High School

Statistics
For Credit
Math Packet

General Rubric for All Standards

4	**Advanced**	*Independently makes accurate, in-depth inferences and applies content/skills to complex, open-ended problems and tasks.*
3	**Proficient**	*Independently uses complex grade-level content/skills with no major errors or omissions.*
2	**Basic**	*Independently uses basic grade-level content/skills with no major errors or omissions. However, major errors or omissions on more complex content/skills*
1	**Emerging**	*Major errors or omissions on basic and more complex grade-level content/skills.*

Packet Contents:

Probability Enrichment Packet for Quarter 3

Learning Targets

- Interpret probability as a long-run relative frequency
- Use simulation to model chance behavior
- Give a probability model for a chance process with equally likely outcomes and use it to find the probability of an event.
- Use basic probability rules, including the complement rule and the addition rule for mutually exclusive events
- Use a two-way table or Venn diagram to model a chance process and calculate probabilities involving two events.
- Apply the general addition rule to calculate probabilities.
- Calculate and interpret conditional probabilities
- Determine if two events are independent
- Use the general multiplication rule to calculate probabilities
- Use a tree diagram to model a chance process involving a sequence of outcomes and to calculate probabilities
- When appropriate, use the multiplication rule for independent events to calculate probabilities.

Topic 1: Randomness, Probability and Simulation

The outcome of something like a coin toss is unpredictable in the short run but has a regular and predictable pattern in the long run.

The **Probability** of something is unpredictable in the short term, but in the long term (after many, many repetitions) it is the proportion of times an outcome of interest occurs. It is always a number between 0 and 1.

The **Law of Large Numbers** says that as we increase repetitions (think flipping a coin more and more times) the proportion of times the successful outcome occurs gets closer and closer to it's probability. (As we flip a coin more and more times, the proportion of heads that occurs gets closer and closer to 50%)

Example: Holy Cow

The Chicago Cubs play their home games at Wrigley Field, located in the Lakeview neighborhood of Chicago. A recent *New York Times* study concluded that the probability that a randomly selected Lakeview resident is a Cubs fan is 0.44.

a) Interpret this probability as a long-run relative frequency.

If you ask many, many Lakeview residents, about 44% of them will be Cubs fans.

b) If a researcher randomly selects 100 Lakeview residents, will exactly 44 of them be Cubs fans? Explain your answer.

Probably not. 100 is still a small number of samples in probability terms, so the number of Cubs fans should be close to 44, but not exactly.

While we think we have good intuition about probability, we really don't. So we need ways to pretend to do many, many trials of something and find the proportion of times the thing we want to know the probability of occurs.

One way to estimate probability is through **simulation.** Let's say you know the probability that someone can ride a bike in your town is 50% and want to simulate asking 1,000 people if they can ride a bike.

Step 1: Find something that has a 50/50 chance of occurring. For this example we'll use a coin.

Step 2: Decide which side will be "can ride a bike" and which side will be "can't ride a bike". Let's say heads will be "can ride a bike".

Step 3: Perform your 1,000 coin flips and record how many heads you get.

Step 4: These results are the answer to your question.

Example: The birthday problem

In a certain Statistics class of 24 students, two of the students discovered they share the same birthday. Surprised by these results, the students decide to perform a simulation to estimate the probability that a class of 24 students has at least two students with the same birthday.

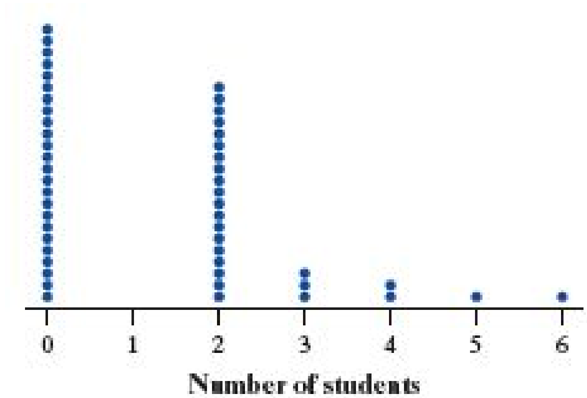
a) Assume that birthdays are randomly distributed throughout the year (and ignore leap years). Describe how you would use a random number generator to carry out this simulation.

