SECTION 2
Statistics and Models

Objectives

- Explain how scientists use statistics.
- Explain why the size of a statistical sample is important.
- Describe three types of models commonly used by scientists.
- Explain the relationship between probability and risk.
- Explain the importance of conceptual and mathematical models.

Key Terms

statistics
mean
distribution
probability
sample
risk
model
conceptual model
mathematical model

Environmental science provides a lot of data that need to be organized and interpreted before they are useful. **Statistics** is the collection and classification of data that are in the form of numbers. People commonly use the term statistics to describe numbers, such as the batting record of a baseball player. Sportswriters also use the methods of statistics to translate a player's batting record over many games into a batting average, which allows people to easily compare the batting records of different players.

How Scientists Use Statistics

Scientists are also interested in comparing things, but scientists use statistics for a wide range of purposes. Scientists rely on and use statistics to summarize, characterize, analyze, and compare data. Statistics is actually a branch of mathematics that provides scientists with important tools for analyzing and understanding their data.

Consider the experiment in which students studied mussels to see if the mussels were harmed by fertilizer in their water. Students collected data on mussel length and phosphate levels during this experiment. Some mussels in the control group grew more than some mussels in the experimental group, yet some grew less. How could the students turn this data into meaningful numbers?

Statistics Works with Populations Scientists use statistics to describe statistical populations. A *statistical population* is a group of similar things that a scientist is interested in learning about. For example, the dwarf wedge mussels shown in **Figure 8** are part of the population of all dwarf wedge mussels on the Ashuelot River.

Figure 8 ▶ Students found these dwarf wedge mussel shells in a muskrat den. These mussels are part of the statistical population of all dwarf wedge mussels on the Ashuelot River.

What Is the Average? Although statistical populations are composed of similar individuals, these individuals often have different characteristics. For example, in the population of students in your classroom, each student has a different height, weight, and so on.

As part of their experiments, the Keene High School students measured the lengths of dwarf wedge mussels in a population, as shown in **Figure 8.** By adding the lengths of the mussels and then dividing by the number of mussels, students calculated the average length of the mussels, which in statistical terms is called the *mean*. A **mean** is the number obtained by adding up the data for a given characteristic and dividing this sum by the number of individuals. For scientists, the mean provides a single numerical measure for a given aspect of a population. Scientists can easily compare different populations by comparing their means. The mean length of the mussels represented in **Figure 9** is about 30 mm.

The Distribution The bar graph in **Figure 9** shows the lengths of dwarf wedge mussels in a population. The pattern that the bars create when viewed as a whole is called the *distribution*. A **distribution** is the relative arrangement of the members of a statistical population. In **Figure 9,** the lengths of the individuals are arranged between 15 and 50 mm.

The overall shape of the bars, which rise to form a hump in the middle of the graph, is also part of the distribution. The line connecting the tops of the bars in **Figure 9** forms the shape of a bell. The graphs of many characteristics of populations, such as the heights of people, form bell-shaped curves. A bell-shaped curve indicates a *normal distribution*. In a normal distribution, the data are grouped symmetrically around the mean.

internet connect

www.scilinks.org
Topic: Statistics in Science
SciLinks code: HE4105
SciLinks Maintained by the National Science Teachers Association

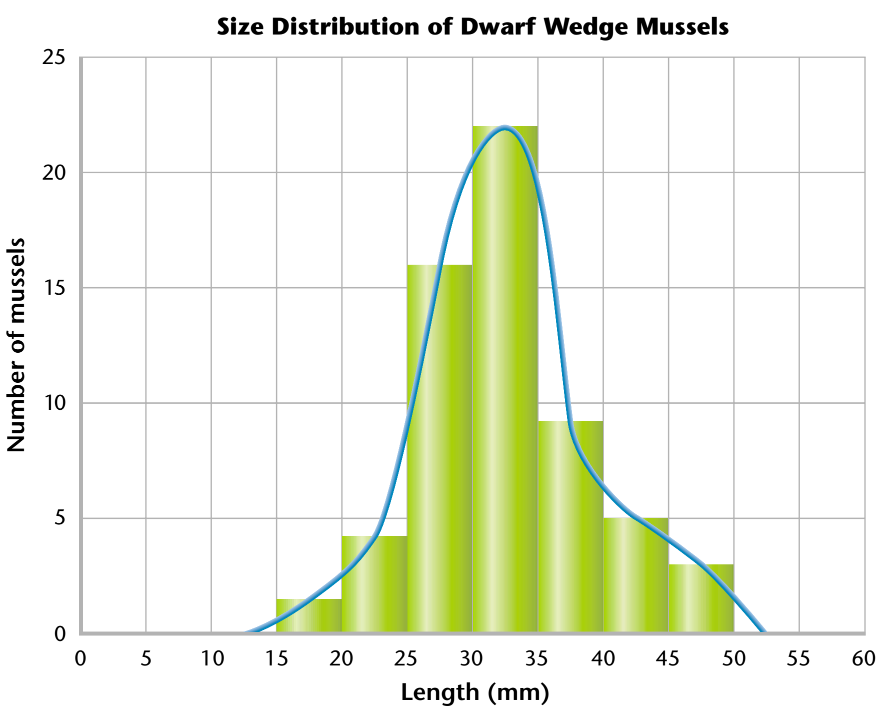
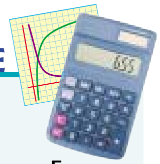


