1

A CASE STUDY OF STATISTICS IN ACTION

Overview

Goals

Although statistics is a mathematical science, it is not a branch of mathematics the way probability is. Statistical thinking differs from mathematical thinking in important ways. The goal of this chapter is to involve students from the start in two important differences present throughout the book: the role of context and the logic of statistical inference. We will begin with examples illustrating these differences in order to give students the longest possible time to work with them.

Role of Context

"Although mathematicians often rely on applied context both for motivation and as a source of problems for research, the ultimate focus in mathematical thinking is on abstract patterns: the context is part of the irrelevant detail that must be boiled off over the flame of abstraction in order to reveal the previously hidden crystal of pure structure. *In mathematics, context obscures structure.* Like mathematicians, data analysts also look for patterns, but ultimately, in data analysis, whether the patterns have meaning, and whether they have any value, depends on how the threads of those patterns interweave with the complementary threads of the story line. *In data analysis, context provides meaning.*" [George W. Cobb and David S. Moore, "Mathematics, Statistics and Teaching," *American Mathematical Monthly,* vol. 104, 1997, pp. 801–824.]

Some students may be tempted to dismiss the story line (and the questions related to it) as something to plow through in a hurry to get to "the real stuff," like formulas or computing algorithms. In statistics, unlike pure mathematics, it is important to encourage students to think carefully about the context for its own sake—not just to ask what a result means in the context of a particular application but also to let the context suggest questions to be answered, even if the method *du jour* does not and cannot provide an answer. To learn to think like a statistician, the student needs to be aware of the questions that a method cannot answer, as well as those that can be answered.

Logic of Statistical Inference

A second important difference between mathematics and statistics has to do with the logic of statistical inference, which many students of statistics (at all levels!) experience as the most difficult part of the subject. The difficulties arise for a number of legitimate reasons. One is the "what if" nature of the reasoning, which is typically expressed using a conditional subjunctive: If this particular hypothesis were in fact true, how likely would it be to get data like ours? You can simplify the grammar (we do that in the text), but there is a warp built into the logic that you can't avoid if you want to do it right. It is verboten to ask the question we really want answered: Given our data, how likely is it that this particular hypothesis is true? No wonder students find inference difficult!

In this first chapter, we try to steer a middle course between the rock of ignoring the essential parts of the logic and the hard place of too much detail. We want students to get an honest look at what's hard in order for students to appreciate what is unique about statistical logic. But we don't want to swamp them with premature rigor. (Rigor can demand a preoccupation with details that fog up an otherwise clear explanation.)

Because some of the concepts and logic of inference are so hard, it is particularly important to do Activity 1.1, which is designed to give students hands-on experience with a concrete and comparatively simple version of the thinking they will be expected to learn and use in later chapters of the book.

Time Required

Traditional Schedule			Block	4 x 4 Block
Section 1.1				
1–2 days	Day 1 Day 2	Overview, graphical displays Exploring data, summary, exercises	1 day	2 short days or 1 long day
Section 1.2				
2–3 days	Day 1 Day 2 Day 3	Activity 1.1, simulation Logic of statistical inference Summary and exercises	2 days	3 short days or 1 long and 1 short day
Review				
1–2 days			1 day	2 short days or 1 long day

Materials

Section 1.1: None

Section 1.2: D13 requires one coin per student. Activity 1.1 requires ten 3×5 cards (or small pieces of paper) per group.

Suggested Assignments

Classwork			
Section	**Essential**	**Recommended**	**Optional**
1.1	D2, D3, D6, D7, D8 P1, P2	D1, D4, D5	
1.2	Activity 1.1 D9, D10, D11 P3	D13 P4	D12

Homework			
Section	**Essential**	**Recommended**	**Optional**
1.1	E1, E2	E4, E5	E3
1.2	E6, E7	E9, E10	E8
Review	E11, E13	E14, E15	E12

1.1 Discrimination in the Workplace: Data Exploration

Objectives

- to explore a set of data related to an actual court case involving alleged age discrimination and practice relating patterns in data to possible meanings in the applied context
- to learn the basic structure of a data table, with cases as rows and variables as columns
- to learn by example what a distribution is and use dot plots to represent a distribution
- to organize categorical data in a two-way table

Important Terms and Concepts

- exploration versus inference
- cases and variables
- variability
- distribution
- dot plot

Lesson Planning

Class Time

One to two days

Materials

None

Suggested Assignments

Classwork		
Essential	**Recommended**	**Optional**
D2, D3, D6, D7, D8 P1, P2	D1, D4, D5	

Homework		
Essential	**Recommended**	**Optional**
E1, E2	E4, E5	E3

Lesson Notes: Exploring Data

Westvaco is pronounced with a long *a,* as in Waco, Texas.

