The Manager's Guide to Statistics

Erol A. Peköz

2009 Edition

Printed in the United States of America.

ISBN: 0-9795704-1-7

To order give your college bookstore the information below:

The Manager's Guide to Statistics
by Erol A. Peköz (2009)
Publisher: www.ProbabilityBookstore.com, Boston, MA
ISBN: 0-9795704-1-7

Data sets and the Excel add-in software can be found at the following website: `smgpublish.bu.edu/pekoz`

Contents

Preface

This book is an introduction to statistics for someone who does not need to know all the details of statistical theory and would just like to know how statistics is commonly used in business. We cover some of the basic statistical tools, some pitfalls to watch out for when using them, and we give some intuitive explanations of how everything works. We try to explain everything in words rather than with Greek symbols or mathematical formulas. You'll also find the technical details and notation in technical notes so as a student you could be comfortable using other statistics books in the future.

It's difficult for students of statistics to learn, at the same, the imposing mathematical notation and the subtle, fragile concepts behind the scenes. If a teacher tries to teach both at the same time, invariably attention paid by students to the concepts goes out the window. The student leaves the course with only a vague recollection of a jumble of Greek letters and of some frustrating time spent on the computer pointing and clicking and coping with error messages. This is why we make such an effort to teach the concepts first without the imposing notation and then include much of the notation and technical details as technical notes afterwards.

We recommend using Excel, preferably with the add-in included with this book, for the statistical calculations. Excel is ubiquitous in business and becoming more comfortable with it is of great value in itself. Though Excel is not the program of choice for professional statisticians, it is the program of choice for business people.

This book is designed to be the textbook for a one semester introductory course for undergraduate business students, MBA students, or other aspiring leaders and decision makers.

About the Author

Erol A. Peköz is Associate Professor in the School of Management at Boston University. He received his B.S. from Cornell University and Ph.D. in Operations Research from the University of California, Berkeley in 1995 and has published more than 25 technical articles in applied probability and statistics. He also has taught at the University of California, Berkeley, the University of California, Los Angeles, and Harvard University. At Boston University he was awarded the 2001 Broderick Prize for Teaching, and is also a co-author of the textbook *A Second Course in Probability* (with Sheldon Ross).

Acknowledgements

I would like to thank Craig Bleyer, Paul Berger, Don DeLand, Andrew Gelman, Janelle Heineke, Bruce Kaplan, Michael Shwartz, Terri Ward, and Mustafa Yilmaz for their valuable contributions.

Chapter 1

Seeing the Real Story

Statistics are tools for seeing and telling a story from data. Computers can usually handle the number-crunching calculations we may need, but interpreting the output and seeing a story is up to us. Instead of starting off by teaching you how to do calculations, we will start with some advice on how to see the real story behind the numbers.

One theme throughout this book is that it is easy to be misled by statistics if you don't know what to watch out for. Even honestly gathered data may appear to be telling one story on the surface, when actually quite the opposite is the real story. But after this chapter you shouldn't just become skeptical of all of statistics. Being overly skeptical is just as bad as being overly gullible: both keep you from the real truth. Our goal for you is to become wise enough to know when to be skeptical and when to believe. In this chapter we will cover a few commonly encountered ways statistics tend to mislead people, so you will know what to watch out for.

1. Comparing Rates: Missing Denominators

Sometimes the ratio of two numbers tells more than either of the numbers do by themselves. Next are a few examples.

How many hospital beds does a city need? If your city has ten times more hospital beds than another city, does this mean your city probably has more beds than it needs? To start to answer this, we really need to know the number of people in

each city. We should not just compare the **count** of beds, but should instead compare the ratio of the number of beds to the number of people in the city. We call this type of ratio a **rate**. For example, if a city of 200,000 people has 100 hospital beds, this corresponds to a rate of one hospital bed for every 2,000 people. Though assessing if a city has more beds than it needs is controversial and complex,[1] we should definitely start off by looking at rates instead of counts.

> Watch out if someone is comparing counts when they should instead be comparing rates.

Are you safer without a seat belt? During the year 2002 in California, 1,524 people wearing seat belts were killed in car accidents but only 1,343 people without seat belts were killed in car accidents.[2] Do these numbers mean you are safer without a seat belt?

