Unit 18: Introduction to Probability

SUMMARY OF VIDEO

There are lots of times in everyday life when we want to predict something in the future. Rather than just guessing, probability is the mathematical way to make these kinds of predictions. Here are some examples that use the language of probability: a 50% percent chance of snow, a 20% percent chance of complications from surgery, a one in one hundred seventy five million chance of winning the lottery. We encounter statements such as these all the time in daily life – but what do they really mean?

These statements are attempts to quantify uncertainty. For example, how likely is it to rain later in the day – the answer to this question helps people decide if they should carry an umbrella. When meteorologist Kevin Skarupa issues his forecast for the residents of New Hampshire, he doesn't know for sure what is going to happen. Weather is an example of a random phenomenon. It is an event with an uncertain outcome, but it does have a regular pattern over time.

Today's meteorologists rely on multiple complicated mathematical models to make their predictions for the public. The models churn through tons of weather related data – from current weather balloon information and surface observations, to historic patterns. These models combine all these weather inputs to create maps predicting what will happen in the next few days based on what has happened in the past when similar scenarios have been observed. Over time, the weather exhibits patterns; but for any one particular instance in the future, the weather is not completely predictable with perfect accuracy. That is why forecasters talk in probabilities – for example, "70% chance of rain."

When a station announces a 70% chance of rain, it generally means that 70% of the viewers, if equally spread out, will see precipitation. In this case, the forecaster feels sure that 70% of the area will see rain that day – so the 70% refers to area coverage. The percentage number is also used to quantify the likelihood of any precipitation at all – the degree of confidence of getting rain. So, when reporting a percentage, meteorologists are usually expressing a combination of degree of confidence and area coverage.

The probability of any event is the proportion, or percentage, of times it would occur in a long series of repetitions. Random phenomena like weather events are not chaotic; they are unpredictable in the short run, but they have a regular pattern in the long run. Take the example of flipping a coin. The toss can come up heads or it can come up tails. Suppose we start flipping a coin and then record the proportion of heads. Say, on the first flip we get tails, the proportion of heads is 0. On the next toss we get heads, and the proportion of heads goes up to 0.5. On the next two flips we get tails, and now the proportion of heads in the four flips is down to 0.25. In the short term, the proportion of heads is quite variable. But, suppose we continue flipping the coin – 50 times, 100 times, 1000 times. Over the long term a pattern emerges – the proportions hover around 0.5 – as can be seen in Figure 18.1.

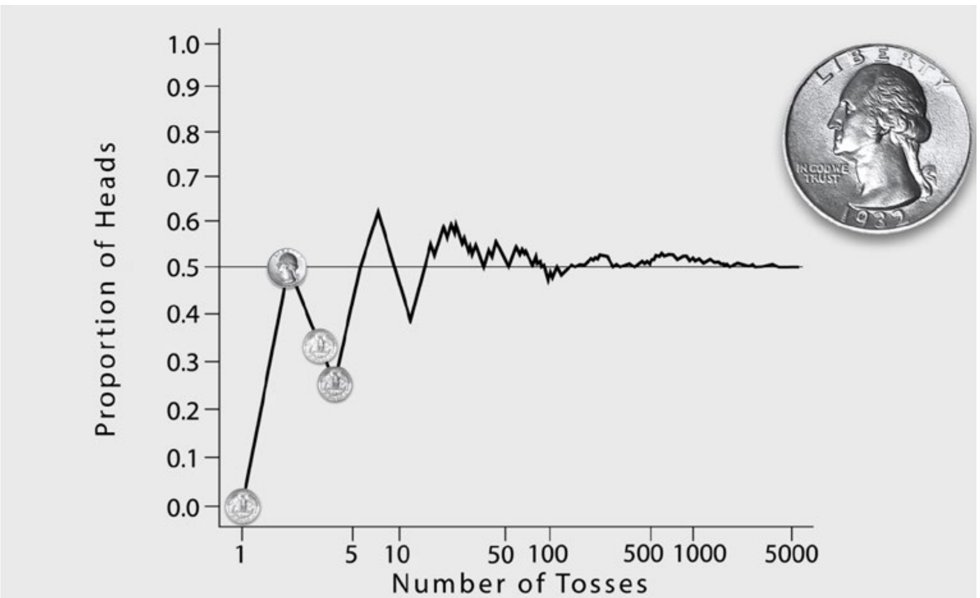

Figure 18.1. Proportion of heads in flipping a coin.

