the Value of a Normal Random Variable

***Objective 1: Find and Interpret the Area under a Normal Curve***

Objective 1, Page 1

1. Suppose that the random variable *X* is normally distributed with mean  and standard deviation . Explain the distribution of the random variable . What is the name for the random variable *Z*?

Objective 1, Page 2

1. Explain how to find the area to the left of *x* for a normally distributed random variable *X*, using Table V.

Objective 1, Page 3

 *Answer the following after watching the video.*

1. Explain how to find the area to the right of *x* for a normally distributed random variable *X*, using Table V.

Objective 1, Page 5

**Example 1 *Finding and Interpreting Area Under a Normal Curve***

A pediatrician obtains the heights of her 200 three-year-old female patients. The heights are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Use the normal model to determine the proportion of the 3-year-old females who have a height less than 35 inches.

Objective 1, Page 6

Note that the proportion of 3-year-old females who are shorter than 35 inches according to the normal model is close to the actual results. The normal curve accurately models the heights.

Because the area under the normal curve represents a proportion, we can also use the area to find percentile ranks of scores.

Objective 1, Page 7

**Example 2 *Finding and Interpreting Area Under a Normal Curve***

A pediatrician obtains the heights of her 200 three-year-old female patients. The heights are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Use the normal model to determine the probability that a randomly selected 3-year-old female is between 35 and 40 inches tall, inclusive. That is, find 

Objective 1, Page 10

 *Answer the following after watching the video.*

1. Summarize the methods for finding the area to the left of *x*, the area to the right of *x*, and the area between  and .

**Area to the Left of *x***

**Area to the Right of *x***

**Area Between  and **

***Objective 2: Find the Value of a Normal Random Variable***

Objective 2, Page 1

Often, we do not want to find the proportion, probability, or percentile given a value of a normal random variable. Rather, we want to find the value of a normal random variable that corresponds to a certain proportion, probability, or percentile. For example, we might want to know the height of a 3-year-old girl who is at the 20th percentile. Or we might want to know the scores on a standardized exam that separate the middle 90% of scores from the bottom and top 5%.

Objective 2, Page 2

**Example 3 *Finding the Value of a Normal Random Variable***

The heights of a pediatrician's 3-year-old female patients are approximately normally distributed, with mean 38.72 inches and standard deviation 3.17 inches. Find the height of a 3-year-old female at the 20th percentile.

Objective 2, Page 4

**Example 4 *Finding the Value of a Normal Random Variable***

The scores earned on the mathematics portion of the SAT, a college entrance exam, are approximately normally distributed with mean 516 and standard deviation 116. What scores separate the middle 90% of test takers from the bottom and top 5%? In other words, find the 5th and 95th percentiles.

Data from The College Board

Objective 2, Page 6

1. What does the notation  represent?

Objective 2, Page 7

**Example 5 *Finding the Value of ***

Find the value of 

Objective 2, Page 9

1. For any continuous random variable, what is the probability of observing a specific value of the random variable?

Since the probability of observing a specific value of a continuous random variable is 0, the following probabilities are equivalent.



**Section 7.3 Assessing Normality**

**Objective**

1. Use Normal Probability Plots to Assess Normality

***Objective 1: Use Normal Probability Plots to Assess Normality***

Objective 1, Page 1

1. What is a normal score?
2. What is a normal probability plot?
3. List the four steps for drawing a normal probability plot by hand.

Objective 1, Page 4

The idea behind finding the expected z-score is, if the data come from a normally distributed population, we could predict the area to the left of each data value.

Objective 1, Page 5

1. If sample data are taken from a population that is normally distributed, how will the normal probability plot appear?

Objective 1, Page 5(continued)

1. Explain how to determine if a normal probability plot is “linear enough”.

Objective 1, Page 6

**Example 1 *Drawing a Normal Probability Plot by Hand***

The data in Table 2 represent the finishing time (in seconds) for six randomly selected races of a greyhound named Barbies Bomber in the 5/16-mile race at Greyhound Park in Dubuque, Iowa. Is there evidence to support the belief that the variable “finishing time” is normally distributed?