This first lesson is a good place to explain the three types of questions used in the text: Discussion Questions, Practice Problems, and Exercises.

Throughout the text, students will be asked to justify their answers to a question. This includes stating assumptions, giving appropriate graphs and computations, and writing a conclusion in context.

Graphical displays are essential in the investigation of data because sometimes patterns in data are not at all obvious without a plot. You can model the importance of plotting data to find patterns by making graphical displays as often as possible.

In all problems, encourage students to relate their statistical work to its real-world context. In mathematics, what each example has in common with others in the same section is often emphasized. Although such patterns of shared structure are important, in statistics it is also essential to think about what makes one example different from another. For example, in the *Martin* case, it is reasonable to ask how other contextual matters, such as years of service at Westvaco, years of education and specialized training, etc. relate to the questions asked.

The answers to many discussion questions may vary depending on students' individual insights. Learning to think like a statistician becomes easier if students can learn to carry on a dialogue between their intuition and their formal learning. Intuition is a notoriously untrustworthy guide in probability problems, but students do better if they come to regard their intuition as a legitimate voice, albeit one that often needs to be educated by a second voice. The alternative—to leave the voice of intuition unattended—invites trouble.

Discussion

D1. Reasonable suggestions from students at this stage include

- compare the average age of those laid off and the average age of those retained;
- compare the proportion of those over, say, age 50 who were laid off with the proportion of those under age 50 who were laid off;
- make a graph that shows both the ages of those laid off and those not laid off. (If students suggest this option, you might take the opportunity to show them how to make a back-to-back stem-and-leaf plot.)

Students may do a lot or a little with this question. In some classes, students may only offer very general ideas. On the other hand, students who have learned to "take charge" in their mathematics classes may want to thoroughly explore the data tables, bar graphs, and averages before continuing with the discussion. At any rate, all students should see the need for some method of summarizing and displaying the data in Display 1.1.

D2. This dot plot shows a possible case of age discrimination. The pattern shows that those laid off were, on balance, older than those who kept their jobs. Of the six workers under age 50, half kept their jobs. Of the eight workers over age 50, only one kept his or her job. This pattern provides some support for Martin's claim. (Without such a pattern, there would be no statistical evidence of discrimination, although there might be other evidence.)

Here are two important points that might be raised, which you can either discuss now or wait until D7 and D8.

- Is there enough evidence? Some students may think there are too few employees involved to make the pattern worth taking seriously. This is a key issue, one that will be addressed in D7 in a preliminary way and more fully in Section 1.2.
- The pattern of the ages says nothing about the reason for the pattern, which may or may not be age bias. (For example, suppose that Westvaco had decided to retain employees with better computer skills and that younger employees tended to have more experience with computers.) This is the focus of D8.

D3. Display 1.4 provides stronger support for Martin's case than Display 1.3. The pattern in the dot plots in Display 1.4 shows that of those laid off, older workers were selected in earlier rounds, younger workers in later rounds. Martin's argument used the fact that older workers were laid off earlier as evidence of age bias: If Westvaco had been able to stop the layoffs after the first two rounds, for example, the difference in ages (between those fired and those kept) would have been even more extreme than it actually was. Westvaco argued that this pattern was irrelevant because, regardless of how the preliminary planning was done, the actual layoffs occurred all at once.

D4. Both groups, hourly and salaried, show similar patterns: older workers were more likely than younger ones to lose their jobs. From the dot plots alone, it is hard to decide which group, hourly or salaried, provides stronger evidence of age bias. The salaried workers are a larger group, although the difference in ages between laid off and retained may be more pronounced for the hourly workers. On balance, though, variability makes it hard to judge the strength of the evidence from the dot plots. Useful as graphs are, they have limitations, and there are times when numerical summaries are helpful. Most of the time, a good analysis involves both graphical displays and appropriate numerical summaries.

D5. Overall, the pattern in Display 1.6 is strikingly similar to its counterpart in Display 1.4. The graphs show that older workers were especially likely to be targeted for layoff in the earlier rounds and that younger workers were more likely to be targeted in Rounds 4 and 5. Both displays also show that most of the layoffs were planned in the first two rounds.

