Big Ideas in Mathematics
for Future Elementary Teachers

Big Ideas in Data Analysis and Probability

John Beam, Jason Belnap, Eric Kuennen, Amy Parrott, Carol E. Seaman, and Jennifer Szydlik
(Updated Summer 2016)
Dear Future Teacher,

We wrote this book to help you to see the structure that underlies elementary mathematics, to give you experiences really doing mathematics, and to show you how children think and learn. We fully intend this course to transform your relationship with math.

As teachers of future elementary teachers, we created or gathered the activities for this text, and then we tried them out with our own students and modified them based on their suggestions and insights. We know that some of the problems are tough – you will get stuck sometimes. Please don’t let that discourage you. There’s much value in wrestling with an idea.

All our best,

John, Jason, Eric, Amy, Carol & Jen
Hey! Read this. It will help you understand the book.

*The only way to learn mathematics is to do mathematics.*

Paul Halmos

This book was written to prepare future elementary teachers for the mathematical work of teaching. The focus of this module is data analysis – and this domain encompasses both statistical and probabilistic ideas. This text is *not* just intended to help you relearn your elementary mathematics; it *is* about teaching you to think like a mathematician and it *is* about helping you to think like a mathematics teacher. The National Council of Teachers of Mathematics (NCTM, 2000) writes:

> Teachers need several different kinds of mathematical knowledge – knowledge about the whole domain; deep, flexible knowledge about curriculum goals and about the important ideas that are central to their grade level; knowledge about the challenges students are likely to encounter in learning these ideas; knowledge about how the ideas can be represented to teach them effectively; and knowledge about how students’ understanding can be assessed (p. 17).

We are going to work toward these goals. (*Read them again. This is a tall order. In which areas do you need the most work?*) Throughout this book, we will ask you to consider questions that may arise in your elementary classroom.

*When you roll a pair of dice, is a one and a three a different outcome than a three and a one?*

*If the chance of rain is 40% on Saturday and 20% on Sunday, what is the likelihood that it will rain this weekend?*

*How many people do you need to sample in order to reasonably find out how many hours of TV second graders watch in a week?*

*What does it **mean** to say that the probability of a head when flipping a coin is ½?*

As mathematicians we will also convey to you the beauty of our subject. We view mathematics as the study of patterns and structures. We want to show you how to reason like a mathematician – and we want you to show this to your students too. This *way of reasoning* is just as important as any content you teach. When you stand before your class, you are a representative of the mathematical community. We will help you to be a good one.

No one can do this thinking for you. Mathematics isn’t a subject you can memorize; it is about ways of thinking and knowing. *You* need to do examples, gather data, look for patterns, experiment, draw pictures, think, try again, make arguments, and think some more. The big ideas of data analysis and probability are not always easy – but they are fundamentally
important for your students to understand and so they are fundamentally important for you to understand.

Each section of this book begins with a **Class Activity**. The activity is designed for small-group work in class. Some activities may take your class as little as 30 minutes to complete and discuss. Others may take you two or more class periods. The **Read and Study, Connections to the Elementary Grades**, and **Homework** sections are presented within the context of the activity ideas. No solutions are provided to activities or homework problems – you will have to solve them yourselves and discuss them with your class.

The mathematics content in this book prepares you to teach the Common Core State Standards for Mathematics for grades K - 8. These are the standards that you will likely follow when you are an elementary teacher, so we will highlight aspects of them throughout the text. In order for you to see how the mathematical work you are doing appears in the elementary grades, we have made explicit connections to **Bridges in Mathematics** from The Math Learning Center. This is the online elementary grades mathematics curriculum adopted by the Oshkosh Area School District. You will often be asked to go to the site **bridges.mathlearningcenter.org** to read or do problems. Your instructor will provide you with a code so that you can access these materials.

On this and the next page you will find the Common Core State Standards (CCSS) for school mathematics that are related to the topics in this book. There are also Standards for each grade related to number and operation, geometry, and measurement – you’ll study those in other courses. **Read the below Standards carefully; they will give you an overview of the elementary grades curriculum in data analysis and probability.**
Common Core State Standards for Measurement & Data

Grade One:
- Organize, represent, and interpret data with up to three categories; ask and answer questions about the total number of data points, how many in each category, and how many more or less are in one category than in another.

Grade Two:
- Generate measurement data by measuring lengths of several objects to the nearest whole unit, or by making repeated measurements of the same object. Show the measurements by making a line plot, where the horizontal scale is marked off in whole-number units.

- Draw a picture graph and a bar graph (with single-unit scale) to represent a data set with up to four categories. Solve simple put-together, take-apart, and compare problems using information presented in a bar graph.

Grade Three:
- Draw a scaled picture graph and a scaled bar graph to represent a data set with several categories. Solve one- and two-step “how many more” and “how many less” problems using information presented in scaled bar graphs. For example, draw a bar graph in which each square in the bar graph might represent 5 pets.

- Generate measurement data by measuring lengths using rulers marked with halves and fourths of an inch. Show the data by making a line plot, where the horizontal scale is marked off in appropriate units—whole numbers, halves, or quarters.

Grade Four:
- Make a line plot to display a data set of measurements in fractions of a unit (1/2, 1/4, 1/8). Solve problems involving addition and subtraction of fractions by using information presented in line plots. For example, from a line plot find and interpret the difference in length between the longest and shortest specimens in an insect collection.

Grade Five:
- Make a line plot to display a data set of measurements in fractions of a unit (1/2, 1/4, 1/8). Use operations on fractions for this grade to solve problems involving information presented in line plots. For example, given different measurements of liquid in identical beakers, find the amount of liquid each beaker would contain if the total amount in all the beakers were redistributed equally.
Common Core State Standards for Statistics & Probability

Grade Six:

Develop understanding of statistical variability.

- Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers. For example, “How old am I?” is not a statistical question, but “How old are the students in my school?” is a statistical question because one anticipates variability in students’ ages.

- Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

- Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

Summarize and describe distributions.

- Display numerical data in plots on a number line, including dot plots, histograms, and box plots.

- Summarize numerical data sets in relation to their context, such as by:
  - Reporting the number of observations.
  - Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
  - Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
  - Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

Grade Seven:

Use random sampling to draw inferences about a population.

- Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.

- Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. For example, estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be.
Draw informal comparative inferences about two populations.

- Informally assess the degree of visual overlap of two numerical data distributions with similar variabilities, measuring the difference between the centers by expressing it as a multiple of a measure of variability. For example, the mean height of players on the basketball team is 10 cm greater than the mean height of players on the soccer team, about twice the variability (mean absolute deviation) on either team; on a dot plot, the separation between the two distributions of heights is noticeable.

- Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. For example, decide whether the words in a chapter of a seventh-grade science book are generally longer than the words in a chapter of a fourth-grade science book.

Investigate chance processes and develop, use, and evaluate probability models.

- Understand that the probability of a chance event is a number between 0 and 1 that expresses the likelihood of the event occurring. Larger numbers indicate greater likelihood. A probability near 0 indicates an unlikely event, a probability around 1/2 indicates an event that is neither unlikely nor likely, and a probability near 1 indicates a likely event.

Grade Eight:

Investigate patterns of association in bivariate data.

- Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities.
- Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.
- Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association, informally fit a straight line, and informally assess the model fit by judging the closeness of the data points to the line.
- Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept. For example, in a linear model for a biology experiment, interpret a slope of 1.5 cm/hr as meaning that an additional hour of sunlight each day is associated with an additional 1.5 cm in mature plant height.
- Understand that patterns of association can also be seen in bivariate categorical data by displaying frequencies and relative frequencies in a two-way table. Construct and interpret a two-way table summarizing data on two categorical variables collected from the same subjects.
- Use relative frequencies calculated for rows or columns to describe possible association between the two variables. For example, collect data from students in your class on whether or not they have a curfew on school nights and whether or not they have assigned chores at home. Is there evidence that those who have a curfew also tend to have chores?
Big Ideas in Data Analysis and Probability

Table of Contents

To be a teacher requires extensive and highly organized bodies of knowledge. 
Shulman, 1985, p. 47

CHAPTER 1: COUNTING

Class Activity 1: Photographs ................................................................. p. 13
  Counting Ordered Groups
  The Multiplication Principle

Class Activity 2: Committees ............................................................... p. 21
  Counting Unordered Groups

Class Activity 3: Patterns in Committee Numbers ......................... p. 25
  Explaining Patterns in Pascal’s Triangle

Class Activity 4: From Photos to Committees .................................... p. 27
  Dealing with Duplicates

CHAPTER 2: PROBABILITY

Class Activity 5: Dice Sums ............................................................... p. 32
  The Language of Probability
  Experimental and Theoretical Analyses

Class Activity 6: Rat Mazes ............................................................... p. 39
  Tree Diagrams
  Compound Events

Class Activity 7: The Maternity Ward .............................................. p. 44
  Thinking about Sample Size
  The Law of Large Numbers

Class Activity 8: Probability Challenge! ......................................... p. 48
  Misconceptions about Chance
CHAPTER 3: STATISTICS and DEALING WITH DATA

Class Activity 9: Student Weight and Rectangles ..........................................................p. 54
   The Language of Sampling
   Random Samples

Class Activity 10: Snow Removal ..................................................................................p. 64
   Analyzing a Survey
   Statistically Speaking

Class Activity 11: Class Survey ......................................................................................p. 70
   Numerical and Categorical Data
   Survey Design

Class Activity 12: Name Games and In the Balance ......................................................p. 71
   Mean, Median and Mode
   Outliers
   Mean as Balancing Point
   Mean as Redistribution

Class Activity 13: Measuring Spread ..........................................................................p. 81
   Range, IQR and Mean Absolute Deviation
   A Criterion for Finding Suspected Outliers
   Boxplots

Class Activity 14: Matching Game ..............................................................................p. 88
   Distributions, Skewness and Symmetry
   Pie Charts, and Bar Graphs

Class Activity 15: Old Faithful Eruptions ...................................................................p. 99
   Forming Questions from Data
   Misleading Displays

CHAPTER 4: RATIOS, RATES, and PROPORTIONS

Class Activity 16: Let’s Be Rational ............................................................................p.104
   Ratios
   Bar Diagrams
Class Activity 17: The Watermelon Problem .................................................................p. 110
  Percents
  Percent Change
  The Language of Percents

Class Activity 18: How Do You Rate? .................................................................p. 117
  Rates
  Unit Conversions

Class Activity 19: Goldfish .................................................................p. 123
  Capture/Recapture
  Proportions

CHAPTER 5: RELATIONSHIPS AMONG SETS OF DATA

Class Activity 20: Shoe Size and Height.................................................................p. 129
  Scatterplots
  Outliers on a Scatterplot

Class Activity 21: Lots about Lines .................................................................p. 134
  Slope and Intercepts
  Recognizing Linear Data

Class Activity 22: Manatee Data .................................................................p. 139
  The Idea of Linear Regression
  Interpolation and Extrapolation

Class Activity 23: Correlation .................................................................p. 146
  The Idea of Correlation
  Causation Cautions

Class Activity 24: Caffeine and Heart Rate.................................................................p. 152
  The Language of Clinical Studies
  Confounding Variables

Class Activity 25: Music and Math.................................................................p. 156
  Clinical Study Analysis

Class Activity 26: Class Survey Revisited.................................................................p. 157
CHAPTER 6: PROBABILITY REVISITED

Class Activity 27: Casino Royale .................................................................p. 159
   Expected Value

Class Activity 28: Bad Bulb Binomial ......................................................... p. 164
   Binomial Probabilities

Class Activity 29: The Problem of Points ...................................................p. 169
   A Brief History of Probability

APPENDICES

Probability Simulation Project................................................................. p. 173

References ................................................................................................p. 177

Glossary ..................................................................................................... p. 178
Chapter One

Counting
Class Activity 1: Photographs

To the complaint, 'There are no people in these photographs,' I respond, 'There are always two people: the photographer and the viewer.'

Ansel Adams

1) How many different ways can you line up five people for a photograph if all five will be in the picture?

2) How many different ways can you line up 10 people for a photo? n people? Explain why your idea makes sense.

3) If there are a total of five people, how many different ways can you line them up 3 at a time for a photo? Two at a time? One at a time?

4) If there are \( n \) total people, how many different ways can you line up \( k \) of them at a time for a picture? (You may need to collect more data to help you think about this.)
Not everything that can be counted counts, and not everything that counts can be counted.

Albert Einstein

It is time to learn to ‘count.’ Basically there are two types of counting situations. The first is like the photographs problem where different orderings of objects are counted as different. In the second type of situation, different groups are counted as different but ordering within a group doesn’t matter. We’ll explore that in the next class activity. Both types of counting are based on a fundamental principle. But before we tell you about that principle, spend about 10 minutes solving the following problem using as many different representations (pictures, charts, etc) as possible.

At Dan’s Ice Cream Parlor you can pick a sugar cone or a regular cone. Then you can choose one of three sizes: large (3 scoops), medium (2 scoops), or small (1 scoop). Finally, you may select any of the five flavors: vanilla, chocolate, strawberry, chocolate chip, or mint.

If you cannot mix flavors (i.e., you can’t have one scoop of chocolate and one scoop of vanilla on your cone), how many different ice cream cones are possible?

We’ll leave you some blank space to work on the problem. Really do it before you turn the page.

Did you notice the italicized questions and requests in the previous paragraphs? We ask those types of questions throughout the text to help you to focus on important ideas. Also, they are part of your homework, so be sure that you answer them as you read.
Okay, there are lots of ways to represent a solution to the ice cream problem.

1) As a decision tree:

Choose a cone  Choose a size  Choose a flavor

- S
  - L
    - mint
    - strawberry
    - vanilla
    - chocolate
    - chocolate chip
  - M
    - mint
    - strawberry
    - vanilla
    - chocolate
    - chocolate chip

So we see that there are 2 sets of 3 sets of $5 = 2 \times 3 \times 5 = 30$ different ice cream cones.
2) As a table of some kind:

<table>
<thead>
<tr>
<th></th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sugar</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Regular</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

So we see there are $5 + 5 + 5 + 5 + 5 = 30$ different ice cream cones. We could see it as $2 \times 3 = 6$ boxes, each having 5 flavors so $6 \times 5 = 30$.

3) In words: You can pick any of two cone types and any of five ice-cream flavors, so that’s 10 different ice cream cones. Now each of those 10 comes in 3 sizes, so that makes a total of 30 possibilities.

Notice that in each case a multiplicative process emerges. That brings us to the fundamental principle of counting: the multiplication principle. It says this: Suppose you want to find the total number ways you can accomplish some task, and that task can be broken down into a sequence of subtasks. If there are $A$ different ways to do the first subtask, and for each of these ways there are $B$ different ways to the second subtask, then the total number of ways to do both subtasks in succession is $A \times B$.

For our problem, in order to make an ice cream cone, we had to select a cone type, a size, and a flavor of ice cream. So we can think the entire process as a sequence of three subtasks. There are 2 different ways to choose a cone type, 3 different ways to choose a size and 5 different ways to choose a flavor, so that is $2 \times 3 \times 5 = 30$ different ways to choose a cone and a size and a flavor. The decision tree on the previous page illustrates why it makes sense to multiply 2, 3, and 5 to find the total number of possible ice cream cones. Since there are 3 sizes and 5 flavors, there are 3 groups of 5, or 15 ways to choose a size and a flavor. Then when we choose one of the 2 kinds of cones, we see there are 2 groups of 15. In other words, the diagram illustrates that there are 2 groups of 3 groups of 5, for a total of $2 \times 3 \times 5 = 30$ different options.

You found that this principle was part of the structure of your solution to the photograph problems. If you think of the photographer as having three chairs she must fill, you can think of her task as a sequence of three subtasks: fill the first chair, then the second chair, and then the third.
If there are a total of five people to seat, then she has 5 choices for who sits in the first chair. For each of these 5 choices of who sits in the first chair, she can then make 4 choices of who sits in the next chair. This gives 20 different ways to put people in the first two chairs. Then for each of these 20 ways, there are three remaining choices for who sits in the last chair. So there will be $5 \times 4 \times 3 = 60$ different poses. *Explain why it makes sense to multiply these numbers together. List out all 60 different poses in a way that shows that there are 5 groups of 4 groups of 3 poses.*

**Connections to the Elementary Grades**

*A major responsibility of teachers is to create a learning environment in which students’ use of multiple representations is encouraged.*  
**NCTM Principles and Standards**

Situations involving representing and modeling multiplicative structures appear throughout the elementary curriculum. We will remind you of some of those models now. *As you read, compare these models to the representations we used when solving the ice cream problem and the photographs problem.*

**Grouping model of multiplication:** This model helps children to think of the product of say $4 \times 3$ as 4 groups with 3 in each:

```
*** *** *** ***
```

Here is a problem that might help children imagine a grouping model for $43 \times 22$: Candy comes in bags of 22 pieces. If you buy 43 bags of candy, how many pieces is that?

**Repeated Additions Model of Multiplication:** Similar to a grouping model, this is a model that helps children to think of multiplication (say $4 \times 3$) as $3 + 3 + 3 + 3$.

Here is a problem that might help children imagine a repeated additions model for $43 \times 22$: Jenni earned an allowance of 22 cents each day. How much will she have after 43 days?
Array Model of Multiplication: This is a model that leads children to think of the answer to a multiplication problem as a rectangular grid. Here is $4 \times 3$ as an array:

```
* * * *
* * * *
* * * *
```

Here is a problem that might help children imagine an array model for $43 \times 22$: I planted 43 rows of sunflowers and each row has 22 sunflowers. How many sunflowers did I plant?

Area Model of Multiplication: Multiplication ($4 \times 3$) can also be modeled as the area of a 4 by 3 rectangle (note that this is really just a special type of array):

```
[Rectangular grid]
```

Here is a problem that might help children imagine an area model for $43 \times 22$: A rectangular blackboard measured 43 inches by 22 inches. How many square inches of area is that?

Tree Model of Multiplication: Finally $4 \times 3$ could be represented with a tree showing four sets of three branches:

```
[Tree diagram]
```

Here is a problem that might help a child picture a tree model for the problem: $4 \times 3$ (Note that a tree is cumbersome if the numbers get too big so we’ll stick with the small example here.) Addie had 4 choices for a shirt to wear. For each shirt, she could choose any of three pairs of pants. How many different outfits could she make all together?

The multiplication principle of counting is an explicit topic for students in grades 4 or 5.
Homework

I've missed over 9,000 shots in my career. I've lost almost 300 games. I've failed over and over and over again in my life. And that is why I succeed.

Michael Jordan

1) Do all the italicized things in the Read and Study section.

2) In elementary school, children are encouraged to solve multiplication rule problems by drawing or physically modeling all the possibilities. Here is a problem for third grade. Solve it the way a child might by drawing all the possibilities.

You have three clean shirts and 2 pairs of clean pants. How many different outfits can you make?

3) A license plates starts with three letters and ends in three digits.
   a) How many plates can be made that start with an “A”?
   b) How many contain exactly two “A”s?
   c) How many start with a vowel and have no digit or letter repeated?

4) Suppose you are allowed to select different flavored scoops of ice cream on a cone at Dan’s Ice Cream Parlor. Now how many different ice cream cones are possible? Carefully explain your thinking and make clear any assumptions you are making.

5) You need to select a president and a vice president for your 12-member club. How many different ways can that be done? Explain how this problem is related to the photograph problem.

6) You are ordering a meal from fancy restaurant that gives you two choices of appetizers, three choices of main courses, and 6 choices of dessert. Explain why it makes sense to multiply 2 times 3 times 6 to figure out how many different meals you can order. Use a tree diagram and also an array model to illustrate the multiplication.

7) List all the different words (they don’t need to mean anything) can be made using all the letters in HAT.

8) How many different words can be made using all the letters in BLANKET? Explain why your answer makes sense.

9) You are ordering a 2-topping pizza from Sam’s Pizza. Sam allows you to choose any of 5 different toppings. How many different 2-topping pizzas are possible? In what ways does this situation differ from that of the photograph situation from the class activity? Explain.
Class Activity 2: Committees

To get something done, a committee should consist of no more than three (people), two of whom are absent.

Robert Copeland

1) List all the different four-member committees can be made with a group of four people. (Note: a committee of four with Ann, Bob, Cat and Don is the same committee as Bob, Cat, Ann and Don. Changing the order of members does not change the committee.)

2) List all the different three-member committees can be made with a group of four people.

3) List all the different two-member committees can be made with a group of four people.

4) List all the different one-member committees can be made with a group of four people.

5) List all the different zero-member committees can be made with a group of four people.

6) Suppose another person (Ernie) joins your group and now you have five total people. How many committees can be made containing zero members? 1 member? 2 members? 3 members? 4 members? 5 members?
Class Activity 2 Continued: Committees

Now we’ll organize this data in a big table. \( n \) is going to represent the total number of people. \( k \) will represent the committee size. The body of the table will show the number of possible \( k \)-sized committees that can be made with \( n \) total people. So, for example, if you have 4 total people and want to make committees of size 2, there will be 6 possible committees. See how much of this table you can complete.

<table>
<thead>
<tr>
<th>( n ) (total people)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Read and Study

Give a man a fish and he eats for a day.
Teach a man to fish and he eats for a lifetime.

Author Unknown

We hate to give much away at this point – so you’ll have to bear with us while we explain why we won’t just tell you how to compute committee numbers from a formula. By the way, if you recall how to do this, but don’t understand why the formula makes sense, then don’t use it. And don’t show anyone in your class how to use it either. Remember that math is not about knowing a formula. It is about making a formula, and it is about understanding why it makes sense. If you get a formula as a ‘gift’ then in some sense you lose the opportunity to own it, to really do some mathematics, and to understand it. We won’t take that from you, and we hope that you won’t take that opportunity away from your future students either.

You may recognize the table that you generated in the last class activity as a form of Pascal’s Triangle. To us this is what Pascal’s Triangle is: an organized chart of the committee numbers. And we can explain many of the features in Pascal’s Triangle in terms of thinking about committees. For example, why does it make sense that we would see symmetry in each line of the triangle?

```
1
1  1
1  2  1
1  3  3  1
1  4  6  4  1
1  5 10 10  5  1
```

Let’s look at a specific case of that symmetry: we see that the fifth row goes

```
1  5 10 10  5  1
```

*What does this row represent in terms of committees? Stop and think about this. Then explain.*
Well, if you have 5 total people: A, B, C, D and E, there is 1 empty committee (Does this bother you, by the way? In mathematics we define a committee in terms of its members. So any two committees with the same membership are, in fact, the same committee. So there is only one committee that contains no members at all - and any other committee with no members is that same committee: the empty committee.) There are 5 one-person committees, 10 two-member committees, etc.

So in order to explain the symmetry, we need to explain why it makes sense that with 5 total people there is the same number of committees of size zero as there is committees of size five, and why there is the same number of committees of size one as there is committees of size four, and why there is the same number of committees of size 2 as committees of size 3.

*Think about that last case now. Look at the group of five total people in the picture. Why does it make sense that there would be the same number of committees of size 3 as committees of size 2? Explain.*

In the next class activity, you will have the opportunity to discuss your thinking about this question in class and to identify and explain other patterns in Pascal’s Triangle as well.

**Homework**

*When angry, count to four; when very angry, swear.*

*Mark Twain*

1) Do all of the italicized things in the *Read and Study* section.

2) Explain how to use the table to answer this question: If you have 8 total people, how many 5-member committees are possible?

3) Explain why it makes sense in terms of committees that the second column of the table is just the pattern of natural numbers.

4) Spend ten minutes on this: look at the committee table and identify at least four more patterns that you see there. Write them down and bring them to class next time.
Art is the imposing of a pattern on experience, and our aesthetic enjoyment is recognition of the pattern.

Alfred North Whitehead

As we have discussed, the Committee Table is full of nice patterns that can be explained when imagined in terms of committees. Your job as a class is to identify patterns in the table (they can be trivial or complex) and then to make arguments for why the pattern exists based on thinking about what the committees the numbers represent.
Homework

And in the end it’s not the years in your life that count. It’s the life in your years.

Abraham Lincoln

1) Suppose there are 8 total people. Carefully explain why it makes sense that there is the same number of committees of size 3 as there are committees of size 5.

2) A portion of Pascal's Triangle is shown below:

```
1
1  1
1  2  1
1  3  3  1
1  4  6  4  1
1  5  10 10  5  1
... ...
```

a) Explain why it makes sense in terms of committees that the last diagonal of the Committee Table is all 1s.
b) Carefully explain why it makes sense in terms of committees that the circled 4 plus the circled 6 equals the circled 10.
c) Carefully explain why it makes sense in terms of committees that each row of the triangle sums to a power of 2.

3) List all the different three-letter "words" (they don’t have to mean anything) that can be made using all the letters in BOA. Now do the same for BOO. See if you can figure out how to deal with repeated letters in this case.

4) List all the different words (they don’t have to mean anything) you can make using all the letters in PAST. Now list all the words you can make using the letters in SASS. See if you can figure out how to deal with repeated letters in this case.
Class Activity 4: From Photos to Committees

A committee can make a decision that is dumber than any of its members.

David Coblitz

The purpose of this activity is to find a shortcut way to compute the number of committees in cases where it might be inconvenient to use the Committee Table and to understand why that method works.

For example, suppose you needed to know how many different 5-card hands could be dealt from an ordinary deck of 52 cards. That is like a committee problem where the total number of objects \( n \) is 52 and the committee size \( k \) is 5. By the way, we would typically write that number of committees as \( 52 \binom{5}{} \). By the end of today, hopefully you will be able to figure that out without having to make a giant Committee Table to do so. In order to do this, we’ll see if we can a relationship between photos and committees …

1) List all the photographs that can be made with four total people if three of them will be in each photo.

2) List all the committees of size three that can be made from a total of four people.

3) Now look at those lists to see if you can figure out how to adjust the number of photos (which is easy to compute) to get the number of committees. (You may need to try a couple more listing-examples. You decide.)
Read and Study

The hardest arithmetic to master is that which allows us to count our blessings.

Eric Hoffer

So you have found that the way to get from Photos to Committees is to divide out the duplicate stuff. This is an important idea, so we are going to spend some time on it. As usual looking at a specific example can help, so let’s take the example of five total people and look first at all the photos of size three that can be made, and then all the 3-member committees:

Here are the five people:

A   B   C   D   E

We already know there will be $5 \times 4 \times 3 = 60$ photos. Explain why we know that in advance of listing them.

Photos:

ABC  ABD  ABE  ACD  ACE  ADE  BCD  BCE  BDE  CDE
ACB  ADB  AEB  ADC  AEC  AED  BDC  BEC  BED  CED
BAC  BAD  BAE  CAD  CAE  DAE  DBC  CBE  DBE  DCE
BCA  BDA  BEA  CDA  CEA  DAE  CDB  CEB  DEB  DEC
CAB  DAB  EAB  DAC  EAC  EAD  DBC  EBC  EBD  ECD
CBA  DBA  EBA  DCA  ECA  EDA  DCB  ECB  EDB  EDC

Notice how we have a systematic way if listing things. That way we are less likely to miss any. Describe our system.

Now have a look at how the committees (in bold) show up in our list:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ABD</td>
<td>ABE</td>
<td>ACD</td>
<td>ACE</td>
</tr>
<tr>
<td>ACB</td>
<td>ADB</td>
<td>AEB</td>
<td>ADC</td>
<td>AEC</td>
</tr>
<tr>
<td>BAC</td>
<td>BAD</td>
<td>BAE</td>
<td>CAD</td>
<td>CAE</td>
</tr>
<tr>
<td>BCA</td>
<td>BDA</td>
<td>BEA</td>
<td>CDA</td>
<td>CEA</td>
</tr>
<tr>
<td>CAB</td>
<td>DAB</td>
<td>EAB</td>
<td>DAC</td>
<td>EAC</td>
</tr>
<tr>
<td>CBA</td>
<td>DBA</td>
<td>EBA</td>
<td>DCA</td>
<td>ECA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADE</td>
<td>BCD</td>
<td>BCE</td>
<td>BDE</td>
<td>CDE</td>
</tr>
<tr>
<td>AED</td>
<td>BDC</td>
<td>BEC</td>
<td>BED</td>
<td>CED</td>
</tr>
<tr>
<td>DAE</td>
<td>CDB</td>
<td>CEB</td>
<td>DEB</td>
<td>DEC</td>
</tr>
<tr>
<td>DEA</td>
<td>CEB</td>
<td>EBC</td>
<td>EBD</td>
<td>ECD</td>
</tr>
<tr>
<td>EDA</td>
<td>DCB</td>
<td>ECB</td>
<td>EDB</td>
<td>EDC</td>
</tr>
</tbody>
</table>

Do you see that each committee of size three can be arranged in 6 ways to make photographs? So that means that in this case, the number of committees $\times 6 = $ number of photos. Using our new notation, it looks like this:

$5C_3 \times 6 = 5P_3$
Or we could write it this way:

\[ \binom{5}{3} = \frac{5!}{3! \cdot 2!} \]

But this is only true in the case of three-member committees and photos. What will we need to divide by in the case of 2-member committees and photos? 4? k? That is the problem you solved today in class. But if you still don’t fully understand it, be sure to ask in your next class or talk to your instructor. Be persistent about this – it is a key idea.