   Label each day of the year with a number from 1 to 365.

   Generate 24 random numbers and match those numbers up with the day of the year from your list.

   Record the number of students in the class with the same birthday.

The dotplot shows the number of students who share a birthday with another student in a class of 24 students in 50 trials. There may be multiple sets of matching birthdays in each simulated class.



b) Explain what the dot at 5 represents.

   One trial in which 5 students had the same birthday.

Example: Names in a hat

Suppose I want to choose a simple random sample of size 6 from a group of 60 seniors and 30 juniors. I want to be sure that my sample has 4 seniors and 2 juniors. To do this, I write each person's name on a slip of equal-size paper and mix the slips in a large hat. I will select names one at a time from the bag until I get 4 seniors and 2 juniors (i.e., after 4 seniors are selected, the remaining seniors will be discarded, and after 2 juniors are selected, the remaining juniors will be discarded).

It takes 14 name selections to satisfy the requirement of getting 4 seniors and 2 juniors. Does this outcome provide convincing evidence that the names were not properly shuffled? To help answer this question, we want to perform a simulation to estimate the probability that it will take 14 or more selections to get 4 seniors and 2 juniors.

a) Describe how to use a table of random digits to perform one trial of the simulation.

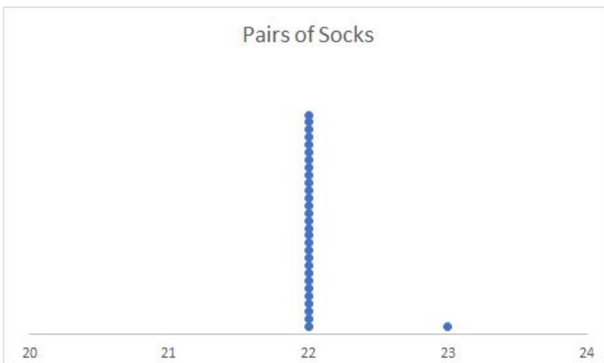
   Label the 60 seniors 01 through 60, and the 30 juniors 61 through 90. Look at pairs of digits from a random number table, ignoring repeats. Look at these digits until you have 4 different labels from 01 to 60 and 2 different labels from 61 to 90. Then count how many pairs of digits you looked at.

b) In 50 trials of the simulation, there was 1 instance for which the number of name selections was 14 or greater. Do these results give convincing evidence that the names were not properly shuffled?

   Yes. There's about a 1/50 = 0.02 = 2% chance that it will take 14 or more selections if names are properly shuffled. Because this probability is small, we have convincing evidence that the names were not properly shuffled.

Topic 1 - Test Your understanding

1. Insurance companies create and use mortality tables based on historical data to determine the price paid for insurance policies. Based on mortality tables, a 20 year-old American female living to age 21 is 0.999.
   a. Interpret this probability as a long run relative frequency
   b. If a researcher randomly selects 1,000 American 20 year-old females, will exactly 999 of them live to age 21?
2. Zoe has 25 pairs of socks, each pair is a different color and / or pattern. Last time she did laundry, she lost two socks (unfortunately, not a pair) and now has two unmatched socks. She wonders what the probability is that if she loses another sock, she will still have 23 pairs of socks.
   a. Assume the third sock is lost at random, describe how you would use a random number generator to carry out the simulation.
   b. The dot plot below shows the number of pairs of socks in 30 trials. What does the dot at 23 represent?