D6. **a.** Yes. Older workers were slightly more likely than younger workers (52% vs. 44%) to lose their jobs. This pattern is consistent with age bias, but the two issues raised in the answer to D2 (and the focus of D7 and D8) are relevant here.

b.

		Laid Off	Retained	Total	% Laid Off
Under 50?	**Yes**	6	10	**16**	**37.5%**
	No	12	8	**20**	**60.0%**
	Total	**18**	**18**	**36**	**50.0%**

With "older" redefined as "50 or older," the evidence here is stronger than in part a. Here older workers were much more likely than younger ones to lose their jobs (60% vs. 38%).

c. The pattern in the two tables taken together shows that workers 40 or older were more likely than those under 40 to lose their jobs, and those 50 or older were substantially more likely than those under 50 to lose their jobs. Thus, the two tables taken together provide stronger evidence in favor of Martin.

Note that the fact that age 40 and over is a protected class does not mean that we must divide the employees using that age only. The work force in Westvaco's Engineering Department was older than many groups of employees, as you can see from the various dot plots. Only 25% were under 40; 44% were under 50. Using 50 as the cutoff age to define "older" gives age groups roughly the same size and is more informative (other things being equal). The message of possible age discrimination may be stronger for the table split at age 50 because Westvaco may not be discriminating against workers aged 40 through 50, but only against workers older than that.

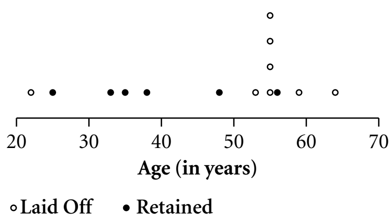
D7. Statistical inference can provide an answer to this question, and Section 1.2 will sketch how this is done. For now, the purpose of the question is to involve students in thinking about what the question asks and why it is important.

Students may be worried that there are too few workers to make conclusions or that the differences are too small to make any conclusions. Try to get them to focus on the crucial question: Do the patterns look like the kind you would expect to occur just by chance? No calculations are to be done; have students only observe the pattern.

D8. Considerations that might explain how older workers might have been laid off disproportionately without their being victims of age discrimination include the following: Older workers tended to hold jobs that had become obsolete. Older workers tended not to have the up-to-date training needed to use Westvaco's computers or other technology.

Practice

P1. This plot is rather striking; the older hourly workers were far more likely to be laid off in Rounds 1 through 3 than were the younger workers.



Students can construct this plot quickly and easily if they rely on the plots in Display 1.4.

This problem offers another opportunity to show students how to make a back-to-back stem-and-leaf plot.

P2. **a.** The two tables appear here. The second table, which uses 50 to define "older," shows stronger evidence of possible age bias.

		Laid Off	Retained	Total	% Laid Off
Under 40?	**Yes**	3	2	**5**	**60.0%**
	No	7	2	**9**	**77.8%**
	Total	**10**	**4**	**14**	**71.4%**

		Laid Off	Retained	Total	% Laid Off
Under 50?	**Yes**	3	3	**6**	**50.0%**
	No	7	1	**8**	**87.5%**
	Total	**10**	**4**	**14**	**71.4%**

b. The two patterns in the tables for hourly workers in part a are similar to those in the tables for the salaried workers. One pattern shows that older workers were more likely than younger ones to lose their jobs. The other pattern shows that the difference in the percentage laid off for the two age groups is much more pronounced when 50 rather than 40 is used to define "older."

Note that age discrimination could still be at work even if the *number* of younger workers laid off was larger than the number of older workers laid off. The next table illustrates this possibility for a hypothetical company and the limitation of using counts as opposed to proportions.

		Laid Off	Retained	Total	% Laid Off
Under 40?	**Yes**	40	20	**60**	**66.7%**
	No	30	10	**40**	**75.0%**
	Total	**70**	**30**	**100**	**70.0%**

Here 30 older workers were laid off compared with 40 younger workers, yet a higher percentage of older workers (75%) were laid off than of younger workers (66.7%).

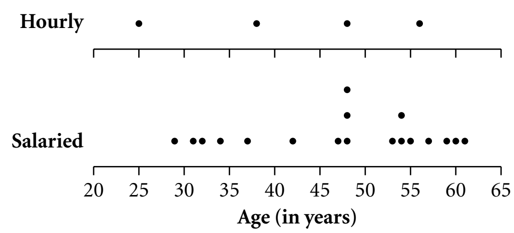
Exercises

E1.