Solution. No, you are not safer without a seat belt. Since most people wear seat belts, we would expect a large number of deaths among people wearing seat belts. To see the benefits of a seat belt you should compare the death rates for the people with and without seat belts. You could do this by dividing the number of deaths for each group by the total number of people (or the total number of accidents) for each group. This rate would turn out to be much higher for the group of people who weren't wearing seat belts. It's also true that many more people are killed riding in cars than riding motorcycles: only 318 motorcyclists were killed in 2002 in California. Even though motorcycles are more dangerous than cars (and the death rate turns out to be higher), there are fewer deaths because there are fewer motorcyclists.

Is New York City more dangerous than Iraq? The death rate for Americans (that is, deaths per thousand Americans per year) in Iraq during the war was actually lower than the death rate in New York City during the same time period.[3] Does this mean it was safer to be sent to Iraq than to New York City?

[1] There is research that claims that the over-supply of hospital beds induces demand for them. A study by E. Peköz and M. Shwartz funded by the *Department of Health and Human Services, Agency for Healthcare Research and Quality* titled "Do More Hospital Beds in an Area Induce Excess Demand?" is investigating the evidence for this.

[2] See *2002 Annual Report of Fatal and Injury Motor Vehicle Traffic Collisions* at `http://www.chp.ca.gov/pdf/2002-sec4.pdf`, page 21.

[3] The New York City mortality rate was around 700 per 100,000 people per year as measured in the year 2000 census. The article "Counting the Dead" by James Dunnigan posted on July 29,

Solution. No, Iraq was much more dangerous. It's true there were fewer American deaths per year in Iraq than in New York City during that time, and this is because there were more Americans in New York City than in Iraq. But that still doesn't explain why there were fewer deaths per thousand Americans in Iraq. The reason for this difference in rates is because Americans in Iraq had a different age range than people in New York City. In Iraq the Americans were primarily young healthy soldiers, while the New York City death rate included the sick and the elderly, groups that typically have high death rates. If you compared Americans in Iraq with people of the same age range in New York City you would find a much higher death rate in Iraq. This type of age difference is called a *confounding factor*; we will talk more about these types of factors in the next section.

Exercises

1. Managers of a manufacturing plant keep track of on-the-job accidents. Last year there were 50 worker injuries that happened on the day shift and only five worker injuries that happened on the night shift. Does this mean the night shift was safer? Or are we comparing the wrong type of numbers?

2. Insurance industry records show the Cadillac Escalade SUV has the highest theft rate of all cars in terms of the number of thefts per thousand vehicles. But in the same records if you just look at the total number of thefts, more Toyota Camrys are stolen each year than any other car—including Escalades.[4] The 1989 Camry, in particular, is the one most often stolen. (a) How can you explain the difference here? (b) If you are considering buying either a Camry or an Escalade, which car would be more likely to get stolen?

3. A newspaper article reports about a dangerous surge in pharmacy prescription errors.[6] The article details how the vast majority of complaints to the Massachusetts Department of Public Health have been lodged against CVS Corporation's pharmacies. The article goes on to say this is particularly troublesome

2004 on strategypage.com reports the figure as 360 per 100,000 troops per year in Iraq. The yearly death rate in New York City for men 20-24 was 120 per 100,000 in 1999-2001 (it was only 40 per 100,000 for women, and for men over 85 it was 15,000 per 100,000). See `http://strategypage.com/dls/articles/200472922.asp`, and `www.nyc.gov/html/doh/pdf/vs/2002sum.pdf`.

[4] `http://www.auto-theft.info/Statistics.htm` and `http://money.cnn.com/2004/02/27/pf/autos/nicb_most_stolen/`

[5] Photos from `http://www.geartekcorporation.com/dailyphoto/2005/toyotacamry.html`, and `http://www.cadillacforums.com/cadillac-models/cadillac-escalade.html`

[6] "Massachusetts pharmacist woes a prescription for peril," by Jessica Heslam, *The Boston Herald,* Thursday, July 14, 2005, page 2.