3. The director of Residence Life at a local university wants to evaluate the satisfaction level of the 816 students (481 females and 335 males) in the new single room dormitory. He writes the room number of each student on identical slips of paper, then draws 50 rooms to send the survey. He surveys 10 female and 40 male students. We wonder if he shuffled the papers sufficiently. To help answer this we want to perform a simulation to estimate the probability of randomly selecting 10 or fewer females from the dormitory.
   a. Describe how to use a table of random digits to perform one trial of the simulation.
   b. In 500 trials of the simulation, there were 3 instances in which the number of females was 10 or less. Do these results give convincing evidence that the papers were not properly suffled?

Topic 2: Probability Rules

A **probability model** is when you list all possible outcomes for something along with the probability of each outcome happening.

The **sample space** is the list of all possible outcomes.

An **event** is any collection of outcomes.

Finding probabilities for equally likely outcomes (when each outcome has the same chance of happening):

$$P(A) = \frac{\textit{number of outcomes in event A}}{\textit{total number of outcomes in sample space}}$$

Example: rock, paper, scissors

There is a website where humans can play paper, scissors, rock with a computer. Irresistibly drawn to it, you play the game 2 times. Assume that the computer is randomly choosing its moves for both games.

a) Give a probability model for the computer's chance process.

   PP, PS, PR, SP, SS, SR, RP, RS, RR

   Because the computer is randomly choosing its moves for both games, the outcomes will be equally likely and the probability of each is 1/9.

b) Define event A as the computer chooses the same move for both games. Find P(A).

   P(A) means "the probability of A". So this problem is asking us to find the probability that the computer chooses PP, SS or RR.

   There are 3 of those outcomes so P(A) = 3/9 = 0.333

Basic Probability Rules

1) If all outcomes in the sample space are equally likely, then the probability that event A occurs is

$$P(A)= \frac{\textit{number of outcomes in event A}}{\textit{total number of outcomes in sample space}}$$

2) The probability of any event is always a number between 0 and 1. (or 0% and 100%)
3) All possible outcomes together must have probabilities that sum to 1 (or 100%)
4) The probability that an event does *not* occur is 1 minus the probability that it does occur. (This is called the complement rule)

The **Complement rule**

A^C is the complement of event A. That is, that A does not occur.

(If there is a 40% chance of rain, what are the chances of no rain?)

$$P(A^C) = 1 - P(A)$$

Mutually Exclusive events

Two events are mutually exclusive if they have no outcomes in common, so they can never occur together.

Addition rule for mutually exclusive events

P(A or B)=P(A) + P(B)

Example: wing night

Buffalo Wild Wings ran a promotion called the Blazin' Bonus, in which every $25 gift card purchased also received a "Bonus" gift card for $5, $15, $25, or $100. According to the company, here are the probabilities for each Bonus gift card:

Blazin' Bonus	$5	$15	$25	$100
Probability	0.890	0.098	0.010	0.002

a) Explain why this is a valid probability model.

   The probability of each outcome is a number between 0 and 1 and the probabilities add up to 1.

b) Find the probability that you don't get a $5 Bonus card.

   $P(\$5^C) = 1 - P(\$5) = 1 - .890 = 0.110$

c) What's the probability that you get a $25 or $100 Bonus card?

   These events are mutually exclusive, so I can use the addition rule for mutually exclusive events.

   P($25 or $100) = P($25) + P($100) = 0.010 + 0.002 = 0.012

Example: subject preference and gender

Do males and females have a different preference for Math or English classes? The two-way table summarizes data about gender and subject preference for a class of 25 Statistics students.

		Gender		
		Male	Female	Total
Prefered Subject	Math	8	12	20
	English	2	3	5
	Total	10	15	25

Suppose we choose a student from the class at random. Define event A as getting a male student and event B as getting a student who prefers math classes.

a) Find P(A).

   P(A) = P(male) = 10/25 = 0.40

b) Find P(A and B). Interpret this value in context.

   P(male and math) = 8/25 = 0.32

   There is about a 32% chance that a randomly selected student from this class is male and prefers math.

c) Find P(A or B).