		Laid Off	Retained	Total	% Laid Off
Pay Status	**Hourly**	10	4	**14**	**71.4%**
	Salaried	18	18	**36**	**50.0%**
	Total	**28**	**22**	**50**	**56.0%**

As the table makes clear, hourly workers were more likely than salaried workers to lose their jobs.

E2. **a.**



Only four hourly workers kept their jobs, making comparison difficult. However, the distribution of salaried workers who kept their jobs generally falls to the right of the distribution of hourly workers who kept their jobs. Indeed, the average age of the hourly workers laid off was 41.75, whereas it was 50.556 for the salaried workers.

b. We cannot conclude this from the dot plots in part a alone. We would have to take into consideration the age distributions for hourly and salaried workers *before* the layoffs began, as well as after. All we can say is that the salaried workers who kept their jobs tended to be older than the hourly workers who kept their jobs.

E3. **a.**

Round	Laid Off	40 or Older	Percentage
1	11	9	82%
2	9	8	89%
3	3	2	67%
4	4	2	50%
5	1	0	0%

b. Most of the layoffs came early. Of the 28 who were chosen for layoff, 20 were identified in the first two rounds. Only 8 were chosen for layoff in the last three rounds altogether.

Early rounds hit the older workers harder; later rounds tended to have higher percentages of younger workers. In the first two rounds, the percentage was very high among those over 40 who were laid off, at 85% (17 of 20). In the later rounds, the percentage for those over 40 dropped to 67% in Round 3 (2 of 3), 50% in Round 4 (2 of 4), and 0% (0 of 1) in Round 5.

The pattern is consistent with what you would expect to see if the department head who planned the layoffs was trying to cut costs by laying off the older, more experienced, and thus possibly more expensive workers first.

E4. The work on this question can be as elaborate and involved as you have time for. One reasonable way to approach such an analysis is to follow the analysis of age, using year of hire as the variable. (This is equivalent to working with a new variable, *seniority,* defined as 1991 − *hire year.*) This dot plot compares the years of hire for those laid off and those retained.

One striking feature is that all 12 of those hired in 1962 or earlier lost their jobs. Of those hired in 1985 or later, 8 of 13, or 62%, lost their jobs. Of those hired in the middle years, 1963–1984, only 8 of 25, or 32%, lost their jobs. The information seems to suggest that "last hired, first fired" was not the policy, except for the most recently hired (1985–1990).

Year of Hire	Laid Off	Retained	Total	% Laid Off
62 or before	12	0	12	100.0%
63–74	5	13	18	27.8%
75–83	3	4	7	42.9%
85–90	8	5	13	61.5%
Total	**28**	**22**	**50**	**56.0%**

E5. • ***Major league baseball standings.*** The cases are the teams. The variables include games won, games lost, winning percentage, and games behind the leader.

• ***New York Stock Exchange.*** The cases are the companies whose stock is listed on the NYSE. Variables include the opening price for the day, the low, the high, the close, and the net change from the previous day's close.

• ***Nutritional summary.*** The cases are the nutritional components, such as fat, carbohydrates, protein, cholesterol, and various vitamins and minerals. Variables might include amount per serving and percentage of the recommended daily allowance.

1.2 Discrimination in the Workplace: Inference

Objectives

- to learn how to use simulation to generate a sampling distribution and estimate a probability
- to use the concept of a sampling distribution and the logic of significance testing to make decisions about data

This section gives an "owner's manual" description of the "vehicle" that we call statistical inference. Students will learn how to drive this vehicle but will not fully understand how it works or know the technical names of all the parts. Chapters 8 through 12 will be more like a "service manual" description. Activity 1.1 is an excellent way to introduce statistical inference in a hands-on experience that is also fun.

Important Terms and Concepts

- summary statistic
- simulation
- chance model
- statistical inference

Lesson Planning

Class Time

Two to three days

Materials

D13 requires one coin per student. For Activity 1.1, each group of students will need ten 3 × 5 cards (or small pieces of paper).

Suggested Assignments

Classwork		
Essential	**Recommended**	**Optional**
Activity 1.1 D9, D10, D11 P3	D13 P4	D12

Homework		
Essential	**Recommended**	**Optional**
E6, E7	E9, E10	E8

Lesson Notes: Inference

Why the Hourly Workers?

Why does the text concentrate on the hourly workers when Bob Martin was salaried? First, Martin's case that Westvaco discriminated against older workers is strengthened by the analyses for hourly workers. Second, there were only 14 hourly workers and 36 salaried workers, so using the hourly workers makes the simulations a reasonable size.