Chapter 7

Introduction to Multiple Regression

1. Introduction

La Quinta Motor Inns is a mid-sized hotel chain headquartered in San Antonio, Texas. As a growing chain, hotel executives spent much time considering where they should open new hotels. One year the chain had consultants develop an equation for forecasting future profitability for potential new hotel sites under consideration. This approach was so successful that the company president said that he did not feel obliged to personally select the new hotel sites anymore.[1] The profit forecasting equation the consultants came up with used only the state population for the potential site, the planned price of rooms at the hotel, the median income in the area, and the number of college students within four miles of the hotel. That's it.

How did the consultants use such limited data to get an accurate profit forecast? To combine several variables together to forecast another variable, they used the statistical tool called "multiple regression." Simple regression uses a single X variable to forecast a Y variable; multiple regression, in contrast, uses multiple X variables to forecast a Y variable. For forecasting hotel profitability, the Y variable was profitability and the X variables were things such income and population in the area, room price, and so forth.

[1] Kimes, S. and J. Fitzsimmons, "Selecting profitable hotel sites at La Quinta Motor Inns," *Interfaces* 20, no. 2 (1990): 12–20.

Our discussion of multiple regression will take place over two chapters. In this chapter, we will discuss where you can use a multiple regression equation and then how to interpret it. We will consider in broad terms what makes a good equation. Once we have this grounding, in the final chapter we will turn to how one can choose the best predictor variables.

2. The Multiple Regression Equation

As we saw in an earlier chapter, simple regression equations are of the form

$$Y = b + mX$$

where Y represents what you are trying the forecast, b is the Y-intercept, and X is the variable used to make your forecast.

Multiple regression forecasting equations, in contrast, look something like

$$Y = b + m_1X_1 + m_2X_2 + m_3X_3 + \cdots$$

where Y represents what you are trying to forecast and X_1, X_2, and X_3 are the different X variables you are using to make your forecast. Here b is still the Y-intercept. The big difference, of course, is that multiple regression equations have more than one X variable in them.

The coefficients m_1, m_2, and m_3 can tell you the marginal relationship between each of the X-variables and the Y variable, and these can reveal managerially relevant facts about your data. They tell you the difference in the Y variable you expect to see associated with a unit difference in a given X variable, when all other X variables stay the same.

> The coefficients in a multiple regression equation tell you the difference in the Y variable you expect to see associated with a one unit difference in a given X variable, when all other X variables stay the same.

It is important to realize that these coefficients measure the association between variables and can be very misleading about the direction of causation. For example,

Table 7.1. An excerpt from the file `brooklinehomes.xls`.

Parcel-Id	Value	Sqft	Age	Bedrooms
046199 001000000	1255300	3900	106	10
046067 000300000	1384400	3229	100	7
046413 001400000	623100	2280	54	4
046085 002800000	1302200	2778	83	4
046224 001100000	1537200	3704	106	6
046395 000200000	1332400	2374	66	3
046111 000400000	1993900	4374	86	7
046433 001900000	1618700	3945	81	6
046029 000900000	333200	2760	96	4
046278 000800000	1626000	2482	46	4

if consultants had used the number of housekeeping employees as one of the X-variables in the hotel profit forecasting equation we mentioned above, they probably would have seen a positive coefficient for this variable. This would tell us that hotels with a larger housekeeping staff tend to have higher profits—but this does not tell us that profit would rise if we hired more housekeepers. The causation probably goes in the other direction: profitable hotels hire more housekeepers to keep up with high demand. The bottom line is that to determine the direction of causation you need to rely on your business common sense—not just the regression coefficients the computer gives you.

Let's get into a specific example.

Home values When we study the value of homes in the town of Brookline, Massachusetts using the computer, we see the data fit the following multiple regression equation:

$$Y = 230{,}040 + 478X_1 - 6{,}400X_2,$$

where Y is the value of the home in dollars, X_1 is the square footage of the home and X_2 is the age of the home in years. You can find the data in the file `brooklinehomes.xls` (an excerpt is in Table 7.1).[2] Programs such as Excel can take data and generate a multiple regression equation and we will show you later how to do this. Before we try to interpret the coefficients, let's first use this equation to make a basic forecast.

[2]Data file from `http://www.town.brookline.ma.us/Assessors/`

Question. Estimate the average value of homes that are 50 years old and have 3,000 square feet.

