

Lesson 3

Probability Distributions

While a prisoner of war during World War II, J. Kerrich conducted an experiment in which he flipped a coin 10,000 times and kept a record of the outcomes. A portion of the results is given in the table below.



Number of Tosses	Number of Heads
10	4
50	25
100	44
500	255
1,000	502
5,000	2,533
10,000	5,067

Source: J. Kerrich. *An Experimental Introduction to the Theory of Probability*. Copenhagen: J. Jorgenson and Co., 1964.

Think About This Situation

Think about Kerrich's results as you answer these questions.

- How many heads would you expect if you tossed a coin 100 times? 32 times? 15 times?
- After how many tosses is the number of heads in Kerrich's table closest to the expected number of heads? Furthest?
- Was the percentage of heads closer to the expected percentage of 50% after tossing 10 times or 10,000 times?
- If six Monopoly players are sent to jail, how many would you expect to get out of jail by rolling doubles on their first try? On their second try?

INVESTIGATION 1 Theoretical Waiting-Time Distributions

In Lesson 1, you explored waiting-time distributions by conducting experiments and simulations. As you probably noticed, two different groups could get quite different histograms when they constructed them from simulations. The two groups then would have different estimates of a probability. In this investigation, you will construct waiting-time distributions theoretically, so everyone should get the same answers to the probability questions related to these distributions.

If you flip a fair coin 10 times, you *expect* to get 5 heads. However, you don't get 5 heads each time you flip a fair coin 10 times. Sometimes you get fewer, as Kerrich did; sometimes you get more. In the long run, however, the average number of heads will be 5.



Expectation or *expected value* is another word for the *theoretical average*. For example, if you flip a coin 5 times, you might say in ordinary language that you expect to get 2 or 3 heads. However, in mathematics you should say that you expect to get 2.5 heads, because *on the average* that is the number of heads you will get.

1. Imagine 36 students are playing modified Monopoly in a class tournament. All are sent to jail. A student must roll doubles to get out of jail. (There is no other way out.)
 - a. How many of the 36 students do you expect to get out of jail by rolling doubles on the first try? (Remember that the word “expect” has a mathematical meaning.) How many students do you expect to remain in jail?
 - b. How many of the remaining students do you expect to get out of jail on the second try? How many students do you expect to remain in jail then?
 - c. How many of the remaining students do you expect to get out of jail on the third try? How many students do you expect to remain in jail then?
 - d. Complete a table like the one at the top of the next page. Round numbers to the nearest hundredth. The first three lines should agree with your answers to Parts a–c.

Rolling Dice to Get Doubles

Number of Rolls to Get Doubles	Expected Number of Students Released on the Given Number of Rolls	Expected Number of Students Still in Jail
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

e. What patterns of change do you see in this table? If possible, describe each pattern using the idea of *NOW* and *NEXT*.

The table you created shows a **theoretical distribution**. The table never really can be completed, however, as the rows should be continued indefinitely.

2. Use the theoretical distribution to reason beyond the table entries.
 - a. How many of the 36 students do you expect to be in jail after 12 tries to roll doubles?
 - b. Add a “13 or more” row to your table and write the expected number of students in the appropriate place.
 - c. If you had started with 10,000 people instead of 36, how many of them would you expect to need 13 or more rolls to get out of jail?
3. Make a histogram of the “expected number of students released on the given number of rolls” from the frequency table you constructed in Activity 1.
 - a. Compare this histogram to the one you constructed following your class’s simulation of this situation in Activity 1 of the first investigation from Lesson 1 (page 457–458).
 - b. Examine your histogram of the theoretical distribution. The height of each bar is what proportion of the height of the bar to its left?
 - c. Using your histogram, estimate the average of the distribution.
 - d. Calculate the average number of rolls of the dice it takes to get doubles. Compare your calculated average to your estimate in Part c.
4. If there were 1,000 people (rather than 36) who had been sent to jail in Activity 1, how would the histogram change? How would the average change?

5. Thirty percent of “M&M’s”[®] Plain Chocolate Candies are brown. Suppose each of 1,000 students removes candies one at a time from a large bag until he or she gets a brown one.
 - a. How many students do you expect to get a brown candy on the first try?
 - b. Make a table like the one in Activity 1. Give the table 10 rows, and complete each row.
 - c. How many of the 1,000 students do you expect to need 11 or more draws to get a brown candy?
 - d. Make a histogram of the distribution shown in your table.
 - e. Estimate the average number of draws to get a brown candy from your histogram. Then calculate the average using an appropriate formula.