   NOTE! THESE ARE NOT MUTUALLY EXCLUSIVE

   P(male or math) = 10/25 + 20/25 – 8/25 = (10+20–8)/25 = 22/25 = 0.88 or (8 + 2 + 12)/25

General Addition Rule

We use the **general addition rule** when events are not mutually exclusive

P(A or B)=P(A)+P(B)-P(A and B)

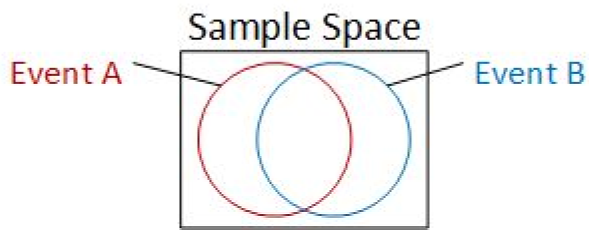
Example: where are the best tacos?

A survey of all students at a large high school revealed that, in the last month, 38% of them had dined at Taco Bell, 16% had dined at Chipotle, and 9% had dined at both. Suppose we select a student at random. What's the probability that the student has dined at Taco Bell or Chipotle in the last month?

P(T or C) = 0.38 + 0.16 – 0.09 = 0.45

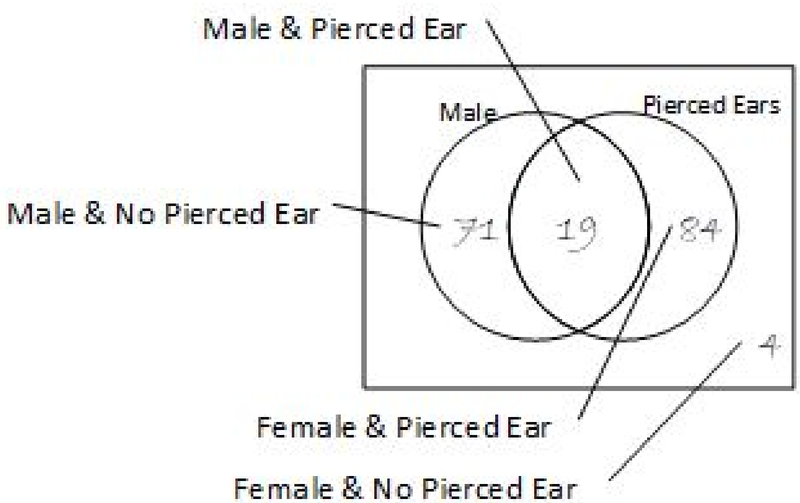
The probability that the student dined at Taco Bell or Chipotle in the last month is 0.45.

Venn diagram



Example: How are Venn diagrams and 2-way tables related?

		Gender		
		Male	Female	Total
Pierced Ear	Yes	19	84	108
	No	71	4	75
	Total	90	88	178



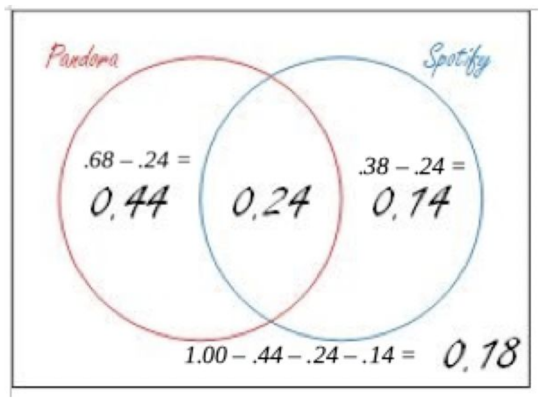
The **intersection** is all outcomes that are common to both A and B

The **union** is all outcomes in either A or B or both.

Example: Pandora or Spotify?

According to a recent report, Pandora and Spotify are the most used music-streaming apps. A group of Statistics students surveyed all the seniors in their school and found that 68% use Pandora, 38% use Spotify, and 24% use both. Suppose we select a senior at random.

a) Make a Venn diagram to display the sample space of this chance process using the events P: uses Pandora and S: uses Spotify.



b) Find the probability that the person uses neither Pandora nor Spotify.