Answer. We plug in 3,000 for X_1 and 50 for X_2 in the equation. This gives us

$$\begin{aligned} Y &= 230{,}040 + 478(3{,}000) - 6{,}400(50) \\ &= \$1{,}344{,}040. \end{aligned}$$

This tells us that the average value of these types of homes is about \$1.3 million. That makes it a pretty expensive town to live in.

> To get forecasts from a multiple regression equation, just plug in X values and what you get out is the forecast.

We mentioned above how the coefficients can give important managerial insights. In this case, we can interpret the coefficient 478 as follows: for houses of the same age, with each extra square foot we see an additional \$478 in value on average. We can interpret the coefficient $-6{,}400$ as follows: in houses having the same square footage, we see a decrease in value of \$6,400 for each additional year in age. This makes sense because we usually expect larger houses to be more valuable and older houses to be less valuable. This equation fits our general intuition of what likely should be the case.

It's a good idea to try to intuitively make sense of the regression coefficients before you start using the equation to make forecasts. Doing this can teach you something new about your data and can sometimes reveal problems with your data. There may be a few extreme outliers or missing values (sometimes the computer treats these as "0") in a data set that can throw all the coefficients off. As we discussed earlier with correlation and simple regression, it may be best to remove such extreme outliers and treat them as special cases.

Suppose we decided to add in another variable, X_3, that equals the number of bedrooms. It seems reasonable that a house with three or four bedrooms is likely to be more valuable to a growing family than a house with only two bedrooms. When we use the computer to re-calculate the regression equation with this new variable we get the following:

$$Y = 231{,}485 + 480X_1 - 6{,}365X_2 - 2{,}682X_3.$$

You may be surprised to see that all the coefficients and the intercept have changed even though we're looking at the same town and the same houses. Why has everything changed?

The answer is that the interpretation of each of the coefficients has now completely changed. To see this, let's first take a look at the coefficient for X_3, the number of bedrooms. The interpretation of the coefficient $-2{,}682$ is that for houses of the same age and the same square footage, with each additional bedroom we see a *decrease* in value of around \$2,700.

But this last interpretation seems puzzling—each extra bedroom *decreases* the value? We stated above that we expected an extra bedroom to definitely *increase* the value. The answer to this puzzle is that the coefficient now tells you what happens to the value when you have an extra bedroom but still have the same age and same square footage overall; houses with an extra bedroom but without any additional square footage have *lower* home value because the other rooms in the house would have to be smaller and the house more crowded. So the negative coefficient makes sense after all. It is important to understand that the coefficient tells you what happens to the Y variable when all the other variables in the equation stay constant and you look at a change in only a single X variable.

In summary, each of the coefficients in this new multiple regression equation now represents something completely different: what happens to the home value when you look at a difference in a variable with the restriction that all the other variables in the equations stay the same. The moral is that you should watch out when adding or removing X variables from a multiple regression—all the coefficients can change in surprising ways as the interpretations change. This also means that the interpretation of the coefficient for some variable depends on which other variables are in the equation.

> When adding or removing X variables from a multiple regression, the coefficients can change in surprising ways as the interpretations change.

Incidentally, to get a better picture of the value of an extra bedroom we could try this alternate approach. When you add a bedroom to your home you naturally expect that the square footage would rise. This means we should leave out square footage from the multiple regression equation and predict home value using only the other variables: the age variable, X_2, and the number of bedrooms, X_3. Running

Chapter 9

Sampling Variability and Standard Error

1. Introduction

USA Today reported that approximately 60% of all TV sets in the country were tuned to the 2006 Super Bowl and that this was an increase of around five percentage points from 2005.[1] This turns out to be about 90 million viewers which makes it one of the most-watched television shows in history. Furthermore, it was reported that women made up approximately 44% of the audience.[2] Because of the size of the audience, the network was selling commercial time for the outrageous sum of almost $6 million per minute—almost $100,000 per second.

But how could anybody possibly know how many TV sets were tuned in to the Super Bowl? And how could they possibly know if viewers were men or women? Nobody from the newspaper came around to my house looking to see who was watching and I doubt anyone came around to your house either. Since the value of television advertising is directly linked to being able to measure the size of the audience, is there any accuracy to these figures? Or are they just pulled out of thin air?