**Homework**

*There are no secrets to success. It is the result of preparation, hard work, and learning from failure.*

_Colin Powell_

1) Suppose we have a class of 20 people.
   a) How many different groups of size four could we make?
   b) How many groups of size four could we make that don’t include Ann?
   c) How many groups of size four could we make that include Ann?
   d) Ann and Bob are not speaking to each other. How many groups of size four could be made if Ann and Bob cannot be in a group together?

2) How many total groups (all sizes) can be made from 8 total people? Explain.

3) How many different “words” (they need not mean anything) can be made
   a) using all the letters in ORANGE?
   b) using all the letters in APPLE?
   c) using all the letters in BANANA?

4) Make sure you understand how to figure out each of the following problems:
   a) How many photos can be made with 7 people if everyone wants to be in every photograph?
   b) How many groups of size 3 can be made from a group of 10 people?
   c) How many different bridge hands can be made from a deck of cards? (There are 52 total cards and a bridge hand contains 13 cards dealt at random.)
   d) How many photographs can be made with 13 people if only 6 people will be in each photograph at a time?

5) Calculate by hand:
   a) \( \binom{11}{3} \)
   b) \( 6P_6 \)
   c) \( 7P_4 \)
   d) \( nC_1 \)
   e) \( nP_n \)
6) Suppose we have 8 total people and want to take all possible photographs that contain 4 people. Carefully explain why it makes sense that the number of photographs = 8 \times 7 \times 6 \times 5. Be sure to explain the multiplication as well as the numbers involved.

7) Suppose there are 8 total people. Carefully explain why it makes sense that the (number of photographs of size 3) = 3! \times (number of committees of size 3).

8) Do the “Challenge Problem” about the dart board in Unit 1, Module 3, Session 1 of the Bridges in Mathematics Grade 2 Home Connections. What are some different ways that a second grader might “show their work” in solving this problem?
Chapter Two

Probability
Class Activity 5: Dice Sums

The excitement that a gambler feels when making a bet is equal to the amount he might win times the probability of winning it.

Blaise Pascal

Roll two dice at once and if the sum is 2, 3, 4, 5, 10, 11 or 12, your group wins. If the sum is 6, 7, 8, or 9, the instructor wins. Is this a fair game? Do this problem two ways. First, actually play the game lots of times and collect data to decide. Then, do a theoretical analysis of the problem to decide. (You will need to begin by deciding what it means that a game of chance is “fair.”)
Games of chance have been around for at least as long as people have recorded history, but it wasn’t until the 16th century that mathematicians first began to study the rules that might help answer questions such as “Is this game fair?” or “How likely am I to draw a royal flush in a game of poker?” We call the area of mathematics that contains these rules probability theory.

Simply put, probability is the study of chance or random events. Suppose we roll a standard six-sided die once. We don’t know the result of the roll in advance, so we say that the outcome is random. However, the outcome is not completely unpredictable. Our die has six sides and so we know the outcome will be one of six numbers: 1, 2, 3, 4, 5, or 6. We also know (assuming we have a fair die) that any one of these outcomes is equally likely; that is, if we roll the die many, many, many times and record all the outcomes, we will get approximately the same percentage of ‘1’s as we do ‘2’s as we do ‘3’s and so forth. Suppose we would like to be able to assign a number to express how likely it is that we get a 5 on one roll. In other words, we want to know the probability of getting a 5 on one roll of a standard 6-sided die. By the way, the use of good mathematical language is a habit. If you want to use mathematical language correctly with your elementary students, you must practice doing so now. You will find underlined vocabulary and ideas defined for you in the glossary. Go there and have a look at the definition of “probability” now.

In order to talk about the answer to this question, we must agree on a way to measure probability. Four main rules have been adopted:

1) The probability of an outcome is a number from 0 to 1. If an outcome is certain to occur (such as, at the equator the sun will rise tomorrow), then its probability is 1. If an outcome cannot happen (such as, the moon will fall into the Pacific Ocean tonight), then its probability is 0. If an outcome is neither certain nor impossible, then the probability will be some number between 0 and 1.

2) The sum of the probabilities of all possible outcomes of a random experiment is 1.

3) The probability that an event will occur is 1 minus the probability that it will not occur. That is, an event will either occur or it will not occur, and the sum of those two probabilities must be 1.

4) If two events are disjoint (they have no outcomes in common), the probability that one or the other will occur is the sum of the probabilities that each will occur by itself.
So we must think about how we can assign a number from 0 to 1 to the outcome “obtaining a 5 when rolling one die” that we will call its probability. To do so mathematicians use the language of sets. Here come the official definitions of random experiment, outcome, event, and sample space. A random experiment is an activity (such as rolling our 6-sided die) where the outcome (1, 2, 3, 4, 5, or 6) cannot be known in advance. A set of all possible outcomes, \{1, 2, 3, 4, 5, 6\}, is a sample space for that experiment and an event is any subset of the sample space (such as “rolling a 5” or “rolling an even number”). Note a subtle distinction here between an outcome and an event: an outcome is an element of the sample space, whereas an event is a subset of the sample space. So if the sample space is \{1, 2, 3, 4, 5, 6\} then ‘3’ is called an outcome, whereas \{3\} is called an event. \{1, 3, 5\} is another event.

List a few more events for this sample space using good curly-bracket set notation.

Does the empty set meet the definition of an “event?” Why or why not?

Okay, now by rule number 1, the probability of each outcome in the sample space will be a number between zero and one, and by rule number 2, the sum of all the probabilities of the outcomes in the sample space must equal one.

So how do we determine the probability of “rolling a 5?” Well that turns out to be pretty easy in the case where all the outcomes are equally likely to occur. (We will assume that our die is evenly balanced and a perfect cube - we call such a die fair – and so we are just as likely to obtain any one of the numbers from 1 to 6.) When outcomes are equally likely, the probability that any one occurs is just the ratio of 1 to the total number of outcomes. In the case of rolling one fair die, the probability of rolling a ‘5’ is \(\frac{1}{6}\).

When an event is more complicated (such as obtaining a sum of 3 when rolling two dice), we can use the same approach but we need to choose our sample space wisely. Suppose we are interested in the sum obtained when rolling two fair dice. If we list the possible outcomes for the sum, we get the sample space of \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}, which contains eleven outcomes. Unlike the earlier example of rolling one die however, these outcomes are not equally likely. For example, there are several ways we could obtain a sum of 7 (1 + 6, 2 + 5, etc.), but only one way we could obtain a sum of 12 (6 + 6). When the outcomes are not equally likely, we need a different way to determine the probability of each outcome. In this example, the probability of each sum is not \(\frac{1}{11}\).
One way to determine the probabilities of each sum is to consider a sample space of the random experiment of rolling two dice where the outcomes are equally likely. For example, we can think of the outcomes in this sample space as ordered pairs \((x, y)\), where \(x\) is the number on the first die and \(y\) is the number on the second die like you may have done when you worked on the class activity.

This sample space is shown in the table below. The advantage of using this sample space is that all outcomes here are equally likely.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1, 1)</td>
<td>(1, 2)</td>
<td>(1, 3)</td>
<td>(1, 4)</td>
<td>(1, 5)</td>
<td>(1, 6)</td>
</tr>
<tr>
<td>2</td>
<td>(2, 1)</td>
<td>(2, 2)</td>
<td>(2, 3)</td>
<td>(2, 4)</td>
<td>(2, 5)</td>
<td>(2, 6)</td>
</tr>
<tr>
<td>3</td>
<td>(3, 1)</td>
<td>(3, 2)</td>
<td>(3, 3)</td>
<td>(3, 4)</td>
<td>(3, 5)</td>
<td>(3, 6)</td>
</tr>
<tr>
<td>4</td>
<td>(4, 1)</td>
<td>(4, 2)</td>
<td>(4, 3)</td>
<td>(4, 4)</td>
<td>(4, 5)</td>
<td>(4, 6)</td>
</tr>
<tr>
<td>5</td>
<td>(5, 1)</td>
<td>(5, 2)</td>
<td>(5, 3)</td>
<td>(5, 4)</td>
<td>(5, 5)</td>
<td>(5, 6)</td>
</tr>
<tr>
<td>6</td>
<td>(6, 1)</td>
<td>(6, 2)</td>
<td>(6, 3)</td>
<td>(6, 4)</td>
<td>(6, 5)</td>
<td>(6, 6)</td>
</tr>
</tbody>
</table>

A possible sample space for a roll of two dice

Notice that there are 36 equally likely outcomes in this sample space and so the probability of each outcome is \(\frac{1}{36}\). We can use this sample space to compute the probability of the event “obtaining a sum of 3” (that’s the event \(((1, 2), (2, 1))\)) by counting the number of possible outcomes that have a sum of 3. We must be careful to make certain that we actually do count all possible outcomes that have a sum of 3. For example, in the sample space, the outcome \((1, 2)\) which gives \(1 + 2 = 3\) and the outcome \((2, 1)\) which gives \(2 + 1 = 3\) are two separate elements. That is, there are two ways to generate the sum of 3: 1) we can roll a 1 on the first die and then a 2 on the second die or 2) we can roll a 2 on the first die and then a 1 on the second die.

To help your upper elementary students think about this we suggest two things. First, have them do an experiment where everyone rolls a pair of dice for five minutes and tallies the number of times they roll a sum of 2 and separately the number of times they roll a sum of 3. In this way, they will see that the sum of 3 is twice as likely to occur. Second, have them each
hold one red die and one green die. Ask that they make a sum of 3. Then ask that they make a sum of three a different way so they can see that the outcome (1, 2) giving \(1 + 2 = 3\) and the outcome (2, 1) giving \(2 + 1 = 3\) are two separate elements in the sample space. Try this yourself. Now make a sum of 2 and notice that there is no other way to do this.

Rolling a 1 on the red die and a 2 on the green die is a different outcome from rolling a 2 on the red die and a 1 on the green die, but there is only one way to have a 1 on each die.

Okay, back to our sample space with 36 outcomes. Using probability rule 4, the probability of rolling two dice with a sum of 3, written \(P(\text{Sum}=3)\), is the sum of the probabilities of each outcome which equals \(\frac{1}{36} + \frac{1}{36} = \frac{2}{36} = \frac{1}{18}\). We can also compute \(P(\text{Sum}=3)\) as the ratio of the number of outcomes that have a sum of 3 to the number of total outcomes since all the outcomes are equally likely. So \(P(\text{Sum}=3) = \frac{2}{36}\).

This probability can be expressed as a fraction \(\frac{1}{18}\), as a decimal (0.056), as a percent (5.6%) or as odds (2 : 34, the chances of having a sum of three to the chances of having a sum other than three). Check all this to be sure it makes sense to you.

Let’s consider another example: What are my chances of drawing a face card from a standard deck of 52 playing cards? Before you read further, take a minute to see if you can figure it out.

Here is our thinking. Pay attention to the language we use as well as to the way we solve the problem. A standard deck of cards contains 52 cards, 13 of each of four suits. Each suit contains 3 face cards, the jack, the queen, and the king, so I have \(4 \times 3 = 12\) ways of drawing a face card and 52 ways of drawing any card. So the probability of drawing a face card from a deck of cards is \(\frac{12}{52}\). This probability can be expressed as a fraction \(\frac{3}{13}\), as a decimal (0.231), as a percent (23.1%) or as odds (12 : 40, the chances of drawing a face card to the chances of not drawing a face card). In this example, the experiment is “drawing a card from a deck of cards,” the event is “drawing a face card,” and the sample space contains 52 equally likely outcomes (the 52 cards in the deck).
Connections to the Elementary Grades

In grades 3 – 5, all students should understand that the measure of the likelihood of an event can be represented by a number from 0 to 1.

NCTM Principles and Standards, p. 176

In both of the Read and Study scenarios we have been discussing what we call theoretical probabilities, that is probabilities that are assigned based on assumptions about the physical uniformity and symmetry of an object (such as a die or a deck of cards). This may have left you with the impression that all random experiments can be analyzed in a theoretical manner. Not so. Consider the experiment of “flipping” a large marshmallow. Now it could land on a flat circular end or on its side (the curved part of the cylinder), so we’ll identify the sample space as the set: {end, side}. Children may think that these two outcomes each have a probability of ½ since there are two choices, but some experimenting with actual marshmallows will show that these are not equally likely outcomes. Furthermore, a theoretical analysis of the probabilities of each outcome would require knowledge of physics and geometry, and might depend on the surface, the temperature, or other factors. In cases like that, it is better to perform an experiment to determine the probability. In other words, it is best to have the children flip the marshmallow a large number of times and to use the data to estimate the probability of each outcome.

Homework

You’ll always miss 100% of the shots you don’t take.  
Wayne Gretzky

1) Do all the italicized things in the Read and Study and Connections sections.

2) If you decided that the game “Dice Sums” was unfair, change the rules so that the game is fair. Is there more than one set of rules that would give a fair game? If so, how many different sets of rules can you make?

3) Consider a new game called “Dice Differences.” Two dice are tossed and the smaller number is subtracted from the larger number. Player I scores one point if the difference is odd. Player II scores one point if the difference is even. (Note: Zero is an even number.) Is this a fair game? If so, make an argument that it is. If the game is unfair, change the rules so that the game is fair.

4) Read the “These Beans Have Got to Go!” activity from Unit 2 Module 1 Session 1 of the Bridges in Mathematics Grade 2 Home Connections, on page 29-34.
a. How would you place your beans if you were playing this game (and wanted to win).

b. How would you expect a typical 2nd grader to place their beans?

c. What answers would you expect that your future 2nd grade students might write to questions 3, 4 and 5 on page 34.

5) A bag contains two white marbles, four green marbles, and six red marbles. The random experiment is to draw a marble from the bag.

   a) Write a reasonable sample space for this experiment. Use set notation.

   b) What is the probability of drawing a white marble? A red marble? A green marble? Explain your answers.

   c) What is the probability of drawing a black marble? Explain.

   d) What is the probability of not drawing a green marble? Explain.

   e) How many marbles of what colors must be added to the bag to make the probability of drawing a green marble equal to 0.5?

6) A bag contains several marbles. Some are red, some are white, and the rest are green.

   The probability of drawing a red marble is $\frac{1}{5}$ and the probability of drawing a white marble is $\frac{1}{3}$.

   a) What is the probability of drawing a green marble? Explain.

   b) What is the smallest number of marbles that could be in the bag? Explain.

   c) Could the bag contain 48 marbles? If so, how many of each color? Explain.

   d) If the bag contains four red marbles and eight white marbles, how many green marbles does it contain? Explain.

7) Decide whether the following game is fair or unfair: Players have three yellow chips, each with an A side and a B side, and one green chip with an A side and a B side. Flip all four chips. Player I wins if all three yellow chips show A, if the green chip shows A, or if all chips show A. Otherwise, Player II wins. If the game is fair, make an argument that it is. If the game is unfair, change the rules so that the game is fair.

8) The 6 outcomes of rolling a fair die are equally likely; so are the 52 outcomes of a random draw of one card from a deck of cards. Name at least three more random experiments that generate equally likely outcomes. Name at least three that produce outcomes that are not equally likely.

9) Consider the random experiment of tossing three coins at once. Write a sample space for this experiment that has equally likely outcomes. Now write another sample space for this same experiment that does not have equally likely outcomes.
Class Activity 6: Rat Mazes

The trouble with the rat race is that even if you win, you're still a rat.  
Lily Tomlin

1) Suppose that a rat is sent into each of the below mazes at the ‘start.’ If the rat cannot go backwards, but otherwise makes all the decisions at random, what is the probability that she finds a cheese in each case? Explain your answers as you would to an upper elementary school student.

2) Make a rat maze to illustrate each of these problems:
   a) You flip a coin and then roll a die. What is the probability that you get a Head and then a multiple of three?
   b) You flip a coin three times, what is the probability that you get exactly two Heads?
Many situations require successive choices – like in the rat maze problems where the rat had to make at least two decisions about which path to take to get to the cheese. Here is another example: getting dressed for school in the morning requires several decisions. I must choose one of two pairs of jeans (only two are clean, a blue pair and a black one), one of three sweaters (red, purple, blue), and one of four pairs of shoes (let’s just call them A, B, C, & D). Let’s say that shoe pairs B and D don’t coordinate with the black jeans, but shoe pairs A & C can be worn with either pair of jeans. 1) What is the probability that I wear blue jeans, a red sweater, and C shoes? 2) What is the probability that my outfit contains the color blue? 3) What is the probability that I wear the B pairs of shoes?

Tree diagrams (or rat mazes) are useful in finding answers to questions like these. Below we have made a tree that counts all the possible outfits (or paths in the maze). In other words, the choices for getting dressed look like this.

Now assuming that all choices are made at random, there are a couple ways to compute the probabilities in questions 1-3 above. Take a minute to do so before reading further.

My personal favorite is the sending-rats-into-the-maze technique. I’ll send, say, 24 rats into the top of the maze. Then I expect 12 to go left down the blue jean choice, 4 of those to go down each sweater color, and 1 to come out of each of the shoe choices A, B, C and D. Of the 12 that will go right down the black jean choice, 4 will go down each sweater route and 2 will emerge from each shoe type A and C. Now we can use this information to answer the above questions. 1) Since 1 out of 24 come out of the blue jeans, red sweater, C shoes path, the
probability of that is \( \frac{1}{24} \). 2) Since all the blue-jeans rats and also the blue-sweater rats contain blue (that’s 12 + 4), \( \frac{16}{24} \) or \( \frac{2}{3} \) is the probability that my outfit contains blue. 3) Since 3 rats end up with B shoes, the probability that I wear those shoes is \( \frac{3}{24} \) or \( \frac{1}{8} \).

Okay, now you might be thinking, but how did she know to send 24 rats into the maze? Does that number matter? To see, you try it. First use 48 rats. Then try 100 to see what happens. Really do it. What did you learn?

Another solution is to use the probability equivalent of the multiplication counting principle to find the probability that we reach the end of each branch. For example, the probability that I wear blue jeans, a red sweater, and shoe pair A is \( \frac{1}{2} \times \frac{1}{3} \times \frac{1}{4} = \frac{1}{24} \). Why do we say we are using the multiplication counting principle here?

Find the probabilities of each of the outfits I could choose. Do you need to calculate each probability individually? Or can you make an argument that all of the outfits with blue jeans are equally likely? What about the outfits with black jeans?

Finally, let’s think about the last two questions we posed in the first paragraph using this second approach: What is the probability that my outfit contains the color blue? My outfit will contain blue if I wear blue jeans or if I wear the blue sweater. Half of the possible outfits include the blue jeans and one third of the remaining half of the outfits use the blue sweater. So a total of \( \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2} + \frac{1}{6} = \frac{2}{3} \) of the outfits contain blue; the probability that I wear blue is \( \frac{2}{3} \). Use this idea to find the probability that I wear the B pair of shoes.
Connections to the Elementary Grades

In grades 6 – 8 all students should compute probabilities for simple compound events, using such methods as organized lists, tree diagrams, and area models.

NCTM Principles and Standards, p. 248

Tree diagrams are part of the elementary grades curriculum because they give children a way to “see” all of the possible outcomes and to keep track of the probabilities of each choice that might be made in a compound event. By the way, it is not enough to know just one way to do things anymore. If you are going to be a teacher, you will need to make sense of the thinking of your students even if their methods are different from yours. Practicing different ways of thinking about and explaining problems will make you a better and more flexible teacher.

Homework

Once you learn to quit, it becomes a habit.

Vince Lombardi

1) Do all the italicized things in the Read and Study section.

2) Decide if each of the following statements is true or false. If it is true, explain why. If it is false, rewrite the statement so that it is true.
   a) If the probability of an event happening is \( \frac{1}{8} \), then the probability that it does not happen is \( \frac{8}{3} \).
   b) The probability that an outcome in the sample space occurs is 1.
   c) If an event contains more than one outcome, then the probability of the event is the sum of the probabilities of each outcome.
   d) If I have a 60% chance of making a first free throw and a 75% chance of making a second, then the probability that I miss both shots is 10%.

3) Make a rat maze to model each of the random experiments.
   a) I have a bag containing three chips. The red chip has a side A and a side B. The yellow chip has a side B and a side C. The blue chip has both sides marked A. I draw one chip at random from the bag and toss it.
   b) Three chips are tossed. The red chip has a side A and a side B. The yellow chip has a side B and a side C. The blue chip has both sides marked A.
4) Suppose the chips in #3 above are all tossed. Use your second rat maze to determine the probability of getting exactly
   a) one C  
   b) one B  
   c) one A  
   d) two Cs  
   e) two Bs  
   f) two As

5) Suppose that three chips are tossed as in #3. Player I scores a point if a pair of As turn up, and player II scores a point if a pair of Bs or a C turns up. Is this a fair game? If so, explain why. If not, change the rules to make it fair.

6) A basketball player has a 70% free throw shooting average. The player goes up for a one-and-one free throw situation (this means that the player shoots one free throw, and only if she makes it, she gets to attempt a second shot). What is the probability the player will make 0 shots? 1 shot? 2 shots?

7) Suppose that I toss a fair coin up to four times or until I get a Head (whichever comes first).
   a) Make a maze to model this random experiment.
   b) Write out the sample space using good notation.
   c) What is the probability that you get TTH?

8) Place 2 black counters and 1 white counter into a paper bag and shake the bag. Without looking, draw 1 counter. Then draw a second counter without putting the first counter back into the bag. If the 2 counters drawn are the same color, you win. Otherwise, you lose. What is the probability of winning? Does the probability of winning change if you replace the first counter before drawing the second counter? Why or why not? What is the probability of winning if you replace?

9) Here’s how you play our lottery: you write down any six numbers from the set \{1, 2, 3, \ldots, 35, 36\}. Then six winning numbers are drawn from that set at random (without replacement – so all of them will be different numbers).
   a) If you match all the winners in order, you win $5,000,000.00. What is the probability of that?
   b) If you match all the winners but not necessarily in order, you win $1,000,000.00. What is the probability of that?
   c) How are these problems related to our Photograph and Committee problems? Explain. Be specific. How would you explain this to someone?
Class Activity 7: The Maternity Ward

When I was a kid, I was surrounded by girls: older sisters, older girl cousins just down the street.... Each year I would wish for a baby brother. It never happened.  

Wally Lamb

Is it more likely that 70% or more of the babies born on a given day are boys in a small hospital or in a large hospital (or doesn’t it matter)? Your class should decide on a simulation to do in order to help you to answer this question.
Read and Study

“I think you’re begging the question,” said Haydock, “and I can see looming ahead one of those terrible exercises in probability where six men have white hats and six men have black hats and you have to work it out by mathematics how likely it is that the hats will get mixed up and in what proportion. If you start thinking about things like that, you would go round the bend. Let me assure you of that!”

Agatha Christie in The Mirror Crack’d

Our intuition often misleads us when we think about probabilities. Initially, you may have thought that the chances of having 70% or more of the babies born be boys would be better in a large hospital because it is more likely that more babies are born in a large hospital on one day and therefore more likely to have more boys. It is more likely that the number of babies born on a given day is larger in a large hospital than it is in a small one, but this does not mean that it is more likely that the percentage of boy babies will be 70% or more in the large hospital.

In fact, the larger the hospital (and therefore the larger the number of babies born on a given day), the more likely it is that the percentage of boys born will be 50% (the approximate probability of having a baby boy in a single birth). Mathematically, we call this result the law of large numbers, which states that in repeated, independent trials of a random experiment, (such as babies being born in a hospital), as number of trials (births) increases, the experimental probabilities observed (the probability that a baby born is a boy) will converge to the theoretical probability (in this case, 50%). In other words, the larger the sample, the more likely you are to get close to the theoretical result. So the more births there are on a given day (the larger the hospital), the more likely it is that the percentage of boys born is near 50%.

The fact that the trials should be independent is important. (Be sure to use the glossary for new mathematical terms; don’t assume that the mathematical meaning of a word is the same as the meaning the word might have in common usage.) For our purposes consider two events to be independent if the occurrence or non-occurrence of one event has no effect on the probability of the other event happening. In the hospital problem, the birth of a baby boy to one mother has no effect on the boy/girl outcome of the next birth in the hospital. The outcomes of the two births are independent of one another. What about the roll of two dice? Are the outcomes on each die independent? What about three flips of a coin? Is the head/tail outcome on each flip independent?

Not all events are independent. Consider the event “it will rain tomorrow” and the event “the wedding will be held outside tomorrow.” The probability that the wedding will be held outside is affected by whether it rains. Write down another example of two events that are not independent.
One way to teach children about the law of large numbers is to run lots of simulations like you did in class with the hospital problem. That is, we modeled the number of boy and girl births on a given day in various size hospitals, calculated the average percentage of boy births for each size hospital and then compared the results. In order to have this simulation work we needed a model (a coin perhaps) that has the approximately the same probability of success as does the event we are investigating. And we need to be certain that the number of ‘births’ we consider large is large enough for the law of large numbers to apply. A practical way to design collecting data on a large number of trials in your classroom is to pool everyone’s data.

You were asked to run a simulation in the class activity. How did you design your simulation? How might you have used the roll of one fair die? How did you decide the number of births that would represent a small hospital and the number of births that would represent a large hospital? What was the effect of pooling all the classes’ data?

Homework

I am always doing that which I cannot do, in order that I may learn how to do it.

Pablo Picasso

1) Do all the italicized things in the Read and Study section.

2) A coin has been tossed five times and has come up heads each time. Which of the following statements are true?
   a) There is an equal chance of coming up heads or tails on the next toss.
   b) The coin is more likely to come up heads on the next toss.
   c) The coin is more likely to come up tails on the next toss.
   d) I question whether the coin is fair.

3) A coin has been tossed five hundred times and has come up heads each time. Which of the following statements are true?
   a) There is an equal chance of coming up heads or tails on the next toss.
   b) The coin is more likely to come up heads on the next toss.
   c) The coin is more likely to come up tails on the next toss.
   d) I question whether the coin is fair.

4) Explain how the law of large numbers is related to your answers to the question in problems 2) and 3) above.

5) Explain how the law of large numbers is related to the determination of experimental probabilities. What does this law tell you about simulations you will do in your elementary classrooms?
6) A fair coin has been tossed a thousand times and the number of heads and tails recorded. Which of the following statements is most likely true?
   a) The number of tails is 601 and the number of heads is 399.
   b) 47% of the tosses were heads.
   c) There were 20 more heads tossed than tails.

7) Suppose we flip a coin 10,000,000,000 times (just suppose) and notice that 5,801,595,601 are heads. What can you say about the coin? Explain.

8) Design a simulation to determine the experimental probability of drawing a spade from a standard deck of cards. How can you be reasonably certain your result is close to the theoretical probability? Carry out your simulation and compare your result to the theoretical probability of drawing a spade from a deck of cards.
Class Activity 8: Probability Challenge!

If you don’t risk anything you risk even more.

Erica Jong

Discuss as a group whether each of the following statements is True or False. In each case, briefly explain your group’s choice.

1) T  F  If there is a 20% chance of rain on Saturday and a 40% chance of rain on Sunday, then there is a 60% chance of getting rain this weekend.

2) T  F  If there is a 20% chance of rain on Saturday and a 40% chance of rain on Sunday, then there is a 30% chance of getting rain this weekend.

3) T  F  The exact sequence HHTHHHTT is more likely that the sequence HHHHHHHH when tossing a fair coin 8 times.

4) T  F  In a family of 8 children, it is more likely that there are half boys and half girls than it is that there are all boys.

5) T  F  If John is a UWO student who is tall and athletic, it is more likely that he is a college basketball player and a business major than it is he is just a business major.

6) T  F  If John is a UWO student who is tall and athletic, it is more likely that he is a college basketball player than it is that he is a business major.

7) T  F  You are more likely to die in a terrorist attack than to die in a car accident.

8) T  F  You are less likely to win the lottery if your picks are the numbers 1, 2, 3, 4, 5, 6, than you are if your picks are 2, 31, 15, 19, 9 and 5.

9) T  F  If you have played and lost the lottery for 300 days in a row, then you are more likely to win tomorrow.
Nothing defines humans better than their willingness to do irrational things in the pursuit of phenomenally unlikely payoffs. This is the principle behind lotteries, dating, and religion.

Scott Adams

Humans are not very good at estimating chance. We fall victim to lots of fallacies (ways of thinking that are faulty) in the realm of probability. For example, in the class activity you talked about that fact that the numbers 1, 2, 3, 4, 5, 6 were as likely to be winning picks for the lottery as the numbers 2, 31, 15, 19, 9 and 5. Most people think that six consecutive numbers are much less likely to win because somehow they seem less random. This is an error of "local" randomness known as the representativeness fallacy. People with this fallacy believe that even small samples should look sufficiently random. For example, many people believe that the sequence HHTHHHTT is more likely than the sequence HHHHHHHH because the first sequence looks like a more typical result than the second, even though both sequences of 8 coin tosses have the same probability. The fallacy is confusing a particular outcome (HHTHHHTT) with a larger class of outcomes that it appears to represent (sequences with a mix of heads and tails).

Some people believe that after a string of losses they are “due” to win in order to balance out wins and losses. That’s also a type of representativeness called the gambler’s fallacy.

Here’s one more flavor of the representativeness fallacy. Consider the following question:

Liem is an extremely athletic looking young man who drives a fast car.
Is Liem more likely a pro-bowl professional football player or an accountant?

