questions. They can use their developing facility with rational numbers and proportionality to refine their observations and conjectures. For example, when considering the relative-frequency histogram in figure 6.27, most students would observe that “the paper plane goes between 15 and 21 feet about as often as it goes between 24 and 33 feet,” but such an observation is not very precise about the frequency. A teacher could press students to make more-precise statements: “The plane goes between 15 and 21 feet about 45 percent of the time.”

Box plots are useful when making comparisons between populations. A teacher might pose the following question about the box plots in figure 6.28:

From the box plots (in fig. 6.28), which type of plane appears to fly farther? Which type of plane is more consistent in the distance it flies?

From the relative position of the two graphs, students can infer that the two-clip plane generally flies slightly farther than the one-clip plane. Students can answer the second question by using the spreads of the data portrayed in the box plots to argue that the one-clip plane is more variable in the distance it travels than the two-clip plane.

Scatterplots are useful for detecting and examining relationships between two characteristics of a population. For example, a teacher might ask students to consider if a relationship exists between the length and the width of warblers’ eggs (activity adapted from Encyclopaedia Britannica Educational Corporation [1998, pp. 104–19]). She might provide the students with data and ask the students to make a scatterplot in which each point displays the length and the width of an egg, as shown in figure 6.29. Most students will note that the relationship between the length and the width of the eggs seems to be direct (or positive); that is, longer eggs also tend to be wider. Many students will also note that the points on this scatterplot approximate a straight line, thus suggesting a nearly linear relationship between length and width. To make this relationship even more apparent, the teacher could have students draw an approximate line of fit for the data, as has been done in figure 6.29. Students could apply their developing understanding of the slope of a line to determine that
the slope is approximately three-fourths and that therefore the ratio of the width to the length of warblers' eggs is approximately 3:4.

Teachers can also help students learn to use scatterplots to consider the relationship between two characteristics in different populations. For example, students could measure the height and arm span for groups of middle school and high school students and then make a scatterplot in which the points for middle school students are plotted with one color and the points for high school students are plotted with a second color. Students can make observations about the differences between the two samples, such as that students in the high school sample are generally taller than those in the middle school sample. They can also use the plots to examine possible similarities. In particular, if students draw an approximate line of fit for each set of points, they can determine whether the slopes are approximately equal (i.e., the lines are approximately parallel), which would indicate that the relationship between height and arm span is about the same for both middle school and high school students.

Because linearity is an important idea in the middle grades, students should encounter many scatterplots that have a nearly linear shape. But teachers should also have students explore plots that represent nonlinear relationships. For example, in connection with their study of geometry and measurement, students could measure the lengths of the bases of several similar triangles and use formulas to find their areas or graph paper to estimate their areas. Creating a scatterplot of the lengths of the bases and the areas will make evident the quadratic relationship between length and area in similar figures.

Teachers should encourage students to plot many data sets and look for relationships in the plots; computer graphing software and graphing calculators can be very helpful in this work. Students should see a range of examples in which plotting data sets suggests linear relationships, nonlinear relationships, and no apparent relationship at all. When a scatterplot suggests that a relationship exists, teachers should help students determine the nature of the relationship from the shape and direction of the plot. For example, for an apparently linear relationship, students could use their understanding of slope to decide whether the relationship is direct or inverse. Students should discuss what the relationships they have observed might reveal about the sample, and they should also discuss whether their conjectures about the sample might apply to larger populations containing the sample. For example, if a sample consists of students from one sixth-grade class in a school, how valid might the inferences made from the sample be for all sixth graders in the school? For all middle-grades students in the school? For all sixth graders in the city? For all sixth graders in the country? Such discussions can suggest further studies students might undertake to test the generality of their conjectures.

Understand and apply basic concepts of probability

Teachers should give middle-grades students numerous opportunities to engage in probabilistic thinking about simple situations from which students can develop notions of chance. They should use appropriate terminology in their discussions of chance and use probability to
make predictions and test conjectures. For example, a teacher might give students the following problem:

Suppose you have a box containing 100 slips of paper numbered from 1 through 100. If you select one slip of paper at random, what is the probability that the number is a multiple of 5? A multiple of 8? Is not a multiple of 5? Is a multiple of both 5 and 8?