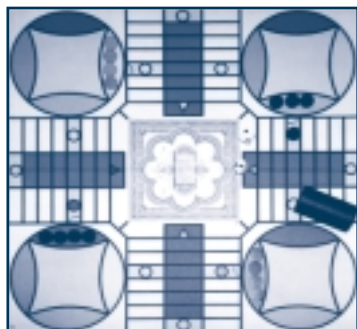
Checkpoint

Summarize key ideas about theoretical waiting-time distributions.

- a. How is the word “expect” used differently in mathematics than in everyday life?
- b. Suppose the probability of an event is p . How many times would you expect this event to happen in a series of n independent trials?
- c. Suppose the probability of an event is p and you have made a histogram of the waiting-time distribution for the event. How is the height of each bar of the histogram related to the height of the bar to its left?

Be prepared to share your ideas with the whole class.

On Your Own



In the game of Parcheesi[®], based on the Indian game pachisi, a player cannot move a pawn for the first time until he or she rolls a five with a pair of dice. The five may be on either die, or the five may be the sum of both dice.

- a. What is the probability a player can move a pawn on the first roll of the dice?
- b. If four players begin a game of Parcheesi, how many do you expect to move a pawn on their first roll of the dice?
- c. How many of the four players do you expect to require two rolls of the dice in order to move a pawn?

MORE

Modeling • Organizing • Reflecting • Extending

Modeling

1. If there is a 40% chance of rain today, it means that it rained on 40% of the days in the past that had weather conditions similar to those today.



- a. On 14 different days, the weather report says there is a 40% chance of rain. On how many of these days do you expect it to rain? On how many of these days do you expect it not to rain?
 - b. On 20 different days, the weather report says there is a 50% chance of rain. It actually rained on 9 of those days. Do you think the meteorologist did a good job of predicting rain? Explain.
2. Tay-Sachs disease results when a baby inherits the Tay-Sachs gene from both parents. (A person who can pass the gene to the next generation is called a carrier.) Since the disease results in early death, no adults carry two Tay-Sachs genes. A college biology textbook says that “about 1 in 30 American Jews is a carrier, which would result in about 1 Tay-Sachs child in 3600 Jewish births.” (Source: Audesirk, G. and T. *Biology: Life on Earth*. New York: Macmillan, 1989.) Explain how the 1 in 3600 figure was computed.
 3. In each Kellogg’s® Cocoa Krispies® cereal box, there was a sticker of either a bird of paradise, a tiger, an African elephant, or a crocodile. Assume that the stickers were placed randomly into the boxes. Thirty-six people are buying boxes, each trying to get a tiger.
 - a. Make a theoretical waiting-time distribution table and histogram of the number of boxes purchased to get a tiger sticker. Place the numbers 1 to 11 in your table followed by a “12 or more” row.
 - b. How many of these people do you expect would need to purchase more than six boxes of the cereal?
 - c. On the average, how many boxes of cereal need to be purchased to get a tiger sticker?

4. Imagine 1,024 students in a school auditorium all standing up. Each student flips a coin. The students who get heads on this flip sit down. Each student who remains standing flips a coin a second time. The students who get heads on this flip sit down. The coin flips continue until all students are seated.
- Make a frequency table that shows what you expect to happen in this experiment.
 - Make a histogram from your frequency table.
 - What is the average number of flips required until a student sits down?
 - What is the probability that it will take two flips or fewer to get a head? What is the probability that it will take more than two flips?
 - How many times would a student have to toss the coin without getting a head before you would say a rare event has occurred?
 - How many rare events do you expect to occur among the 1,024 students?
5. The player with the highest field goal percentage in the history of the National Basketball Association (NBA) is Artis Gilmore. In his career in the NBA, Gilmore attempted 9,570 field goals and made 5,732 of them.
- What was Gilmore's field goal percentage?
 - During a typical game, Gilmore might attempt 25 field goals. In a typical game, how many field goals would you expect Gilmore to make?