**Table 2**

|  |  |
| --- | --- |
| 31.35 | 32.52 |
| 32.06 | 31.26 |
| 31.91 | 32.37 |

Data from Greyhound Park, Dubuque, IA

Objective 1, Page 8

Typically, normal probability plots are drawn using either a graphing calculator with advanced statistical features or statistical software such as StatCrunch.

Objective 1, Page 9

**Example 2 *Drawing a Normal Probability Plot Using Technology***

Draw a normal probability plot of the Barbies Bomber data in Table 2 using technology. Is there evidence to support the belief that the variable “finishing time” is normally distributed?

**Table 2**

|  |  |
| --- | --- |
| 31.35 | 32.52 |
| 32.06 | 31.26 |
| 31.91 | 32.37 |

Data from Greyhound Park, Dubuque, IA

Objective 1, Page 10

**Example 3 *Assessing Normality***

The data in Table 4 represent the time 100 randomly selected riders spent waiting in line (in minutes) for the Demon Roller Coaster. Is the random variable “wait time” normally distributed?

**Table 4**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 7 | 3 | 5 | 107 | 8 | 37 | 16 | 41 | 7 | 25 | 22 | 19 | 1 | 40 | 1 | 29 | 93 |
| 33 | 76 | 14 | 8 | 9 | 45 | 15 | 81 | 94 | 10 | 115 | 18 | 0 | 18 | 11 | 60 | 34 |
| 30 | 6 | 21 | 0 | 86 | 6 | 11 | 1 | 1 | 3 | 9 | 79 | 41 | 2 | 9 | 6 | 19 |
| 4 | 3 | 2 | 7 | 18 | 0 | 93 | 68 | 6 | 94 | 16 | 13 | 24 | 6 | 12 | 121 | 30 |
| 35 | 39 | 9 | 15 | 53 | 9 | 47 | 5 | 55 | 64 | 51 | 80 | 26 | 24 | 12 | 0 |  |
| 94 | 18 | 4 | 61 | 38 | 38 | 21 | 61 | 9 | 80 | 18 | 21 | 8 | 14 | 47 | 56 |  |

**Section 7.4 The Normal Approximation to the Binomial Probability Distribution**

**Objective**

1. Approximate Binomial Probabilities Using the Normal Distribution

***Objective 1: Approximate Binomial Probabilities Using the Normal Distribution***

Introduction, Page 1

 *Answer the following after watching the video.*

1. What are the three criteria for a binomial probability experiment?
2. Under what conditions will a binomial random variable be approximately normally distributed?

Objective 1, Page 1

1. If the binomial random variable *X* is approximately distributed, state the formulas for its mean and standard deviation.

Objective 1, Page 2

1. If *n* = 40 and *p* = 0.5, we can use a normal model because . Compute  and .

Objective 1, Page 4

To approximate the probability of a specific value of the binomial random variable, such as P(18), find the area under the normal curve from x = 17.5 to x = 18.5. We add and subtract 0.5 from x = 18 as a correction for continuity because we are using a continuous density function to approximate a discrete probability.

Objective 1, Page 5

 *Watch the video that summarizes the various corrections for continuity.*

Objective 1, Page 6

1. What is the continuity correction in each of the following cases?
2. *P*(*X* = *a*)
3. 
4. 
5. 

Objective 1, Page 9

**Example 1 *The Normal Approximation to a Binomial Random Variable***

According to the American Red Cross, 7% of people in the United States have blood type O-negative. What is the probability that in a simple random sample of 500 people in the United States fewer than 30 have blood type O-negative?

Objective 1, Page 10

Note that the approximate result using the normal model is only off by 0.0007 from the exact probability computed using technology. Also, notice the shape of the distribution in the StatCrunch output.

Objective 1, Page 12

**Example 2 *A Normal Approximation to the Binomial***

According to the Gallup Organization, 65% of adult Americans are in favor of the death penalty for individuals convicted of murder. Erica selects a random sample of 1000 adult Americans in Will County, Illinois, and finds that 630 of them are in favor of the death penalty for individuals convicted of murder.