If you said Liem was more likely a pro-bowl professional football player, you fell for the base rate fallacy. The characteristics provided (extremely athletic looking young man and drives a fast car) may seem stereotypically more representative of professional football players than of accountants, causing people to ignore the rates at which pro football players and accounts occur in the population. In fact that there are far fewer professional football players around than there are accountants – so you should bet that Liem is the latter.
Kahneman and Tversky used the following problem to study another type of faulty logic called the **conjunction fallacy**.

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. Is it more likely that Linda is a bank teller, or both a bank teller and a feminist? Why?

*Answer the question above before reading the discussion that follows.*

It is more likely that she is just one thing (a bank teller) than that she is both things (a bank teller and a feminist). (*And* is a conjunction – hence the name of the fallacy.) However, the description of Linda given in the problem fits the stereotype of a feminist, whereas it doesn’t fit the stereotypical bank teller. Because it is easy to imagine Linda as a feminist, people often misjudge that she is more likely to be both a bank teller and a feminist than a bank teller (feminist or not).

Here is another way people reason incorrectly about chance: People tend to judge the probability of certain types of events by how easy it is to remember, or to imagine, the event occurring. We call this the **availability fallacy**. When an unusual event is presented vividly to our minds, it becomes more available; and in becoming more available, it seems more likely. In the aftermath of the events of September 11, 2001, many people chose to drive rather than to fly when making a long trip, because they judged it much more likely that they would be involved in a plane hijacking than in an automobile accident.

Availability can be very subtle and pervasive. Many parents in the US for example are afraid to let their children play outside or walk to school because they fear stranger abductions (due in part to the fact that stranger abductions always make national news and so are heard by and available to all of us). Ironically stranger abduction is far less likely to occur than children contracting diseases (like type II diabetes or problems related to obesity) due to sedentary lifestyles. And it is far more dangerous for most children to ride in a car to school than it is for them to walk. In fact, according to the Centers for Disease Control and prevention, the following are the top five most likely causes of childhood injury: car accidents, homicide (almost always at the hands of someone the child knows), child abuse, suicide, and drowning. What do parents fear the most? According to surveys conducted by the Mayo Clinic, they fear kidnapping, school snipers, terrorists, dangerous strangers, and drugs (as reported in *The New York Times*, September 19, 2010).
Here is an organization of the fallacies we have discussed in this section:

- Representativeness
  - (Base Rate Fallacy and Gambler’s Fallacy are types of representativeness)
- Conjunction Fallacy
- Availability

Of course there are many faulty ways of thinking about chance that don’t have fancy names. For example, consider the first two problems in the class activity about computing the chance of rain on the weekend if there is a 20% chance of rain on Saturday and a 40% chance of rain on Sunday. Based on this information, all we can really say is that the probability that it will rain on the weekend is somewhere between 40% and 60%. If we assume that every time it rains on Saturday, it also rains on Sunday, then the probability that it rains on the weekend would be just 40%.

*Under what assumption would the probability be 60%?*

In both of these extreme cases “Rain on Saturday” and “Rain on Sunday” would be dependent events. *Why?* If instead we assume that “rain on Saturday” and “rain on Sunday” are independent events, then we could calculate that the probability that it rains on the weekend is 52%.

*Make a tree diagram to show that this calculation is correct.*

---


---

**Homework**

*There are things which seem incredible to most men who have not studied mathematics.*

*Aristotle*

1) Do all the italicized things in the Read and Study section.
2) Bill is intelligent but unimaginitive. In school, he was strong in mathematics but weak in social studies and humanities. True or False? It is more likely that Bill is an accountant that plays jazz for a hobby than it is that Bill plays jazz for a hobby. Explain your reasoning and describe the type of fallacy this example represents, if any.

3) Suppose Mary is a university student who loves to play the piano. Is it more likely that she is a music major or an education major? Explain your reasoning and describe the type of fallacy this example represents, if any.

4) True or False? There are more words that end in “ING” than there are words whose second to the last letter is N. Explain your reasoning and describe the type of fallacy this example represents, if any.

5) True or False? If I flip a coin 7 times, I am just as likely to get the sequence TTTTTTT as I am to get the sequence TTHTTHT. Do you agree or disagree with this statement? Explain your reasoning and describe the type of fallacy this example represents, if any.

6) True or False? If I flip a coin 7 times, I am just as likely to get exactly two heads as I am to get exactly no heads. Explain your reasoning and describe the type of fallacy this example represents, if any.

7) True or False? I just flipped 7 tails in a row so now I am more likely to get a head to even things out. Explain your reasoning and describe the type of fallacy this example represents, if any.

8) A population of students has an average height of 68 inches. True or False? It is more likely that one person chosen at random from the population is taller than 72 inches than it is that a group ten people chosen at random has an average height of more than 72 inches. Explain your reasoning and describe the type of fallacy that this example represents, if any.

9) Find some innocent college student to answer questions 2) – 8) above. Did they show any of the misconceptions described in this section? Which ones? Bring the results to class. (Really do this.)

10) A baseball team facing a double-header has a 70% chance of winning the first game and an 80% chance of winning the second. Assume winning the second game is independent from winning the first game. What is the probability that they win exactly one of those games?
Chapter Three

Statistics and Dealing With Data
Class Activity 9: Student Weights and Rectangles

*The mathematics is not there ‘til we put it there.*

*Sir Arthur Eddington* (MQS)

What is the average weight of all the students registered for classes at your school this semester? In your groups, make a plan for some way to gather information to answer this question. Be specific about what you will do and when you will do it. At this point, time and cost is no object.

(Wait to discuss this question as a class before you turn the page.)
Class Activity 9 (continued)

Now, have a look at the population of 100 rectangles on the next page. Notice that each one has an area (a size). For example, rectangle number 74 has size 6 square units because it is composed of 6 small squares. Carefully select a sample of ten rectangles that you think will have an average size that is close to average size of all the rectangles on the page. Make sure to personally choose each of the ten. Then find the average size of your self-selected ten. Your class will then make a frequency graph showing everyone’s averages.

Next, each person should select 10 rectangles at random. (Your instructor will have a plan for doing this.) Then find the average size of your randomly-selected ten. Your class should make another frequency graph showing everyone’s averages.

Finally, estimate, based on the frequency graphs, what is the average size of all 100 rectangles on the page? Explain your answer. What does all this have to do with the student-weight problem?
A Population of Rectangles
Read and Study

*He uses statistics as a drunken man uses lamp posts – for support rather than illumination.*

Andrew Lang

This rectangle size problem contains a huge lesson in statistics. Huge. It is this: humans are no good at selecting samples. If you want unbiased information, random samples are the way to go. In order to talk more about this, we are going to introduce the language of sampling. Remember that we want you to learn these definitions and to use them.

First, a **sample** refers to a subset of a population. In our student weight problem, the **population** of interest was all students registered for classes at your school this semester. In an ideal world, we just would have every single member of the population come to get weighed and we would calculate the average weight. Unfortunately, this would likely prove impossible. Lots (most?) of the students would not show up to be weighed, and forcing them to do so would be unethical. Worse, those who would be willing to show up might be different than those who would refuse. People who were weight-conscious or overweight might be more likely to avoid stepping on your scale. So in the end, you would end up with a biased sample.

Another solution might be to send out anonymous surveys to the entire population asking each person to record his or her weight. But would you return that survey? Most people would not (mass surveys mailed out like that tend to have very low response rates) and those people who return them are again different from the population. Additionally, when a survey relies on participants to self-report their own data (like weight), sometimes that data isn’t accurate.

Okay, well we might try to call everyone or go door-to-door to collect information on student weights. That type of sampling yields much higher response rates but could prove very expensive and take a lot of time (particularly if you happen to have thousands of students in the population).

So, what are we to do? Here is the statistician’s solution. Forget about gathering data (information) from each and every member of the population and focus your time, energy and resources instead on getting information from just a random sample of the population. That way you can afford to go door-to-door and get data from most members of the sample. It still won’t be perfect. But it turns out that this is the very best you can do in a free society.

Why should the sample be random? Why not simply collect data on the weights of your friends, or maybe everyone in your dorm, or why not stand in front of the campus library and ask everyone who walks by for their weights? Because of the threat of ... **sampling bias**.
A sampling procedure is biased if it tends to over-sample (or under-sample) population members with certain characteristics. This will most surely happen if you self-select your sample – after all, the students you know have different characteristics than the population of your campus. What are some of those differences? Stop and write down at least three.

But your sample will also be biased, in a more subtle way, if you stand in front of the library and ask people “at random.” This is because you are not really asking random people. You are asking the people who walk by the library at that time of day (maybe oversampling people who study often or under-sampling those who have a class at that time). You are asking people who have a look at you and decide they want to answer your question (unwittingly oversampling people like yourself). Give another reason why the library sample might be biased.

Sampling bias is sneaky. It creeps in based on when you sample and where you sample and how you sample (as well as who you sample). Watch for it as you analyze studies.

So the solution is to generate a random sample and then to use your resources to get the best possible response rate from them. Here are two types of random sampling that are considered acceptable. The first is the best. It is called simple random sampling and it is mathematically equivalent to pulling names out of a hat. Every member of the population has the same chance of being in the sample as any other member of the population and any subset of the population has the same chance of being in the sample as any other subset of the same size. Stop. Read that last sentence again and think about it. If you had your registrar generate a list of all students enrolled in classes this semester and you assigned each student a number and then had your calculator generate 100 random numbers and used those numbers to get names, then that would be simple random sampling. If I lined up your class in a row and then took every third person to make my sample, would that be simple random sampling? Explain.

There are other forms of acceptable sampling. For example, cluster sampling is random sampling in stages. For example, if you randomly chose ten page numbers from the registrar’s list of students and then randomly chose 10 people from each of those pages, then that would be cluster sampling.

The U.S. Census has made use of cluster sampling when it has attempted to find those people who did not return their census forms. We’ll tell that story in a later section, for now let us say that a census means that you gather information from each and every member of a population. The U.S. Census is but one example of an attempt at a census. If you tried to get weight information from every student registered at your campus, then you too would be attempting a census. It is pretty much impossible to conduct a census with a large population because you never get data from everyone. If you collect data from a subset of the population, you are really conducting a survey.
Okay, here are two more terms to know: the numeric information that you want to know about a population (if you could get it) is called a **parameter**; the numeric information you get from a sample is called a **statistic**. In our example, the parameter we were after was the average weight of all the students registered for classes at your school. If we could do a census of that population, we could find that parameter. However, we have discussed some of reasons why we’re not likely to get that parameter. Instead, we will settle for performing a survey of a sample of the population, and what we will get is a statistic.

Here is our illuminating picture:

![Population vs Sample](image_url)

Ask everyone, it’s called a census. Ask a subset, it’s called a survey.

# computed from the population ↔ parameter.  # computed from the sample ↔ statistic.

The difference between the parameter and the statistic is called **sampling error**. Sampling error is an *idea*. In practice, when conducting a survey, you aren’t going to know what it is because all you will have is a statistic. You’ll probably never know the exact parameter. But we still use these words to talk about the ideas.

One source of sampling error we have already discussed: sample bias. Another source of sampling error is chance error. **Chance error** is error due to the diversity of the members of the population and the fact that you are just collecting data from a subset of them. If the population is diverse (different) in their weights, for example, then if you pick a random sample of them, you may end up with an average weight that is a bit different from the population weight. You may even, due to chance alone, get some people who are very thin or very heavy in the sample. Think back to your random sample of ten rectangles from the Class Activity. The average size of your random rectangles was likely close to, but not exactly, the average size of all 100 rectangles on the sheet. The difference was due to chance error.
Now think back to the average size of your ten self-selected rectangles. The difference between that number and the average size of all 100 rectangles on the sheet is due to both chance error and sample bias.

How do you reduce sample bias? Make your sample random.

How do you reduce chance error? Make your sample bigger.

So just how big does a sample have to be? It seems like you would need some significant sampling rate (like maybe 25% of the population), but it turns out that good surveys can be done with quite small samples. For example, national polling to predict the outcome of the presidential election is often done with only a few thousand people. That’s a few thousand out of more than 100 million likely voters.

**Sampling rate** [(# in the sample) ÷ (# in the population)] can be so small because people are not all that diverse from a sampling perspective. Think of it this way: suppose you are going to sample a big pot of soup to see how it tastes. As long as that soup is stirred up really well, and the chunks of meats and vegetables in there aren’t too big or too different, all you need is a small bite to know the flavor. By the way, what was the sampling rate of your random sample from the rectangle sheet? Figure it out.

---

**Connections to the Elementary Grades**

_In grades 3-5 all students should – propose and justify conclusions and predictions that are based on data and design studies to further investigate the conclusions and predictions._

_National Council of Teachers of Mathematics Principles and Standards for School Mathematic, p. 176_

Ideas of sampling are surprisingly complex – but groundwork for them can be laid in the elementary school. The NCTM Standards advocates that students in grades 3-5 should

“…begin to understand that many data sets are samples of larger populations. They can look at several samples drawn from the same population, such as different classrooms in their school, or compare statistics about their own sample to known parameters for a larger population… they can think about the issues that affect the representativeness of a sample – how well it represents the population from which it is drawn – and begin to notice how samples from the population can vary (p. 181).”

Notice the use of the technical language. What do they mean by “known parameters for a larger population?” Give some examples.
Suppose that your third graders have collected data from their class on the number of times last week that they eat hot lunch at school and found the following data displayed below as a frequency plot (each X represents 1 child in the class):

```
X   X
X   X
X   X
X   X
X   X
X   X
X   X
X   X
X   X
X   X
```

# of Hot Lunch Days last week

Of course there are many questions you might ask your students about these data – but there are some questions particularly related to sampling that should be asked and discussed.

Questions 1: Was there anything special about last week that might have made these data different than if we’d asked the same question next week or do you think last week was a typical week?

*In what ways might they respond to that question? What points about sampling could you make in each case?*

Question 2: How do you think the data might have been different if we’d asked the fifth graders the same question?

*In what ways might they respond to that question? What points about sampling could you make in each case?*

Questions 3: Would it be okay to label this graph “Number of Days Each Week that Children Eat Hot Lunch at Our School?” Why or why not?

*In what ways might they respond to that question? What points about sampling could you make in each case?*
Notice that these questions help children think about whether the data is representative of a population larger than the sample of 20 third-graders shown on the graph, and it helps them to begin to think critically about what the data says and what it does not say. Sampling becomes an explicit topic in the middle grades.

**Homework**

*We know that polls are just a collection of statistics that reflect what people are thinking in ‘reality.’ And reality has a well known liberal bias.*

Stephen Colbert

1) Do all the italicized things in the Read and Study and Connections sections.

2) Explain the difference between a census and a survey.

3) Explain the difference between a parameter and a statistic.

4) True or False? The bigger the sample the smaller the sample bias. Explain your thinking.

5) Sometimes a website or publication will advertise a poll asking its viewers to ‘call in’ or ‘click to respond’ to questions on a topic. In this case we say that the sample is self-selected. What types of biases are likely in this type of survey? Explain.

6) It is possible to bias results of a survey by asking questions using “charged” language. Consider these questions from Republican National Committee’s 2010 Obama Agenda Survey, and from a Democratic State Senator’s district survey. (Can you tell which is which?) Explain how each question might be asked in a less leading manner.

   a) “Do you believe that the best way to increase the quality and effectiveness of public education in the U.S. is to rapidly expand federal funding while eliminating performance standards and accountability?”

   b) “Do you support the creation of a national health insurance plan that would be administered by bureaucrats in Washington, D.C.?”

   c) “Do you believe that Barack Obama’s nominees for federal courts should be immediately and unquestionably approved for their lifetime appointments by the U.S. Senate?”

   d) “Do you support or oppose legislation to restore indexing of the homestead tax credit, which would boost our economy and help people to stay in their homes?”

   e) “Do you support or oppose limiting local control and weakening environmental standards to encourage iron ore mining in Wisconsin?”
7) The Columbus City Council wanted to know whether the 65,000 citizens of the city approved of a proposal to spend $1,000,000 to revitalize the downtown. The following study was conducted for this purpose. A survey was mailed to every 10th person on the list of registered voters in the city (it was mailed to 3500 people) asking about the proposal. Citizens were asked to complete the survey and return it to the Council by mail in a self-addressed, stamped envelope. 3100 people responded; of these, 75% supported the expenditure and 25% did not.

a) The population for this study is
   A) all registered voters in the city
   B) all citizens of the city of Columbus
   C) the 3500 people who were sent a survey
   D) the 3100 people who returned a survey
   E) none of the above

b) What was the sampling rate for this survey?

c) The "75%" reported above is a
   A) parameter
   B) population
   C) sample
   D) statistic
   E) none of the above

d) This survey suffers primarily from
   A) sampling bias
   B) chance error
   C) having a sample size that is too small
   D) non-response bias

e) True or False? The results of this survey may be unreliable because registered voters may not be representative of all the citizens. Explain your thinking.

f) True or False? The method of sampling in this survey could best be described as simple random sampling. Explain your thinking.

g) True or False? Sample bias could be reduced by taking a larger sample of registered voters. Explain your thinking.
Class Activity 10: Snow Removal

Errors using inadequate data are much less than those using no data at all.
Charles Babbage

In order to study how satisfied the 120,000 citizens of Green Bay, WI are with snow removal from city streets, I conducted the following study. I stood in front of the Green Bay Walmart on a Monday morning in February from 9:00 until noon and I asked every third person who walked in whether they were satisfied with snow removal from streets in Green Bay. One hundred and eight people said they were satisfied. Twenty-five people said they were not, and 18 refused to answer my question.

Answer the below question in your groups.

1) Describe the population for the study.

2) Did I conduct a census or a survey? Explain.

3) What was the sampling rate? (Write this in several different ways.)

4) What was the response rate? (Write this in several different ways.)

5) True or False? This study suffers primarily from non-response bias. Explain.

6) True or False? This study suffers primarily from sampling bias. Explain.

7) True or False? The results of my study may not be valid because city employees may have been part of the sample. Explain.

8) True or False? A main problem with my study is that my sample size is too small. Explain.

9) True or False? People should conclude from my study that Green Bay citizens are generally happy with snow removal from city streets.

10) Describe at least three ways to improve my study.
“Four out of five dentists surveyed recommend sugarless gum for their patients who chew gum.” You’ve heard that claim. It is perhaps the most well remembered statistical ‘fact’ in the history of marketing research. But what does it tell you? The makers of sugarless gum hope this statistic convinces you to purchase their product. But before you run out for a pack of sugarless gum, you should ask a few questions about this ‘fact.’ How were the dentists in the sample chosen? What was the response rate? What was the sample size? What was the exact question to which the dentists were asked to respond? And even if it turns out that the survey used exemplary methodology, note that the result does not suggest you that you should take up chewing sugarless gum. It only suggests that if you already chew gum, these dentists suggest you chew sugarless gum. All surveys are not created equal. In fact some are downright misleading. It is important for you to think critically about information based on survey data.

Surveys are meant to provide information about a population. They can help companies understand how to market products; they can help governmental agencies understand constituents; they can help campaigns predict public reactions to ideas or events; they can help second grade teachers understand which movies their students prefer. Here is an example of one real survey.

The United State Census Bureau uses surveys to understand the population of our country. (By the way, as teachers you can get good data for your students to analyze from the U.S. Census Bureau website. Keep that in mind.) For example, they conduct a survey of U.S. housing units every other year in order to make decisions about federal housing projects. In their words, their aim is to “[p]rovide a current and ongoing series of data on the size, composition, and state of housing in the United States and changes in the housing stock over time.” Here is a figure showing some of the results of the 2007 American Housing Survey (AHS). Take a few minutes to study it.
Figures 1, 2, and 3 highlight some findings from the 2007 American Housing Survey.

Figure 1.  
**Selected Features of Occupied Homes: 2007**  
(Percent of occupied units. The numbers in parentheses show table numbers where more data are available)

<table>
<thead>
<tr>
<th>Feature</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move in before 1950 (2-9)</td>
<td>0.5</td>
</tr>
<tr>
<td>Lacking some or all plumbing facilities (2-4)</td>
<td>1.1</td>
</tr>
<tr>
<td>Incomplete kitchen (2-4)</td>
<td>1.6</td>
</tr>
<tr>
<td>Seven or more stories in structure (2-2)</td>
<td>1.8</td>
</tr>
<tr>
<td>More than one person per room (2-3)</td>
<td>2.3</td>
</tr>
<tr>
<td>Three generations (2-9)</td>
<td>2.9</td>
</tr>
<tr>
<td>Sprinkler system inside home (2-4)</td>
<td>3.9</td>
</tr>
<tr>
<td>Home built in last 4 years (2-1)</td>
<td>5.2</td>
</tr>
<tr>
<td>One adult with children (2-9)</td>
<td>6.2</td>
</tr>
<tr>
<td>Manufactured/mobile homes (2-1)</td>
<td>6.3</td>
</tr>
<tr>
<td>Condominiums and cooperatives (2-1)</td>
<td>6.4</td>
</tr>
<tr>
<td>Homes built before 1920 (2-1)</td>
<td>6.9</td>
</tr>
<tr>
<td>No cars, trucks, or vans (2-7)</td>
<td>7.8</td>
</tr>
<tr>
<td>With Hispanic householders (2-1)</td>
<td>11.4</td>
</tr>
<tr>
<td>With Black-alone householders (2-1)</td>
<td>12.5</td>
</tr>
<tr>
<td>Below poverty (2-1)</td>
<td>12.8</td>
</tr>
<tr>
<td>Moved in past year (2-1)</td>
<td>16.2</td>
</tr>
<tr>
<td>One-person households (2-9)</td>
<td>27.1</td>
</tr>
<tr>
<td>Renter occupied (2-1)</td>
<td>31.7</td>
</tr>
<tr>
<td>Working carbon monoxide detector (2-4)</td>
<td>32.6</td>
</tr>
<tr>
<td>Owner-occupied units valued at $200,000 + (3-14)</td>
<td>32.7</td>
</tr>
<tr>
<td>Household with children (2-9)</td>
<td>34.2</td>
</tr>
<tr>
<td>Usable fireplace (2-7)</td>
<td>34.5</td>
</tr>
<tr>
<td>Separate dining room (2-7)</td>
<td>48.9</td>
</tr>
<tr>
<td>Two or more complete bathrooms (2-3)</td>
<td>49.7</td>
</tr>
<tr>
<td>Three or more bedrooms (2-3)</td>
<td>63.3</td>
</tr>
<tr>
<td>Dishwasher (2-4)</td>
<td>64.0</td>
</tr>
<tr>
<td>Garage or carport (2-7)</td>
<td>65.3</td>
</tr>
<tr>
<td>Owner occupied (2-1)</td>
<td>68.3</td>
</tr>
<tr>
<td>Single-family structures (2-1)</td>
<td>76.3</td>
</tr>
<tr>
<td>Public sewer (2-4)</td>
<td>80.2</td>
</tr>
<tr>
<td>Washing machine (2-4)</td>
<td>82.5</td>
</tr>
<tr>
<td>Air conditioning (2-4)</td>
<td>86.4</td>
</tr>
<tr>
<td>Working smoke detector (2-4)</td>
<td>92.4</td>
</tr>
<tr>
<td>Phone available (2-7)</td>
<td>97.8</td>
</tr>
</tbody>
</table>


---

These results drive national policy, so we hope that the AHS is well designed. Let’s have a look at the methodology. Read the following paragraph based on the information from the US Census Bureau website:

The AHS has sampled the same 53,000 addresses in each odd-numbered year since 1985. Those addresses were chosen in the following manner.

“First the United States was divided into areas made up of counties or groups of counties and independent cities known as primary sampling units (PSUs). A sample of these PSUs was selected. Then a sample of housing units was selected within these PSUs” (U.S. Census Bureau, 2007, p. B-1).

The AHS includes both occupied and vacant housing units and all people in those housing units. Data is collected using computer-assisted telephone and personal-visit interviews. When a housing unit was vacant, information was gathered from neighbors, rental agents, or landlords. The AHS is a voluntary survey meaning that households can decline to answer questions. Interviews for the data presented above were conducted from mid-April to September 2007.

Okay, should we believe the AHS results? Let’s talk through an analysis of the survey design.

First, what is the population for this survey and what is the sample?

According to the 2001 census, there are about 200,000,000 housing units total in the United States – that is the approximate population. The sample contained about 53,000 of these homes. That gives a sampling rate of about 0.0027%. This might seem small, but the sample was carefully selected using cluster sampling and so the sampling error will be fairly small too. (Why was this cluster sampling? Explain why based on the definition.) In other words, there is reason to believe that the households in the sample are representative of households in the United States.

Let’s go one step further in thinking about the quality of the data. It could be that the sample was well-chosen, but that many households gave faulty information (or simply refused to give information at all). Luckily the Census Bureau used the best possible method for collecting data – they called first, and if they could not reach a household by phone, they sent someone to the address to conduct an interview with either the occupants, or in the case of an empty unit, a neighbor or landlord or real estate agent. This likely produced a fairly high response rate. (In fact, the Bureau reports an 89% response rate. In the other cases either no one was at home even after repeated visits, the occupants refused to be interviewed, or the address could not be located.)

So, our take is that the methodology for the AHS is sound and the results are believable. We want you to learn to think like this whenever you need to judge the believability of a claim based on data.
Homework

*USA Today has come out with a new survey -- apparently 3 out of every 4 people make up 75% of the population.*

David Letterman

1) Do all the italicized things in the *Read and Study* section.

2) Spend a few minutes looking at the figure on occupied homes from the 2007 AHS. Write a paragraph highlighting the things that you find most interesting or notable about these results.

3) In order to find out how its readers felt about its coverage of the 2008 Presidential Election, a local newspaper (with a circulation of about 10,000) ran a front-page request in the Sunday paper that readers go to their website and complete an online anonymous survey that contained the following two questions:

   I. What is your political affiliation?
      ___ Democrat
      ___ Republican
      ___ Independent
      ___ Other

   II. How would you rate our coverage of the 2008 presidential election?
       ___ slanted left
       ___ slanted right
       ___ unbiased
       ___ no opinion

Six hundred readers responded. Of those, 70% of Democrats, 40% of Republicans, 66% of Independents, and 20% of “Other” said that the coverage was unbiased.

   a) Analyze the sampling method. Is the sample of respondents likely to be representative of the population the paper wants to represent? Give specific reasons why or why not.
   b) What does the data tell us about the paper’s coverage of the 2008 Presidential Election? Explain.
   c) Describe some ways to improve this survey.

4) Go to the United States Census Bureau website ([http://www.census.gov/schools/](http://www.census.gov/schools/)) and find the Teachers and Students link. Spend 15 minutes browsing the activities there. Then design a mathematics lesson for Grade 3 using the State Facts for Students link.
5) A researcher wants to predict the outcome of the Frostbite Falls mayoral election. In order to do this, she mailed a survey to every tenth person on a list of all 40,980 registered voters in the city. One thousand fifty-six people returned the survey. 34% of respondents said they supported Snidely Whiplash; 48% said they supported Dudley DoRight; the remaining 18% were undecided.

I. The population for this study is
   A) all mayoral-race voters in the city of Frostbite Falls
   B) all citizens of the city of Frostbite Falls
   C) the people who were sent a survey
   D) the people who returned a survey
   E) none of the above

II. The response rate was ____________.

III. The sampling rate was ____________.

IV. The "48%" reported above is a
   A) parameter
   B) population
   C) sample
   D) statistic
   E) none of the above

V. This survey suffers primarily from
   A) selection bias
   B) nonresponse bias
   C) chance error
   D) having a sample size that is too small

VI. True or False? The results of this survey may be unreliable because members of the city government may have been included in the sample. Explain your thinking.

VII. True or False? The method of sampling in this survey could best be described as simple random sampling. Explain your thinking.

VIII. True or False? Bias could be reduced by taking every 5th person on the list of registered voters rather than every 10th. Explain your thinking.
Class Activity 11: Class Survey

Everyone takes surveys. Whoever makes a statement about human behavior has engaged in a survey of some sort.

Andrew Greeley

Suppose your group has been hired by the American Federation of Teachers. You are charged with designing a six-item survey to be given to the students in your class for the purpose of better understanding some aspect of teacher preparation at your school. The survey should be written around a theme (cost, content, methods, quality of instruction, opportunities for work with children, job-outlook – or any other relevant theme chosen by your group). Three of your questions should request categorical information, and three should request numerical information. You are going to display and analyze this data in a future activity so write carefully-worded questions, and if you provide response options to some items, think carefully about those too. Once your survey is turned in to the instructor, he or she will reproduce it verbatim for your class to take; you will not have the opportunity to clarify items at the time the survey is given.

In what ways is your class representative of pre-service teachers at your school? In what ways might you not be representative of all pre-service teachers at your school?
Class Activity 12a: Name Games

I feel like a fugitive from the law of averages.

William H. Mauldin

1) Each person in the class should write his or her first name on sticky notes so that there is one letter on each of the person’s notes. (So for example, I would write J on one note, E on another, and N on the third.) Put your names in stacks on the board like we did in the example below. Now your class needs to rearrange your classes’ sticky notes (or even parts of sticky notes) so that every stack ends up with the same number of letters. Do that now.

Ami  Jen  Mo  Omar  ...

The mean (or average) of a set of numerical observations is the sum of all their values divided by the number of observations. Explain why it makes sense that your method just computed the mean name length for your class.