Figure 9 ▶ This bar graph shows the distribution of lengths in a population of dwarf wedge mussels. The location of each bar on the x-axis indicates length. The height of each bar represents the number of mussels as shown on the y-axis. For example, the second bar indicates that there are four mussels between 20 and 25 mm long.

MATHPRACTICE



Probability Probability is often determined by observing ratios or patterns. For example, imagine that you count 20 pine trees in a forest and notice that four of those trees have pine cones. What is the probability that the next pine tree you come across will have pine cones?

What Is the Probability? The chance that something will happen is called **probability.** For example, if you toss a penny, what is the probability that it will come up heads? Most people would say "half and half," and they would be right. The chance of a tossed penny coming up heads is ½, which can also be expressed as 0.5 or 50%. In fact, probability is usually expressed as a number between 0 and 1 and written as a decimal rather than as a fraction. Suppose the penny comes up heads 7 out of 10 times. Does this result prove that the probability of a penny coming up heads is 0.7? No, it does not. So what is the problem?

The problem is that the *sample size*—the number of objects or events sampled—is too small to yield an accurate result. In statistics, a **sample** is the group of individuals or events selected to represent the population. If you toss a penny 10 times, your sample size is 10. If you continue tossing 1,000 times, you are almost certain to get about 50% heads and 50% tails. In this example, the sample is the number of coin tosses you make, while the population is the total number of coin tosses possible. Scientists try to make sure that the samples they take are large enough to give an accurate estimate for the whole population.

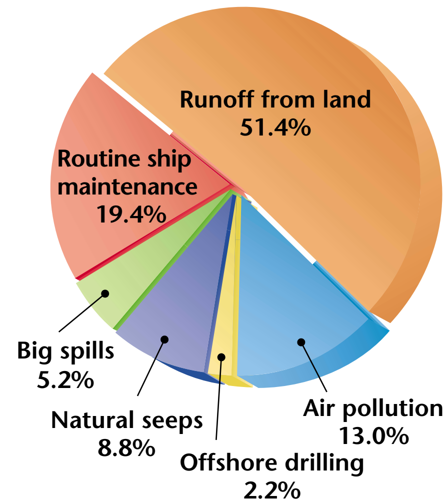
Statistics in Everyday Life

You have probably heard, "There is a 50 percent chance of rain today." **Figure 10** shows an example of a natural event that we often associate with probability—a thunderstorm. You encounter statistics often and use them more than you may think. People are constantly trying to determine the chance of something happening. A guess or gut instinct is probably just an unconscious sense of probability.

Figure 10 ▶ Most people are familiar with statistics regarding the weather, such as the chance, or probability, of a thunderstorm.

Understanding the News The news contains statistics every day, even if they are not obvious. For example, a reporter may say, "A study shows that forest fires increased air pollution in the city last year." We could ask many statistical questions about this news item. We might first ask what the average amount of air pollution in the city is. We could gather data on air pollution levels over the past 20 years and graph them. Then we could calculate the mean, and ask ourselves how different last year's data were from the average. We might graph the data and look at the distribution. Do this year's pollution levels seem unusually high compared to levels in other years? Recognizing and paying attention to statistics will make you a better consumer of information, including information about the environment.

Figure 11 ▶ People may worry about the risk of a big oil spill, but as the graph below shows, most of the oil polluting the oceans comes from ordinary sources.