* 1. According to the Gallup Organization, 65% of adult Americans are in favor of the death penalty for individuals convicted of murder. Erica selects a random sample of 1000 adult Americans in Will County, Illinois, and finds that 630 of them are in favor of the death penalty for individuals convicted of murder.

Does the result from part (A) contradict the Gallup Organization's findings? Explain

**Chapter 8 – Sampling Distributions**

**OUTLINE**

1. Distribution of the Sample Mean
2. Distribution of the Sample Proportion

**Putting It Together**

In chapters 6 and 7, we learned about random variables and their probability distributions.

In this chapter, we continue our discussion of probability distributions where statistics, such as , will be the random variable. Statistics are random variables because the value of a statistic varies from sample to sample. For this reason, statistics have probability distributions associated with them. For example, there is a probability distribution for the sample mean, sample proportion, and so on. We use probability distributions to make probability statements regarding the statistic. In this chapter, we examine the shape, center, and spread of statistics such as .

**Section 8.1 Distribution of the Sample Mean**

**Objectives**

1. Describe the Distribution of the Sample Mean: Normal Population
2. Describe the Distribution of the Sample Mean: Non-normal Population

Introduction, Page 1

 *Watch the video for an overview of where we have been and where we are going in the course.*

In Chapters 1 through 4 we learned how to identify the research objective (Step 1 of the statistical process) as well as collect (Step 2) and describe data (Step 3). In Chapters 5 through 7 we developed the skills that allow us to perform inference (Step 4). Because it is difficult to gain access to populations, the data found in Step 2 is often from a sample. Sample data are used to make inferences about the population. For example, we might compute the mean of a sample and use this information to draw conclusions regarding the population mean. The rest of this course focuses on how sample data are used to draw conclusions about populations.

Introduction, Page 2

 *Watch the video for an overview of the material presented in this chapter.*

A random variable is a numerical measure of the outcome of a probability experiment. Statistics such as the sample mean, , are random variables. Statistics are random variables because the value of a statistic varies from sample to sample. For this reason, statistics have probability distributions associated with them. For example, there is a probability distribution for the sample mean and the sample proportion.

Introduction, Page 3

The sample mean will vary from sample to sample. Our goal in this section is to describe the distribution of the sample mean. Remember, when we describe a distribution, we do so in terms of its shape, center, and spread.

Introduction, Page 4

1. What is the sampling distribution of a statistic?

Introduction, Page 4 (Continued)

1. What is the sampling distribution of the sample mean ?
2. List the three steps for determining the sampling distribution of the sample mean.

Once a particular sample is obtained, it cannot be obtained a second time.

***Objective 1: Describe the Distribution of the Sample Mean: Normal Population***

Objective 1, Page 1

*Answer the following after watching the video.*

1. Describe the shape of the distribution of the sample mean as the sample size increases.

Objective 1, Page 1 (Continued)

1. What does the mean of the distribution of the sample mean, *,* equal?
2. As the sample size *n* increases, what happens to the standard deviation of the distribution of the sample mean?

Objective 1, Page 2

1. List the formulas for the mean and standard deviation of the sampling distribution of .
2. What is the standard error of the mean?

Objective 1, Page 3

Note, in both simulations, the standard error of the mean was close to the approximate standard error.

Objective 1, Page 4

1. Describe the shape of the sampling distribution of  if the random variable *X* is normally distributed.

Objective 1, Page 7

**Example 1 *Finding Probabilities of a Sample Mean***

The IQ, *X*, of humans is approximately normally distributed with mean  and standard deviation . Compute the probability that a simple random sample of size *n* = 10 results in a sample mean greater than 110. That is, compute .

***Objective 2: Describe the Distribution of the Sample Mean: Non-normal Population***

Objective 2, Page 1

 *Answer the following after Activity 1: Sampling Distribution of the Sample Mean: Non-normal Population*

1. As the sample size increases, describe the effect on the center and spread of the distribution.

Objective 2, Page 2

 Watch the video to help reinforce the concepts from Activity 1.