2) Now we will make a different type of graph, a bar graph. Each person should get one sticky note to place on the graph based on the length of his or her first name. For example if there were eight people in the class with names Mo, Ami, Jen, Omar, Karen, Steve, Sienna, and Juanita the bar graph would look something like this:

<table>
<thead>
<tr>
<th>frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>2  3  4  5  6  7</td>
</tr>
</tbody>
</table>

Name Lengths

Figure out a way to rearrange the notes to help you to find the mean name length on a graph like this. Why does it make sense that your method just computed the mean name length for your class? How is this method different from the method you used in question 1? Discuss this.

(This activity continues on the next page.)
Class Activity 12a (continued):

3) The **median** of a set of numerical observations is the middle observation when the data is placed in numerical order. In the case where there are two “middle” observations, the median is the average of them. Decide how you could use each type of graph above to help you find the median name length.

4) Which is likely to be larger for your class? The mean number of siblings or the median number? Explain. Then, collect the data and see if you are correct.
Class Activity 12b: In the Balance

*So divinely is the world organized that every one of us, in our place and time, is in balance with everything else.*

*Johann Wolfgang von Goethe*

Your task is to put weights on the numbered pegs of a balance scale to keep the system balanced.

1) Find several different arrangements of 3 weights so that the system will balance. Make a general conjecture about all the arrangements that will balance the system.

2) If you had one weight 1 to the right and two weights 2 to the left, predict where you would need to place a fourth weight to balance the system.

3) Find several ways to balance four weights where only one weight is on the left side. Make a general conjecture about all the arrangements of this type that will balance the system.

4) Predict some arrangements that will balance for 5 weights. Make a general conjecture about all the arrangements of 5 weights that will balance the system.

5) Suppose there are 4 cats with mean weight of 10 pounds. You know the weights of 3 of the cats. Those weights are: 4 pounds, 7 pounds, and 15 pounds. How can you use the balance to figure out the weight of the 4th cat? (Don’t do a calculation. Figure it out with the balance only.)

6) Use the balance to figure out this problem: You have taken four math quizzes. Your scores were 82, 67, 85, and 72. What do you need to score on the 5th quiz for your average quiz score to be 75? (Don’t do a calculation. Figure it out with the balance only.)
The average adult laughs 15 times a day; the average child, more than 400 times.

Martha Beck

There are three statistics that are commonly used to describe the center of a distribution of numerical data: mean, median and mode. The mean is computed by summing the values and dividing by the number of them. But a better way for children to think about mean is using the idea of “evening off.” The mean is the average value, the value each person would get if all the data was shared evenly. For example, suppose the picture below shows the cookies that belong to each child.

Keesha’s cookies:

Andi’s cookies:

Tim’s cookies:

Myosia’s cookie:

Carlos’ cookie:

As you saw in the class activity, in order to find the mean number of cookies, the children could imagine sharing cookies (or even parts of cookies) so that they each child has exactly the same number. By the way, is this data display like the one in problem 1) or the one in problem 2) from the Class Activity? Explain.
Keesha's cookies:

Andi's cookies:

Tim's cookies:

Myosia's cookies:

Carlos' cookies:

So the children see that the mean number of cookies is 3. In this case we didn’t need to break up cookies in order to share them, but you certainly should do examples like that with your students.

What we have illustrated here is the redistribution concept of the mean: that the mean value of a set of numbers is what each observation would get if the data was redistributed to that each observation had the same amount. This concept is essentially the same as the partitive concept of division: taking a total amount and sharing it equally into a number of groups, and seeing how much each group gets.

In class activity 12 b, we also discovered the balance point concept of the mean, namely, that the mean is the point in a distribution where the sum of the distances that the data is above the mean balances with the sum of the distances that the data is below the mean.

The mode of a set of data is the value that occurs most frequently. What is the mode for the cookie data? Sometimes a data set will have many modes because lots of values will occur with the same maximum frequency.

The idea of median is the idea of middle or center value. So we saw that the number of cookies held by the five people is 4, 3, 7, 1, and 1. If we put these values in numerical order, we get:
Why do we need to put them in numerical order first? How would you explain this to a child?

We see that the middle value is 3 (we’ll write $M = 3$.) In the case of an even number of values, we have no one middle value, and so the median is the average of the two middle values.

Find the mean, mode, and median of my bowling scores:

\[
\begin{array}{cccccccc}
112 & 114 & 120 & 123 & 127 & 127 & 130 & 167 & 220 \\
\end{array}
\]

Why does it make sense that the mean is larger than the median for these data? Explain.

We will define $Q_1$ (the first quartile) as the median of the data positioned strictly before the median when the data is placed in numerical order. For example using my bowling scores, 127 is the median.

\[
\begin{array}{cccccccc}
112 & 114 & 120 & 123 & | & 127 & | & 127 & 130 & 167 & 220 \\
\end{array}
\]

So the scores 112, 114, 120 and 123 are positioned before the median. The median of those scores is $(114 + 120)/2 = 117$. So $Q_1 = 117$.

$Q_3$ (the third quartile) is the median of the data positioned strictly after the median when the data is placed in numerical order: 127, 130, 167 and 220. $Q_3 = 148.5$.

Now we can report something we call a \textbf{five-number summary} for these scores:

Minimum = 112, \hspace{1em} Q_1 = 117, \hspace{1em} M = 127, \hspace{1em} Q_3 = 148.5, \hspace{1em} and Maximum = 220
By the way children may want to know if duplicates in the data can be thrown out. What would you say to them? Why? What about values of zero – could they be discarded? What would you say to the children about that?

We used a small data set for the purpose of illustrating these ideas. We note that for large data sets, these statistics (mean, median, mode, quartiles, and percentiles) are typically computed using a calculator or computer. We’ll give you the opportunity to use technology - if you have it - and work with a larger data set in a later class activity.

As a teacher, you will need to discuss with parents the meaning of their child’s score on a standardized exam, so let’s spend a few minutes on the ideas that you might find useful in that context. Typically these scores are reported as percentiles. If a child in your class scores in the $n^{th}$ percentile this means that “$n$ percent” of the children who took the test scored at or below that score. For example, if a child scores in the 80$^{th}$ percentile on a test, that means that 80% of the children who took the test scored at or below her score. Another way to think about this is to imagine lining up 100 children according to their test scores, this child would be number 80 from the bottom in the line (or number 20 from the top). In this context the median is the 50$^{th}$ percentile. What percentile is $Q_3$?

**Connections to the Elementary Grades**

*In grades 3-5 all students should use measures of center, focusing on the median, and understand what each does and does not indicate about the data set.*

_National Council of Teachers of Mathematics_  
_Principles and Standards for School Mathematics, p. 176_

Children in grades 3-5 can easily understand and compute median and mode, but typically the *idea* of mean is a more difficult topic. This is not because it is hard to *calculate* a mean; rather it is because it is more difficult to understand the *idea* of a mean.

Below is an example of an activity appropriate for children in grade 4 that represents an opportunity for the teacher to discuss mean as a *redistribution process* rather than as a computation.

*Explain how you could use sticky notes to have children use evening-off to find the mean of the below data without ever doing the standard computation. Really do it.*
The family sizes for children in our class are shown below. Make a bar graph for these data and then find the mean number of family members.

<table>
<thead>
<tr>
<th>Family</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jenson</td>
<td>8</td>
</tr>
<tr>
<td>Lewis</td>
<td>3</td>
</tr>
<tr>
<td>McDuff</td>
<td>7</td>
</tr>
<tr>
<td>Earles</td>
<td>4</td>
</tr>
<tr>
<td>Seaman</td>
<td>5</td>
</tr>
<tr>
<td>Ortiz</td>
<td>4</td>
</tr>
<tr>
<td>Peters</td>
<td>6</td>
</tr>
<tr>
<td>Todd</td>
<td>5</td>
</tr>
<tr>
<td>Miene</td>
<td>3</td>
</tr>
</tbody>
</table>

We know that the standard computation (add the values, then divide by the number of values) is quicker and ‘easier’ to do. But as an elementary mathematics teacher, your job is to help your children understand big ideas — and not to give them shortcuts that allow them to get right answers without understanding the ideas. Children must learn that mathematics makes sense, and that thinking (instead of memorizing) is the way we do mathematics.

**Homework**

*We cannot teach people anything; we can only help them discover it within themselves.*

*Galileo Galilei*

1) Do all the italicized things in the *Read and Study* section.

2) Do all the italicized things in the *Connections* section.

3) Look at the “Fulcrum Investigation” in Unit 8, Module 1, Session 3 of the 4th grade *Bridges in Mathematics* curriculum. In this activity, students are given a pencil, a 12 inch ruler and several “tiles” (weights or pennies) all of the same weight. By placing 6 tiles on “0” on the ruler, they can simulate a 60-pound fourth grade student (Sarah) sitting on the end of a seesaw. Another weight is placed on the other end of the seesaw (at 12 inches). If the fulcrum is in the middle of the pencil (at 6”), then it would take 6 tiles (representing 60 pounds) at the other end to “lift” or balance Sara. Using an actual pencil, ruler, and weights (such as pennies), predict then experiment how
much weight it takes to balance Sarah if the fulcrum is placed at the 4 inch mark, the 5 inch mark, the 7 inch mark, and the 8 inch mark.

a. Answer questions 1-6 in the “Fulcrum Investigation”.
b. How can you use the concept of the Mean to answer questions 5 and 6 in this investigation?
c. A “Challenge Question” in the next session asks “If Sarah can’t move the fulcrum from the middle of the seesaw, what can she do to balance with the 40-pound first grader? Why does this method work?” Answer this question. How do you think fourth graders would be able to this out? How can you use the concept of the mean to figure this out?
d. Repeat part c, but now suppose she was to balance with the 100-pound 7th grader without moving the fulcrum.

4) Here is a data set of the number of years members of the math department have been employed at my school: 7, 25, 7, 15, 8, 23, 27, 6, 1, 3, 28, 22, 24, 9, 15, 14, 18, 8. Compute the mean, median and mode for these data. What do these data tell you about the group of mathematics faculty at this school?

5) Decide whether each of the following is True or False. In each case, argue that you are right.

a) T F If ten people took an exam then it is possible that all but one of them scored less than the mean.
b) T F If ten people took an exam then it is possible that all but one of them scored less than the median.
c) T F If ten people took a 100 point exam, it is possible that the mean and median are 90 points apart.
d) T F If ten people took a thousand point exam, it is possible the mean and the median are 90 points apart.
e) T F The mean income of your town is larger than the median income.

6) Suppose that ten people graduated with a degree in computer science from your school and now their mean annual income is $75,000. If money is your only concern, does this convince you to major in computer science? Explain.

7) Suppose that ten people graduated with a degree in computer science from your school and now their median annual income is $75,000. If money is your only concern,
does this convince you to major in computer science? Is it more convincing or less convincing than the scenario in the previous problem? Explain.

8) Suppose that one hundred people graduated with a degree in computer science from your school and now their mean annual income is $75,000. If money is your only concern, does this convince you to major in computer science? Explain. Is this more or less convincing than the scenarios in the previous two problems? Explain.

9) Ann needs a mean of 80 on her five exams in order to earn a B in her class. Her exam scores so far are 78, 90, 64 and 83. What does she need to get on her fifth exam to earn the B? Do this problem in at least two different ways. One way should use the redistribution concept of the mean. Another should use the balance point concept of the mean.

10) A way to think about the median of a set of numbers is that it is the value that splits the distribution into two equal ‘counts.’ Consider this frequency graph of 18 scores on a quiz:

```
Score on a Quiz (out of ten possible points)

1  2  3  4  5  6  7  8  9  10
X  X
X  X  X
X  X  X  X
```

a) Find the median score and explain, as you would to a student, why it makes sense.
b) We say that this distribution is skewed left because it has an asymmetric tail (or even outliers) to the left. In skewed left data, do you expect the mean to be higher or lower than the median? Explain your thinking.
c) Now imagine that the distribution above shows 18 equal weights arranged on a seesaw. If you wanted the seesaw to balance, you would need to place the fulcrum at the mean. In other words, you can think about mean as the balancing point of the distribution. Find the mean and see if it looks like the balancing point. Why does it make sense that the mean is smaller than the median for these data?

11) A child in your fourth-grade class scored a 200 out of 500 on a standardized mathematics assessment in which the average score was 240. This student is reported as scoring in the 35th percentile. You need to explain to the parents exactly what this means. What do you tell them? Are you concerned about this child’s performance? Explain your thinking.
Class Activity 13: Measuring the Spread

It’s not that Good doesn’t triumph over Evil, it’s that the point spread is too small.

Bob Thaves

Here are the exam scores for two students, Andreas and Belle:

Andreas: 45, 57, 69, 69, 80, 94

Belle: 54, 70, 71, 72, 72, 77

Which of these students had better scores? Explain how you decided before you read further.

Notice that these students’ scores are very similar based on the measures of center (mean and median) – but they are not so similar based on the spread of their scores. Spread is an idea that is captured by different statistics. Here are three common measures of spread:

1) **Range** = maximum - minimum. Compute the range for each student’s scores. What is it that range measures? Be as specific as you can.

2) **Inter-Quartile Range (IQR)** = $Q_3 - Q_1$. Compute the IQR of each student’s scores. What is it that the IQR measures? Be as specific as you can.

3) **Mean Absolute Deviation (d)** = \( \frac{1}{n} \left[ |(x_1 - \overline{y})| + |(x_2 - \overline{y})| + |(x_3 - \overline{y})| + \ldots + |(x_n - \overline{y})| \right] \). In this formula \( n \) is the number of values, \( \overline{y} \) is the mean, and each “\( x \)” is a specific value. Recall that, for example \(|-3|\) means “the absolute value of -3.” Compute the mean absolute deviation of Andreas’ scores and then Belle’s as a group. What is it that the mean absolute deviation measures? Be as specific as you can.

(This activity continues on the next page.)
Class Activity 13 continued:

Now we have the machinery we need to identify an outlier in the data. First, a definition:

An **outlier** is a specific observation that lies well outside the overall pattern of the data.

You might ask, what does it mean to be “well outside” the overall pattern of the data? Glad you asked. We have a criterion for that. It is called the **1.5 × IQR criterion**: a value is considered an outlier if it falls more than $1.5 \times \text{IQR}$ above $Q_3$ or if it falls more than $1.5 \times \text{IQR}$ below $Q_1$.

Let’s use this criterion to test Andreas’ scores for outliers: $45, 57, 69 \mid 69, 80, 94$

$M = (69 + 69)/2 = 69$ and sits in the position held by the bar in the above data set, so the scores 45, 57 and 69 are positioned below the median (M). $Q_1 = 57$ and likewise, 69, 80 and 94 are positioned above the median so $Q_3 = 80$.

$$\text{IQR} = Q_3 - Q_1 = 80 - 57 = 23 \quad 1.5 \times \text{IQR} = 34.5$$

So an value is an outlier if it is more than 34.5 points above $Q_3$ or more than 34.5 points below $Q_1$. In other words, outliers in Andreas’ data are bigger than $80 + 34.5 = 114.5$ or smaller than $57 – 34.5 = 22.5$. Since there are no scores that big or that small in his data, he has no outliers based on the criterion.

Check Belle’s scores for outliers using our criterion.

How can it be that Belle has an outlier among her scores when Andreas does not? Explain.
In the Class Activity you worked with three different statistics that describe the spread of the data. The range shows you the entire spread of the data set. The IQR is a measure of the spread of the middle half of the data. The mean absolute deviation measures how spread out the data is (on average) around its mean.

We will use two primary statistics for determining the center of a distribution of data: mean and median, and two primary measures of spread of the data: mean absolute deviation and interquartile range (IQR). So, how do they fit together and why would we choose to report one over another? Our first consideration is this: does the data contain outliers? See, we know that an outlier can have a big affect on the mean (but not on the median). For example, consider these test scores with mean 75.44 and median 78:

\[67, 68, 68, 71, 78, 79, 80, 83, 85\]

Before you go any farther, check these scores for outliers using the 1.5 × IQR criterion.

Now, suppose that the score of 67 was replaced by a score of 12. (Notice that’s an outlier.) This change has no effect on the median (which is still 78) but it does affect the value of the mean. Compute the new mean to check.

We say that the median is “resistant to outliers” and that the mean is “affected by outliers.” If your data contains an outlier (or several outliers) that affect the mean, then the median is often the better (more honest?) measure of center for you to report for those data.

Which of the measures of spread is resistant to outliers and which is affected by them? Take a few minutes to figure it out.
Now we are going to introduce a display of the five-number-summary of a data set: the boxplot. Recall that the five-number-summary consists of the minimum, \( Q_1 \), the median, \( Q_3 \), and the maximum value when the data is arranged in numeric order.

Here is some fictitious data on the number of years mathematics faculty members have been in our department. Take a minute to find the median and the quartiles for these data.

\[1, 3, 6, 7, 7, 8, 8, 9, 14, 15, 15, 18, 22, 23, 24, 25, 27, 28\]

Here is an example of a boxplot display of the data on number of years mathematics faculty members have been in our department.

Number of Years of Service

Notice that a boxplot is nothing more than a picture of the five-number summary. The “box” part shows \( Q_1 \), \( M \), and \( Q_3 \), and the “arms” stretch out the reach the min and max on either side. When making a boxplot, be sure that your number line has a consistent scale, that you label your display, and that you provide information about the number of values (in this case \( N = 18 \)) shown on the plot.

**Homework**

*Oh who can tell the range of joy or set the bounds of beauty?*  
*Sara Teasdale*

1) Do all the italicized things in the *Read and Study* section.
2) Answer all of the questions in the “Mean, Mode & Range” problem in Unit 4, Module 4, Session 2 in the Bridges in Mathematics Grade 4 Student Book. Then, as a future teacher, consider the following questions: Why might a child answer 17 inches on question 6? Why a child answer 36 inches? Why might a child answer 42 inches? Why might a child answer 52 inches?

3) Consider again this line graph of 18 scores on a quiz:

```
X X
X X X
X X X X
```

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Score on a Quiz (out of ten)

a) What is the range of these data?
b) What is the IQR?
c) True or False? (Do this without performing a calculation.) The mean absolute deviation of these data is about 5. Explain your thinking.
d) Are either of the scores 1 or 2 an outlier? Use our criterion to check.
e) What measure of center and what measure of spread would you report for these data and why?
f) Make a boxplot of these data.
g) How could you use the boxplot to quickly spot outliers in the data? Explain.

4) Here are those bowling scores once again:

```
112 114 120 123 127 127 130 167 220
```

a) Which (if any) of these scores are outliers based on our criterion?
b) Now go to the introduction to this text and read the Common Core State Standards for Grade 6. Notice that they ask that children “summarize numerical data sets in relation to their context.” Do the things they suggest for the set of bowling scores.

i. Reporting the number of observations.
ii. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
iii. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
iv. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.”
5) Editors of *Entertainment Weekly* ranked every single episode ever made of *Star Trek: The Next Generation* from best (ranked #1) to worst (ranked #178). Then they compiled the rankings based on the season in which the episode aired. Below are boxplots showing the rankings of episodes in each of the seven seasons (from *Workshop Statistics*, p. 89):

![Star Trek Rankings plotted by Season](image)

a) Which was the best overall season and which was the worst based on these rankings? Make an argument in each case.
b) Do any of these seasons have rankings that are outliers (for that season)? Explain.
c) Which season(s) has a distribution of rankings that is skewed left? Explain.
d) Create a series of good questions you could ask upper elementary students about these boxplots; then answer them yourself.

6) This is a mean absolute deviation contest. You may use only numbers in the set \{1, 2, 3, 4, 5\} (You can use the same number as many times as you like).

a) Select four numbers with the lowest mean absolute deviation. (Actually compute it). Are those the only four numbers that give the lowest mean absolute deviation?
b) Select four numbers with the highest mean absolute deviation. (Actually compute it). Are those the only four numbers that give the highest mean absolute deviation?
7) Which is likely bigger?
   a) The mean new-house cost in Seattle, WA or the median new-house cost? Explain.
   b) The mean or the median in a skewed-left distribution? Explain.
   c) The range of a set of numbers or the mean absolute deviation of the set? Explain.
   d) The second quartile or the median? Explain.
   e) Now, give an example of a set of 8 numbers whose mean is 1000 bigger than its median or explain why it is impossible to do so.
Class Activity 14: The Matching Game

Let the punishment match the offense.

Cicero

On the following page you will find bar graphs displaying distributions for eight different sets of exam scores. The numerical summaries for those data sets are listed below.

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Mean</th>
<th>Median</th>
<th>Mean Absolute Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>44.5</td>
<td>44.0</td>
<td>7.7</td>
</tr>
<tr>
<td>B</td>
<td>37.0</td>
<td>36.0</td>
<td>4.0</td>
</tr>
<tr>
<td>C</td>
<td>32.5</td>
<td>31.5</td>
<td>12.0</td>
</tr>
<tr>
<td>D</td>
<td>40.0</td>
<td>41.5</td>
<td>25.5</td>
</tr>
<tr>
<td>E</td>
<td>11.5</td>
<td>3.0</td>
<td>20.0</td>
</tr>
<tr>
<td>F</td>
<td>40.0</td>
<td>38.5</td>
<td>11.0</td>
</tr>
<tr>
<td>G</td>
<td>33.5</td>
<td>34.5</td>
<td>10.0</td>
</tr>
<tr>
<td>H</td>
<td>20.5</td>
<td>20.5</td>
<td>10.0</td>
</tr>
</tbody>
</table>

1) The same number of students took each exam. How many was that? Explain how you know.

2) Match each summary with a distribution. Make sure to write explanations for your choices.

3) One of the bar graphs is missing. Figure out which data set has the missing bar graph, then create a possible bar graph for that data set.

4) What are things that an upper elementary student could learn by working on this activity? Be specific.
Insert bar graphs from p. 23 of old manual here.
Read and Study

An apprentice carpenter may want only a hammer and saw, but a master craftsman employs many precision tools.

Robert L. Kruse

In this section we are going to talk specifically about some ways to display data and about features of data distributions. First, a **distribution** of data is simply a display that shows the values of the values and their frequencies (or relative frequencies). Often when we refer to a distribution, we mean the pattern of the data when it is made into a bar graph or a line graph showing values of the data set on the horizontal axis, and either frequency of values or relative frequency (percents of values) on the vertical axis.

All three of the distributions you are about to see were found at the *Quantitative Environmental Learning Project* website (http://www.seattlecentral.edu/qelp/Data_MathTopics.html). For example the bar graph below shows the distribution of ozone concentration in San Diego in the summer of 1998 as measured by the California Air Resources Board. The EPA has established the standard that ozone levels over 0.12 parts per millimeter (ppm) are unsafe.

![Summer 1998 San Diego Ozone](image)

**Approximately how many days are displayed all together?**

**In how many days was the ozone level in San Diego classified as unsafe by the EPA?**

**Is the ozone distribution skewed right or skewed left or neither? Explain.**
Now take a minute to make sense of the reservoir data below.

**Reservoir Levels 12/31/98**

*About how many reservoirs are shown on this bar graph?*

*What percent of reservoirs are filled to at least 90% capacity?*

*Do you expect the mean of the reservoir data to be higher than the median or the other way around? Explain.*

The graphs includes reservoirs in both Oregon and Arizona. *Where do you think Arizona’s reservoirs are likely appearing on the bar graph? Why?*
Finally, below we have one more graph showing a fairly symmetric distribution of the number of hurricanes occurring in each two-week period beginning June 1.

![Graph showing US Atlantic Hurricanes: 1851-2000](image)

**On which dates are hurricanes most likely based on these data?**

In a **bar graph** the vertical axis represents a count (frequency) or a percent (relative frequency) and the horizontal axis could represent values of a categorical variable (like colors), values of a numeric discrete variable (like shoe size), or values of a continuous variable (like heights).

If you have collected data on the number of seats won by various parties in parliamentary elections, you could make a bar graph of that data by listing parties along the horizontal axis and making the bar-heights represent either frequency or relative frequency of your samples’ responses. This bar graph (data from wikipedia.org) shows both the results of a 2004 election, and those of 1999 election. **Take a moment to make sense of these data.**
Connections to the Elementary Grades

*Instructional programs from prekindergarten through grade 12 should enable students to formulate questions that can be addressed with data and collect, organize, and display relevant data to answer them.*

National Council of Teachers of Mathematics
Principles and Standards for School Mathematics, p. 176

Children in elementary school learn to make and interpret pictograms, frequency plots (sometimes called line graphs), bar graphs, and pie charts. On order to help you distinguish among these types of displays, let’s begin with an example of each. The data below show how children in Mr. Jensen’s class travel to school.

<table>
<thead>
<tr>
<th>Student</th>
<th>Transportation</th>
<th>Student</th>
<th>Transportation</th>
<th>Student</th>
<th>Transportation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby</td>
<td>Bus</td>
<td>Sasha</td>
<td>Foot</td>
<td>Xia</td>
<td>Foot</td>
</tr>
<tr>
<td>Winton</td>
<td>Bicycle</td>
<td>Lucy</td>
<td>Bus</td>
<td>Kim</td>
<td>Foot</td>
</tr>
<tr>
<td>Dexter</td>
<td>Car</td>
<td>Joe</td>
<td>Car</td>
<td>Ben</td>
<td>Car</td>
</tr>
<tr>
<td>Trevor</td>
<td>Car</td>
<td>Aiden</td>
<td>Bus</td>
<td>Leah</td>
<td>Car</td>
</tr>
<tr>
<td>Audrey</td>
<td>Car</td>
<td>Justin</td>
<td>Foot</td>
<td>Jeffery</td>
<td>Foot</td>
</tr>
<tr>
<td>Mavis</td>
<td>Foot</td>
<td>Campbell</td>
<td>Car</td>
<td>Quezal</td>
<td>Car</td>
</tr>
<tr>
<td>Kate</td>
<td>Foot</td>
<td>Sonja</td>
<td>Car</td>
<td>Cheyenne</td>
<td>Bicycle</td>
</tr>
<tr>
<td>Lim</td>
<td>Bicycle</td>
<td>Jen</td>
<td>Bus</td>
<td>Steve</td>
<td>Bicycle</td>
</tr>
</tbody>
</table>
First, here is a pictogram of the data.

**Transportation Pictogram for Mr. Jensen’s Class**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bike</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Car</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foot</td>
<td>= 1 child on foot</td>
<td>= 1 child driven to school</td>
<td>= 1 biker</td>
</tr>
</tbody>
</table>

A pictogram represents each value as a meaningful icon. The display can be oriented either vertically or horizontally (as can line plots and bar graphs). Notice how the graph is carefully labeled and that a key is provided for the icons. Make sure to help children get in the habit of labeling and defining their symbols. Also stress that it is important that all the icons be the same size or the graph could be misleading. A pictogram can be understood by children as young as those in kindergarten.

A **frequency plot** (sometimes called a line plot if the horizontal axis is a number line) of the same data is shown below. Note that this plot is a precursor to the bar graph in the sense that it displays frequency of response in a more abstract manner than does the pictogram.

**Transportation Frequency Plot for Mr. Jensen’s Class**

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>———</td>
<td>———</td>
<td>———</td>
<td>———</td>
</tr>
<tr>
<td>Bus</td>
<td>Bike</td>
<td>Car</td>
<td>Foot</td>
</tr>
</tbody>
</table>
Create a bar graph for the data from Mr. Jensen’s class.

Create a pie graph for the data from Mr. Jensen’s class. Explain how you decided how big to make each sector.

**Homework**

*The difference between a successful person and others is not a lack of strength, not a lack of knowledge, but a lack of will.*

*Vince Lombardi*

1) Do all of the italicized things in the Read and Study section.

2) Do all the problems in the Connections section. What are the exact percentages for the pie chart? How can you compute the angles to the nearest degree? Do it and then explain your work.
3) Here is a table of data showing the pets owned by children in Ms. Green’s third grade class.

<table>
<thead>
<tr>
<th>Student</th>
<th>Pets in the House</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>None</td>
</tr>
<tr>
<td>James</td>
<td>1 cat and 1 dog</td>
</tr>
<tr>
<td>Cleo</td>
<td>3 dogs</td>
</tr>
<tr>
<td>Violet</td>
<td>1 cat</td>
</tr>
<tr>
<td>Andre</td>
<td>6 fish</td>
</tr>
<tr>
<td>Jose</td>
<td>7 fish, 1 dog, 2 cats</td>
</tr>
<tr>
<td>Aidos</td>
<td>None</td>
</tr>
<tr>
<td>Yvonne</td>
<td>None</td>
</tr>
<tr>
<td>Tyrise</td>
<td>1 hamster</td>
</tr>
<tr>
<td>Ellen</td>
<td>2 dogs</td>
</tr>
<tr>
<td>Tomas</td>
<td>1 dog</td>
</tr>
<tr>
<td>Yani</td>
<td>3 fish and 2 cats</td>
</tr>
<tr>
<td>Owen</td>
<td>None</td>
</tr>
<tr>
<td>Audrey</td>
<td>1 cat</td>
</tr>
</tbody>
</table>

a) Make a frequency plot of these data with children’s names on the horizontal axis.
b) Make a line plot with ‘number of pets’ on the horizontal axis.
c) Make a bar graph of these data with ‘number of pets’ on the horizontal axis.
d) Make a pie chart of these data.
e) What problems might children encounter in trying to represent these data in the above ways?