While we don't know what is going to happen on any one toss, over time we can predict that we will get a proportion of heads around 0.5.

Like the proportions on the graph in Figure 18.1, probabilities are between zero and one. Events with a probability closer to zero are less likely and those with a probability closer to one are more likely to happen. Our probability of 0.5 for heads in a coin toss means that either outcome is equally likely – 50/50.

Probabilities are assigned to more than just coin tosses. NASA's Near Earth Object Program is closely monitoring asteroids with the potential to do serious damage to our planet. While "near-Earth space" is home to over 9,000 known asteroids, only about half of them are large enough and have orbits that come close enough to Earth to classify them as PHAs – Potentially Hazardous Asteroids. Scientists track these PHAs, collecting data on them, so they can determine the likelihood that any might be on a collision course with Earth. As more and more data come in about an asteroid's orbit, scientists refine their predictions of its orbit, which allows them to start work on computing the probability that it will collide with Earth.

The poster child for near-Earth objects is an asteroid called Apophis. It is coming very close to Earth in 2029. When it was first discovered, scientists didn't know how close it would come to Earth. In fact, the uncertainty region during its passage by Earth was so large, that Earth was right in the middle of it. (See Figure 18.2.)

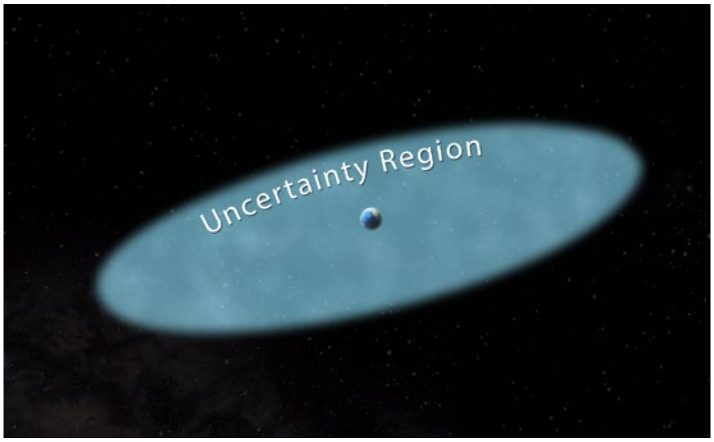

Figure 18.2. Initial uncertainty region.

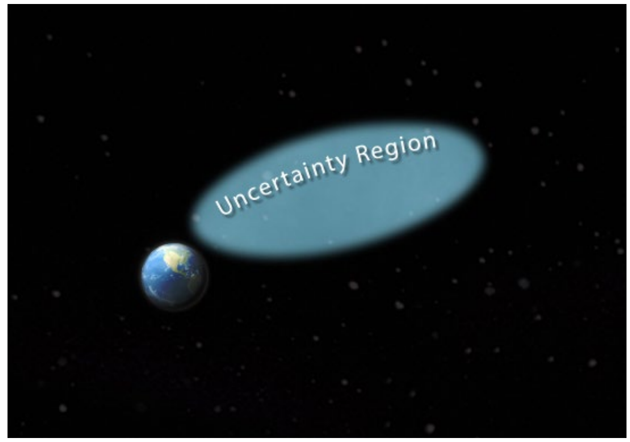

Figure 18.3. Refined uncertainty region

As we got more and more observations on the asteroid, scientists were able to refine their projections of its path and shrink the uncertainty region so that it no longer intersects with Earth. (See Figure 18.3.) So, we now know that Apophis will pass by Earth in 2029 and get very close, but the probability that it will hit Earth is essentially zero. However, Apophis' close encounter with Earth's gravity in 2029 will bend its trajectory, which makes the job of predicting where it will go from there much more difficult. Apophis' next passage by Earth will be in 2036 and scientists will once again be collecting data, predicting its orbit, and assessing the likelihood that Apophis is on a collision course with Earth.