Objective 2, Page 3

1. What is the mean of the sampling distribution of the sample mean equal to? What is the standard deviation of the sampling distribution of the sample mean equal to?

Objective 2, Page 3 (continued)

1. What happens to the shape of the sampling distribution of the sample mean as the sample size increases?
2. State the Central Limit Theorem.

Objective 2, Page 4

How large does the sample size need to be before we can say that the sampling distribution of  is approximately normal? The answer depends on the shape of the distribution of the underlying population. Distributions that are highly skewed will require a larger sample size for the distribution of  to become approximately normal.

Objective 2, Page 5

Notice that even for a highly skewed population of household incomes for a town, the distribution of the sample mean is approximately normal for *n* = 25.

Objective 2, Page 6

1. State the rule of thumb for invoking the Central Limit Theorem.

Objective 2, Page 9

**Example 2 *Weight Gain during Pregnancy***

The mean weight gain during pregnancy is 30 pounds, with a standard deviation of 12.9 pounds. Weight gain during pregnancy is skewed right. An obstetrician obtains a random sample of 35 low-income patients and determines that their mean weight gain during pregnancy was 36.2 pounds. Does this result suggest anything unusual?

Objective 2, Page 11

 *Watch the video for a summary of the shape, center, and spread of the distribution of the sample mean for both normal and non-normal populations.*

**Section 8.2 Distribution of the Sample Proportion**

**Objectives**

1. Describe the Sampling Distribution of a Sample Proportion
2. Compute Probabilities of a Sample Proportion

***Objective 1: Describe the Sampling Distribution of a Sample Proportion***

Objective 1, Page 1

1. Define the sample proportion, .

Objective 1, Page 2

**Example 1 *Computing a Sample Proportion***

The Harris Poll conducted a survey of 1200 adult Americans who vacation during the summer and asked whether the individuals planned to work while on summer vacation. Of those surveyed, 552 indicated that they planned to work while on vacation. Find the sample proportion of individuals surveyed who planned to work while on summer vacation.

Objective 1, Page 5

Because the value of the sample proportion, , varies from sample to sample, it is a random variable and has a probability distribution.

Objective 1, Page 6

 *Answer the following after watching the video.*

1. As the sample size increases, describe what happens to the shape of the sampling distribution of the sample proportion.
2. What does the mean of the sampling distribution of the sample proportion equal?
3. As the sample size increases, describe what happens to the standard deviation of the sampling distribution of the sample proportion.

Objective 1, Page 7

 Answer the following after Activity 1: Sampling Distribution of the Sample Proportion.

1. What is the mean of the distribution in all three cases?
2. What role does sample size play in the standard deviation?

Objective 1, Page 7 (Continued)

1. What role does sample size play in the shape of the sampling distribution of ?

Objective 1, Page 8

1. Under what conditions is the shape of the sampling distribution of  approximately normal?
2. State the formulas for the mean and standard deviation of the sampling distribution of .

The sample size, *n*, can be no more than 5% of the population size, *N*. That is, 

Objective 1, Page 10

**Example 2 *Describing the Sampling Distribution of the Sample Proportion***

Based on a study conducted by the Gallup Organization, 77% of Americans believe that the state of moral values in the United States is getting worse. Suppose we obtain a simple random sample of n = 60 Americans and determine which of them believe that the state of moral values in the United States is getting worse. Describe the sampling distribution of the sample proportion for Americans with this belief.

***Objective 2: Compute Probabilities of a Sample Proportion***

Objective 2, Page 1

Now that we know how to describe the sampling distribution of the sample proportion, we can compute probabilities involving sample proportions.

Objective 2, Page 2

**Example 3 *Computing Probabilities of a Sample Proportion***

According to the National Center for Health Statistics, 15% of all Americans have hearing trouble.

1. In a random sample of 120 Americans, what is the probability that at most 12% have hearing trouble?
2. Suppose that a random sample of 120 Americans who regularly listen to music using headphones results in 26 having hearing trouble. What might you conclude?