4) Answer the questions in the “Favorite Books” problem in Unit 2 Module 4 Session 1 in the Grade 3 *Bridges in Mathematics* curriculum. How can you use this problem to discuss the similarities and differences between picture graphs and bar graphs?

5) Answer the questions in the “Gift Wrap Fundraiser” problem in Unit 2 Module 4 Session 2 in the Grade 3 *Bridges in Mathematics* curriculum. Here’s some additional questions to think about as a future teacher:
   a) Explain why a child might give the answer of “5” to question 2 about how many students sold 7 rolls of gift wrap.
   b) Explain why a child might give the answer of “7” to question 4 about how many rolls of gift wrap Sarah sold.
   c) Why is the answer to question 5 NOT the total number of Xs on the line plot?
   d) Make a frequency plot for the same data where each X represents a roll of gift wrap. Then explain to use the graph to answer questions 1-5.

![Housefly Wing Lengths](image)

a) How would you describe the shape of this distribution in terms of skewness and symmetry?
b) Approximately how many flies were measured for this study? Explain.
c) True or False? The mean of these data is equal to the median. Explain.
d) True or False? About 90% of flies have wings smaller than 51.5 (× 0.1mm). Explain.
e) Make up a good question to ask an elementary school child about this data set.

7) Here is a graph showing 100 years of data on Worldwide Earthquakes bigger than 7.0 on the Richter Scale from the *Quantitative Environmental Learning Project* (data from the US Geological Survey). So for example, there were 4 years when the number of global quakes (above 7.0) was about 35.
a) What does this graph tell you? Be specific.
b) Estimate the median annual number of quakes above 7.0.
c) How would you describe this shape of the distribution in terms of skewness or symmetry? What does that shape tell you about earthquakes?
Class Activity 15: Old Faithful Eruptions

*It is better to know some of the questions than all of the answers.*  
*James Thurber*

Old Faithful geyser in Yellowstone National Park unleashes an impressive column of water and steam every hour or so. Below you will find a sample of wait times (read horizontally) between eruptions (in minutes) for the geyser. The job for your group is to analyze these data using any techniques that make sense to you. You are welcome to use technology if available. Feel free to explore whatever questions arise in your group you look at the data — and to make whatever displays make sense to help you to answer those questions. After about 30 minutes of group work, you will present your questions, graphs, and findings to your class.

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>71</td>
<td>57</td>
<td>80</td>
<td>75</td>
<td>77</td>
<td>60</td>
<td>86</td>
</tr>
<tr>
<td>56</td>
<td>81</td>
<td>50</td>
<td>89</td>
<td>54</td>
<td>90</td>
<td>73</td>
<td>60</td>
</tr>
<tr>
<td>65</td>
<td>82</td>
<td>84</td>
<td>54</td>
<td>85</td>
<td>58</td>
<td>79</td>
<td>57</td>
</tr>
<tr>
<td>68</td>
<td>76</td>
<td>78</td>
<td>74</td>
<td>85</td>
<td>75</td>
<td>65</td>
<td>76</td>
</tr>
<tr>
<td>91</td>
<td>50</td>
<td>87</td>
<td>48</td>
<td>93</td>
<td>54</td>
<td>86</td>
<td>53</td>
</tr>
<tr>
<td>52</td>
<td>83</td>
<td>60</td>
<td>87</td>
<td>49</td>
<td>80</td>
<td>60</td>
<td>92</td>
</tr>
<tr>
<td>89</td>
<td>60</td>
<td>84</td>
<td>69</td>
<td>74</td>
<td>71</td>
<td>108</td>
<td>50</td>
</tr>
<tr>
<td>57</td>
<td>80</td>
<td>61</td>
<td>82</td>
<td>48</td>
<td>81</td>
<td>73</td>
<td>62</td>
</tr>
<tr>
<td>54</td>
<td>80</td>
<td>73</td>
<td>81</td>
<td>62</td>
<td>81</td>
<td>71</td>
<td>79</td>
</tr>
<tr>
<td>74</td>
<td>59</td>
<td>81</td>
<td>66</td>
<td>87</td>
<td>53</td>
<td>80</td>
<td>50</td>
</tr>
<tr>
<td>51</td>
<td>82</td>
<td>58</td>
<td>81</td>
<td>49</td>
<td>92</td>
<td>50</td>
<td>88</td>
</tr>
<tr>
<td>93</td>
<td>56</td>
<td>89</td>
<td>51</td>
<td>79</td>
<td>58</td>
<td>82</td>
<td>52</td>
</tr>
<tr>
<td>52</td>
<td>78</td>
<td>69</td>
<td>75</td>
<td>77</td>
<td>53</td>
<td>80</td>
<td>55</td>
</tr>
<tr>
<td>53</td>
<td>85</td>
<td>61</td>
<td>93</td>
<td>54</td>
<td>76</td>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>86</td>
<td>78</td>
<td>71</td>
<td>77</td>
<td>76</td>
<td>94</td>
<td>75</td>
<td>50</td>
</tr>
<tr>
<td>82</td>
<td>72</td>
<td>77</td>
<td>75</td>
<td>65</td>
<td>79</td>
<td>72</td>
<td>78</td>
</tr>
<tr>
<td>79</td>
<td>75</td>
<td>78</td>
<td>64</td>
<td>80</td>
<td>49</td>
<td>88</td>
<td>54</td>
</tr>
<tr>
<td>51</td>
<td>96</td>
<td>50</td>
<td>80</td>
<td>78</td>
<td>81</td>
<td>72</td>
<td>75</td>
</tr>
<tr>
<td>87</td>
<td>69</td>
<td>55</td>
<td>83</td>
<td>49</td>
<td>82</td>
<td>57</td>
<td>84</td>
</tr>
<tr>
<td>84</td>
<td>73</td>
<td>78</td>
<td>57</td>
<td>79</td>
<td>57</td>
<td>90</td>
<td>62</td>
</tr>
<tr>
<td>78</td>
<td>52</td>
<td>98</td>
<td>48</td>
<td>78</td>
<td>79</td>
<td>65</td>
<td>84</td>
</tr>
<tr>
<td>83</td>
<td>60</td>
<td>80</td>
<td>50</td>
<td>88</td>
<td>50</td>
<td>84</td>
<td>74</td>
</tr>
<tr>
<td>65</td>
<td>89</td>
<td>49</td>
<td>88</td>
<td>51</td>
<td>78</td>
<td>86</td>
<td>65</td>
</tr>
<tr>
<td>77</td>
<td>69</td>
<td>92</td>
<td>68</td>
<td>87</td>
<td>61</td>
<td>81</td>
<td>55</td>
</tr>
<tr>
<td>53</td>
<td>84</td>
<td>70</td>
<td>73</td>
<td>93</td>
<td>50</td>
<td>87</td>
<td>77</td>
</tr>
<tr>
<td>72</td>
<td>82</td>
<td>74</td>
<td>80</td>
<td>49</td>
<td>91</td>
<td>53</td>
<td>86</td>
</tr>
<tr>
<td>79</td>
<td>89</td>
<td>87</td>
<td>76</td>
<td>59</td>
<td>80</td>
<td>89</td>
<td>45</td>
</tr>
<tr>
<td>72</td>
<td>71</td>
<td>54</td>
<td>79</td>
<td>74</td>
<td>65</td>
<td>78</td>
<td>57</td>
</tr>
<tr>
<td>72</td>
<td>84</td>
<td>47</td>
<td>84</td>
<td>57</td>
<td>87</td>
<td>68</td>
<td>86</td>
</tr>
<tr>
<td>73</td>
<td>53</td>
<td>82</td>
<td>93</td>
<td>77</td>
<td>54</td>
<td>96</td>
<td>48</td>
</tr>
<tr>
<td>63</td>
<td>84</td>
<td>76</td>
<td>62</td>
<td>83</td>
<td>50</td>
<td>85</td>
<td>78</td>
</tr>
<tr>
<td>81</td>
<td>78</td>
<td>76</td>
<td>74</td>
<td>81</td>
<td>66</td>
<td>84</td>
<td>48</td>
</tr>
<tr>
<td>47</td>
<td>87</td>
<td>51</td>
<td>78</td>
<td>54</td>
<td>87</td>
<td>52</td>
<td>85</td>
</tr>
<tr>
<td>88</td>
<td>79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These data come from *A Handbook of Small Data Sets* edited by D.J. Hand, F. Daly, A.D. Lunn, K.J. McConway and E. Ostrowski from Chapman and Hall. These data can be found at [http://www.stat.ncsu.edu/sas/sicl/data/geyser.dat](http://www.stat.ncsu.edu/sas/sicl/data/geyser.dat).
Read and Study

There are three kinds of lies: lies, damned lies, and statistics.

Benjamin Disraeli

It is important not only that children learn to display data but also that they learn to make sense of charts and graphs they find in books and other media, and to be critical of misleading displays. Consider the examples below:

Amy claims that the children in her kindergarten class like cats more than dogs. She has made the following graph of her classmates’ pet preferences to support her claim. Why is this pictogram misleading? Explain. What would you say to this child?

Children who like cats

Children who like dogs

Tuff Truck Company claims that their trucks are the most reliable and shows the below graph to make their case. Why is this bar graph misleading? Explain. (By the way, this is a real-life example. Only the names have been changed.)

In fact, critical analysis of data displays is part of the curriculum for children in upper elementary school.
Homework

Smoking is one of the leading causes of statistics.

Fletcher Knebel

1) Answer the italicized question in the Read and Study section.

2) Review questions: Carefully define the terms and, if appropriate, describe the difference requested. Some examples might help, but examples alone are not enough.

   a) What is meant by the term **descriptive statistics**?
   b) What is the difference between a **numerical variable** and a **categorical variable**?
   c) What is the difference between a **discrete variable** and a **continuous variable**?
   d) What is an **outlier**?
   e) What do we mean when we stay a statistic is **resistant to outliers**?

3) A **stem and leaf plot** (or simply **stemplot**) is a listing of all the data typically arranged so the tens place makes the stem and the ones are the ‘leaves.’ For example, here are scores on a Geometry Final Exam arranged as a stemplot:

   **Exams Scores in Geometry**
   
   2 | 0
   3 | 2, 7
   4 | 
   5 | 6, 8, 9, 9
   6 | 2, 4, 5, 5, 7, 8
   7 | 0, 0, 2, 4, 5, 6, 7
   8 | 0, 1, 1, 2, 4, 4
   9 | 2, 4, 5
   10| 0

   We would read the scores as 20, 32, 37, 56, 58, 59, 59, 62 etc.

   a) Describe the shape of the distribution of scores in terms of skewness and symmetry.
   b) What is the median score?
   c) Give some examples of data sets for which a stem plot would be impractical.
   d) Make a boxplot of these data.

4) Consider this graph on the next page making the case that we are having a national ‘crime wave!’ What do you think? Does the graph provide compelling evidence? Explain.
Reported burglaries in the United States, 2001-2006
Source: Statistical Abstract of the United States, 2008
Chapter Four

Ratios, Rates, and Proportions

1.33 or 4:3
Standard aspect ratio and standard-definition video

1.66:1
Aspect ratio used for most European theatrical showings

1.78:1 or 16:9
Standard aspect ratio for high-definition video

1.85:1
Aspect ratio used for most U.S. theatrical showings since the 1980s

2.39:1
Aspect ratio of current anamorphic (wide-screen) theatrical showings

2.75:1
Aspect ratio of Ultra-Panavision 70

4.00:1
Rare use of Polyvision (three 35mm: 1.33:1 projected side by side)
only used in Napoleon (1927)

Layered comparison of different aspect ratios
Class Activity 16: Let’s Be Rational

*Those who refuse to do arithmetic are doomed to talk nonsense.*

John McCarthy, artificial intelligence pioneer

Below are some problems involving ratios. In your group, first let every one spend a few minutes by themselves thinking about each problem and how they would solve it. Then share your solution methods in your group. Try to find several ways that children could think about these problems. For each problem, think of a way you could illustrate the problem with pictures or a diagram.

1) Suppose the ratio of boys to girls in Mr. Gauss’ fifth-grade class is 5:3. How many students are in the class?

2) Suppose the ratio of boys to girls in Mr. Gauss’ fifth-grade class is 5:3, and that there are 12 more boys than girls. How many students are in the class?

3) Calvin and Hannah each have a collection of toy cars. The ratio of the number of Calvin’s cars to the number of Hannah’s cars is 3:7. If Calvin gives 1/6 of his cars to Hannah, what will be the new ratio of the number of Calvin’s cars to Hannah’s?

4) Before they went shopping together, Jack had the same amount of money as Karen. After Jack spent $34 and Karen spent $16, the two noticed that Jack had ¼ the amount of money that Karen had. How much money did each have to start?
Read and Study

The ratio of We's to I's is the best indicator of the development of a team.  
Lewis B. Ergen

A ratio is a comparison of two quantities that have the same unit. Specifically a **ratio** between two quantities is \(A : B\) if there is a unit so that the first quantity measures \(A\) units and the second quantity measures \(B\) units.

Let’s consider an example in detail to really understand this definition. Suppose Joe is 6 feet tall, and his younger brother Ben is 4 feet tall. Then we can say the ratio of Joe’s height to Ben’s height is 6 : 4.

A powerful way to visualize ratios is through a **bar diagram**, in the two quantities being compared are each represented as a bar with length measured in the same units. So to illustrate a ratio of \(A:B\), we would make a bar of length \(A\) units and a bar of length \(B\) units, and then we can “see” the ratio as a visual comparison of the length of the two bars. For example, here is a bar diagram to represent the ratio 6 : 4.

<table>
<thead>
<tr>
<th>Joe’s Height</th>
<th>Ben’s Height</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notice how we made each unit the same size, to emphasize that we are measuring both quantities with the same unit.

Let’s look further at the definition we gave for a ratio. When we say that “the ratio of Joe’s height to Ben’s height is 6 : 4”, the two quantities we are comparing are Joe’s height and Ben’s height, and that there is some unit so that Joe’s height is 6 units and Ben’s height is 4 units. So what is that unit? In this case, the unit is 1 foot, so we can think of each small rectangle in our diagram above as being one foot.

But what if we choose a different unit of measurement? Let’s say we use “1 inch” as our unit of measurement. Then what is the ratio of Joe’s height to Ben’s height? Using inches, Joe is 72 inches tall, and Ben is 48 inches tall, so the ratio of Joe’s height to Ben’s height is 72 : 48. We could make a bar diagram illustrating the ratio 72:48 by dividing each of the units in the diagram above into 12 smaller equal sized units.

Write the ratio of Joe’s height to Ben’s height if the unit of measurement is 1 yard.
Write the ratio of Joe’s height to Ben’s height if the unit of measurement is “Ben’s height”?

Write the ratio of Joe’s height to Ben’s height if the unit of measurement is “Joe’s height”?

If the ratio of Joe’s height to Ben’s height is 3:2, then what is the unit of measurement being used?

Notice how we can have many different ways of writing the same ratio. In this last example, the ratio of Joe’s height can be written as 6 : 4, or 72 : 48, or 1.5 : 1, or 1 : 2/3. We say that two ratios are equivalent if one is obtained from the other by multiplying all measurements by the same non-zero number. Show that each ratios in the previous sentence is equivalent to 6 : 4 by finding the number to multiply the measurements by.

Let’s look at the first problem from the Class Activity. Suppose the ratio of boys to girls in Mr. Gauss’ fifth-grade class is 5 : 3. How many students are in the class? There are many possible answers to this question. It depends on what the unit is in the given ratio of 5 : 3. In the bar diagram shown below illustrating the ratio 5 : 3, each unit shown could represent 1 student, in which case there would only be 8 students in the class (5 boys and 3 girls). But if each unit is 2 people, then there are 16 students (10 boys and 6 girls). If each unit is 3 people, then there are 24 students (15 boys and 9 girls), and so on.

| Number of Boys |   |   |   |   |   |
| Number of Girls |   |   |   |   |

Without any further information, the only thing we can say is that there is some multiple of 8 students in Mr. Gauss’ class. However, in question 2, we were given some further information that there were 12 more boys than girls. With this, we can then figure out that there must be 30 boys and 18 girls in the class, for a total of 48 students (poor Mr. Gauss).
You might have figured this out by thinking of ratios that are equivalent to \(5 : 3\) where the difference between the two numbers is 12. For example

\[
\frac{5}{3} = \frac{10}{6} = \frac{15}{9} = \frac{20}{12} = \frac{25}{15} = \frac{30}{18}
\]

So the ratio \(5 : 3\) is equivalent to \(30 : 18\), and if the unit were 1 person, this would be 30 boys and 18 girls, which is the required 12 more boys than girls.

You may have reasoned that the difference between boys and girls in the ratio \(5 : 3\) is \(5 - 3 = 2\) units. Since this difference is 12 people, 2 units must equal 12 people, so each unit is 6 people. The bar diagram above gives a powerful way of visualizing this. We can see the two extra units that the boys have, and identify these two units with the 12 more boys than the girls.

So it turns out the ratio we were given for the number of boys to girls, namely \(5 : 3\), was using a unit of 6 people. Once we figure out this information, we could add it to our bar diagram, labeling each unit as being 6 people. Then we can easily see the 30 boys and 18 girls.

These examples highlight a Big Idea about ratios: **ratios are unitless**, so a ratio by itself does not tell you the actual quantities involved. If you want to determine the actual quantities, you will need to know or figure out exactly what unit is. In the homework we will give you several more problems where you will need to figure out what the unit is in order to solve the problem.

Ratios can be expressed in colon notation, as fractions, as decimals, as percents, or in words. Let’s go back to Joe and Ben, where Joe is 6 feet tall and Ben is 4 feet tall. The comparison between their heights can be expressed in many equivalent ways:

- The ratio of Joe’s height to Ben’s height is 6:4
• Joe’s height is 6/4 of Ben’s height.
• Joe’s height is 3/2 of Ben’s height
• Joe’s height is 1.5 times Ben’s height
• Joe’s height is 150% of Ben’s height

We could also change the order of the comparison, and compare Ben’s height to Joe’s height:

• The ratio of Ben’s height to Joe’s height is 4:6.
• Ben’s height is 4/6 of Joe’s height.
• Ben’s height is 2/3 of Joe’s height
• Ben’s height is about .67 times Joe’s height
• Ben’s height is about 67% of Joe’s height

**Homework**

*Success is following the pattern of life one enjoys most.*  

*Al Capp*

1) Do all the italicized things and unfinished examples in the **Read and Study** section.

2) Read the “More Multiplicative Comparisons at the Giants Door” activity in Unit 1, Module 3, Session 3 of the *Bridges in Mathematics* Grade 4 Student Book. Answer all of the questions 1-7 in that activity. Then also answer these:
   a) The hammer is ________ as tall as the dog.
   b) The rake is ________ times as tall as the saw.
   c) The saw is ________ as tall as the rake.
   d) The shovel is ________ as tall as the wooden door.
   e) The wooden door is ________ times as tall as the shovel.

3) In a sample the ratio of men to women is about 5 : 8. If the sample contains 200 people, how many men and how many women are in the sample? Explain your thinking.

4) Read the “Facts & Fish” problem in Unit 6, Module 1, Session 5 of the *Bridges in Mathematics* Grade 1 Home Connections. Answer #6 in that activity. In what way does this problem involve ratios? What are some ways a first grade child might “show their work” in solving this problem?

5) Consider this problem: *Mia makes 12 out of every 20 shots she takes from the free throw line. How many shots does she expect to make from the line if she takes 50 total*
shots? Suppose the children in your sixth-grade class explained their thinking about the problem is these ways. In each case, make sense of the child’s thinking and decide whether the child is correct.

a) Joe’s solution: “12 out of 20 is the same as 6 out of 10. So I did 6 times 5 to get 30 out of 50.”
b) Omar’s solution: “I took 12 out of the first 20 and then 12 more out of a second 20 and that’s 24. Then we need ten more shots, and she’d make 6 of those and so that’s 30 shots that she’d make.”
c) Gabrielle’s solution: “20 times 2.5 makes 50. So I took 12 times 2.5 and that is 30 shots she’d make.”
d) Terrell’s solution: “12/20 is 60%, and that is 60 out of 100 or 30 out of 50.”

6) It’s time to make the connection between ratios and fractions explicit:
   a) If the fraction of a cake that is remaining is 3/5, interpret the meaning of the ratio 3:5.
   b) If the ratio of boys to girls in class is 3:5, interpret the meaning of the fraction 3/5.
   c) In general, what is the meaning of the fraction 3/5? Explain precisely how this fraction is related to the ratio 3:5.

7) Discuss the ways in which proportional reasoning (thinking about equivalent ratios) can help students understand converting fractions to decimals.

8) Show how to solve the following problems by using a bar diagram.
   a) The ratio of apple juice to cranberry juice in a juice cocktail is 8:3. If a pitcher contains 24 cups of apple juice, how much juice cocktail is there in the pitcher?
   b) At the beginning of the week, the ratio of Katie’s money to Sasha’s money as 4:7. Then Katie spent half of her money, and Sasha spent $60, so that at the end of the week, Sasha had twice as much money as Katie. How much money did Katie start with?
   c) A box of cookies has three different kinds of cookies. The ratio of oatmeal cookies to sugar cookies is 2:3, and the ratio of sugar cookies to chocolate chip is 6:13. What is the ratio of oatmeal cookies to chocolate chip?
Class Activity 17: The Watermelon Problem

They say that ninety percent of TV is junk. But, ninety percent of everything is junk.
Gene Roddenberry

A 100-pound watermelon, initially 99% water (by weight), was left sitting in the sun. Some of the water evaporates. Afterwards, it is only 98% water. How much does it weigh now?
Let’s talk about ratios that are expressed as percents, as this is perhaps the most common way in which ratios are expressed. A percent is a ratio where the second number is 100. The following are equivalent:

\[ A \% \quad A : 100 \quad \frac{A}{100} \]

Let’s consider an example. Suppose Abby has $4 and Bob has $5. We can illustrate this situation with a bar diagram with a unit of $1.

Abby’s Money

[Diagram of Abby’s money]

Bob’s Money

[Diagram of Bob’s money]

Take a few minutes to answer the following questions:

- *Abby’s money is what percent of Bob’s money?*

- *What percent of Abby’s money does Bob have?*

- *Abby has what percent less money than Bob?*

- *Bob has what percent more money than Abby?*

Since percents are ratios where the second number is 100, we could answer these questions by finding equivalent ratios. Since we know the ratio of Abby’s money to Bob’s money is 4:5, this can also be expressed as “the ratio of Abby’s money to Bob’s money is 80:100”, so Abby’s money is 80 percent of Bob’s money. Similarly, we can reverse the order of comparison, and say that “the ratio of Bob’s money to Abby’s money is 5 : 4”, which is equivalent to “the ratio of Bob’s money to Abby’s money is 125 : 100”, so Bob’s money is 125 percent of Abby’s money.
The bar diagram can help us to visualize these comparisons. When asked: “Abby’s money is what percent of Bob’s money?” we are asked to compare Abby’s money to Bob’s money. We can see that Abby has 4 units, and Bob has 5 units, so Abby only has 4/5, or 80% of Bob’s money. In the fraction 4/5, the whole that has been split into 5 equal parts is the whole of Bob’s money. Notice how the whole in the fraction (Bob’s money) is the same whole we are finding the percent of. Since “Bob’s Money” is the whole in this situation, we are thinking of the whole of Bob’s money being the 100%, which we could also indicate on the bar diagram. Since 5 units is 100% Bob’s money, then each unit is 20% of Bob’s money, so Abby’s money is 80% of Bob’s money.

<table>
<thead>
<tr>
<th>Abby’s Money</th>
<th>Bob’s Money</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20% 20% 20% 20% 20%</td>
</tr>
<tr>
<td></td>
<td>100 percent of Bob’s Money</td>
</tr>
</tbody>
</table>

Similarly, if we change the order of the comparison and compare Bob’s money to Abby’s, we can see that Bob has 5 units to Abby’s 4, so he has 5/4, or 125% of Abby’s money. In the fraction 5/4, the whole that has been split into 4 equal parts is the whole of Abby’s money, and this is the same whole we are finding the percent of when we say that Bob has 125% of Abby’s money. In the diagram, if we think of 100% as representing the whole of Abby’s money, then each unit is 25% of Abby’s money. Since Bob has 5 units, Bob has 125% of Abby’s money.

<table>
<thead>
<tr>
<th>100 percent of Abby’s Money</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abby’s Money</td>
</tr>
<tr>
<td>25% 25% 25% 25%</td>
</tr>
</tbody>
</table>

In the other two questions we asked above, we are asking you to pay attention to the difference between the two quantities being compared. Since Abby has 4 units and Bob has 5 units, Abby has one fewer unit than Bob. This one fewer unit is 1/5 or 20% of Bob’s money, and so we can say that Abby has 20% less money than Bob.

Now here is a cool thing. The fact that Abby has 20% less money than Bob does not mean that Bob has 20% more money than Abby. Why not? Really think about this. Then explain.
The fact that Abby has 1 unit fewer than Bob does mean that Bob has 1 unit more than Abby. However, we now want to compare this 1 unit difference to Abby’s money. This 1 unit more is 1/4 or 25% of Abby’s money, so we say that Bob has 25% more money than Abby.

In general, to compute a fractional change, or a percent change, we always compare the amount of change to the original amount.

Here is another set of examples to consider. Take some time now to answer these questions. For each, make a bar diagram comparing the two numbers of credits, and then compare the amount of increase or decrease in the bars to the original amount. We left you some space to make your diagrams.

- Last semester Karen took 8 credits. This semester she is taking 12 credits. By what percent did her credit load increase?

- Last semester Loren took 12 credits. This semester she is taking 16 credits. By what percent did her credit load increase?

- Last semester Moe took 16 credits. This semester he is taking 12 credits. By what percent did his credit load decrease?

Have a look at the answers to the three questions above. Notice how in each example, the amount of the change was 4 credits. However, this same amount, 4 credits, corresponds in these examples to 50%, 33%, or 25%, depending on what this 4 credits is being compared to. This is a really big idea about fractions and percents: that the same fixed quantity can be expressed as different percents, depending on what whole that quantity is being compared to.

Here’s another Big Idea: since percents are ratios, percents are unitless. By working on the watermelon problem we hope that you now see how important it is to realize this fact. We know it is very tempting to think of a percent as being a unit, after all, it sure sounds like one. When we hear a phrase such as “30 percent”, that sure sounds as if “percent” is a unit, just like in the phrase “30 dollars”, a dollar is a unit, or in “30 feet”, a foot is a unit. But the distinction between a “percent” and units such as “dollars” or “feet” is huge. Consider this: 30
feet is always twice as long as 15 feet. And 30 dollars is always twice as much as 15 dollars. But 30 percent may not be twice as much as 15 percent. Read that last sentence again. For example, compare 30 percent of the population of Wisconsin with 15 percent of the population of California. Give another example when 30 percent is not twice as much as 15 percent.

Here’s another illustration: 50 feet is always 10 feet longer than 40 feet. And 50 dollars is always 10 dollars more than 40 dollars. However, 50 percent is not 10 percent more than 40 percent. In fact, if all of the percentages involved were based on the same whole, then 50 percent would be 25 percent more than 40 percent. Read that last sentence again, and fight to understand it.

It may be hard to believe that 50 percent could be 25 percent more than 40 percent, but we hope you will be able to see this after thinking through some more examples, and in the homework, we will ask you to give your own example of this.

**Connections to the Elementary Grades**

*Say what you mean and mean what you say.*

Anonymous

“Then you should say what you mean," the March Hare went on.
"I do," Alice hastily replied; "at least--at least I mean what I say--that's the same thing, you know."
"Not the same thing a bit!" said the Hatter. "You might just as well say that "I see what I eat" is the same thing as "I eat what I see"!"

Lewis Carroll

We will close this section with some thoughts about language. When ratios are expressed in words or with colon notation, the two quantities being compared are usually stated explicitly. For example, if we say “the ratio of Abby’s money to Erin’s money is 4:5” we know we are comparing the quantity of “Abby’s money” to the quantity of “Erin’s money”.

Especially when ratios are expressed as percents, then depending on the language used, the two quantities being compared might not be expressed explicitly. When we say that “Abby’s money is 80 percent of Erin’s money” this means that we are comparing “Abby’s money” to “Erin’s money” and that “Erin’s money” has been expressed as the number 100.

However, it is common to not express things so explicitly. For example, the watermelon problem stated that the watermelon is 99% water (by weight). Notice that this statement doesn’t explicitly state “percent of” to make it clear what is being considered the whole 100.
In particular, the phrase “99 percent water” does not mean “99 percent of the water”. We need to interpret this statement and figure out that is means that the water’s weight is 99 percent of the watermelon’s weight.

We hope this all illustrates the importance of paying particular attention to language and being precise in our use of language. Furthermore, it means that can not reduce mathematical problem solving to merely translating “English” into “Math” as many teachers and students are tempted to do, without careful thought. A big part of the Common Core mathematical practice standard of “attending to precision” is paying attention to the precise meanings of the language that we use. Formulating what we mean using precise language is a powerful tool in problem solving and an important mathematical practice.

**Homework**

*I've got a theory that if you give 100 percent all of the time, somehow things will work out in the end.*

*Larry Bird*

1) Do all the italicized things and unfinished examples in the *Read and Study* section.