0.18

Topic 2 - Test Your Understanding

1. Sickle-Cell Anemia is a serious inherited disease that is about 3 times more likely to occur in African-American babies than in other babies. A person with two sickle-cell genes will have the disease, but a person with just a single sickle-cell gene will be a carrier of the disease. Two parents are carriers of sickle-cell anemia plan on having a child.
   a. Create a probability model for the sickle-cell gene(s) of the child.
   b. Define event A as the child has sickle-cell anemia. Explain what P(A) means, then find it.
2. The distribution of colors for Plain M&M's™ is given in the following table:

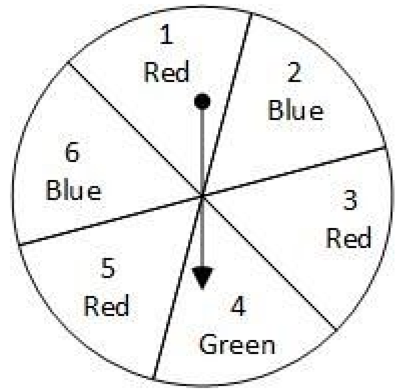
Color	Blue	Brown	Green	Orange	Red	Yellow
Probability	0.24	0.13	0.16	0.20	0.10	0.17

   a. Explain why this is a valid probability model.
   b. When selecting a candy at random, what is the probability it is not blue?
   c. When selecting a candy at random, what is the probability that you get one of the two least likely colors?
3. According to the U.S. Government statistics, mononucleosis (mono) is four times more common among college students than the rest of the population. Blood tests for the disease are not 100% accurate. The following table of data was obtained regarding students who came to a local college's health center complaining of tiredness, a sore throat and a slight fever. Define event A as having mononucleosis and event B as having a positive blood test. Suppose we choose a student at random.

	Has Mononucleosis	Does Not Have Mononucleosis	Total
Positive Blood Test	72	4	76
Negative Blood Test	8	56	64
Total	80	60	140

   a. Find P(A)
   b. Find P(A and B). Interpret this value in context.
   c. Find P(A or B). Interpret this value in context.
   d. Create a Venn Diagram to display the sample space of this situation.
   e. Find the probability of the student has a negative blood test and does not have mononucleosis.

5. Find each probability using the spinner to the right
    a. P(4)
    b. P(Blue)
    c. P(7)
    d. P(Yellow)
    e. P(not Red)
    f. P(Red or Blue)
    g. P(Red, Blue, and Green)
    h. P(1, 2, 3, or 5)
    i. P(1, 2, 3, 4, 5, or 6)
    j. P(Blue or 4)
    k. P(not Blue or Green)



6. According to the National Center for Health Statistics, in December, 2012, 60% of U.S. Households had a traditional Landline, 89% of households had cell phones, and 51% had both. Suppose we randomly selected a household in December, 2012. What's the probability that the household has a traditional landline telephone or a cellphone?

Topic 3: Conditional Probability and Independence

		Gender		
		Male	Female	Total
Pierced Ear	Yes	19	84	103
	No	71	4	75
	Total	90	88	178

P(male) = 90/178

P(pierced ear) = 103/178

P(male and pierced ear) = 19/178

P(male or pierced ear) = 90/178 + 103/178 – 19/178 = (90 + 103 – 19)/178 = 174/178 = 0.998
or (19 + 71 + 4)/178 = 174/178 = 0.998

If we know that a randomly selected student has a pierced ear, what is the probability that the student is male? (Conditional probability! The denominator changes.) 19/103 = 0.184

If we know that a randomly selected student is male, what's the probability that the student has a pierced ear? 19/90 = 0.211

Conditional probability

The probability that something happens once we already know something else has happened.

The probability of A GIVEN we know B has occurred.

Example: Olympic medals

In the 2016 Summer Olympics, the United States and China won the most medals. Suppose we randomly select a medal from the 191 that are represented in the two-way table. Define events G: gold medal, U: United States, and B: bronze medal.