Why It Is Important to Do Activity 1.1

Turning probability questions into simulations has a couple of advantages. Students can use probabilities without getting bogged down in the (often delightful but centrifugal) intricacies of particular computations, and it also reinforces the definition of probability as the proportion of "successes" in the long run. This activity gives students a sense of what the distribution should look like if layoffs were determined at random.

A Technical Summary

Because our introductory example deliberately avoids technical terms, we provide the following short summary that links the formal language of statistics to the example.

The *Martin* case example is typical of many examples of *statistical tests of hypotheses.* In this kind of test, we want to test whether an observed difference could reasonably be due just to chance. In the *Martin* case, the *null hypothesis* is that workers were chosen at random, without regard to age, for layoff. So in Activity 1.1, we set up a model that chooses three workers at random, with every worker given the same chance of being chosen. To test this hypothesis, we choose a *test statistic,* a summary number based on the data, such as the average age of those laid off. An extreme value of the test statistic (large average age) is evidence against the null hypothesis. We use *simulation* (or sometimes mathematical theory) to find the chance of getting a value for the test statistic as extreme as or more extreme than the one we actually observed if the null hypothesis is true. (This number is called the *P-value* or *significance level.*) If values as extreme as the test statistic are reasonably likely, we conclude that the observed effect could be due just to chance. If, on the other hand, a value that extreme is very unlikely, we *reject the null hypothesis* and conclude that the observed effect is "real."

As students use statistical inference to make decisions, they can refine their mathematical thinking in a critical way. Most high school students develop their concepts of proof from the structured arena of Euclidean geometry. Decision making that involves considering uncertainty and variability is a very new and challenging task for students whose prior experience was limited to right answers and two-column proofs. In the case of inference, we build an argument based on the likelihood of a sample yielding results that are very different from the values expected under the null hypothesis. Such simulations do provide evidence of the results but do not prove or disprove the conjecture in the "pure" sense of proof.

E10 introduces systematic counting as an alternative to simulations for finding probabilities.

Discussion

D9. No, it is not likely. To get an average age of 58 or greater, you couldn't select any worker under age 55—and half the workers are under age 55.

D10. It favors Martin if it is unlikely to lay off, just by chance, workers as old as those laid off by Westvaco.

Activity 1.1: By Chance or by Design?

The activity should only take about 15 minutes and can be shortened or lengthened depending on the number of samples you have each group draw before you make a dot plot of the results.

1. Ideally, students should work in groups of two. Each group of students will need ten 3×5 cards or ten small identical pieces of paper. They can cut these themselves, or you might use a paper cutter ahead of time and provide each group with a handful of congruent rectangles. Students can draw either from a box or from a paper bag, but be sure to encourage them to mix the pieces thoroughly each time before they draw.

2–4. Sample results appear in the dot plot in Display 1.8 on page 13 of the student text. Using a dot plot when collecting the results from your class works well, giving students a sense of the distribution and where the average age of 58 would fall in that distribution.

If your class has had experience with stem-and-leaf plots, you may want to collect the data that way.

```
Average age

3 |
4 |
5 | 2 3
6 |

5 | 2 3 represents an average of 52.3 years.
```

5–6. To estimate the probability, divide the number of times the average age satisfies the condition (of being as extreme as or more extreme than 58) by the total number of samples collected in your class. The true probability of getting an average age of 58 or more is $6/120 = .05$ (Your class's estimate won't necessarily be close to .05, though). With $n = 50$ (5 groups of students doing 10 replications each), the margin of error is about 0.06. With $n = 200$ (20 groups doing 10 replications each), the margin of error is half that big. If your class's estimate is a lot larger than .05, talk informally about the relationship between the sample size and the precision of the estimate before going on to obtain a better estimate by increasing the number of repetitions.

Students will do this problem theoretically in E14, but it is still important that they learn to design and carry out a simulation and to understand a sampling distribution like that in Display 1.8.