ACNielsen, the company that computes these figures, is able to very accurately es-

[1] See `http://www.usatoday.com/life/television/news/2006-02-06-super-bowl-rating_x.htm`

[2] `http://www.forbes.com/2006/01/31/advertising-super-bowl-cz_af_0201ads_super06.html`

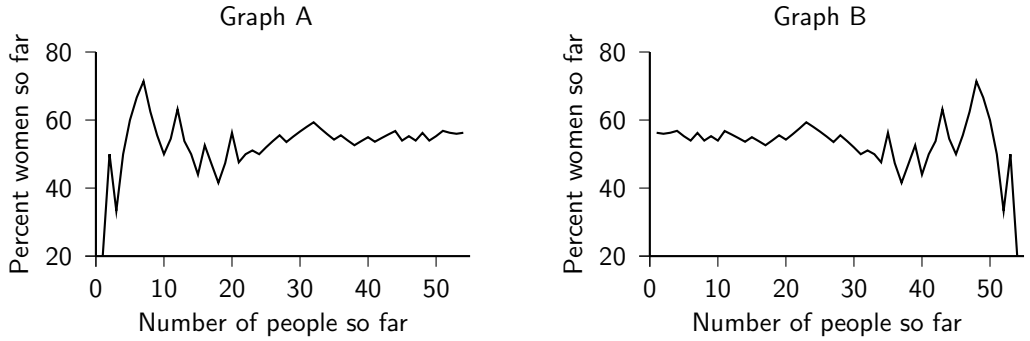

Figure 9.1. Which graph would we see?

timate these percentages simply by monitoring a very small sample of households—usually only around a thousand households spread across the entire country. The same principle applies with political polls: even if the pollster doesn't speak to every American directly, a small sample can be very representative. In this chapter we will discuss how this process works and exactly how accurate such small samples tend to be. There is a simple but strong theoretical underpinning to estimating the accuracy of such samples and the purpose of this chapter to describe it and apply it to different situations.

2. The Law of Averages

Let's start by introducing, in an intuitive way, the simple principle upon which this entire chapter rests. Suppose we stand on a street in downtown Boston and watch the people walk by. As each person walks by, we graph the percentage of people we've seen so far who are women. Which graph in Figure 9.1 would most likely represent what we would see—Graph A or Graph B?

If you said Graph A, you would be correct. Graph A starts out with large fluctuations and then starts to stabilize at around 50%. Graph B is the mirror image; it starts out with small fluctuations and then has larger fluctuations later on. The principle in action here is the idea that as the sample gets larger, the percentage of women in the sample tends to stabilize very close to the true population percentage (somewhere around 50% here in Boston). In a small sample, there is more likely to be a lot of random variation. This is because in a small sample any single person can change the overall percentage quite a bit, but as the sample gets larger each individual person has very little impact on his or her own. This also means larger samples become

more representative of the population. This principle is called the **Law of Averages**, because the same thing also applies to a sample average fluctuating around the population average.

> The **law of averages** says that with a large sample, a percentage observed in your sample will tend to be very close to the true percentage in the population. With a small sample, you are more likely to observe a large fluctuation away from the true population percentage.

Another thing to notice about Graph A in Figure 9.1 is that the percentage quickly stabilizes even after only 50 people. This means that a sample as small as 50 people from a city like Boston with millions of residents can be reasonably representative. I will discuss later in this chapter and in the next chapter how to determine the sample size you should use.

Example. In the example above, would you be more likely to see more than 70% women in a random sample of 20 people or 200 people?

Answer. Twenty people. With 200 people you are very likely to see close to 50 percent women; in a small sample of only 20 people it is more likely you might see a fluctuation away from this number.

Example. In the example above, would you be more likely to see more than 40% women in a random sample of 20 people or 200 people?

Answer. This time the answer is 200 people. In a large sample you are very likely to see close to 50% and thus more than 40%.

Example. Which gives you a larger chance of coming out ahead financially: playing a slot machine two times or 50 times?