STUDENT LEARNING OBJECTIVES

A. Be able to identify random phenomena affecting everyday life.

B. Understand that it is not possible to predict with certainty short-term behavior of random phenomena but it is possible to predict long-run patterns.

C. Be able to calculate proportions (or relative frequencies).

D. Be able to express probabilities as percentages.

CONTENT OVERVIEW

Toss a coin or choose a simple random sample (SRS) from a population. The results can't be predicted in advance. When we flip a coin, we know we will get a head or a tail. We expect both outcomes to be equally likely, but we don't know for certain which outcome will occur the next time we flip the coin. An instructor chooses a random sample of four students each day to put homework problems on the board. Because the selection is random, each possible size-four sample is equally likely to be selected. One day Joe comes to class without having done his homework. If the class is small, his chances of getting chosen are pretty good. If his class is large, he is less likely to be in the selected sample. Joe won't know if he will be caught without his homework until the sample is actually drawn.

Flipping a coin and choosing a random sample are both examples of **random phenomena.** Other examples include: the outcome of rolling a die, the gender of the next person passing through a turnstile at a subway, the growth of a child in one month, the color of the next car that exits the parking lot, and whether guessing on a true-false question will result in a correct answer. In each of these cases, the next outcome is uncertain but, over the long run of many repetitions, a pattern emerges.

We use **probability** to assess the likelihood that a random phenomenon has a particular outcome. Probability is a number between 0 and 1. If A represents a particular outcome or set of outcomes of a random phenomenon, we write $P(A)$ to denote the probability that event A will occur. The closer $P(A)$ is to 0, the less likely it is for event A to occur. The closer $P(A)$ is to 1, the more likely it is for event A to occur. Probabilities are often expressed as percentages. For example, the probability of flipping two heads in a row is 0.25 or 25%.

Next, we look at three ways that probabilities can be assigned to events. First, suppose a sports reporter predicts that the Yankees have a 75% chance of beating the Red Sox in their next game. In this case, the reporter is most likely giving his or her professional assessment of the likelihood that the Yankees will win. That assessment is based on his or her knowledge of the players, whether the team is playing at their home field, past interactions between these two teams, and a whole host of other factors. So, we could classify this type of probability assignment as informed intuition.

Second, a large medical laboratory developed its own test for a person's vitamin D level. Too much vitamin D can be toxic, while insufficient vitamin D is linked to certain illnesses. After complaints that the test might be giving erroneous results, the company randomly selected a sample of patients and retested their vitamin D levels. Suppose that the sample consisted of 200 patients and that 18 of the initial tests were determined to be erroneous. The probability of an erroneous test result can be estimated from the proportion of erroneous tests found in the sample. In this case,

Probability of erroneous test = (frequency of erroneous test)/(number of tests)
= 18/200 = 0.09 or 9%.

Third, suppose a student did not study for a multiple choice exam. There were five choices for answers, (a) – (e), and only one correct answer. The student guessed the answer to each question without even reading the question. Here there is an underlying assumption that the student is equally likely to answer (a), (b), (c), (d), or (e) and that the same is true for the correct answer. To assess his probability of getting a correct answer, he used the probability formula for equally likely outcomes.

$$\text{Probability of an event} = \frac{\text{number of outcomes in the event}}{\text{total number of outcomes}}$$

$$\text{Probability of a correct answer} = \frac{\text{number of correct answers per question}}{\text{total number of ways to answer the question}} = \frac{1}{5} \text{ or } 20\%.$$

We have shown three ways to assign probabilities: informed intuition (educated guess), proportion (relative frequency), and a formula used when outcomes are equally likely. Remember that probabilities are always between 0 and 1. Anytime you calculate a probability and get something like 1.4 or -0.2, go back and check your calculations because you have made a mistake.

KEY TERMS

The outcome of a **random phenomenon** in any single instance is uncertain. However, if the phenomenon is repeated over and over, a regular pattern to the outcomes emerges over the long run.

Probability is a measure of how likely it is that something will happen or something is true. Probabilities are always between 0 and 1. Events with probabilities closer to 0 are less likely to happen and probabilities closer to 1 are more likely to happen.

THE VIDEO

Take out a piece of paper and be ready to write down answers to these questions as you watch the video.