Thinking About Risk In scientific terms, **risk** is the probability of an unwanted outcome. For example, if you have a 1 in 4 chance of failing a class, the risk is ¼, or 0.25. **Figure 11** shows a well-publicized environmental problem—oil spills. But as you can see in the pie graph, the risk of pollution from large oil spills is much smaller than the risk of oil pollution from everyday sources.

The most important risk we consider is the risk of death. Most people overestimate the risk of dying from sensational causes, such as plane crashes, and underestimate the risk from common causes, such as smoking. Likewise, most citizens overestimate the risk of sensational environmental problems and underestimate the risk of ordinary ones, as shown in **Table 2.**

Table 2 ▼

Perceptions of Risk by Experts and Ordinary Citizens

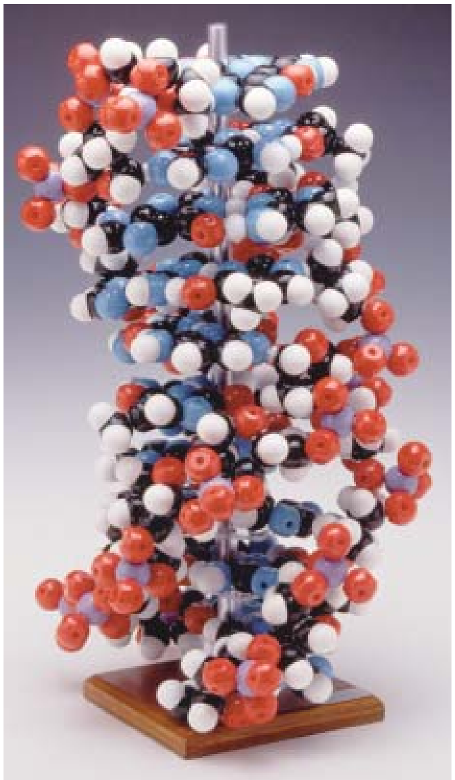
	High risk	Low risk
Experts	ozone depletion; global climate change	oil spills; radioactive materials; water pollution
Citizens	ozone depletion; radioactive waste; oil spills	global climate change; water pollution

Source: U.S. Environmental Protection Agency.

Connection to Law

Oil Tankers The Oil Pollution Act of 1990 was a direct response to the *Exxon Valdez* oil spill. The controversial bill had been debated for 14 years; it passed swiftly in the aftermath of the disaster. Under the law, all oil tankers operating in United States waters must be protected with double hulls by 2015.

Figure 12 ▶ This plastic model of a DNA molecule is an example of a physical model.



Models

You are probably already familiar with models. Museums have models of ships, dinosaurs, and atoms. Architects build models of buildings. Even crash-test dummies are models. **Models** are representations of objects or systems. Although people usually think of models as things they can touch, scientists use several different types of models to help them learn about our environment.

Physical Models All of the models mentioned above are physical models. *Physical models* are three-dimensional models you can touch. Their most important feature is that they closely resemble the object or system they represent, although they may be larger or smaller.

One of the most famous physical models was used to discover the structure of DNA. The two scientists that built the structural model of DNA knew information about the size, shape, and bonding qualities of the subunits of DNA. With this knowledge, the scientists created model pieces that resembled the subunits and the bonds between them. These pieces helped them figure out the potential structures of DNA. Discovering the structure of DNA furthered other research that helped scientists understand how DNA replicates in a living cell. **Figure 12** shows a modern model of a DNA molecule. The most useful models teach scientists something new and help to further other discoveries.

Graphical Models Maps and charts are the most common examples of *graphical models*. Showing someone a road map is easier than telling him or her how to get somewhere. An example of a graphical model is the map of the Denver, Colorado, area in **Figure 13.** Scientists use graphical models to show things such as the positions of the stars, the amount of forest cover in a given area, and the depth of water in a river or along a coast.

Geofact

Fossil-Fuel Deposits Fossil fuels, such as coal and oil, are often buried deep underground in particular rock formations. We find kinds of fossil fuels by drilling for rocks that indicate the presence of fossil fuels and then we make models of where the coal or oil is likely to be found.