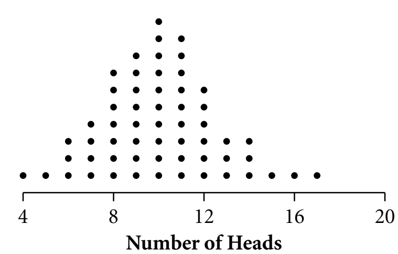
Discussion

D11. The probability of getting any specified average age is quite small. For example, if Westvaco laid off the 35-, 38-, and 48-year-olds, the average age would be 40.3 and there certainly would have been no suspicion of age discrimination. However, if you look at Display 1.8, you see that the probability of getting an average age of exactly 40.3 is actually smaller than getting an average age of 58. Westvaco shouldn't be under suspicion because the average age of its laid-off workers was exactly 58, and getting exactly 58 just by chance is very unlikely; Westvaco should be under suspicion because the average age it got is in the extreme upper part of the distribution of all possibilities. That is, *it is the location of age 58 in the distribution that is important, not the probability of getting that exact number.*

D12. **a.** No, the probability is larger than .025.
b. If the *Martin* probability had been .01, it would have met this requirement. If it had been .10, it would not have.

Note about significance levels: Using any one choice of threshold for deciding statistical significance is somewhat arbitrary, even though fixed-level testing using a significance level of .05 is so common. There is no best value (except in a strict decision-theoretic setting where you know the costs of Type I and Type II errors), but there is an unambiguous ordering: the smaller the probability, the stronger the evidence against the null hypothesis. Though students may legitimately disagree about how low a value they personally would require in order to rule out the chance model, they should all come to recognize that a lower significance level means stronger evidence is required.

D13. **a.** If the coin were fair, it would be very unlikely to get 19 heads or more in 20 tosses. (The probability is $\approx .00002$.)
b. The chance model is that the coin is fair. Students should flip a fair (or assumed to be fair) coin 20 times and count the number of heads. The number of heads is the summary statistic. This process should be repeated as many times as you have time for in your class. You may display the distribution on a dot plot. A typical distribution for 500 repetitions is given here.



NOTE: When making this dot plot, Minitab prints "Each dot represents 10 points." However, there are 58 dots here for 500 sets of flips, so some dots represent fewer than 10 flips. Whenever the number of flips above a particular tick mark isn't a multiple of 10, Minitab uses a dot for the excess. For example, suppose the value 12 occurred 53 times. Then Minitab would use 7 dots above the 12. ■

In this simulation of 500 repetitions, the largest number of heads ever to appear was 17. The probability of getting 19 or more heads is very close to 0. If the coin is fair, it is extremely unlikely to get 19 or 20 heads, so we conclude that the coin is not fair.

Summarizing, here's the logic: Assume the model is correct that your friend's coin is fair. Question: Is the actual data (19 heads in 20 tosses) easy to get from this model? Answer: No, it's almost impossible to get 19 heads in 20 tosses of a fair coin. Conclusion: Your friend's coin is not fair.

Emphasize to your students that you can estimate the chance of getting 19 or more heads from a fair coin, but you cannot estimate the chance that your friend's coin is fair. If you assume the model is correct (the coin is fair), you can estimate the probability you need by simulation. Notice, however, that it doesn't work to assume that outcome (19 heads) and then use simulation to estimate the chance of a fair coin.

Practice

P3. **a.** The average age is now 52.67. On the dot plot of Display 1.8, this is the column that is just a bit more than half-way between 50 and 55. That is, it is the tallest column in the group of seven columns of dots *between* 50 and 55. The number of repetitions out of 1000 that gave an average age of 52.67 or larger is 5(31) = 155. Thus, the estimated probability of getting an average age of 52.67 or larger is .155.

b. There is no evidence of age discrimination because an average age of 52.67 or larger is relatively easy to get just by chance.

P4. **a.** The average of 33 and 35 is 34.

b. The chance model is that two workers were selected at random for layoff in Round 3 from six workers with ages 25, 33, 35, 38, 48, and 56. Place those ages on six cards and draw two at random. Repeat many times, computing the average age each time. Make a dot plot showing the distribution of these average ages and compute the proportion of times you get an average age of 34 or larger.

Simulations will vary. For example, a student might first get the ages 25 and 38, for an average of 31.5.

c. Answers will vary according to the simulation, but the chance is quite large, well above 50%. In fact, it is equal to 12/15, or 80%.

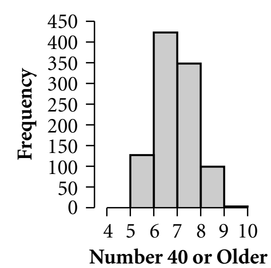
No. If you look just at this one round, the average age of people laid off is about what you would expect by chance alone.

Exercises

E6. **a.** Simulations will vary. For example, a student might draw the following 10 ages: 22, 25, 33, 48, 53, 55, 55, 55, 56, and 59. The summary statistic would then be 7 because 7 of the 10 are age 40 or older.