1. What is a random phenomenon?

2. Explain why weather is an example of a random phenomenon.

3. What does it mean when a weather reporter says that there is a 70% chance of rain tomorrow?

4. If we flip a fair coin repeatedly, what can be said about the proportion of heads in the short run? What can be said about the proportion of heads in the long run?

5. What can you say about an event whose probability is close to one compared to an event whose probability is close to zero?

6. In 2029 the asteroid Apophis is predicted to pass close to Earth. According to current models predicting its path, what is the probability that it will collide with Earth?

UNIT ACTIVITY:
OBSERVING RANDOM PHENOMENA

In this activity, you will observe two random phenomena – flipping a coin and tossing a tack.

Part I: Flipping a Coin

1. a. What does it mean to say you are flipping a fair coin?

b. A run is a string of the same outcome in a row. If you flip a fair coin 100 times, estimate the length of the longest run you would expect to observe.

2. a. Flip a coin 100 times. Record the outcome of each flip.

b. What is the length of the longest run (either heads or tails)? Is it longer or shorter than what you expected?

c. Calculate the proportion of heads in the first 10 flips, in the first 20 flips, in the first 50 flips, and in all 100 flips.

d. Based on the results from 100 flips, do you think you were flipping a fair coin? Explain.

3. a Combine the data from the class. Calculate the proportion of heads.

b. Does your proportion in (a) give you reason to believe that the coins students were flipping were not fair? Explain.

Part II: Tossing a Thumbtack

Figure 18.4. Tacks sitting point down and point up. Photo by Tomasz Sienicki.

4. When you toss a thumbtack, it can land point up or point down. For flipping a coin, we expect the two outcomes, heads or tails, to be equally likely. But is the same true for tossing tacks? Your task in this question is to collect data on tossing a thumbtack and then to use your data to assign probabilities to the two possible outcomes.

a. Collect data on the outcomes of tossing a thumbtack. You decide how many repetitions you will need. How many times did the tack land point up?

b. Use your data from (a) to assign probabilities to landing point up or point down.

c. What is the sum of your probability assignments from (b)?

EXERCISES

1. Identify five random phenomena that occur in your life.

2. Random phenomena can't be predicted for certain in the short term, but exhibit regular patterns in the long term. Which of the data sets in (a – d) do not appear to be from the random phenomena of coin tossing? Explain.

a. T T T T T T T T T T T T T T T T T T T T

b. H H T H H T T H H H H T H T H H T T T T H H H H

c. H T H T H T H T H T H T H T H T H T H T H T H T

d. T H H T T H H T T H H T T H H T T H H T T H H T

3. In a class experiment, 20 students each flipped a coin 50 times. Their results appear in Table 18.1.

Student	1	2	3	4	5	6	7	8	9	10
Number Heads	27	27	27	28	27	28	20	24	19	28
Student	11	12	13	14	15	16	17	18	19	20
Number Heads	22	28	21	26	21	26	25	25	28	24

Table 18.1. Results from flipping coin 50 times.

a. Table 18.2 presents the cumulative results of the student data from Table 18.1 – starting with student 1, next combining the results from students 1 and 2, next combining the results from students 1, 2, and 3 and so forth. Make a copy of Table 18.2 and complete the table. Round proportions to three decimals.

Number Flips	Number Heads	Proportion	Number Flips	Number Heads	Proportion
50	27	0.540	550		
100	54	0.540	600		
150	81		650		
200			700		
250			750		
300			800		
350			850		
400			900		
450			950		
500			1000		

Table 18.2. Cumulative results of student coin flipping data.

b. Plot the proportion (vertical axis) versus the number of flips (horizontal axis). Connect the points with line segments. The long run proportion of heads for a fair coin is 0.5. Add a horizontal line at 0.5.

c. Based on the plot you drew for (b), do you think that the coin used to produce the data in Table 18.1 was a fair coin? Explain.