		Medals			
		Gold	Silver	Bronze	Total
Country	United States	46	37	38	121
	China	26	18	26	70
	Total	72	55	64	191

Find P(B | U). Interpret this value in context.

Read this as "The probability of B given U". The | is read as "given" which represents conditional probability.

38/121= 0.314

Independent events

Events are independent if the fact that one event happened does not change the likelihood that the other event will happen.

General multiplication rule

The probability that A and B both occur can be found with

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

Example: Hot Coffee

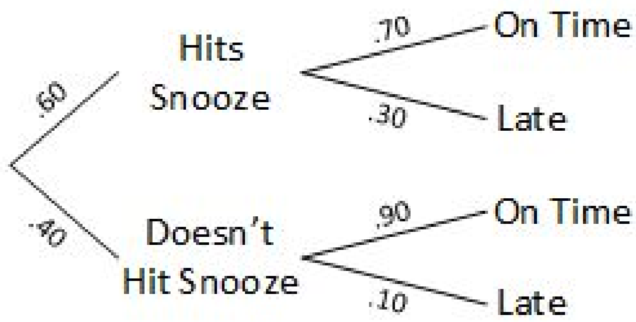
Students who work at a local coffee shop recorded the drink orders of all the customers on a Saturday. They found that 64% of customers ordered a hot drink, and 80% of these customers added cream to their drink. Find the probability that a randomly selected Saturday customer orders a hot drink and adds cream to the drink.

P(hot drink and adds cream) = P(hot drink) · P(adds cream|hot drink) = (.64)(.80) = 0.512

Sometimes we have a lot of information in a problem with a lot of conditional probabilities. In these cases, it's easiest to organize what you know with a **tree diagram**. Then we can use the multiplication rule.

Example: Snooze or On Time

Shannon hits the snooze button on her alarm on 60% of school days. If she hits snooze, there is a 0.70 probability that she makes it to her first class on time. If she doesn't hit snooze and gets up right away, there is a 0.90 probability that she makes it to class on time. Suppose we select a school day at random and record whether Shannon hits the snooze button and whether she arrives in class on time.



What is the probability that Shannon hits the snooze button and is late for class on a randomly selected school day?

P(hits snooze and is late) = P(hits snooze) · P(is late|hit snooze) = (.60)(.30) = 0.18

What's the probability that Shannon is late to class on a randomly selected school day?

Think: how many different ways can Shannon be late? She can be late and hit the snooze OR be late and not hit the snooze.

P(late) = P(hits snooze and late) + P(doesn't hit snooze and late) = (.60)(.30) + (.40)(.10) = 0.22

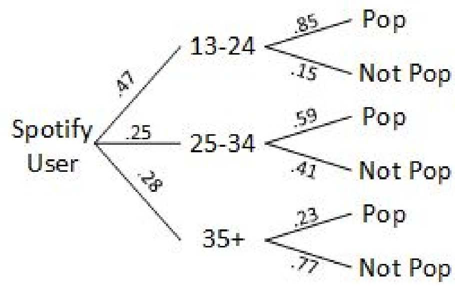
Suppose that Shannon is late for class on a randomly chosen school day. What is the probability that she hit the snooze button that morning?

$$P(\text{hits snooze}|\text{late}) = \frac{P(\text{hits snooze \& late})}{P(\text{late})} = \frac{0.18}{0.22} = 0.818$$

Example: Who prefers pop music?

In 2015, Spotify revealed that about 47% of its users are aged 13–24, 25% are aged 25–34, and 28% are 35 or older. Suppose that for the 13–24 year olds 85% identified pop as their favorite music genre, for the 25–34 year olds 59% identified pop as their favorite music genre, and for the 35 or older group 23% identified pop as their favorite music genre. Suppose we select one 2015 Spotify user at random and record his or her age and whether his or her favorite music genre is pop.

(a) Draw a tree diagram to model this chance process.