Note that instead of writing the ages on slips of paper, for the purposes of this problem students could just write "40 or over" or "under 40" for each person. Also, note that you can save time by utilizing the complementary event—drawing 10 workers out of 14 who are laid off is equivalent to drawing only 4 out of 14 who are retained. Students can then analyze the remaining 10 slips of paper. Some students may not understand that this is legitimate.

Simulations will vary, but the number of workers in the protected class should center around 6.4, which is the average number of workers aged 40 or more that you would get when drawing 10 workers to lay off from the 14 hourly workers.



b. Student estimates will vary depending on their simulations. The theoretical probability of drawing 10 ages from the 14 and getting 7 or more who are aged 40 or older is $455/1001 \approx .45$.

c. It is quite likely to randomly draw 10 ages from the 14 and get 7 or more who are aged 40 or older. Thus, we reasonably can attribute the Westvaco result to chance. This test provides no evidence for Martin.

d. A general rule in statistics is that you should not throw away information that might be relevant. Only using whether the person is 40 or older throws away information—the person's exact age. For example, suppose Westvaco had laid off everyone over 60 and kept everyone under 60. That would be age discrimination but might not show up in an under/over age 40 analysis because all those aged 40–59 were kept.

Here is a specific example of that principle: In the *Martin* case, there were 10 hourly workers involved in the second round of layoffs; 4 of these workers were under 40. Three were chosen to be laid off; all 3 were 40 or older. To see whether this result could reasonably be due just to chance, consider drawing 3 tickets at random from a box of 10 tickets, 4 of which say "Under 40" and 6 of which say "40 or older." A natural summary is the number of tickets out of 3 that say "40 or older." Simulating this process a large number of times shows that the chance that all 3 tickets say "40 or older" is about 1/6. (In fact, it is possible to count the possible outcomes. There are

$$\binom{10}{3} = 120$$

possible outcomes, of which

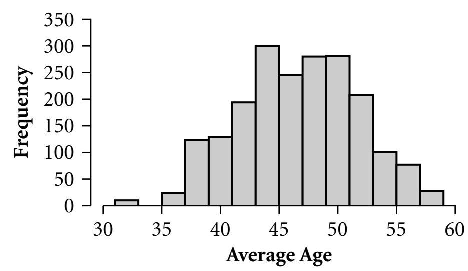
$$\binom{6}{3} = 20$$

have all 3 tickets saying "40 or older," so 1/6 is the exact probability.)

This probability is large enough that we cannot reasonably rule out the possibility that such data could arise by chance alone. Thus, this conclusion differs from the conclusion on the actual ages, which students analyzed in Activity 1.1. Because the analysis based on the actual ages uses more of the available information, it is more informative and trustworthy. The analysis based on the actual ages (the mean) uses more of the available information. On the other hand, it might be easier to convince a judge or jury by using 40 or over because that is the federal protected class.

E7. **a.** A calculator can be used to do the simulation where 4 ages are drawn out of the 10 hourly workers. The ages of the workers can be numbered 1 through 10, and random drawings of four integers between 1 and 10 (that will correspond to the ages) generated using the command **RandInt(1,10,4)**. Students can then find the average age of the 4 workers whose ages correspond to the four random integers. As an alternative, write the ages of the 10 hourly workers on identical cards, mix the cards up, and draw four at random. Then find the average age of those 4 workers. This chance model reflects a situation where chance alone decides who is laid off. Repeat this process many times, displaying the average ages on a dot plot. Decide whether it is reasonably likely to get an average age of 57.25 or greater.

b. Simulations will vary. For example, a student might draw the four ages 33, 35, 56, and 64. The average age would then be 47. The distributions should center around 46.5.



Conclusions may vary with the simulation, but students should realize that it is unlikely to get an average age as large as 57.25 just by chance. The theoretical probability of getting an average age of 57.25 or more is only 4/210 ≈ .019. That is, getting an average age this high would happen less than 2 times out of every 100, if selections were done randomly. This is strong evidence in Martin's favor.

E8. **a.** The first two: sum of the ages, difference of the average ages. These are equivalent to the mean because you will get the same probability of a value as extreme as or even more extreme than the sum of the ages and the difference of the average ages as you do for the mean. To see this, if you have the sum of the ages, divide by 3 to get the average age. If you have the difference of the averages, *D*, it is an algebraic challenge to get the average age of those laid off: multiply *D* by the number of workers retained, add to this the sum of all the ages, and then divide by the total number of workers.

b. It would be reasonable to use any of the possible choices to get an indication of how the ages of those laid off compare with those that would occur by chance alone.