Students should be able to use basic notions of chance and some basic knowledge of number theory to determine the likelihood of selecting a number that is a multiple of 5 and the likelihood of not selecting a multiple of 5. In order to facilitate classroom discussion, the teacher should help students learn commonly accepted terminology. For example, students should know that “selecting a multiple of 5” and “selecting a number that is not a multiple of 5” are complementary events and that because 40 is in the set of possible outcomes for both “selecting a multiple of 5” and “selecting a multiple of 8,” they are not mutually exclusive events.

Teachers can help students relate probability to their work with data analysis and to proportionality as they reason from relative-frequency histograms. For example, referring to the data displayed in figure 6.27, a teacher might pose questions like, How likely is it that the next time you throw a one-clip paper airplane, it goes at least 27 feet? No more than 21 feet?

Although the computation of probabilities can appear to be simple work with fractions, students must grapple with many conceptual challenges in order to understand probability. Misconceptions about probability have been held not only by many students but also by many adults (Konold 1989). To correct misconceptions, it is useful for students to make predictions and then compare the predictions with actual outcomes.

Computer simulations may help students avoid or overcome erroneous probabilistic thinking. Simulations afford students access to relatively large samples that can be generated quickly and modified easily. Technology can thus facilitate students' learning of probability in at least two ways: With large samples, the sample distribution is more likely to be “close” to the actual population distribution, thus reducing the likelihood of incorrect inferences based on empirical samples. With easily generated samples, students can focus on the analysis of the data rather than be distracted by the demands of data collection. If simulations are used, teachers need to help students understand what the simulation data represent and how they relate to the problem situation, such as flipping coins.

Although simulations can be useful, students also need to develop their probabilistic thinking by frequent experience with actual experiments. Many can be quite simple. For example, students could be asked to predict the probability of various outcomes of flipping two coins sixty times. Some students will incorrectly expect that there are three equally likely outcomes of flipping two coins once: two heads, two tails, and one of each. If so, they may predict that each of these will occur about twenty times. If groups of students conducted this experiment, they could construct a relative-frequency bar graph from the pooled data for the entire class. Then they could discuss whether the results of the experiment are consistent with their predictions. If students are accustomed to reasoning from and about data, they will understand that
discrepancies between predictions and outcomes from a large and representative sample must be taken seriously. The detection of discrepancies can lead to learning when students turn to classmates and their teacher for alternative ways to think about the possible results of flipping two coins (or other similar compound events). Teachers can then introduce students to various methods—organized lists, tree diagrams, and area models—to help them understand and compute the probabilities of compound events.

Using a problem like the following, a teacher might assess students’ understanding of probability in a manner that includes data analysis and reveals possible misconceptions:

For the one-clip paper airplane, which was flight-tested with the results shown in the relative-frequency histogram (in fig. 6.27), what is the probability that exactly one of the next two throws will be a dud (i.e., it will travel less that 21 feet) and the other will be a success (i.e., it will travel 21 feet or more)?

To solve this problem, students would need to understand the data representation in figure 6.27 and use ratios to estimate that there is about a 45 percent chance that a throw will be a dud and about a 55 percent chance that it will be a success. Then they would need to use some method for handling the compound event and deal with the fact that there are two ways it might occur. Students who understand all that is required might produce a tree diagram like the one in figure 6.30 to show that the total probability is \( \frac{198}{400} \), or .495, since each of the two possibilities—"dud first, then success" and "success first, then dud"—has a probability of \( \frac{99}{400} \).

![Tree Diagram](image)

**First throw**  
- less than 21 feet  
  - \( \frac{9}{20} \)
  - \( \frac{11}{20} \)

**Second throw**  
- less than 21 feet  
  - \( \frac{9}{20} \times \frac{11}{20} = \frac{99}{400} \)

- 21 feet or more  
  - \( \frac{11}{20} \times \frac{9}{20} = \frac{99}{400} \)

\[ \frac{99}{400} + \frac{99}{400} = \frac{198}{400} = .495 \]