(b) Find the probability that the person identifies his or her favorite music genre as pop.

$$P(pop)=(.47)(.85)+(.25)(.59)+(.28)(.23)=0.6114$$

(c) Suppose the chosen person identifies his or her favorite music genre as pop. What's the probability that he or she is aged 13–24?

$$P(13\text{-}24 \mid pop)=\frac{P(13\text{-}24 \text{ and } pop)}{P(pop)}=\frac{.3995}{.6114}=.6534$$

Example: Multiplication rule if the events are independent

Suppose that Pedro drives the same route to work on Monday through Friday. His route includes one traffic light. The probability that the light will be green when Pedro arrives is 0.42, yellow is 0.03 and red is 0.55.

What's the probability that the light is green on Monday and red on Tuesday?

These events are independent, so in this case we can just multiply the appropriate probabilities.

$(0.42)(0.55) = 0.231$

What is the probability that Pedro finds the light red on Monday through Friday?

$(0.55)(0.55)(0.55)(0.55)(0.55) = (0.55)^5 = 0.0503$

Example: All Blondes

In the United States, 28% of the population has hair that is blonde. A couple in the United States is expecting quadruplets. Does this mean there is a $(.28)^4 = 0.0006$ probability that the quadruplets will all have hair that is blonde?

No, it is not appropriate to multiply the four probabilities because each child's hair color is not independent of the hair colors of the other children. Knowing the hair color of one child will help predict the hair color of another because hair color is a trait inherited from parents.

Topic 3 - Test Your Understanding

1. This table is from problem 2 from Topic 2. Define event A as having mononucleosis and event B as having a positive blood test. Suppose we choose a student at random.

	Has Mononucleosis	Does Not Have Mononucleosis	Total
Positive Blood Test	72	4	76
Negative Blood Test	8	56	64
Total	80	60	140

   a. Find $P(A|B^c)$. Interpret this value in context.

   b. Find $P(B|A)$. Interpret this value in context.

2. Zoë is taking part in a lottery for a room in one of the two new dormitories at her college next year. She will draw a card randomly to determine which room she will receive. Each card has the name of a dormitory, (X)avier or (Y)oung, and also a one person room number, a two person room number or an apartment number. Thirty percent of available spaces are in Xavier, with 20% of the spaces being one person rooms and 60% of the spaces being two person rooms. In Young, 5% of the spaces are one person rooms and 35% of the spaces are two person rooms.

   a. Create a tree diagram to model this chance process.

   b. Find the probability Zoë receives a one person room.

   c. Suppose Zoë receives a one person room. What is the probability the room is in Young?

   d. Find the probability Zoë receives an apartment if she gets into Xavier?

3. The quarterback of the Wisconsin Weasels throws incomplete passes 50% of the time, completed passes for ten or less yards 33% of the time and completed passes for more than ten yards 17% of the time. In a game, he throws 5 passes.

   a. What is the probability the first two passes are incomplete?

   b. What is the probability the first two passes are both completed for more than 10 yards?

   c. What is the probability he completes all five passes for ten or less yards each?

4. In baseball, a perfect game is when a pitcher doesn't allow batters to reach base in all nine innings. Historically, pitchers throw a perfect inning, which is an inning where no hitters reach base, about 40% of the time. Find the probability that a pitcher throws 9 perfect innings in a row or a perfect game. (Assume that the pitcher's performance in any one inning is independent of his performance in other innings.)