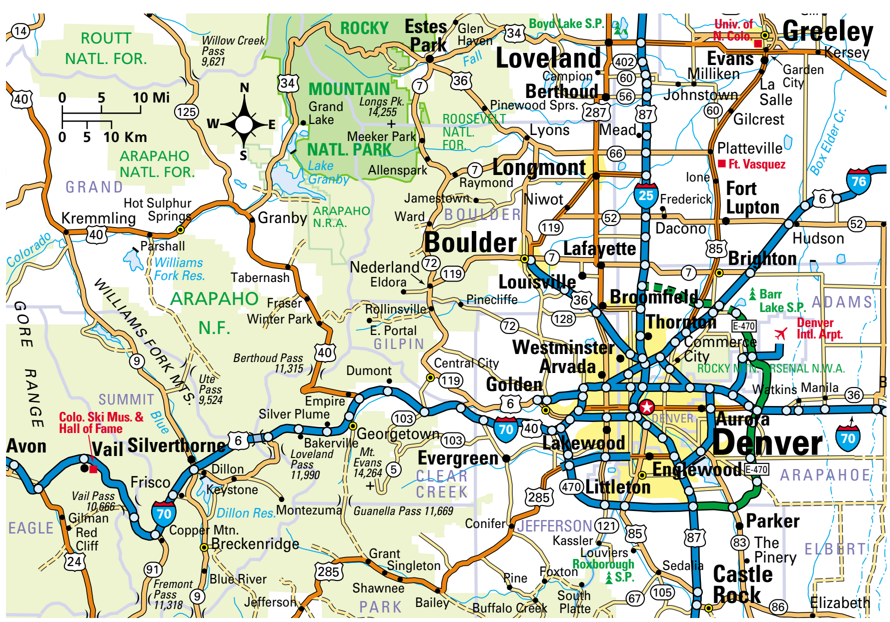


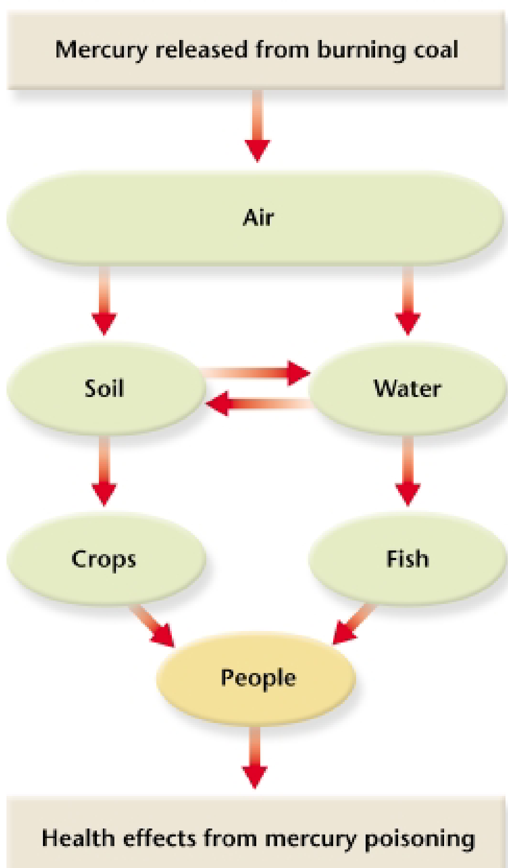
Figure 13 ▶ This map of the Denver, Colorado, area is an example of a graphical model.

Conceptual Models A **conceptual model** is a verbal or graphical explanation for how a system works or is organized. A flow-chart diagram is an example of a type of conceptual model. A flow-chart uses boxes linked by arrows to illustrate what a system contains and how those contents are organized.

Consider the following example. Suppose that a scientist is trying to understand how mercury, a poisonous metal, moves through the environment to reach people after the mercury is released from burning coal. The scientist would use his or her understanding of mercury in the environment to build a conceptual model, as shown in **Figure 14.** Scientists often create such diagrams to help them understand how a system fits together—what components the system contains, how the components are arranged, and how they affect one another.

Conceptual models are not always diagrams. They can also be verbal descriptions or even drawings of how something works or is put together. For example, the famous model of an atom as a large ball being circled by several smaller balls is a conceptual model of the structure of an atom. As this example shows, an actual model can be more than one type. An atomic model made of plastic balls is both a conceptual model as well as a physical model.