The NBA player with the highest lifetime free throw percentage is Mark Price. Price had a free throw "percentage" of 0.904. He made a total of 2,135 free throws.

- Why do you think the word *percentage* is in quotation marks above?
- How many free throws did Price attempt?
- How many free throws would you expect Price to make in 50 attempts?
- Write an equation that relates the number of free throws T expected for a player who makes A attempts and whose free throw percentage is p .



Organizing

1. Refer to the waiting-time distribution for a brown candy in Activity 5 of Investigation 1 (page 488).
 - a. Make a box plot of these waiting times.
 - b. What is the median number of candies drawn to get the first brown one?
 - c. How does the shape of the box plot reflect the shape of the histogram? In which direction is the distribution skewed?
2. When would the mean be a good measure of the center of a waiting-time distribution? When would the median be a good measure?
3. Explain why the histogram of a waiting-time distribution does not have line symmetry.
4. Refer to Activity 1 (page 486) of Investigation 1 for this task.
 - a. Make a scatterplot of the (*number of rolls to get doubles, expected number of students who get out of jail on this roll*) data.
 - b. Find an algebraic model of the form “ $y = \dots$ ” that is a good fit for these data.
 - c. Let *NOW* be the number of people who are expected to get doubles on a roll, and let *NEXT* be the number of people who are expected to get doubles on the next roll. Write an equation relating *NOW* and *NEXT*.
5. Make a box plot for your distribution from Modeling Task 3 (getting a tiger sticker). Using the same number line, make box plots for the distribution from Modeling Task 4 and for the distribution from Activities 1 and 2 (rolling doubles) of Investigation 1. Write a few sentences about what you can learn about waiting-time distributions from examining these box plots.

Reflecting

1. It is quite common for a person’s first guess about a probabilistic situation to be wrong. Part of learning probability is learning to be wary of your first reaction and learning how to check whether your first reaction is correct. What have you found most surprising so far in this unit?
2. List some businesses or jobs in which the manager should be interested in waiting-time distributions.
3. Marina and Jamie are playing a game in which the first person to roll doubles wins. Marina has had 10 turns and hasn’t rolled doubles yet. Marina says, “I’m due to get doubles on my next roll.”
 - a. Explain what Marina means by this statement.
 - b. Is Marina correct? Why or why not?
 - c. Design an experiment to show Marina that she is no more likely to roll doubles on her 11th roll than she was on her 1st roll.

4. Board games involving chance have a long history of providing recreation for people from many different cultures.
 - a. Investigate the history of backgammon, Parcheesi[®], Senet (an Egyptian game), or a similar game that you have played with your family or friends.
 - b. Review the rules of the game.
 - c. Write several questions about the probabilities involved in the game.
 - d. Answer one of your questions either theoretically or by using a simulation to estimate the probability.

Extending

1. What are some characteristics that theoretical waiting-time distributions for independent trials have in common?
2. In this task you will find the average of another type of frequency distribution. This distribution is called a **binomial distribution**. “Binomial” means “having two names.”



On each flip of a coin, you can describe the outcome using one of two names: “heads” or “tails.” In a waiting-time distribution, you might count the *number of trials* until you get the first head. In a binomial distribution, you count the *number of heads* in a fixed number of independent trials.

- a. Flip 11 coins and count the number of heads. Add the result to a copy of the frequency table below.

Number of Heads	Frequency
0	1
1	0
2	4
3	5
4	16
5	18
6	27
7	15
8	5
9	3
10	1
11	0
Total	<input type="text"/>

4. The formula below gives the probability of getting exactly x heads if a coin is flipped n times.

$$P(x) = \frac{n!}{x!(n-x)!} \left(\frac{1}{2}\right)^n$$

The symbol $n!$, read “ n factorial,” is defined as follows:

$$0! = 1$$

$$1! = 1$$

$$2! = 2 \cdot 1 = 2$$

$$3! = 3 \cdot 2 \cdot 1 = 6$$

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

and so on.

- If you flip a coin 11 times, what is the probability of getting exactly 4 heads?
- Use the formula to complete the table below for the experiment of flipping a coin 11 times.

Number of Heads	Probability
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	

- Make a graph of the distribution in this table. How does the graph compare to your histogram from Extending Task 2?
- How does the formula incorporate the Multiplication Rule?