2) Let’s reconsider Mr. Gauss’s class, where the ratio of boys to girls in the class is 5:3. Answer the following questions:
   a) What percent of the class is girls?
   b) What percent of the number of boys is the number of girls?
   c) What percent of the class is boys?
   d) What percent of the number of girls is the number of boys?


4) Suppose each year at ESU, there are 2000 freshmen. Last year 25% of the freshmen at ESU live off campus. This year 30% of freshmen live off campus. What is the percent increase in the number of freshmen living off campus? Make a bar diagram to support your answer.

5) Last year, CRI research conducted a telephone survey of 500 adult Americans and found that 45% were in favor of the president. This year they again surveyed 500 adult Americans and found that the number of people that were in favor of the president fell by 30%. What percent of the people they surveyed this year were in favor of the president? Make a bar diagram to support your answer.
6) Last year 8% of the city of Illville had dysentery. This year 10% of the city had dysentery. By what percent did the number of people in Illville with dysentery increase? Justify your answer.

7) What assumption did you make in order to solve the previous problem? Is this a realistic assumption? Can you figure out how there could be no change (in other words, a 0 percent change) in the number of people in Illville with dysentery and still have the percent with dysentery increase from 8 percent to 10 percent of the city?

8) Acme sells cans of mixed nuts. In the past, 5% of the nuts were cashews, but now Acme is advertising that their new and improved can of mixed nuts has 40% more cashews. What percent of the can of new and improved mixed nuts is cashews? What assumptions are you making about the old and new cans in your solution?

9) Make up your own example of a problem where 50% is 25% more than 40%.

10) Make up your own example of a problem where 60% is the same quantity as 30%.

11) Five years ago, a governor cut a university’s budget by 20%. This year, she increased the budget by 5%, and announced that she gave back a quarter of the budget cut. Is this accurate? Explain.

12) Bob buys and sells rare stamps on Ebay for fun and profit. Usually, Yesterday Bob sold two stamps. Each sold for $60. However, he had paid different amounts for the stamps originally, so that the first stamp ended up being sold at a 25% profit, and the second was sold at a 25% loss. Did Bob end up making money, losing money, or did he break even? Justify your answer.
Class Activity 18: How Do You Rate?

Nothing travels faster than the speed of light, with the possible exception of bad news, which obeys its own special laws.

Douglas Adams

1) It would take Biff 60 minutes to mow a lawn, and it takes Charlie 30 minutes to mow that same lawn. If they work together, how long would it take them to mow that lawn?

2) Suppose you commute each day between Abbotsford and Belleview. In the morning, you average 60 mph. On the return trip home, you hit rush hour and only average 40 miles per hour. What is your average speed for the entire commute?

3) A car is travelling 60 miles per hour. How fast is this in feet per second?
A *rate* is a comparison of two quantities, each made with a different specified unit. In the previous section we defined a ratio as a comparison of two quantities that have the same unit. So a rate is like a ratio, except that in a rate, the two quantities being compared have different units. The result of this is that while ratios are “unitless”, rates do have units. A speed, such as how fast a car is traveling, is a good example of a rate, and some familiar units for speeds are miles per hour, or feet per second, etc.

We assume that you found the first two problems in the Class Activity quite tricky. The reason why they are tricky is that we have gotten used to rates that are expressed where the second quantity in the comparison is one. When we have a speed of 60 miles per hour, this is a rate comparing 60 miles with 1 hour. This means for every 60 miles we would travel 1 hour. This rate is equivalent to travelling 120 miles in 2 hours, or 30 miles in \( \frac{1}{2} \) hour, or 1 mile in \( \frac{1}{60} \) of an hour. There are an infinite number of rates that are equivalent to this speed, but our standard way of expressing this rate is 60 miles in 1 hour, where the second quantity in the comparison is 1.

So in the commuting problem, the rates we are given are both comparing distances to 1 hour. So we are fooled into thinking that both parts of the commute are 1 hour. However, we aren’t traveling the same amount of time in the morning as in the afternoon. *(Why not?)* Instead, we are travelling the same distance in the morning as in the afternoon. So to make sense of the problem, we should instead think of equivalent rates where the *distances* are the same. For example, in the morning, we are traveling at a rate of 120 miles in 2 hours, and in the afternoon we are traveling at a rate of 120 miles in 3 hours. So now we can think if the entire commute were 120 miles there and 120 miles back, it would take us 5 hours to go a total of 240 miles, which is equivalent to 48 miles per hour. Or we could just think about how long it would take to travel one mile of our commute. In the morning it would take 1 minute to travel one mile, and in the afternoon it would take 1.5 minutes to travel one mile. So combined that is 2.5 minutes for 2 miles, which is equivalent to 48 miles per hour. There are many good ways to think about this problem, as long as we realize that it is the distances traveled that are the same from morning to afternoon, not the times traveled.

Similarly, in the lawn-mowing problem, the rates we are given are 60 minutes per lawn and 30 minutes per lawn. Both rates are expressed with the same amount of lawn (namely 1 lawn). However, Biff and Charlie when working together will not mow the same amount of lawn. *(Why not?)* Instead, they will mow the same amount of time. So we should find equivalent rates for each where the *time* is the same. *(Identifying what is changing and what is staying the same is a big theme in many of the more challenging problems we’ve looked at*
recently. For instance, in solving the watermelon problem, it was important to realize that the amount of water was changing, but the amount of rind was staying the same. There are many ways to express equivalent rates. One we’ve seen already is to change the number of each unit proportionally. For example 60 miles per hour is the same as 120 miles in 2 hours. We can also change the order of comparison. 60 miles per hour is the same as 1 hour per 60 miles. Yet another way to find equivalent rates is to express one or both of the quantities in different units. For example, 60 miles per hour is the same as 60 miles per 60 minutes.

Consider the following equivalent rates:

\[
\frac{1 \text{ mile}}{1 \text{ hour}} = \frac{4 \text{ miles}}{4 \text{ hours}}
\]

To obtain the second rate from the first, you might imagine that the first rate has been multiplied by the factor of 4/4.

\[
\frac{1 \text{ mile}}{1 \text{ hour}} \times \frac{4}{4} = \frac{4 \text{ miles}}{4 \text{ hours}}
\]

Now consider this pair of equivalent rates:

\[
\frac{4 \text{ miles}}{4 \text{ hours}} = \frac{21120 \text{ feet}}{4 \text{ hours}}
\]

*How has the second rate been obtained from the first? Really think about this!*

One way is to think of this as a simple substitution. Since 1 mile is 5280 feet, 4 miles is 21120 feet, and so we have just substituted one quantity for an equivalent quantity. Another way is to think of the first rate being multiplied by a rate of 5280 feet per mile, so that

\[
\frac{4 \text{ miles}}{4 \text{ hours}} \times \frac{5280 \text{ feet}}{1 \text{ mile}} = \frac{21120 \text{ feet}}{4 \text{ hours}}
\]

*Why does it make sense that you can get an equivalent rate by multiplying by 4/4?*

*Why does it make sense that you can get an equivalent rate by multiplying by \( \frac{5280 \text{ feet}}{1 \text{ mile}} \)?*
An understanding of, and a facility with rates, ratios, and proportionality are important tools students use in learning many areas of the school curriculum. Below you will find the Common Core state Standards for Mathematics in the domain of ratio and proportion. Read them.

### Common Core State Standards for Ratios and Proportional Relationships

**Grade Six:**

- Understand the concept of a ratio and use ratio language to describe a ratio relationship between two quantities. For example, “The ratio of wings to beaks in the bird house at the zoo was 2:1, because for every 2 wings there was 1 beak.” “For every vote candidate A received, candidate C received nearly three votes.”

- Understand the concept of a unit rate a/b associated with a ratio a:b with b ≠ 0, and use rate language in the context of a ratio relationship. For example, “This recipe has a ratio of 3 cups of flour to 4 cups of sugar, so there is 3/4 cup of flour for each cup of sugar.” “We paid $75 for 15 hamburgers, which is a rate of $5 per hamburger.”

- Use ratio and rate reasoning to solve real-world and mathematical problems, e.g., by reasoning about tables of equivalent ratios, tape diagrams, double number line diagrams, or equations.
  - Make tables of equivalent ratios relating quantities with whole-number measurements, find missing values in the tables, and plot the pairs of values on the coordinate plane. Use tables to compare ratios.
  - Solve unit rate problems including those involving unit pricing and constant speed. For example, if it took 7 hours to mow 4 lawns, then at that rate, how many lawns could be mowed in 35 hours? At what rate were lawns being mowed?
  - Find a percent of a quantity as a rate per 100 (e.g., 30% of a quantity means 30/100 times the quantity); solve problems involving finding the whole, given a part and the percent.
  - Use ratio reasoning to convert measurement units; manipulate and transform units appropriately when multiplying or dividing quantities.
**Homework**

*If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.*

John Louis von Neumann

1) Do all of the italicized things in the Read and Study section.

2) In the Connections section, in the Common Core State Standards box the following problem was posed: “If it took 7 hours to mow 4 lawns, then at that rate, how many lawns could be mowed in 35 hours? At what rate were lawns being mowed? Answer those questions. How can you use a bar diagram to illustrate this situation?

3) Three apples cost $2. How much should 7 apples cost? Explain why this problem demands proportional reasoning. Where does this problem fit within the Common Core State Standards?

4) Do problem #3 in “Estimate & Reason with Water” Unit 8, Module 2, Session 3 from *Bridges in Mathematics* Grade 3 Student Book.
   a) Answer questions a and b in the problem.
   b) Make a list of all the rates involved in this problem, and illustrate each rate with a bar diagram.
   c) Suppose Jhong also has a 100 gallon wading pool. Make a bar diagram that illustrates that it will take 8 1/3 minutes to fill up the pool with the garden hose.

5) The speed of a bullet as it leaves a modern firearm is about 1000 meters per second. How fast is this in miles per hour?
   a) Show how you can solve this problem by setting up a single calculation that multiplies appropriate rates together. In your calculation, how many different rates are involved? Which of these rates are equivalent to the speed of the bullet?
   b) Show how you can solve this problem by finding a sequence of rates where each rate is equivalent to the speed of the bullet.

6) Gas in Germany costs about 1.50 euros per liter. How much is this in dollars per gallon?

7) In a previous homework, you solved the following problem: The ratio of apple juice to cranberry juice in a juice cocktail is 8:3. If a pitcher contains 24 cups of apple juice, how much juice cocktail is there in the pitcher? By using units of “cups of apple juice”, “cups of cranberry juice” and “cups of juice cocktail”, show how to solve this problem setting up a single calculation that multiplies appropriate rates together.
8) The temperature at which water freezes is 0 degrees Celsius, and 32 degrees Fahrenheit. The temperature at which water boils is 100 degrees Celsius, and 212 Fahrenheit. Using this information, if a thermometer reads 15 degrees Celsius, what is this temperature in Fahrenheit? If a thermometer reads 18 degrees Fahrenheit, what is the temperature in Celsius? In what way does this problem involve rates?

9) Here's an old puzzle that we like: If a hen and a half lay an egg and a half in a day and a half, how long does it take one hen to lay an egg?

10) If 3 people can make 3 chairs in 3 days, how long does it take one person to make one chair? (Does this make you reconsider your answer to the hen puzzle?).
Class Activity 19: Goldfish

Fishing is boring, unless you catch an actual fish, and then it is disgusting.  
Dave Barry

We would like to count the number of gold fish in a large pond. Here is our plan:

Step 1: We’ll capture a sample of goldfish from the pond and tag them. Call that number \( n_1 \).

Step 2: We’ll release the fish back into the pond and let them swim around.

Step 3: A few days later, we’ll return and capture another sample of goldfish. Call that number \( n_2 \).

Step 4: We’ll find percent of goldfish in our second sample that are tagged.  
We’ll call that \( p\% \).

We claim that this is enough information to estimate the population of goldfish in the pond.  
Can you figure out how to do it and explain why your idea makes sense?

Under what circumstances (what types of environments, sampling methods, or populations of animals) will your method produce reliable results?
The method you used to count fish in the pond is called capture-recapture, and it is based on the assumption that both samples are representative of the population. The method isn’t very good if that is not the case. For example, this method wouldn’t work for counting animals that migrate away from the area or animals that don’t get sufficiently mixed in among the others. Likewise, it wouldn’t work well if the samples were chosen in a biased way.

You might be expecting a capture-recapture “formula” in this section. But you won’t need one because you are going to understand the method, and that is much better than knowing a formula. If you understand the idea, you won’t get confused by changes in wording or notation in a problem. So here it is:

If the samples are both representative of the fish in the pond, then the proportion of tagged fish in the whole pond should be about the same as the proportion of tagged fish in the second sample. But you know the proportion of tagged fish in the second sample. You also know the total number of tagged fish in the pond. From this information, you can compute an estimate for the total number of fish in the pond. Think this through using the data you collected in class, and be sure to talk with your group or instructor if you don’t see how to think about this.

Capture-recapture isn’t used only for counting fish; in fact, this sampling method has sometimes been used to ‘fix’ the U.S. Census data. As we mentioned in a previous section, it is tough to conduct a census, and the U.S. Census is particularly problematic. For one thing, we are a country of approximately 300 million people (we think). For another thing, there is a significant bunch of people who are homeless or ill or mistrustful of the government, and many of these people either do not receive census forms or do not return them. In past years the government has attempted to correct for these types of undercounting by estimating the number of people “missed” in a certain city or region using the method of capture-recapture. Since that is the same method that you used to estimate the number of fish in the pond, we’ll talk about it now.

Here is the idea: After census forms are returned for a region, a team of census workers will go door-to-door (and in the cases of the homeless, also to the places they might be found) in that region. They will ask everyone whether they have returned a census form. And use that percent to adjust the count.

Let’s do a specific example. Suppose the census bureau wants to count the population of downtown North Gary, Indiana and they have 2100 census forms returned from that region. A team of census workers uses cluster sampling to select several city blocks from the region. They then go to those blocks and try to find everyone who lives there and talk to them.
Suppose they find that only 70% of the people they talk to had returned the form. They will assume that for every 70 people who returned the form, 30 did not and they will use this to correct their estimate of the population of downtown North Gary from 2100 up to 3000.

*Work through that computation to be sure that we’re correct.*

### Connections to the Elementary Grades

*Facility with proportionality develops through work in many areas of the curriculum, including ratio and proportion, percent, similarity, scaling, linear equations, slope, relative-frequency histograms, and probability. The understanding of proportionality should also emerge through problem solving and reasoning, and it is important in connecting mathematical topics and in connecting mathematics and other domains such as science and art.*

*National Council of Teachers of Mathematics
Principles and Standards for School Mathematics, page 212*

You have noticed that the type of mathematical thinking required to understand capture-recapture involves ratios. **Proportional reasoning** is thinking about equivalent ratios. If we have two equivalent ratios and write this down using fractions, we get an equation of the type \( \frac{a}{b} = \frac{c}{d} \). In school algebra, an equation of this type if often called a **proportion**.

Solving problems by setting up a proportion is a common technique taught in school. However, we have found that many students will write down a proportion without thinking about what the two fractions they have written are meant to represent, or whether the two fractions they have written ought to be equal. (Perhaps you experienced this yourself when working on the Watermelon Problem.) As a teacher, if your students are going to use a proportion to help solve a problem, insist that they understand the ratios they are using, and be able to justify why the two ratios should be equal.

Many students in school learn a technique called “cross multiplication” to solve proportions. In the authors’ experience, this can have several big unintended consequences. First, this technique bypasses and seemingly violates a fundamental concept of equations, namely that to preserve an equality, whatever is done to one side ought to be done to the other side as well. Second, students come to think of “cross multiplication” as some kind of mysterious separate kind of operation. Third, many students then apply this mysterious operation whenever they multiply fractions, resulting in mistakes such as \( \frac{3}{4} \times \frac{2}{5} = \frac{8}{15} \).

The “cross multiplication” technique is simply a shortcut that allows one to skip one step in solving an equation. But in skipping that one step, we argue that we are also shortcutting an
understanding of equality and of ratios. In the homework, we will ask you to find other ways of thinking about and explaining how to solve proportions that do not bypass these important concepts.

**Homework**

*Life shrinks or expands in proportion to one’s courage.*  
*Anais Nin*

1) Do all the italicized things in the Read and Study section.

2) Explain how the US Census uses the capture-recapture method. What is the population? What is the initial sample? The recaptured sample?

3) This summer in Minnesota the DNR captured and tagged 240 small-mouth bass and then released them back into Lake Wissota. Two weeks later, they returned to the lake and caught 360 small-mouth bass. Of those, 81 were tagged.
   
   a) Estimate the number of small-mouth bass in the lake and briefly explain the idea behind this procedure.
   
   b) What assumptions must be made to use capture-recapture for this example?

4) What does capture-recapture have to do with proportional reasoning? Explain.

5) Make up your own capture-recapture problem and then solve it.

6) Consider the watermelon problem again,
   
   a) Explain why the proportion 99/100 = 98/x is not an appropriate proportion that can be used to solve the watermelon problem. Try to give at least two reasons why this proportion doesn’t make sense for this problem.
   
   b) Write an appropriate proportion that could be used to solve the watermelon problem.

7) Solve the following problems. In which of these could you set up a proportion to find the solution? In which of these is it not appropriate to write and solve a proportion? Explain.
   
   a) When Bob runs at 4 miles an hour, it takes him 20 minutes to run the path around the park. If he ran instead at 5 miles an hour, how long would it take him to run that path around the park?
   
   b) If it takes Clara 10 minutes to assemble 3 chairs, how long will it take her to assemble 5 chairs?
c) If it takes 4 people 3 hours to paint a fence, how long would it take 6 people to paint that fence?

d) A cellphone plan has a fixed monthly fee and then a per-minute use charge. If Bob talked for 100 minutes and was charged $40, how much would his phone bill have been if he had talked for 200 minutes?

c) The area of a circle with radius 10 feet is about 314 square feet. What is the area of a circle of radius 20 feet?

8) Show two ways to explain how solve the proportion \( \frac{15}{4} = \frac{75}{x} \) for \( x \). Both of your explanations should support a conceptual understanding of ratios and equality.

9) Sue and Julie were running around a track.

a) Suppose Sue and Julie ran equally fast, and Sue started first. When Sue had run 9 laps, Julie had run 3 laps. When Julie completed 15 laps, how many laps had Sue run?

b) Suppose Sue ran twice as fast as Julie, and Sue started first. When Sue had run 9 laps, Julie had run 3 laps. When Julie completed 15 laps, how many laps had Sue run?

c) Rewrite the first sentence of the previous problems so the answer to how many laps Sue had run would be the solution to the proportion \( \frac{9}{3} = \frac{x}{15} \).
Chapter Five

Relations Among Sets of Data
Class Activity 20: Shoe Size and Height

*Arithmetic is being able to count up to twenty without taking off your shoes.*

*Micky Mouse*

Your class is going to make a huge living scatterplot on the floor of your classroom to help you predict a person’s shoe size based on his or her height. You will use masking tape for your axes and sticky notes for labeling. Then, everyone will stand on the data point representing their height and shoe size to observe the pattern of the relationship.

First some definitions:

An **explanatory variable** is one that is used to explain variations in a **response variable**. It is not required that one variable *cause* the other, just that one be used to predict variation in the other. When making a scatterplot it is standard practice to use the explanatory variable on the x-axis and the response variable on the y-axis.

You will need to discuss issues that might arise in working with these data. How will you deal with the fact that women’s shoes are sized differently than men’s shoes? What scale and units should be used on the x-axis and what scale and units should be used on the y-axis? (When making a scatterplot, it is helpful to spread the data out as much as possible.) Answer these questions as a class and make your plot. When you are done, return to your groups, but leave a sticky note on the floor where you stood.

1) Are there any outliers in the data? Look up the definition and use it to make an argument.

2) Is it possible for there to be an outlier on a scatterplot that is not an outlier in either the explanatory or the response variable? Explain.

3) Is it possible to have data point that is an outlier in both the explanatory and response variables yet it is not an outlier on the scatterplot? Explain.

4) Come up with some ideas about how to use the plot to predict a person’s shoe size based on her height. If a man is 6’8” tall, what do you predict his shoe size to be based on the pattern of the relationship?

5) Place a string or tape across the data so that the string captures as best you can the relationship. Find the equation of that line. (Make sure everyone in your group understands how to do this.) Use the equation to predict the shoe-size of a 6’8” man. How confident are you in this prediction? Explain your thinking.
**Read and Study**

*It is a capital mistake to theorize before one has data.*

*Sir Arthur Conan Doyle in Scandal in Bohemia*

Sometimes the purpose of data collection is to examine a relationship between two variables. There are several ways to display a relationship like that. We could, in some cases, make side-by-side boxplots (like we did in the Star Trek activity to compare the quality of the seasons) or we might make side-by-side bar graphs (like this one below that compares stream temperatures under two conditions: under a closed canopy of vegetation and in open sky with no shielding vegetation).

![WA State Stream Temperatures](image)

1999 data from the Center for Urban Water Resources Management (CUWRM) and the Center for Streamside Studies (CSS) at the University of Washington in Seattle

We’ll have you consider some of those displays as part of your homework. In this section we will focus on the scatterplot. A **scatterplot** is a graph of numeric data using a Cartesian coordinate system (an x-axis and a y-axis) so that each case (or value) is represented by a point having an x-coordinate and a y-coordinate. In the class activity you made a scatterplot comparing height and shoe size for students in your class. Each person (case) was represented on the plot first as themselves and then as sticky note. This allowed you to study the pattern of the relationship and to make predictions. You likely found a relatively linear pattern in your height-shoe data, but other types of patterns are possible.
For example, have a look at the Carbon Dioxide level measured each month in the atmosphere at the volcano Mauna Loa in Hawaii. This data set shows both a linear trend and a periodic trend. The periodic behavior can be explained by the yearly cyclic behavior of growth and decay of vegetation. How would you explain the linear trend? What are the two variables being compared in these data?


---

**Homework**

In great attempts it is glorious even to fail.  
*Vince Lombardi*

1) Do all the italicized things in the Read and Study section.
2) When an increase in the explanatory variable generally corresponds to an increase in the response variable, we say that the relationship is a **positive association**. When an increase in the explanatory variable generally corresponds to a decrease in the response variable, we say the relationship is a **negative association**.

   a) Is the association between height and shoe size positive or negative (or neither)? Explain.
   b) Is the association between time and carbon dioxide over Mauna Loa positive or negative (or neither)? Make a case.
   c) What kind of association is it if a *decrease* in the explanatory variable corresponds to an increase in the response variable? Explain.
   d) True or False? We would expect the variables of ‘average number of cigarettes smoked each day’ and ‘longevity’ (lifespan) to have a positive association.
   e) Make up two new variables that you would expect to have a positive association.

3) Below is a set of data showing the position of a glacier over time. (Where “0” was its position in 1985.)

   a) Make a scatterplot using time as the explanatory variable and position as the response variable.
   b) Describe the pattern of the relationship.
   c) Use the pattern to predict the glacier’s position in 1997.
   d) Is the association negative or positive (or neither).

<table>
<thead>
<tr>
<th>Rainbow Glacier shrinkage, North Cascades, WA</th>
</tr>
</thead>
<tbody>
<tr>
<td>data courtesy of Mauri Pelto, North Cascades Glacier Climate Project, Nichols College</td>
</tr>
<tr>
<td>position of glacier's snout relative to 1985 position</td>
</tr>
<tr>
<td>negative values = glacier retreats/shortens</td>
</tr>
<tr>
<td>year</td>
</tr>
<tr>
<td>1985</td>
</tr>
<tr>
<td>1986</td>
</tr>
<tr>
<td>1987</td>
</tr>
<tr>
<td>1988</td>
</tr>
<tr>
<td>1989</td>
</tr>
<tr>
<td>1990</td>
</tr>
<tr>
<td>1991</td>
</tr>
<tr>
<td>1992</td>
</tr>
<tr>
<td>Year</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>1993</td>
</tr>
<tr>
<td>1994</td>
</tr>
<tr>
<td>1995</td>
</tr>
<tr>
<td>1996</td>
</tr>
<tr>
<td>1997</td>
</tr>
<tr>
<td>1998</td>
</tr>
<tr>
<td>1999</td>
</tr>
<tr>
<td>2000</td>
</tr>
</tbody>
</table>
Class Activity 21: Lots About Lines

I think it's the duty of the comedian to find out where the line is drawn and cross it deliberately.

George Carlin

This activity is designed to provide you with a review of linear relationships. There may be people in your group who haven’t thought about these things in a few years, so if this is familiar to you, you will have the chance to practice being a good teacher.

1) The slope of a line is the rate of the change in \( y \) per one unit change in \( x \). Explain why the formula \( m = \frac{y_2 - y_1}{x_2 - x_1} \) makes sense as a way to compute the slope of a line. What pictures might you draw to illustrate this?

2) Find the equation of the line that passes through the points (5, 2) and (-1, 12). Make sure everyone in your group understands this.

3) Sketch the line \( y = \frac{1}{3}x - 12 \). Then use the graph to discuss how you would explain the meaning of the \( \frac{1}{3} \) and the -12 in the context of the graph.

4) Postage for a package costs $4.50 for the first pound and $1.50 for each additional pound.
   a) Model the postage price as a function of weight. In what ways is this relationship linear? In what ways is it not really an ideal line?
   b) Carefully explain each ‘number’ in your line equation.
   c) How much will it cost to mail a 12 lb package? Figure it out in three different ways.

5) How can you recognize linear data from a table? Explore these three sets of data collected on populations of three experimental colonies on three planets to see what patterns in the table show you linear data. If one of these turns out to be fairly linear, find a reasonable equation for the line that describes the relationship between time and population. What do you expect the population of that colony to be when \( t = 15 \)?

<table>
<thead>
<tr>
<th>Colony I</th>
<th>Colony II</th>
<th>Colony III</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t ) (decades)</td>
<td>Population</td>
<td>( t ) (decades)</td>
</tr>
<tr>
<td>1</td>
<td>56</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>88</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>110</td>
<td>7</td>
</tr>
<tr>
<td>9</td>
<td>139</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>174</td>
<td>11</td>
</tr>
</tbody>
</table>
Read and Study

Whenever I see a slippery slope, my instinct is to build a terrace.

John McCarthy

There are many relationships that can best be modeled by lines: total earnings if an employee is paid a fixed amount per hour; shoe size and height; and total distance traveled when driving a constant speed, to name a few. In order to get ready to use a line to describe a relationship on a scatterplot, we are going to remind you about linear equations in general. If you remember all of this, you can focus as you read on thinking about how you will help your future students to understand these ideas.

Okay, so imagine any old line drawn on a plane (this piece of paper). Recall that a plane and a line are mathematical objects, and as such are ideal objects and not real objects. So image that this plane is infinite and so is the line: it stretches perfectly straight and forever in both directions.

Mathematicians have created ready-made notation for us to use to describe a line like this. First think of our plane as having a set of axes on it: we call the horizontal axis the “x-axis” and the vertical axis the “y-axis.” This gives our plane a structure. Now points in the plane can be described based on an x-coordinate and a y-coordinate as measured from the origin (the place where the axes cross).

So if a point on the line is located back 3 units and up 5, then it has the name (-3, 5) and it looks like this on the plane.
In fact this line can be described as a whole set of points that all satisfy an equation of the form

\[ y = mx + b \]

where \( m \) and \( b \) are constants (numbers we call parameters). You need only two pieces of information to make the equation of a line: \( m \) (the slope) and \( b \) (the \( y \)-intercept). And if you don’t have these directly, you can get them from any two points on the line. Here’s how:
Suppose that (-3, 5) and (4, 1) are points on the line based on the way we’ve put down our axes. So that means that when we move from -3 to 4 on the horizontal axis (that’s 7 units), we move down 4 units on the vertical axis. Now slope (m) is defined to be the change in y for a one-unit change in x, so for every one unit we go right we go down \( \frac{4}{7} \) of a unit: \( m = \frac{4}{7} \).

Notice that negative slopes correspond to negative associations and positive slopes to positive associations.

So far our equation is: \( y = \frac{4}{7} x + b \)

To get the value of \( b \), we’ll now reuse a point. Since (4, 1) is on the line, it satisfies our equation, so

\[
1 = \frac{4}{7} \times 4 + b
\]

And so we find that \( b = 1 + \frac{16}{7} \) or \( b = \frac{7}{7} + \frac{16}{7} = \frac{23}{7} \). *Do this calculation yourself.*

The final equation for our line is \( y = \frac{4}{7} x + \frac{23}{7} \).

You might be wondering whether we could have used the point (-3, 5) instead of (4, 1) to find our \( b \). *Try that now and see that you do get the same value of b. Really do it.*

So far, all of this has been pretty abstract. In the *Homework* and in the next section we’ll look at some meaningful situations involving linear relationships.

**Connections to the Elementary Grades**

*In grades 3-5 all students should investigate how a change in one variable relates to a change in a second variable.*

*National Council of Teachers of Mathematics Principles and Standards for School Mathematics*
If a slope is defined to be the rate of change in the y value for a one-unit change in the x value, then what are the units of slope? Well, of course that depends on the units on the x and y variables, so let’s look at a specific example.