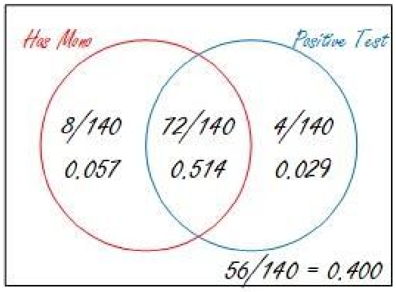
Answers to Test Your Understanding

Topic 1

1. a) If you watch many, many 20 year-old American females, about 99.9% of them will live to 21 years of age.
   b) No. 1,000 20 year olds American females is a relatively small sample and it is unlikely that exactly 999 of them would live to 21.
2. a) Assign each of the remaining socks a number from 1 to 48. Generate a random number. If the number is a 1 or a 2, an unmatched sock was lost, otherwise another pair is broken up.
   b) The dot represents a trial where an unmatched sock was lost, resulting in keeping 23 pairs of matched socks.
3. a) Label the female student rooms 1 through 481 and males from 482 through 816. Select three digits at a time, ignoring repeats and numbers larger than 816 until you have 50 numbers. Count how many numbers are less than or equal to 481.
   b) Yes. There is about a 3/500 = 0.6% chance that if the papers were properly mixed there would be 10 or fewer females in the sample. Because this probability is small, we have convincing evidence that the papers were not properly mixed.

Topic 2

1. a) Let s = sickle-cell gene and n = normal gene
   ss, sn, ns, nn :where in each pair, the first is the gene passed by the father and the second by the mother.
   b) P(A) means the probability of the child having sickle-cell anemia. So we are looking for the gene pair ss. This means P(A) = 0.25
2. a) The probability of each outcome is a number between 0 and 1 and the probabilities add up to 1.
   b) P(NOT blue) = 1.00 – 0.24 = .76
   c) P(two least likely) = P(red or brown) = 0.10 + 0.13 = 0.23
3. a) P(A) = 80 / 140 = 0.571 = 57.1%
   b) P(A and B) = 72 / 140 = 0.514 = 51.4%
   There is a 51.4% chance that a randomly chosen student has a mononucleosis and has a positive blood test.
   c) P(A or B) = 80/140 + 76 / 140 - 72 / 140 = (80 + 76 –72)/140 = 0.6 = 60%
   P(A or B) = (72 + 4 + 8)/140 = 0.6 = 60%
   There is a 60% chance that a randomly chosen student has mononucleosis, a positive blood test or both.

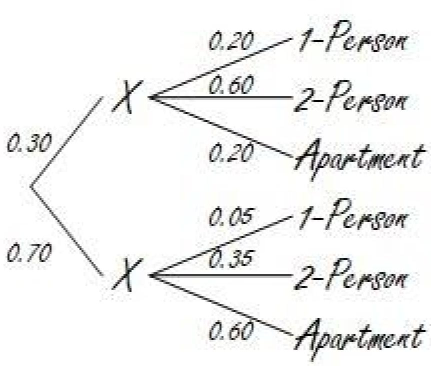


d)
e) 0.4 = 40%

4. a) ⅙
   b) 2/6 = ⅓
   c) 0
   d) 0
   e) 3/6 = ½
   f) ⅚
   g) 6/6 = 1
   h) 4/6 = ⅔
   i) 6/6 = 1
   j) 3/6 = ½
   k) 3/6 = ½
5. P(A or B) = P(A) + P(B) - P (A and B) P(A) = 60% or .60, P(B) = 89% or .89, P(A and B) = P(Both) = 51% or .51. .60 + .89 - .51 = .98
   .98 or 98% of all households have a Landline or a Cellphone.

Topic 3

1. a) $P(A|B^c) = 8/64 = 0.125 = 12.5\%$
   b) $P(B|A) = 72/80 = 0.9 = 90\%$



2. a)
   b) P(1) = 0.30(0.20) + 0.70(0.05) = 0.635
   c) P(Y|1) = 0.70(0.05)/0.635 = 0.035/0.635 = 0.055
   d) P(A|X) = 0.30(.20)/0.30 = 0.06/0.30 = 0.20

3. a) (0.50)(0.50) = 0.25
   b) (0.17)(0.17) = 0.029
   c) (0.33)(0.33)(0.33)(0.33)(0.33) = 0.004
4. A pitcher must throw 9 perfect innings to achieve a perfect game. Each inning is independent and for each of those 9 innings, the typical pitcher will have a 40% or .40 probability of pitching a perfect inning.

1		2		3		4		5		6		7		8		9		
.40	x	.40	x	.40	x	.40	x	.40	x	.40	x	.40	x	.40	x	.40	=	(.40)9 or .000262