Conceptual Model of Mercury Contamination



FIELD ACTIVITY

Conceptual Model Accompany your class outdoors. Observe your surroundings, and write down observations about what you see. In your ***Ecolog,*** draw a conceptual model of something you observe. Your model should be of a system with components that interact, such as a small community of organisms.

internet connect

www.scilinks.org
Topic: Using Models
SciLinks code: HE4119
SCILINKS. Maintained by the National Science Teachers Association

Figure 14 ▶ This conceptual model shows how mercury released from burning coal could end up reaching people, where it could cause poisoning.

Mathematical Models A **mathematical model** is one or more equations that represents the way a system or process works. You can represent many common situations using math models.

For example, suppose that the grapes in a fruit basket at home are getting moldy. You notice that every day the mold covers two more grapes. Here is a mathematical model for the number of moldy grapes on Tuesday:

$$M_{Tue} = M_{Mon} + 2, \text{ where } M = \text{number of moldy grapes}$$

Mathematical models are especially useful in cases with many variables, such as the many things that affect the weather.

Because mathematical models use numbers and equations, people may think the models are always right. But weather models, for example, sometimes predict rain on dry days. In fact, people are the ones who interpret data and write the equations. If the data or the equations are wrong, the model will not be realistic and so will provide incorrect information. Like all models, mathematical models are only as good as the data that went into building them.

People may think of mathematical models as being confined to blackboards and paper, but scientists can use the models to create amazing, useful images. Look at the image of the San Francisco Bay Area in **Figure 15.** This is a "false color" digital satellite image. The satellite measures energy reflected from the Earth's surface. Scientists use mathematical models to relate the amount of energy reflected from objects to the objects' physical condition.

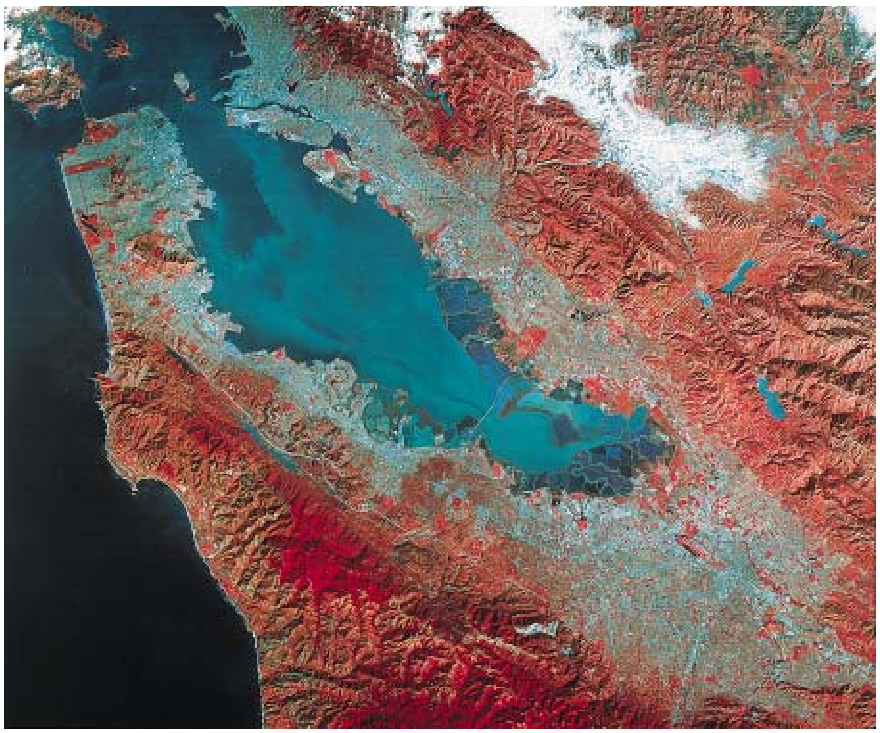


Figure 15 ▶ This is a satellite image of the San Francisco Bay Area. Scientists use mathematical models to understand the terrain from the way objects on the surface reflect light. In this image, healthy vegetation is red.

SECTION 2 Review

1. **Explain** why sample size is important in determining probability.
2. **Explain** what "the mean number of weeds in three plots of land" means.
3. **Describe** three types of models used by scientists.

CRITICAL THINKING

4. **Analyzing Relationships** Explain the relationship between probability and risk.
5. **Applying Ideas** Write a paragraph that uses examples to show how scientists use statistics. WRITING SKILLS
6. **Evaluating Ideas** Why are conceptual and mathematical models especially powerful?