Answer. Unfortunately, you are more likely to come out ahead if you play the slot machine two times (and probably even more likely to come out ahead if you play it only once). If you play the machine a very large number of times you are likely to

lose money, since the odds are against you on all slot machines. If you are hoping for a fluctuation away from these unfortunate odds (as all gamblers of course do), you should play as few times as possible.

In this section we developed an understanding of how larger samples tend to be more representative of the population compared with smaller samples, due to the law of averages. In the next section we will examine how this can be quantified and how this can help you estimate how large of a sample you need in a given business situation.

Exercises

1. Medical errors are a problem of increasing public concern and hospital executives are eager to understand which management practices could reduce them. Analysts studying one particular medical condition looked at hospital records to see if hospital caregivers actually correctly gave patients all the treatments their doctors prescribed for them. For this medical condition, overall only about 80% of patients nationwide correctly received all the treatments prescribed by their doctors. This percentage varied widely from one hospital to another and analysts noticed that smaller hospitals tended to show much more variability in this percentage compared with the larger hospitals. This is illustrated by the funnel shape in Figure 9.2, in which each dot corresponds to one hospital. An analyst says the funnel shape may indicate that large hospitals have better standardization procedures in place that reduce variability. Can you think of another explanation for the funnel shape?

2. Since the amount of money a web site can charge for advertising is closely linked to the daily number of hits, most sites keep track of such numbers. To look for trends, one web site graphed the daily number of hits averaged over each week, as well as the daily number of hits averaged over each month. One graph in Figure 9.3 shows a point for each of 26 months, and the other graph shows a point for each of 26 weeks. Which graph is which? Justify your answer.

3. An airline has a policy of selling 5% more tickets than it has seats on each airplane, because experience has shown that about 10% of people who buy tickets never actually show up for their flights. In the event too many people show up, the airline must 'bump' customers to other flights and pay them compensation. Assuming the flights are sold-out, is it more likely the airline will have to 'bump'

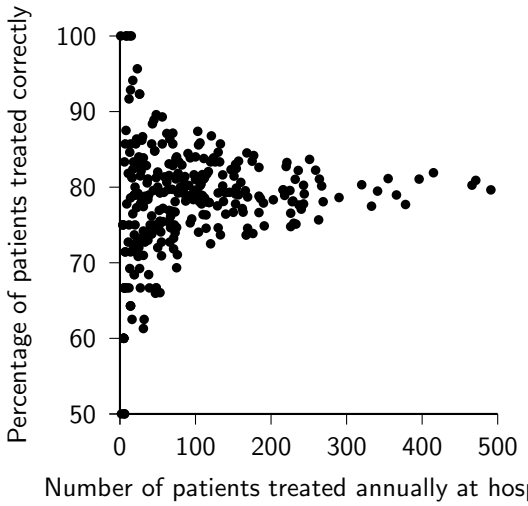

Figure 9.2. The percentage of patients treated correctly versus the number of patients treated annually at a selection of hospitals. What could explain the funnel shape?

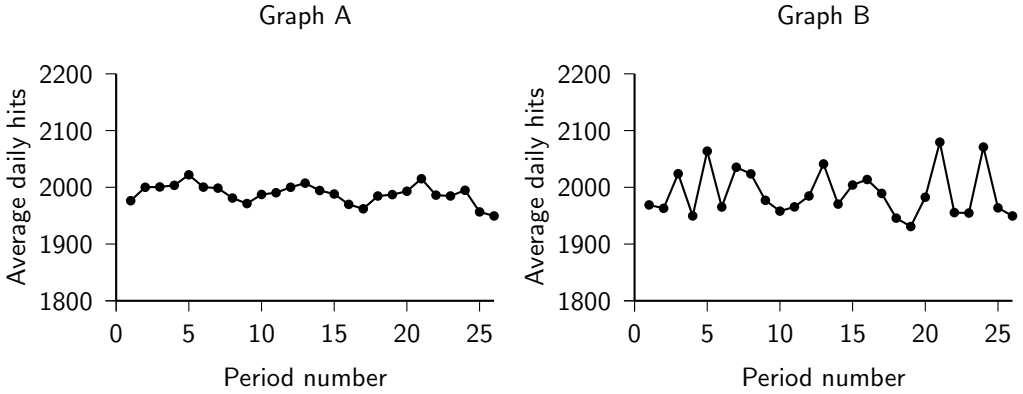

Figure 9.3. Average number of web page hits over time.

people from a large airplane or a small airplane? Answer and explain briefly how the law of averages applies.