However, reducing a set of numbers to a summary statistic may involve loss of information, and typically there will be several choices, sometimes with no obvious best choice. ("Which is best" turns out to be a surprisingly complicated question. The answer is based on which summary statistic is more likely to "catch" a company that does, indeed, discriminate on the basis of age. Using the average age (or an equivalent summary statistic) tends to make better use of the information in the sample than the others.)

E9. "If it is July 4, it is very unlikely to be snowing in Kansas. Therefore, this probably isn't July 4." The hypothesis is that it is July 4, so we begin by assuming that. We then examine the data (it's snowing in Kansas), and because it's unlikely to be snowing on July 4 in Kansas, we reject the hypothesis that it is July 4.

E10. **a.** There are $\binom{10}{2} = 45$ possible pairs. The partial listing given in the problem suggests computing this as $9 + 8 + \ldots + 3 + 2 + 1 = 45$.

b. Only four pairs give an average age of 59.5 or older:

25 33 35 38 48 **55** 55 55 56 **64**
25 33 35 38 48 55 **55** 55 56 **64**
25 33 35 38 48 55 55 **55** 56 **64**
25 33 35 38 48 55 55 55 **56** **64**

c. Thus, the probability of getting an average of 59.5 or more is $4/45 \approx .09$.

d. The evidence of age bias is somewhat weak—weaker than in the *Martin* example, for which the probability was .05.

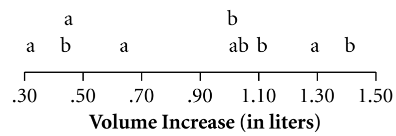
Review

Homework	
Essential	E11, E13
Recommended	E14, E15
Optional	E12

Review Exercises

E11. **a.** The data for Compound B include more of the larger values, but the two data sets are pretty well mixed together. Students will argue both ways; some will say Compound B is better because it has more of the higher values, and some will say that there is no obvious winner.

b.

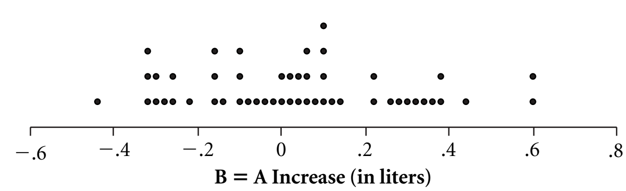


The points for Compound B tend to fall more on the right side of the plot, so it appears that Compound B gives larger measurements than Compound A.

c. Compound A mean = 0.746; Compound B mean = 1.002, for a difference of 0.256. Because the Compound B mean exceeds the Compound A mean, some students will think of this as sufficient evidence to claim that Compound B is better at opening the lungs than Compound A. They are forgetting about variability, indicated by the overlap in the distributions!

E12. To explore the effects of poverty, the variables you might look at include literacy rate, fertility rate, rate of malnutrition, percentage of dwellings with indoor plumbing, number of people per room, percentage of children who have been vaccinated for various diseases, and infant mortality.

E13. **a–b.** Here are the results of 50 trails of a simulation to find the difference between the means.



c. The observed difference of 0.256 was exceeded 11 times out of the 50. This is a high fraction and points to a conclusion that the 0.256 value could well be the result of chance alone.

E14. **a.** There are six sets of three that give an average of 58 or more:

25 33 35 38 48 **55** **55** 55 56 **64**

25 33 35 38 48 **55** 55 **55** 56 **64**

25 33 35 38 48 **55** 55 55 **56** **64**

25 33 35 38 48 55 **55** **55** 56 **64**

25 33 35 38 48 55 **55** 55 **56** **64**

25 33 35 38 48 55 55 **55** **56** **64**

b. The probability of getting an average age of 58 or more by choosing three ages at random is 6/120 = .05.

c. Answers will vary, depending on the results of Activity 1.1. The probabilities differ because their result from Activity 1.1 was from a simulation and so is an estimate of the exact answer computed in part b. Unless the number of repetitions is quite large, there can be a considerable difference.

E15. There are still 120 possible subsets of three, but there are many more ways to get an average as large as the observed average of 55, so the probability is substantially higher than in the actual case, and therefore the evidence of age bias is much weaker. (There are in fact 10 possible ways to get an average age of 55, so the exact probability is 10/120≈ .83.)