4. Assess the probabilities of the following outcomes. Decide if the probability that the outcome will occur is low (between 0 and 1/3), moderate (between 1/3 and 2/3), or high (between 2/3 and 1). Justify your answers.

a. Alex is in a class of 20 students. For each class, the instructor selects 3 students to put homework on the board. To make the selection random, the instructor places the names of the students in a container, mixes them, and then asks a student to draw three names. Alex is unprepared. Assess the chances that he will be called.

b. It is cloudy outside and quite humid and warm. The temperature is expected to drop as even more clouds roll in. Assess the chances for rain.

c. After shuffling a standard deck of cards, you draw a card. Assess the chances of drawing an ace.

d. Without putting the first card drawn back into the deck, you reshuffle the deck. Then you draw a second card. Assess the chances of drawing a red card (a heart or a diamond).

REVIEW QUESTIONS

1. Probability is a measure of how likely an event is to occur. Match each of the probabilities below with one of the statements (a) – (d).

0   0.0002   0.5   1

a. Not playing the lottery but still winning.

b. Drawing a black card (club or spade) from a shuffled deck of 52 playing cards.

c. The sun will come up tomorrow morning (even if it is cloudy and you can't see it).

d. Getting struck by lightning in your lifetime.

2. Amanda writes a letter to her local television station telling them to fire their meteorologist. Her evidence was that out of the ten days that the weather reporter stated there would be a 70% chance of rain, it only rained five times. She had carried an umbrella to work on all ten days expecting that with such a high probability, it definitely was going to rain.

a. Explain to Amanda why a 70% chance of rain does not mean that it will definitely rain.

b. It only rained 5 out of 10 days that the weather reporter forecasted a 70% chance of rain. Was Amanda right that the meteorologist was doing a poor job of predicting the weather? Explain.

3. A perfectly balanced spinner is pictured in Figure 18.5. When you spin the spinner, it can stop on any sector: 1, 2, 3, 4, or 5. In Figure 18.5, the spinner has landed on sector 4.

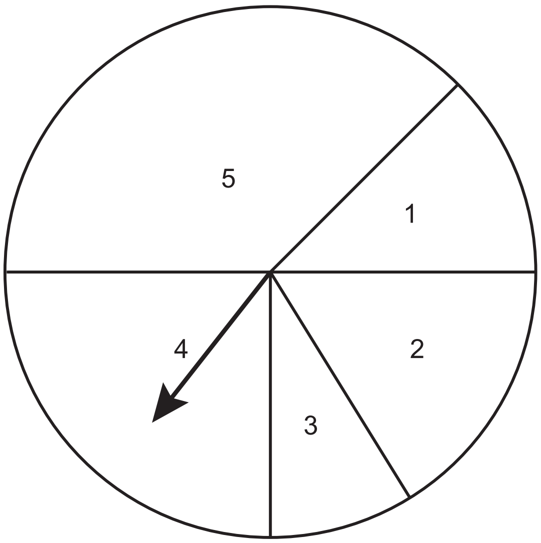

Figure 18.5. Perfectly balanced spinner.

Answer questions that follow. Explain how you arrived at each of your answers.

a. Imagine spinning the spinner shown in Figure 18.5. On which number is it most likely to land?

b. Suppose you spin the spinner 1000 times. How many times would you expect it to land in sector 4? Do you think that what you expect to get would be exactly what you would get if you performed this experiment?

c. Approximately how many times more likely is it for the spinner to land on sector 2 than on sector 3?

d. Estimate the probability of landing on an even number.

4. Each year the study *Monitoring the Future: A Continuing Study of American Youth* surveys students on a wide range of topics, including family background. One of the questions on the survey, including the possible responses, follows.

Did your mother have a paid job (half-time or more) during the time you were growing up?

• No
• Yes, some of the time when I was growing up
• Yes, most of the time
• Yes, all or nearly all of the time

The survey was administered to a large sample of 12th grade students. Care was taken to ensure the sample was representative of all 12th grade students. Responses to this question are summarized in Table 18.3.

Response	Frequency	Probability
No	1845	
Yes/Some	2637	
Yes/Most	2648	
Yes/Nrly All	7148	

Table 18.3. Survey results to question on mother's job.

a. How many students answered this question?

b. Use the data in Table 18.3 to estimate the probabilities associated with mothers' job patterns. Round your estimates to four decimals. Enter your probabilities into a copy of Table 18.3.

c. What is the sum of the probabilities?

d. A randomly selected 12th grade student is asked to answer this question. What is the probability that the student will give a response different from No? Explain how you determined your answer.