4. A stock index fund invests in a very large number of different stocks, so the return you get from it is roughly the average return for the stocks in the index. (a) Which of these two choices would give you a better chance of tripling your money within a month: investing all your money in the index fund, or investing all your money in a few randomly chosen stocks that make up the index? (b) Which of these two choices would give you a better chance of losing almost

all of your money within a month: investing all your money in the index fund, or investing all your money in a few randomly chosen stocks that make up the index?

5. Suppose you stand in downtown Boston and count people who walk by with red hair. Would you be more likely to see that more than half the people have red hair in a random sample of 20 people or 200 people? Explain.

3. The Standard Error for a Sample Percentage

We have seen that if we stand on a corner in downtown Boston and take a sample of the population, as our sample gets larger the sample percentage will approach the population percentage because of the law of averages. Applying this reasoning to a political poll, we can imagine that a poll will come close to reflecting the opinions of all Americans as the sample becomes larger. In other words, all U.S. citizens do not have to be contacted for the poll to be accurate. But the law of averages only suggests a direction: use large samples. It does not tell us how large the sample needs to be. And it does not tell us how confident we can be that various sample sizes will give us an accurate reading. Answering this question is our next step.

Suppose you want to estimate the percentage of people in a population who support a given political candidate. Since there are always fluctuations in a sample percentage, how accurately can we confidently estimate the percentage based on a sample size of, say, only 100 people? And how would our accuracy change if we increased the sample size to 1,000 people?

To answer this question, we must use a statistical formula that quantifies the amount of variation, or error, you would expect to see in a sample percentage. This expected amount of error is called the **standard error for the sample percentage** (SE for %), and calculating it depends on the sample size and roughly what you expect the true percentage to be. For now I will just show you the formula and how to use it, and at the end of the next section we will see why the formula is reasonable. The formula is

$$\text{SE for \%} = \sqrt{\frac{\text{true percentage} \times (100\% - \text{true percentage})}{\text{size of sample}}}$$

where the "true percentage" is the true percentage in the population that you're sampling from—or your best rough estimate of it so far. This formula applies only to what is called a **simple random sample**, where every possible group of people

in population has the same chance of getting selected to be in the sample. In other words, this type of random sampling is just as if you pulled random names on slips of paper out of a hat. (Most large surveys and polls use a slightly different way of choosing the sample, and they must therefore use a slightly different formula for the error. There are entire books written on the many different standard error formulas for each common type of sampling method.)

It may seem like circular reasoning here to use a rough estimate of the "true percentage" in order to compute the accuracy of that estimate itself, but it turns out to work well in practice. If you absolutely have no idea at all about the true percentage you can plug in 50% to get a conservative overestimate of the standard error (this will give the largest possible standard error in the formula—plug in a few numbers and see it for yourself). This type of circular reasoning is called a **bootstrap estimate** due to the old saying that you can help yourself by "pulling yourself up by your bootstraps."

But before you start using this formula, I must emphasize that one standard error does not give a realistic picture of the range of possible errors you might see from a sample—it only gives the typical error. Just as the standard deviation only gives the typical distance of numbers to the average, you get a more realistic picture of the range of a data set if you go two or three standard deviations in each direction from the average. For this reason, people traditionally double the standard error to get what is called the **margin of error**, and this can be viewed as a rough estimate of the largest error you would reasonably expect to see. Of course it is still possible for the error to come out even larger than the margin of error, but this only happens rarely. In the next chapter I will discuss how rarely this happens. To be extra conservative you could triple the standard error, but the common convention is to double it.

The **standard error for a sample percentage** (SE for %) estimates the typical error you expect to see in a sample percentage:

$$\text{SE for \%} = \sqrt{\frac{\text{true percentage} \times (100\% - \text{true percentage})}{\text{size of sample}}}$$

The **margin of error** is twice the standard error, and estimates the largest error you reasonably expect to see:

$$\text{ME for \%} = 2 \times (\text{SE for \%})$$