Suppose x is time in minutes and y represents the depth of rising water in a tank in inches, then slope gives a change in water depth for a unit of time. In other words, it is a rate and its units are inches/minute. If the relationship between water depth and time is linear, then that rate is constant (the slope is the same for all values of x). If the tank is filling slower and slower, then the relationship will not be linear (the graph will look curved), and the rate will change with time (and so will the slope of the graph).

**Homework**

*Inches make champions.*  
*Vince Lombardi*

1) Do all the italicized things in the Read and Study section.

2) Do the problems and italicized things in the Connections section. What could your students learn about slopes by doing the *Everyday Mathematics* problems? Make an equation expressing profit as a function of time for each girl.

3) The depth of water in a tank begins at 3 feet and is filling at a rate of 2 feet/hour. Make an equation for the depth of the water as a function of time. Now graph the relationship where x is measured in hours and y is measured in feet. How does the 3 feet show up on the graph? The 2 feet/hour?

4) Here is a table showing the prices for fudge at a candy store:

<table>
<thead>
<tr>
<th># of ounces</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$3.20</td>
</tr>
<tr>
<td>8</td>
<td>$6.00</td>
</tr>
<tr>
<td>16</td>
<td>$11.00</td>
</tr>
<tr>
<td>32</td>
<td>$20.00</td>
</tr>
<tr>
<td>48</td>
<td>$29.00</td>
</tr>
<tr>
<td>64</td>
<td>$38.00</td>
</tr>
</tbody>
</table>
a) Is the price of fudge a linear function of weight? Explain how you can tell from the table alone.

b) Now graph this data. What do you notice? Find an equation for the price of fudge as a function of weight when you purchase more than 16 ounces. What exactly does the slope tell you in this case?

c) Is the association between weight and price positive or negative? Explain.

5) Give examples of a pair of variables that you would expect to be linearly related. Explain your thinking. Is the association between your variables positive or negative? Why?

6) Find the equation of the line that passes through the points (2, -5) and (4, -1). Explain your work as you would to a student.

7) Graph the line that has slope \( m = \frac{2}{3} \) and passes through the point (-1, -5). Explain your work as you would to a student.

Class Activity 22: Manatee Data

Barbara Manatee, you are the one for me.

Sung by Larry the Cucumber, Veggie Tales

The manatee is on the endangered species list and so its population is carefully tracked. A researcher believes that manatee deaths in Florida could be caused by boating. In order to make a case, she found the following data on boat registrations (Florida Department of Motor Vehicles website gives data for 2000-2005. The rest are estimates.) and known manatee deaths (from the Florida Fish and Wildlife Research Institute website) for the past several years. Make a scatterplot of these data.

1) Is there a relationship between the number of boat registrations and the number of manatee deaths? Explain.

2) Use your scatterplot to predict the number of Manatee deaths when the number of boaters is 900,000. You just interpolated (look that term up in the glossary). How many deaths will there be when the number of registrations reaches 2,000,000? You just extrapolated.

3) Based on these data, would you be willing to conclude that boating is causing the deaths of manatees? Explain your thinking.

4) Find the equation of a line that fits the data. Interpret the meaning of the slope and y-intercept of your line in the context of these data.

138
<table>
<thead>
<tr>
<th>Year</th>
<th>FL Boat Registrations (thousands)</th>
<th>No. of Manatee Deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>598</td>
<td>174</td>
</tr>
<tr>
<td>1992</td>
<td>628</td>
<td>163</td>
</tr>
<tr>
<td>1993</td>
<td>643</td>
<td>145</td>
</tr>
<tr>
<td>1994</td>
<td>712</td>
<td>193</td>
</tr>
<tr>
<td>1995</td>
<td>730</td>
<td>201</td>
</tr>
<tr>
<td>1996</td>
<td>785</td>
<td>415</td>
</tr>
<tr>
<td>1997</td>
<td>800</td>
<td>242</td>
</tr>
<tr>
<td>1998</td>
<td>820</td>
<td>232</td>
</tr>
<tr>
<td>1999</td>
<td>852</td>
<td>269</td>
</tr>
<tr>
<td>2000</td>
<td>880</td>
<td>272</td>
</tr>
<tr>
<td>2001</td>
<td>943</td>
<td>325</td>
</tr>
<tr>
<td>2002</td>
<td>961</td>
<td>305</td>
</tr>
<tr>
<td>2003</td>
<td>978</td>
<td>276</td>
</tr>
<tr>
<td>2004</td>
<td>983</td>
<td>396</td>
</tr>
<tr>
<td>2005</td>
<td>1010</td>
<td>417</td>
</tr>
</tbody>
</table>
Read and Study

*Facts are stubborn things, but statistics are more pliable.*

Mark Twain

In this section we will discuss the *idea* of fitting a best line to data. Typically in practice your software will actually perform the fit (called a ‘least squares regression’ or simply a ‘linear regression’); what you need to do is understand what it all means. First the purpose of fitting a line (or another function) to model data is to be able to make predictions about values of the response variable. For example let’s think about this data table and the below plot presented by the *Quantitative Environmental Learning Project* at the website: [http://www.seattlecentral.edu/qelp/sets/028/028.html](http://www.seattlecentral.edu/qelp/sets/028/028.html).

Source: US Geologic Survey Water Resources Investigations Report 89–4004 Water Quality and Supply on Cortina Rancheria, Colusa County, California

<table>
<thead>
<tr>
<th>Gauging Station</th>
<th>Estimated mean annual rainfall (inches)</th>
<th>Drainage Area (mi²)</th>
<th>Mean annual flow (ft³/sec)</th>
<th>average annual runoff per unit area (in./area)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Fork Cottonwood Creek near Ono</td>
<td>40</td>
<td>244</td>
<td>254</td>
<td>14.1</td>
</tr>
<tr>
<td>Red Bank Creek near Red Bluff</td>
<td>24</td>
<td>93.5</td>
<td>44.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Elder Creek at Gerber</td>
<td>30</td>
<td>136</td>
<td>96.2</td>
<td>9.6</td>
</tr>
<tr>
<td>Thomas Creek at Paskenta</td>
<td>45</td>
<td>194</td>
<td>304</td>
<td>21.2</td>
</tr>
<tr>
<td>Grindstone Creek near Elk Creek</td>
<td>47</td>
<td>156</td>
<td>193</td>
<td>16.8</td>
</tr>
<tr>
<td>Stone Corral Creek near Sites</td>
<td>21</td>
<td>38.2</td>
<td>6.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Bear Creek near Rumsey</td>
<td>27</td>
<td>100</td>
<td>44.3</td>
<td>6.0</td>
</tr>
</tbody>
</table>

It makes sense that a community would want to predict the amount of water runoff based on rainfall to better understand flooding, nutrient loss in the soil, or downstream pollutants from farms. The first thing to do then might be to make a scatterplot of these data to examine the pattern.
Have a look at the data points. Notice that each point represents a measurement from a stream or creek. Also notice that this set of data shows a fairly linear pattern and that there are no outliers among these data. *If there was an outlier on the scatterplot, should it be included when fitting a best line to the data? Explain your thinking.*

What does it mean to for a line to be a ‘best fit?’ Glad you asked. Mathematically a **best fit line** (or **regression line**) is one that minimizes the sum of the squares of the vertical distances of the points to the line. Hmm. *Read that definition again – it is a tough one.* We will spend the next paragraphs explaining just what that means and this is a big idea so we want you to work hard to understand it.

Imagine that the software program on your calculator or computer puts down a line on the above scatterplot (any line) like below. Then it computes the distance that each point is from the line (measured vertically) and squares each of those distances. Those squares are represented geometrically on our graph. Finally the computer sums the areas of all the squares and returns a number.
Now imagine the computer does all this again with a different line. *Have a close look at the next picture to be sure you can see how it shows how the new line creates different squares.*
This second line is a ‘better’ line because the sum of the squares (the final gray square) is smaller.

Now imagine that your software does this process with all possible lines and then tells you which one of them gave the very smallest final sum. That line is called the least squares regression line (see how the name fits the process?) or the line of best fit. Spend some time studying our explanation so that you understand it well enough to explain it to someone.

The way your computer will return this information to you is probably in the form of a slope and a y-intercept – in other words you’ll get the equation of the line. You may also get a number (r) called the correlation coefficient that tells you how good the fit really is – but that is a topic for the next section.

For the water runoff data, the line of best fit is given by the equation: \( y = 0.619x - 9.82 \) where \( y \) is annual water runoff per unit area (in inches) and \( x \) is annual rainfall (in inches). The 0.619 is the slope of the line, and so it tells us that if we are standing on the line, for every unit we move horizontally in the positive direction, we need to move up 0.619 units to stay on the line. Or in the case of these data, it tells us that for every 1 inch increase in the annual rainfall, we get an increase of 0.619 inches per unit area of runoff. Make sure that you understand that last sentence – it is important because it shows how to make sense of the numbers in the equation in terms of the physical situation.

In this context the y-intercept of -9.82 means that if there was no rainfall at all \( (x = 0) \), we’d get negative runoff (which suggests that the data looks different for very small amounts of rain). It is important that you think about the meaning of your line of best fit.

We said that a regression (or best fit) line is found so that we can make predictions about values of the response variable (in this case water runoff in streams) given values of the explanatory variable (in this case the annual rainfall).

Use the equation of the best fit line to predict the runoff in a stream that gets 60 inches of annual rainfall. Did you just interpolate or did you extrapolate? Why? How confident are you in your prediction? Explain.
Homework

The only thing that comes to us without effort is old age.

Gloria Pitzer

1) Do all the italicized things in the Read and Study section.

2) Decide whether each of the following is True or False and explain your thinking.
   a) Outliers should be removed from a data set before finding a line of best fit.
   b) The computer will find a best line even if the data is not linear.
   c) A positive slope for a best fit line through data means the explanatory and response variable have a positive association.
   d) A positive association between the explanatory and response variables means there will be a positive slope for the line of best fit.
   e) The better the line fits the data, the better its predictive power.
   f) The better the line fits the data, the more likely the explanatory variable causes variations in the response variable.

3) A linear relationship was found between the number of hours students studied for an exam ($t$) and the score of the exam ($s$). The least squares regression line for the data is given by the following equation: $s = 6.5t + 52$.
   a) What exactly does the 6.5 tell you in terms of this situation? What are its units?
   b) What does the 52 tell you and what are its units?
   c) If a student studied for 4 hours, what is her predicted score on the exam?
   d) If a student scored a 90 on the exam, how long would you predict she studied?
   e) Give at least two different possible explanations for the association between hours studying and exam score.

4) The data for the gold medal performances for each year since 1900 is provided in inches in the table on the next page. For this problem you should use a calculator or an online applet for computing regression line (ask your teacher to recommend one).
   a) For each data set, enter the data into the Applet with time as the $x$-variable and record performance as the $y$-variable.
   b) Before you do the regression describe the shape of the data. Is the relationship linear?
   c) If the data is linear, do the regression and use the line of best fit to predict the Olympic performance for the year 2012. How confident are you in that prediction? Explain.
<table>
<thead>
<tr>
<th>YEAR</th>
<th>Long Jump (in)</th>
<th>High Jump (in)</th>
<th>Discus (in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1900</td>
<td>282.9</td>
<td>74.8</td>
<td>1418.9</td>
</tr>
<tr>
<td>1904</td>
<td>289.0</td>
<td>71.0</td>
<td>1546.5</td>
</tr>
<tr>
<td>1908</td>
<td>294.5</td>
<td>75.0</td>
<td>1610.0</td>
</tr>
<tr>
<td>1912</td>
<td>299.2</td>
<td>76.0</td>
<td>1780.0</td>
</tr>
<tr>
<td>1920</td>
<td>281.5</td>
<td>76.3</td>
<td>1759.3</td>
</tr>
<tr>
<td>1924</td>
<td>293.1</td>
<td>78.0</td>
<td>1817.1</td>
</tr>
<tr>
<td>1928</td>
<td>304.8</td>
<td>76.4</td>
<td>1863.0</td>
</tr>
<tr>
<td>1932</td>
<td>300.8</td>
<td>77.7</td>
<td>1948.9</td>
</tr>
<tr>
<td>1936</td>
<td>317.3</td>
<td>79.9</td>
<td>1987.4</td>
</tr>
<tr>
<td>1948</td>
<td>308.0</td>
<td>78.0</td>
<td>2078.0</td>
</tr>
<tr>
<td>1952</td>
<td>298.0</td>
<td>80.3</td>
<td>2166.9</td>
</tr>
<tr>
<td>1956</td>
<td>308.0</td>
<td>83.3</td>
<td>2218.5</td>
</tr>
<tr>
<td>1960</td>
<td>319.8</td>
<td>85.0</td>
<td>2330.0</td>
</tr>
<tr>
<td>1964</td>
<td>317.8</td>
<td>85.8</td>
<td>2401.5</td>
</tr>
<tr>
<td>1968</td>
<td>350.5</td>
<td>88.3</td>
<td>2550.5</td>
</tr>
<tr>
<td>1972</td>
<td>324.5</td>
<td>87.8</td>
<td>2535.0</td>
</tr>
<tr>
<td>1976</td>
<td>328.5</td>
<td>88.5</td>
<td>2657.4</td>
</tr>
<tr>
<td>1980</td>
<td>336.3</td>
<td>92.8</td>
<td>2624.0</td>
</tr>
<tr>
<td>1984</td>
<td>336.3</td>
<td>92.5</td>
<td>2622.0</td>
</tr>
<tr>
<td>1988</td>
<td>343.3</td>
<td>93.5</td>
<td>2709.3</td>
</tr>
<tr>
<td>1992</td>
<td>342.5</td>
<td>92</td>
<td>2563.8</td>
</tr>
</tbody>
</table>

5) Now add an extra point various places on the graph of the above data and use the Applet to see how the position of an outlier affects the position of the line. Print out two of your pictures to bring to class.

6) Use an online applet or your calculator to find the line of best fit for the manatee problem. Use your line to predict the number of manatee deaths when there are 700,000 boat registrations. How confident are you in the prediction? What does the slope of the line tell you in terms of this situation? What does the y-intercept tell you? Explain.
**Class Activity 23: Correlation**

*The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents.*

H. P. Lovecraft

Consider the scatterplots below. In each case describe how the variables on the x and y-axes appear to be related. What exactly do you think “r” is measuring?

Correlation (r) = 0.05

Correlation (r) = 0.99

Correlation (r) = 0.70

Correlation (r) = -0.70

(This activity is continued on the next page.)
The **correlation coefficient** \( r \) is a statistic that measures the strength of the linear relationship between two variables. It is defined by the following equation where \( n \) is the number of values, \( m_x \) is the mean of the \( x \) variable, \( m_y \) is the mean of the \( y \) variable, \( s_x \) is something called the standard deviation of the \( x \) variable, and \( s_y \) is the standard deviation of the \( y \) variable. (Don’t worry - you won’t need to compute standard deviations or correlation coefficients. Just have a look at the equation to see what clues it gives you):

\[
r = \frac{1}{n} \left( \frac{\sum_{i=1}^{n} x_i m_x}{s_x} \frac{y_i m_y}{s_y} + \frac{\sum_{i=1}^{n} x_i m_x}{s_x} \frac{y_i m_y}{s_y} + \ldots + \frac{\sum_{i=1}^{n} x_n m_x}{s_x} \frac{y_n m_y}{s_y} \right)
\]

1) Does it matter which variable is the explanatory variable and which is the response when computing correlation? Why or why not?

2) Suppose \( x \) is time (in months) and \( y \) is height of a plant (in inches). What units does \( r \) have? Explain.

3) Does the value of \( r \) change when we change the units of measurement for \( x \) or \( y \)? Why or why not?

4) Is \( r \) resistant to outliers? Why or why not?

5) True or False? If \( r = 0 \), that means there is no relationship between the variables \( x \) and \( y \). Explain your thinking.
People use the term ‘correlated’ when they should say ‘associated.’ In mathematics, two numerical variable are highly correlated if have a strong linear relationship. If the variables are not measured in some numerical manner, then they cannot be correlated. If the variables show some other relationship (besides linear) they might be very strongly associated, but not very highly correlated. Read the quotes that head the Class Activity, Read and Study, and Homework sections about correlation. Do the authors mean to use the term? Or would ‘association’ be better? Explain. As mathematics teachers you will need to practice careful mathematical language.

Correlation is captured by a statistic called the correlation coefficient, \( r \), that can vary between \( r = 1 \), meaning that the data has a perfect linear relationship with a positive slope to \( r = -1 \), meaning that the data shows a perfect linear relationship with a negative slope. When \( r = 0 \), there is no linear relationship at all.

We will not compute a correlation ourselves – our calculator or computer will do that – but what we will do is interpret it. Below you will see lots of ‘clouds’ of data (each little picture is a data set) along with their correlation coefficients (from Wikipedia.com). Study them. Note that those with a positive slope have a positive value of \( r \) and those with negative slope have a negative value of \( r \). Notice that lots of clouds that show a relationship between the variables have an \( r \) of zero (check out that bottom row).
We have told you that the correlation coefficient measures the strength of a linear relationship between two variables – but we have to add a word of caution. The number $r$ should be computed only when it makes sense to do so. We will use the four data sets below to show you what we mean. All of these data sets have a correlation coefficient of 0.8, yet in only the first case (Data Set 1) does the value make sense. It makes no sense to fit a line to Data Set 2 when that data has a nonlinear shape. Data Sets 3 and 4 both have outliers that probably should have been discarded before fitting a line to the data – after that, the linear relationship is likely to be much stronger than 0.8.

In this class, any set of *nice linear-type* data (like that in Case 1 above) data with a correlation coefficient bigger than 0.75 will be considered linear data, but different people who use statistics will have different criteria for what constitutes a good line.

One more point. Even a strong correlation does *not* imply a causal relationship between the variables. Here’s an example: there is a strong correlation between teachers’ raises and alcohol sales in the state of Wisconsin. The bigger the raise, the most alcohol is sold.

Is this a causal relationship? Maybe. Perhaps teachers celebrate their raises by having parties. But more likely it is the case that in years when there is a good economy teachers get higher raises *and* alcohol sales are higher. So it may be that the economy is a *lurking variable* here – a hidden variable that causes *both* effects.
So how do we determine a causal relationship? We perform a clinical study. We’ll talk more about that in the next section.

**Homework**

*If I had to select one quality, one personal characteristic that I regard as being most highly correlated with success, whatever the field, I would pick the trait of persistence.*

*Richard M. DeVos*

1) Do all the italicized things in the Read and Study section.

2) Decide whether each of the following is True or False. Explain your choice.

   a) If the correlation coefficient is 1, then that means that one variable causes the other.
   b) If the data has a correlation coefficient of zero, then that means the variables are not related.
   c) We could measure the correlation between favorite color and personality.

3) Here is a data set that shows data for 12 perch caught in a lake in Finland:

   | Length (cms) | 8.8 | 19.2 | 22.5 | 23.5 | 24.0 | 25.5 | 28.7 | 30.1 | 39.0 | 41.4 | 42.5 | 46.6 |
   | Weight (grams) | 6 | 100 | 110 | 120 | 150 | 145 | 300 | 300 | 685 | 650 | 820 | 1000 |

   a) Make a careful scatterplot of length versus weight by hand for these data and describe the pattern. Are there any outliers? Explain.
   b) Use a calculator or online Applet to perform a linear regression and to find the correlation coefficient. What does $r$ tell you? Use this line to predict the weight of a fish that is 12 cm long. 30cm long. 60 cm long. Would you trust these predictions? Explain.
   c) Does it make sense that the weight of fish would be a linear function of length or would another curve better describe the relationship? Explain your thinking.
4) In each case, give a possible lurking variable that may explain the association described:

   a) As the number of boating permits increases in FL so do the number of manatee deaths. A researcher claims that boaters are killing the manatees. Can you argue with her?

   b) There is a strong positive association between height and reading ability among elementary children. Why?

   c) People who use artificial sweeteners tend to be heavier than those who use sugar. A researcher claims that eating artificial sweeteners causes people to gain weight. Can you argue with him?

   d) The bigger the hospital, the higher the percentage of deaths/per admission. Does this mean larger hospitals give poorer care?

   e) A researcher finds that people who listen to classical music have a higher IQ than those who do not. Based on this finding, is it reasonable to assume that listening to classical music boosts one’s IQ? Explain your thinking.
Today in class some of you will be participating in a clinical study to determine whether consuming 8 ounces of caffeine raises the heart rate of adults. Those of you who choose not to participate will compose the research team. Your instructor has brought two bottles of soda that are exactly alike except that one contains caffeine and the other does not. Neither you nor the research team will know which bottle is which until the experiment is complete. In order to participate, you must be willing to drink 8 ounces of soda (that may or may not contain caffeine) during the next ten minutes. Those who do not participate will pour and serve soda, and run and record heart rate sessions. Decide now if you will participate in the clinical study.

I) Participants should be assigned randomly to two groups. (Eventually one group will consume soda from Bottle I and the other will consume soda from Bottle II). The research team should decide how to do this.

II) All participants in both groups will take a one-minute resting heart rate.

III) Researchers should record heart rates for each participant and record which group he or she is in (I or II). The participants will drink 8 ounces of soda from their designated bottle.

IV) After participants are finished drinking and 15 minutes have elapsed (during which you will discuss in groups the below questions), everyone should once again take a 1-minute heart rate.

V) Finally the research team should measure changes in heart rate and report results of the raw data.

Questions for small-group discussion:

1) Does drinking 8 ounces of caffeine raise heart-rate? What statistics might you compute in order to answer this question?

2) In the case where one group was different from the other, how might you decide if that difference is just due to random chance?

3) What factors might have affected the results? How might you improve the design of this study?
In this section we’ll discuss the language involved in doing a clinical study. First, the purpose of a clinical study is to try to determine a cause and effect relationship between two variables. The caffeine and heart-rate study you did in the class activity was an example of such a study. The goal there was to see whether drinking some caffeine raised heart-rate. In the world of clinical studies, good ones are randomized-controlled – meaning that there is a treatment group and a control group (to which participants were randomly assigned) and the only difference between those two groups is the treatment variable. Even better are the blind randomized controlled studies - meaning that the participants do not know whether they are getting the treatment. In that case, usually a ‘fake treatment’ is given called a placebo. The very best clinical study design is one that is a double-blind randomized controlled study – in that case neither the participants nor the researchers know who is getting the treatment. The study you did in class was of this last type. Think that through and make sure you agree. What was the placebo in that study?

Just because a study is a randomized-controlled, double-blind clinical study doesn’t necessarily mean it will yield good results. As you may have found in the class activity, the conditions of the classroom may not have been ideal for measuring heart rate, there may not have been sufficient time to metabolize the caffeine, or perhaps the sample of participants was flawed (too small or too homogeneous – all the difficulties with sampling can come up in designing clinical studies too). Finally there is the problem of deciding whether there was really an ‘effect’ or whether the difference (if any) found between the two groups was simply due to random chance. Let’s consider an example of a real clinical study: The Salk Polio Vaccine Trial of 1954. It is an amazing clinical study because it is unprecedented in scope and scale. Here’s the story (as summarized from a paper by Dawson, L. (2004). The Salk polio vaccine trial of 1954: risks, randomization and public involvement, Clinical Trials, 1, pp. 122-130.)

In 1954, a terrifying disease called polio was a leadingcrippler of children across the globe. Just two years before, a breakthrough had occurred: Jonas Salk and his team had developed a vaccine using a killed-form of the polio virus. The National Foundation for Infantile Paralysis (NFIP) sponsored the study in which 1.8 million children participated. The study was to be double-blind and randomized- controlled (although in the end, some of the children in certain cities did not receive the placebo). The study needed to be huge – only about 5 children per 10,000 typically got polio in any given year – and it needed to be randomized-controlled because the number of cases varied a lot from year to year (so just comparing how many children had polio in years before to how many got in 1954 would not give enough evidence for effectiveness of the vaccine).
Think of the risk. What if the vaccine wasn’t safe? What if some live virus remained in the vaccines? What if there were unpredictable side-effects? Still, parents were so afraid of the disease, and the NFIP so effective at promoting the trial, that over 60% of parents signed up their children to participate. Children were inoculated with either a placebo or a vaccine designed to protect against three forms of polio between April 26 and June 15, 1954.

About one year later on April 12, 1955, more than one hundred and forty reporters gathered at the University of Michigan to hear the results of the trial: protection was 68% against type I, 100% against type II, and 92% against type III infection. The vaccine was licensed by the US Public Health Service that same day. Today polio is all but eradicated in the US.

**Homework**

You know how it is when you go to be the subject of a psychology experiment, and nobody else shows up, and you think maybe that’s part of the experiment? I’m like that all the time.

*Steven Wright*

1) Do any italicized things in the Read and Study section.

2) The following (fictitious) study was conducted to determine whether a new reading program “Read On” was more effective at teaching children reading than a traditional approach using phonics. Three elementary schools in San Diego that had been using the phonics approach for several years agreed to participate in the study. The “Read On” approach was used by all first grade teachers in 2002 and reading scores from the 120 first graders at the end of 2002 were compared with the data from the 112 first graders who had used the phonics approach in 2001.

b) What was the population for this study?

c) What was the sample?

d) True or False? This study is an example of a randomized controlled study. Explain.

e) True or False? This study was blind. Explain.

f) Give two good examples of possible sources of selection bias for this study.

g) If this study showed that the “Read On” students scored significantly higher on tests of reading than did the phonics students, would you buy the argument that “Read On” is the more effective program? Explain.

h) A **confounding variable** in a clinical study is a variable that was not controlled in the experiment and may have caused the observed effect. Give an example of such a variable for this study.
3) A researcher believes that a new diet supplement will improve marathon performance. In order to test her theory, she contacted by email 500 people who had registered early for the Green Bay Marathon. 42 people agreed to participate in her study. Half (21 runners) were selected at random to take her supplements daily for two months prior to the event. The remaining half trained as usual.

In the treatment group, 16 runners completed the race. Here are their finishing times (in minutes):

302, 288, 276, 265, 256, 230, 220, 212, 210, 210, 201, 200, 170, 150, 148, 146

In the control group, 15 runners finished the race. Here are their finishing times (in minutes):

476, 310, 278, 266, 264, 230, 210, 210, 208, 202, 200, 180, 176, 146, 142

b) What is the population for this study?
c) Find the mean of each set of finishing times.
d) Find the 5-number-summary of each set and make boxplots to help you compare these data.
e) Are there any outliers in either set? Use our criterion to decide.
f) The researcher claims that her supplement lowered finishing times for runners in the treatment group. Do you agree? Make a good argument.
g) The researcher claims that her supplement lowers finishing times in marathon runners. Do you agree? Make another good argument.
h) Now put both sets of data together and make a side-by-side bar graph. What does this presentation of the data show?
Class Activity 25: Music and Math

Where words fail, music speaks.

Hans Christian Anderson

A research team believes that taking piano lessons will improve elementary school children’s spatial reasoning abilities. In order to test this conjecture, they selected at random 20 elementary schools in Winnebago County, WI and then selected 20 first and second graders at random from each school. Of these 400 children who were identified, 360 agreed to participate. Half of the children (again chosen at random from the 360) were given 1 year of piano lessons for one hour each week. The other half were given 1 year of computer lessons for one hour each week. All children were tested both before and after the year of lessons using a standardized test of spatial visualization. The piano group made significantly better gains on the test after 1 year than did the computer group.

Decide whether each of the following is True or False and briefly justify your group’s choice.

1) The study described uses cluster sampling.

2) The population is elementary school children in 20 Winnebago County schools.

3) The sampling rate is 90%.

4) The study described is a randomized controlled study.

5) The study described is blind.

6) The study described is double blind.

7) This study uses a placebo.

8) This study may be biased because children in Wisconsin may not be representative of all children.

9) Based on this study, your group would conclude that taking piano lessons does cause improvement in the spatial-visualization of elementary school children.

Now, give examples of some possible confounding variables for this study.
**Class Activity 26: Class Survey Revisited**

*Statistics: The only science that enables different experts using the same figures to draw different conclusions.*

*Evan Esar*

In an earlier activity, you designed a class survey in order to examine teacher preparation at your school. Today, you will take all those surveys as a class, and then your group will be given the raw data generated by your questions.

1) Your group should decide how to display each data set for your report to the American Federation of Teachers using graphical representations we have discussed in this chapter. You will need to decide how to handle any difficulties that arise. Remember to label your graphs. Here is a reminder of some possible types of displays:

- Pictograms
- Stem and Leaf plots (Stemplots)
- Line plots
- Bar graphs
- Histograms
- Boxplots

2) What cautions regarding the data do you want to highlight in your report to the AFT?

3) What did you learn about writing survey items from this activity? Be specific.

4) Select two of your items that might be related (e.g., Number of hours of studied per week and College GPA) and make a display that would help you analyze the relationship.
Chapter Six

Probability Revisited
Class Activity 27: Casino Royale

All the evidence shows that God was actually quite a gambler, and the universe is a great casino, where dice are thrown, and roulette wheels spin on every occasion.

Stephen Hawking

The Casino Royale game at the County Fair allows you to spin a wheel containing 16 spaces numbered 1 through 16. If you spin any odd number, you lose and get nothing. If you spin a multiple of 4 (the numbers 4, 8, 12, or 16), you win $3. If you spin anything else you win the amount of money specified by the number (e.g., spin a 2, get $2; spin a 6, get $6, etc.). If you play this game many times, what can you expect to win on average?

A fair game is one in which your expected net winnings would be $0. What should a ticket cost to make this a fair game?

Suppose the ticket cost $3 and the rules of the game do not change. How can you change the payouts to make it a fair game? Is there more than one set of payoffs that would work? Explain.
In the Casino Royale class activity you were asked to determine your expected winnings. As with all mathematical terms, “expected” does not mean exactly what you might expect it to mean. If you are a very optimistic person, you may expect to win the lottery every time you play (after all, someone has to win – why not you?). When the star of your school’s basketball team steps to the free throw line, you may expect him or her to make the shot. However, when a mathematician talks about the expected value of an event she means the average value of the event over a very large number of repeated trials.

For example, suppose we roll a fair die. What is the expected value of the outcome?

Think about this question for a minute – what value makes sense as the expected value of the roll of a die?

Notice that expected value cannot mean “the most probable” outcome, since each of the possible six outcomes (1, 2, 3, 4, 5, 6) is equally likely. Let’s set up an experiment to determine the expected value as an average. Suppose we roll the die 10,002 times and record the outcome of each roll. Now we add up all the outcomes and divide by 10,002 to find the average value. (If you are inclined to carry out the experiment, do so now and bring your results to class.)

Instead, suppose we prefer to reason theoretically (which we do). In imagining the 10,002 rolls we would expect a 1 to occur 1/6\(^{th}\) of the time, a 2 to occur 1/6\(^{th}\) of the time, etc., since each outcome is equally likely. So we would expect about 1667 1’s, 1667 2’s etc. So the theoretical average value would be

$$
\frac{1667 \times 1 + 1667 \times 2 + 1667 \times 3 + 1667 \times 4 + 1667 \times 5 + 1667 \times 6}{10002} = 3.5
$$

We can also compute the expected value as the sum of the products of each outcome with the probability of that outcome:

$$
1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3.5.
$$
Explain how the two calculations are related mathematically. Are they really two different ways of finding the theoretical average value? Why or why not?

So we can determine an expected value in two ways: 1) the average value of the outcome over many, many trials (experimental), or 2) the probability-weighted sum of the outcomes (theoretical). Practically speaking, since we need a very, very large number of trials to find expected value experimentally, we use the theoretical method in almost all cases. We can generalize the computation of expected value to give the following definition. Here each \(x\) is an outcome and \(P(x)\) is the probability of that outcome. Study this formula to see how it fits with our second computation above.

\[
\text{Expected value of } x = x_1 \times P(x_1) + x_2 \times P(x_2) + \ldots + x_n \times P(x_n)
\]

Notice that the expected value (since it is an average) can be a value that is not one of possible outcomes.

**Connections to the Elementary Grades**

*There was no rational reason why monkeys might prefer one of these options over the other because, according to the theory of expected value, they’re identical.*

Michael Platt

In the *Middle Grades Mathematics Project* unit on Probability, the authors suggest that students use area models to analyze events in the course of determining the average payoff over a very long run of trials (the expected value). For example, have a look at the below spinner. The red sector contains \(1/4\) of the area of the circle, the blue sector contains \(1/3\) of the area, and green sector contains the balance. So the probability of landing in red is \(1/4\), the probability of landing in blue is \(1/3\), and the probability of landing in green is \(5/12\). *Color the spinner and then check that last probability for yourself.* Notice that it doesn’t make sense to compute the expected value (color?) of the spinner. But if you won $5 for landing on Red, $1 for landing on Blue and nothing for landing Green, then you could compute your expected winnings like this:

\[
\frac{1}{4} \times ($5) + \frac{1}{3} \times ($1) + \frac{5}{12} \times ($0) \approx $1.58.
\]

*What does the $1.58 really mean? Explain it as you would to your upper elementary students.*
Now consider a problem about a basketball game. Suppose Terry is a 60% free-throw shooter and is in a one-and-one situation. (Recall that a team is in a one-and-one situation when the opposing team has collected a total of seven fouls; each time the team goes up for a free throw, they will get a second free throw only if they make the first one.) The children are asked to investigate the following three questions:

1) On any given trip to the free throw line (trial) is Terry most likely to get 0 points, 1 point, or 2 points?
2) Over many trials, say 100, about how many total points would Terry expect to make?
3) Over the long run what is the average number of points per trip (expected value)?

Children are first asked to predict answers to each question *Take time now to make your predictions.*

Next children are asked to determine the theoretical probabilities for question 1. *Take time to do this (a rat maze might help). What are the theoretical probabilities of 0, 1, or 2 points?*

Finally, the children are asked to design a simulation using a spinner to experimentally determine the probability of Terry getting 0, 1, or 2 points on any one given trip to the free throw line. *Why would the spinner above prove useful in this regard? How could you improve the accuracy of the simulation?*

---

**Homework**

*One of the healthiest ways to gamble is with a spade and a package of garden seeds.*

Dan Bennett

1) Do all the italicized things in the *Read and Study* and *Connections* sections.
2) Answer the three questions in the *Connections* section for a 50% free-throw shooter, an 80% shooter, and a 90% shooter. For each case see if you can come up with more than one simulation model and more than one theoretical method to determine the answers.

3) Explain how the law of large numbers is related to the concept of expected value.

4) What are your expected winnings in a fair game of chance? Explain.

5) Every day at your job you are offered the following choices for an hourly rate: $7 an hour, or the amount of money you select from a bag containing one $20 bill, two $5 bills and a $1 bill. Which plan would you choose and why? Use expected value to help you decide.

6) What is the expected value of the sum of a roll of two fair dice? What is the expected value of the difference of the two dice?

7) Suppose the premium of a home insurance policy is $300 and in case of a fire the insurance company will pay out $200,000. Suppose the probability of a fire is 0.0002. If the insurance company issues 10,000 such policies, what is its expected profit?

8) Here is a fun game! You roll a die two times. If you roll exactly one six, you win $10. If you roll exactly two sixes, you win $50. Otherwise you win nothing. What are your expected winnings if you play this game once? Explain how to interpret that number.
1) I just received a shipment of 4 light bulbs. If each bulb has a 15% chance of being defective. What is the probability that none of the bulbs are defective? Exactly one? Exactly two? Exactly 3? Exactly 4? Make a maze, and use it to see.

2) I just received a shipment of 30 light bulbs. What is the probability that exactly 4 are defective?

3) If I flip a fair coin 12 times, what is the probability that I get fewer than three Heads?
Read and Study

Music is the pleasure the human mind experiences from counting without being aware that it is counting.

Gottfried Leibniz

In the Class Activity you explored the behavior of a whole family of probability problems: those that have a binomial distribution. In order for a problem to fit into this family, you must have a binomial setting and this means that four criteria must be met:

1) You are performing a fixed number of independent random trials (like selecting thirty bulbs). Let’s call it that number \( n \).

2) For each trial there are only two possible outcomes (like good and bad). Let’s call one of those outcomes “success” and the other “failure.” (This is the condition that puts the “bi” in binomial.)

3) The probability of success is the same for every trial. If we call that probability \( p \), then the probability of failure will be \( (1 - p) \). (Why?)

4) You are interested in computing the probability of a number of successes (or failures). (For example, trying to answer, “What is the probability that you get exactly 2 bad bulbs.)

Explain why the coin problem in the Class Activity meets each criterion. What is \( n \) for that problem? What is \( p \)?

The good news is that if you can identify a probability problem as a binomial setting, then you can solve it. In the Class Activity you built the tool for solving these types of problems. Let’s look more closely at a mini version of that bulb problem to make sure you see what we mean. Here’s that problem:

I just received a shipment of 3 light bulbs. If each bulb has a 15% chance of being defective. What is the probability that none of the bulbs are defective? Exactly one? Exactly two? Exactly three?

The first thing we do when faced with a probability problem is to begin a tree diagram. Note that the first bulb has a 0.15 probability of being defective and a 0.85 probability of being good. So does the second. And the third. So our tree looks like this:
Using the tree, the probability of no bad bulbs is \( (0.85) \times (0.85) \times (0.85) = (0.85)^3 = 0.614125 \)

We’ll write it like this: \( P(\# \text{ bad} = 0) = 0.614125 \)

Now to get the probability that exactly one bulb is bad, we need to add up the final probabilities of each branch where there is exactly one bad bulb. There are three paths like that:

\[
\begin{align*}
\text{BGG} & \quad \text{GBG} & \quad \text{GGB} \\
P(\# \text{ bad} = 1) & = 3 \times (0.15) \times (0.85) \times (0.85) = 3 \times (0.15)^1 \times (0.85)^2 = 0.325125 \\
\end{align*}
\]

Likewise the probability of exactly two bad bulbs is

\[
P(\# \text{ bad} = 2) = 3 \times (0.15) \times (0.15) \times (0.85) = 3 \times (0.15)^2 \times (0.85)^1 = 0.057375
\]

Why does the three appear in the above equation? Explain.

Finally, there is only one way to get all bad bulbs. Compute that probability now.

When you are done, add up \( P(\# \text{ bad} = 0) + P(\# \text{ bad} = 1) + P(\# \text{ bad} = 2) + P(\# \text{ bad} = 3) \) and see what you get. Then explain why that answer makes sense.
So you might have noticed that all these binomial probabilities are going to be products involving the probability of success $p$ to a power and the probability of failure $(1 - p)$ to a power along with a “coefficient” representing the number of paths that contain that many bad bulbs.

Now what if the number of trials is bigger? What if we have a shipment of 40 bulbs and want to know that probability that exactly 5 are bad?

Well we know that we are interested in paths that contain 5 bad bulbs and 35 good ones. Each path like that will have probability $(0.15)^5 \times (0.85)^{35}$, but how many paths are there like that?

To answer this question, we need to think back to the counting sections of this text. What we really need to count here is the number of strings that contain 35 good bulbs and 5 bad bulbs.

Here is one such string:

$$BBGGGGGBGGBGGGGGBGGGGGBGGBGGGGGBGGBGGGGG$$

So out of 40 blanks we must choose where to place the Bad bulbs (then the rest must be Good).

What is the answer to that question and why? If you know, then carefully explain it. If you don’t, then spend ten minutes checking smaller examples in order to figure it out and bring your work to class.

---

**Homework**

*Mathematics is as much as aspect of culture as it is a collection of algorithms.*

*Carl Boyer*

1) Do all of the italicized things in the *Read and Study* section.

2) Did you really work on that last *Read and Study* question about the G’s and B’s? If not, do it now; it’s important.

3) Decide whether each of the following is a binomial situation. In each case explain your thinking.
a) I roll a die 30 times. What is the probability that I get exactly 5 twos?
b) I flip a coin until I get a head. What is the probability that I need to flip it 4 times?
c) I have a dish containing 12 red candies and 6 green candies. I select 5 candies one by one at random without replacing them. What is the probability that I get exactly 1 green candy?
d) I roll a die 20 times. What is the probability that I get at least 7 twos?

4) Use a tree diagram to solve this problem: A spinner has a 25% chance of landing “red” and a 75% chance of landing “blue.” If I spin it 4 times, what is the probability of getting exactly 2 “reds“?

5) A spinner has a 25% chance of landing “red” and a 75% chance of landing “blue.” If I spin 30 times, what is the probability that I get exactly seven “reds”? Less than seven “reds”? (You can do that second question – stop and think about it for awhile.)

6) If I choose a sample of 100 people at random from a large population containing 37% men and 63% women, what is the probability that my sample contains exactly 37 men? (Note that this situation isn’t quite binomial – explain why not. But we still can get a really good estimate of the desired probability – explain why that is so.)
Class Activity 29: The Problem of Points

Build up your weaknesses until they become your strong points.

Knute Rochne

Ann and Bob (who are both equally-skilled players) are playing a game of chance for high stakes when suddenly the game is indefinitely interrupted. Each round of the game is worth a point. Bob still needs 3 points to win. Ann needs only two points to win. If each player is just as likely to win any given point as the other, how should the $10,000 prize money be divided between them? Argue that you are correct.
Read and Study

God not only plays dice. He also sometimes throws the dice where they cannot be seen.

Stephen Hawking

Sometimes new mathematics is created which seems to have no “real world” application at the time, but which later plays an important role in a developing area, such as computer graphics. At other times, new areas of mathematics are created specifically to solve real world problems. Such is the case with the study of probability, which arose to answer questions that surfaced in the worlds of gaming and gambling.

The systematic study of the theory of probability is thought to have originated in the 17th century with the solution of a “problem of points” by two French mathematicians: Pascal and Fermat. In 1654 they wrote letters back and forth in which they solved the general case (for two players) of the exact same problem you solved in the Class Activity. Not only did they solve this problem, but they also laid the groundwork for the study of probability as a mathematical discipline. Their main ideas were popularized by Huygens, in his *De ratiociniis in ludo aleae*, published in 1657.

During the century that followed their work, other mathematicians, including James and Nicholas Bernoulli and De Moivre, developed more powerful mathematical tools in order to calculate odds in more complicated games. De Moivre, and others also used the theory to calculate fair prices for annuities and insurance policies.

Though probability was a mathematical theory by 1750, the applications of the theory were still only to questions of fair-distribution. No one had considered how to use probability in data analysis. This was true even in the work on annuities and life insurance. It was work on combining observations in astronomy and geodesy that finally brought data analysis and probability together. This work inspired Laplace's invention of the method of inverse probability (which we know as Baysian method of statistical inference), and culminated in Legendre's publication of the method of least squares in 1805. These ideas were brought together in Laplace's great treatise on probability, *Théorie analytique des probabilités*, published in 1812. (Shafer, 1993)

You may wonder why probability is such a young mathematical discipline. We did. After all, geometry is 2500 years old – so why is probability only around 350 years old? There are several theories about that, but the main one is that “chance” itself may be a fairly new human idea; until recently, cultures tended to ascribe all outcomes to the will of God or gods. You rolled a die and won? That is because God willed it. You were struck by lightning? You must have done something really bad. You can imagine that in such a culture, it would make no sense to talk about “random” events, and so it would make no sense to study probability.
Today probability has evolved a rich branch of pure mathematics, and its role as a foundation for mathematical and applied statistics is only one of its many roles in the sciences.


**Homework**

*You’ll always miss 100% of the shots you don’t take.*  
*Wayne Gretzky*

1) Suppose Bob had been winning 19 to 17 when they he and Ann were interrupted. How should the $10,000 be divided? What if he were winning 19 to 16? 18 to 16?

2) Make a conjecture: Does the fair division of the money depend on the number of points the leading player is leading by? Does it depend on the number of games the leading player has already won? Does it depend on the number of games the leading player still needs to win? Is the number of games the losing player has won/needs to win also matter? Collect enough data to make a conjecture and then argue that your conjecture is true.

3) Suppose you had come upon them earlier and noticed that they were tied 15 to 15. You leave and come back 5 flips of the coin later. What’s the probability of Bob being ahead 19 to 16?
Appendices
Probability Simulation Poster Project

Project: For this group project you will analyze a game of chance using both simulation (to compute an experimental probability) and the structure of the problem (to compute a theoretical probability). Each group will be assigned to study a different game below and then create a poster describing the work. Your group’s poster should be visually interesting and include four sections:

1) **Introduction to the Game.** Provide a description of the game and explanations of the questions you will address (Convince us that you understand the problem(s) and define ambiguous terms or ideas.)

2) **Simulation Information.** When you played the game, what did you look for? What did you keep track of? Append the tables of data you collected.

3) **Solution.** What are the answers to your question(s)?

4) **Justification.** Why does your solution make sense mathematically? Argue that it is complete and correct.

### The Games:
1) Pass the Pigs
2) Roulette
3) Dice Game
4) Snake
5) Sum it Up
6) Double Up
7) Random Walk

Rules: This is a project for your group, so you will work together on thinking about the questions and creating the poster. You may not look up anything (on the web, in any text other than our textbook, or from someone else’s old work...) about your specific game or project questions.
Pass the Pigs
(See the game for instructions on how to play and score.)

1) Is the scoring for this game appropriate based on experimental probabilities? If not, what do you suggest for changing the scoring and why?

2) Based on the rules as given, how long should you roll (to what point accumulation) each round to maximize your chances of winning? Argue that you are correct.

Roulette

A Roulette wheel contains the numbers 1 through 36, 0 and 00. Thus it has 38 places where the ball can stop. You choose a number or combination of numbers, the wheel spins, and if the ball stops on your number, you win. If you pick a single number and if the ball stops on your number, you get $36 back for every $1 that you bet. If you pick odd or even and you win, you get double your money back. If you pick one third of the numbers (1, 2, … 36), you win triple your money back if one of your numbers comes up.

The question: I have a strategy for winning at roulette! You first bet $1 on the odd numbers. If you win, you have made a dollar. If you lose the first time, you bet $2 on the odd numbers the second time. If you win this time, you receive $4 back and it cost you $3, so you have again made a dollar. If you lost the second time, you bet $4 on the odd numbers. If you win this time, you receive $8 and the cost has been $7, so you have still made a dollar! Continue this process… Since the ball cannot go on avoiding the odd numbers forever, sooner or later you get ahead by a dollar and then you start the whole process over again! What do you think? Will this strategy work? Why or why not? Can you prove it?

Winning Ticket

A box contains 12 tickets. Each ticket has a different number written on it. Other than that, we know nothing about them. The numbers can be positive or negative, integers or decimals, large or small – anything goes. Among all the tickets in the box there is, of course, one that bears the biggest number of all. That’s the winning ticket. If we turn that ticket in, we win a prize. If we turn any other ticket in, we get nothing. The ground rules are that we can draw a ticket out of the box, look at it, and if we think it’s the winning ticket, turn it in. If we don’t, we get to draw again, but first we must tear the other ticket up (once we pass on a ticket, we cannot use it again). We can continue drawing tickets this way until we find one we like or we run out of tickets… What strategy gives us the very best probability of winning? Can you prove it?
Dice Game

In this game, the player’s first roll of a pair of dice is very important. If the first roll is a 7 or an 11, the player wins. If the first roll is a 2, 3, or 12 the player loses. If the first roll is any other number (4, 5, 6, 8, 9, 10), this number is called the player’s “point.” The player continues to roll until either the point reappears, in which case the player wins, or until a 7 shows up before the point in which case the player loses. What is the probability of winning this game?

Snake

The object of the game is to accumulate as many points as possible in 5 rounds (the rounds are names by the letters in the word “Snake”). In each round a pair of dice is rolled. If neither of the dice shows a 1, then the total value of the roll is recorded, and a decision is made whether to roll the pair of dice again. If either die shows a 1, then the person’s turn is over and all points accumulated by that player in that round are lost. If snake eyes (double ones) is rolled then all points accumulated by the player (in all previous rounds) is lost as well.

1) Describe a safe strategy and then a risky strategy for playing Snake and see what kinds of distributions of scores each produces.
2) What is the optimal strategy for winning at Snake? Can you prove it?

Sum it Up

To play this game, you roll a pair of dice and always take their sum. Each time you roll a sum of 7, you win $2 and may continue rolling. If you roll any other odd sum, the game is over. If you roll doubles, you get $1 and can continue rolling. If you roll any other even sum, the game is over. How much should the game cost to play to make it a fair game? Can you prove it? (A game is fair if the expected return on each $1 is $1.)

Double Up

This is a game played by rolling a pair of dice and taking their sum. If you roll an odd sum, the game is over. If you roll doubles of any number, you win that value (for example, if you roll double fives, you win $5) and the game is over. If you roll an even sum (other than doubles) you may continue rolling. A casino needs to set a price for playing this game so that they make a profit in the long run. How much should the game cost to play?
Random Walk

Random walk is a game involving the board below and a single die. A player may start from position A or position B. The player rolls the die, and if a 1 or 2 appears, the player moves one space horizontally (toward the prize diagonal). If a 3, 4, 5, or 6 appears, the player moves one position vertically (toward the prize diagonal). The player continues to roll and move until he or she lands on the prize diagonal and wins the prize indicated.

1) Is it better to start at position A or B? Give me your evidence.

2) What should it cost to play this game in order for the game to be “fair?” (A game is fair if the expected return on a $1 bet is $1. The expected return is defined to be the product of the probability of the outcome and the amount of money you will receive if you win.) Prove you are correct.
References

Give credit where credit is due.

Author unknown


- *Mathematical Quotations Server (MQS)* at math.furman.edu.


**Glossary**

**Availability Fallacy**: The tendency to make decisions regarding chance based on the how easily certain events can be called to mind.

**Bar graph**: A data display in which the vertical axis represents a count (frequency) or a percent (relative frequency) and the horizontal axis represents values of a categorical variable (like colors), values of a numeric discrete variable (like shoe size), or values of a continuous variable (like heights), grouped into intervals.

**Base Rate Fallacy**: The tendency to make decisions regarding chance that neglect the rates at which relevant characteristics appear in the population.

**Binomial Setting**: A random experiment in which there are a fixed number of independent random trials, each with two possible outcomes (success and failure). The probability of success must be the same for every trial, and the computation of interest is the probability of a number of successes (or failures).

**Blind**: A clinical study is blind if the participants do not know whether or not they have received the treatment.

**Categorical Data**: Data for which it makes sense to place an individual observation into one of several groups (or categories).

**Chance Error**: Error in sampling due simply to the fact that a sample is not exactly the same as a population due to chance alone.

**Census**: Any data collection method in which data is collected from each and every member of the population.

**Clinical study**: Any study designed to determine a cause and effect relationship between two variables.

**Cluster sampling**: Random sampling in stages.

**Confounding variable**: Any variable that was not controlled in the experiment and may have caused the observed effect.

**Conjunction Fallacy**: The tendency to make decisions regarding chance without regard to the fact that A and B both occurring is less likely (or equally likely) than A occurring alone.

**Control group**: In a controlled study, the control group is a group of participants who does not receive the treatment under study and is used for comparison with the treatment group.
**Correlated**: two numerical variables are said to be correlated if have a strong linear relationship.

**Correlation coefficient** \((r)\) is a statistic that measures the strength of the linear relationship between two numerical variables.

**Data**: Any information collected from a sample.

**Distribution of data**: a display that shows the values the data takes along with the frequency (or relative frequency) of those values.

**Disjoint (events)**: Two events are disjoint if they have no outcomes in common.

**Double-blind**: A clinical study is double-blind if neither the participants nor the researchers know who has received the treatment (until after the study is over).

**Equally likely outcomes**: All the outcomes of a random experiment are said to be equally likely if, in the long run, they all occur with the same frequency (they all have the same probability of occurring).

**Event**: any subset of the sample space of a random experiment.

**Expected value (of a numerical event)**: the average value of the event over a very large number of trials.

**Experimental probabilities** are probabilities that are based on data collected by conducting experiments, playing games, or researching statistics.

**Explanatory variable** is one that is used to explain variations in another variable, called the response variable.

**Extrapolation**: Making a prediction outside the bounds of the data.

**First quartile** \((Q_1)\): the median of the data positioned strictly before the median when the data is placed in numerical order.

**Five Number Summary**: of numerical data consists of the following five statistics: Minimum value, \(Q_1\), Median, \(Q_3\), Maximum value.

**Gambler’s Fallacy**: The tendency to make decisions regarding chance based on the (mistaken) belief that after a long string of losses, a win becomes more likely. (This is a type of representativeness fallacy.)
**Histogram**: A data display in which the vertical axis represents a *rate* and the horizontal axis is actually the x-axis - in other words, it is a continuous piece of the real number line and as such it contains all real numbers in an interval. Values of the variable are broken into interval ‘classes.’ (Note: many texts use terms histogram and bar graph interchangeably. We will not do so in this text.)

**Independent (events)**: Two events are said to be independent if knowing that one occurs (or knowing it does not occur) does not change the probability of the other occurring.

**Interpolation**: Making a prediction within the bounds of the data.

**Inter-Quartile Range (IQR)** = $Q_3 - Q_1$. It is the spread of the middle half of the data.

**Law of large numbers**: In repeated, independent trials of a random experiment, as number of trials increases, the experimental probabilities observed will converge to the theoretical probability. In other words, the average of the results obtained from a large number of trials should be close to the expected value, and will tend to become closer as more trials are performed.

**Mean**: The mean (or average) value of a set of numeric data is the sum of all the values divided by the number of values.

**Mean Absolute Deviation** ($d$) = $\frac{1}{n} \sum |(x_1 - \mu)| + |(x_2 - \mu)| + |(x_3 - \mu)| + \ldots + |(x_n - \mu)|$. It is the average distance of the $n$ data points to their mean.

**Median**: The median of a set of numeric values is the middle value when the data is placed in numerical order. In the case where there are two “middle” values, the median is the average of the two.

**Mode**: the value that occurs most frequently.

**Multiplication principle**: If you have $A$ different ways of doing one thing and $B$ different ways of doing another (and one choice does not affect the other) then the total number of ways to do both $A$ and $B$ is $A \times B$.

**Negative association**: Two variables measured on the same individuals have a negative association if increases in one variable tend to correspond to decreases in the other variable.

**Nonresponse bias**: Biased information that results from the fact that some groups in the sample were more likely to respond than others.

**Numerical Data**: number data for which it makes sense to perform arithmetic operations (such as averaging).
**Outcome:** Any one thing that could happen in a random experiment.

**Outlier:** a specific value that lies well outside the overall pattern of the data.

**Parameter:** Some numerical information that is desired from an entire population. Usually a parameter is an unknown value that is estimated by a statistic.

**Percentile:** A score in the $n^{th}$ percentile means that “$n$ percent” of the people who took the test scored at or below that score.

**Placebo:** a fake treatment given to the control group in a clinical study.

**Population:** the largest group about which the researcher or surveyor would like information.

**Positive Association:** Two variables measured on the same individuals have a positive association if increases in one variable tend to correspond to increases in the other variable.

**Probability:** refers to the proportion of times an event would occur if the random experiment was performed a very large number of times.

**Proportion:** A proportion is a statement that two ratios are equivalent.

**Range** = maximum value – minimum value.

**Random Experiment:** any experiment where the outcome depends on chance and cannot be known beforehand. While in a random experiment, the specific outcome cannot be known, there is nonetheless a regular distribution of outcomes after a very large number of trials.

**Randomized-controlled:** a clinical study is said to be randomized-controlled if participants are assigned to treatment and control groups at random.

**Ratio:** a comparison of two counts or measures that have the same unit.

**Representativeness Fallacy:** The tendency to make decisions regarding chance based on the (mistaken) belief that even small samples should be representative of population.

**Regression line:** the line that minimizes the sum of the squares of the vertical distances of the points to the line on a scatterplot.

**Response Rate:** The ratio of the number of respondents (people who actually took part in the study) to the number who were invited to participate.

**Response variable:** a variable whose variation is explained by another variable (called an explanatory variable).
**Sample**: a subset of a population from which data is collected

**Sample bias**: Biased (inaccurate) information that results from a poorly chosen sample.

**Sampling error**: the difference between the value of a parameter and the value of the statistic that represents it. Sampling error can include chance error, sampling bias, and nonresponse bias.

**Sampling rate**: \[\left(\frac{\# \text{ in the sample}}{\# \text{ in the population}}\right)\]

**Sample space**: The sample space of a random experiment is the set of all possible outcomes.

**Scale factor**, a value by which we multiply each original dimension to find the new lengths.

**Scatterplot**: A display that shows the relationship between two numerical variables measured on the same sample of individuals. The values of one variable are shown on the horizontal axis, and values of the other variable are shown on the vertical axis. Each individual is plotted as a point representing an ordered pair (variable 1, variable 2).

**Self-selected sample**: a sample in which respondents volunteered to participate.

**Similar**: Two geometric figures are similar if there is some scale factor we could use to make them coincide. (Similar figure have the same shape, but are not necessarily the same size.)

**Simple Random Sample (SRS)**: A sample in which every subset of the population has the same chance of being in the sample as any other subset of the same size.

**Skewed left**: A distribution of data is skewed left if it has an asymmetric tail to the left.

**Skewed right**: A distribution of data is skewed right if it has an asymmetric tail to the right.

**Slope** (\(m\)): the slope of a line is the change in \(y\) for a one unit change in \(x\)

**Standard Deviation** (\(s\)) = \[\sqrt{\frac{1}{n}\left(x_1 - \mu\right)^2 + \left(x_2 - \mu\right)^2 + \ldots + \left(x_n - \mu\right)^2}\] where \(\mu\) is the mean of the values, \(n\) is the number of values, and each \(x_i\) is an individual value. Standard deviation is a statistic that measures spread around the mean.

**Statistic**: Any numerical information computed from a sample. (For example, a mean calculated from a sample is a statistic, which can be used as an estimate for the population mean, which would be a parameter.)
**Stem and leaf plot** (or simply **stemplot**): is a listing of all the data typically arranged so the tens place makes the stem and the ones are the ‘leaves.’

**Survey**: Any data collection method in which data is collected from just a subset of the population.

**Symmetric Distribution**: A distribution that is symmetric about the median.

**Theoretical Probabilities**: probabilities assigned based on assumptions about the physical uniformity and symmetry of an object (such as a die or coin or bag of candies).

**Third quartile** \((Q_3)\): the median of the data positioned strictly after the median when the data is placed in numerical order.

**Treatment group**: in a controlled study, the treatment group refers to the participants who are given the treatment that is being tested to see if it causes a certain effect.