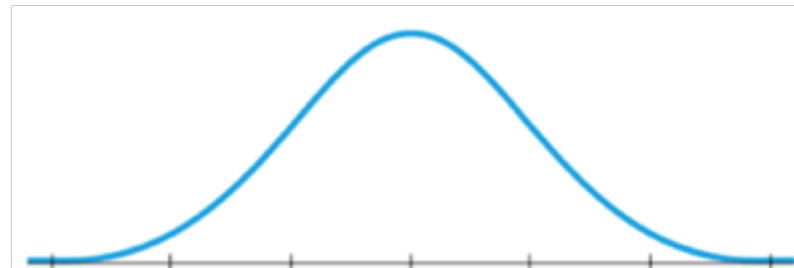Agenda: #58F AP Stats Ch18B

Title: Confidence Interval

Objectives: Find confidence interval by using TI-Calculate codes.

1) W.Up : (Worksheet 15min)

Use invNorm to calculate the real z-score(s) for 90%.

2) Coral Reef Question

(Reading+ Activity 20 min)

Find SE (standard error) and Conclude Statement
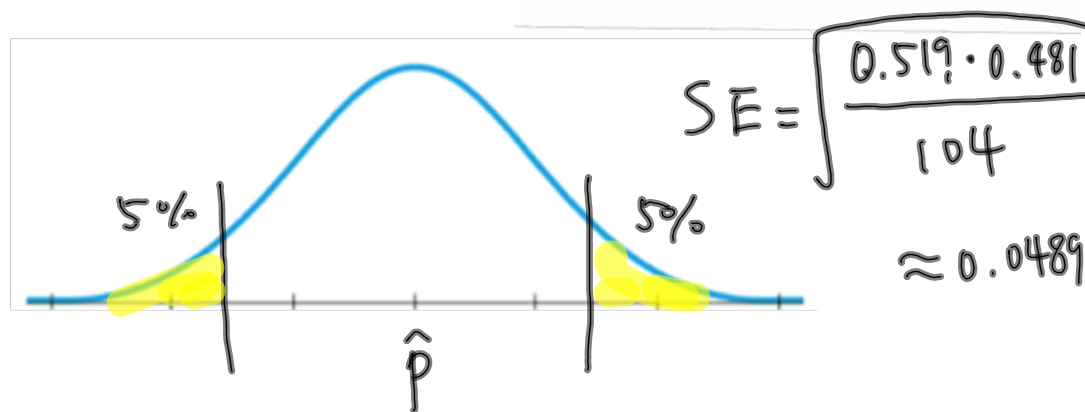
3) Margin of Errors (Reading)

4) Global Warming Questions & Closing Activities (15min)

Our confidence interval had the form

$$\hat{p} \pm 2\, SE(\hat{p}).$$

The extent of the interval on either side of $\hat{p}$ is called the **margin of error** (*ME*).

$$\hat{p} \cdots \frac{59}{104} \quad \text{or} \quad 51.9\%$$

$$SE = \sqrt{\frac{0.519 \cdot 0.481}{104}}$$

5% 5%

$$\approx 0.0489$$

$\hat{p}$

$$Z^* = 1.645$$

(90%)

$$invNorm(0.05) \approx 1.645$$

$$CI = \hat{p} \pm Z^*(SE)$$

$$= .519 \pm (1.645)(0.0489)$$

$$= .519 \pm 0.0804$$

$$(.439 , .599)$$

We are 90% confident
that true p exists between

.439 and .599.

Each confidence interval discussed in the book has a name. You'll see many different kinds of confidence intervals in the following chapters. Some will be about more than *one* sample, some will be about statistics other than *proportions,* and some will use models other than the Normal. The interval calculated and interpreted here is sometimes called a **one-proportion** *z*-**interval.**[4]

TI-84 <mark>one-proportion z-interval.[4]</mark>

Stats → Tests →

1 - Prop Zint

$x : 54$

$n : 104$

C-Interval : .90

The extent of the interval on either side of $\hat{p}$ is called the **margin of error** (*ME*).

We'll want to use the same approach for many other situations besides estimating proportions. In fact, almost any population parameter—a proportion, a mean, or a regression slope, for example—can be estimated with some margin of error. The margin of error is a way to describe our uncertainty in estimating the population value. We'll see how to find a margin of error for each of these values and for others.

For all of those statistics, regardless of how we calculate the margin of error, we'll be able to construct a confidence interval that looks like this:

$$Estimate \pm ME.$$

The margin of error for our 95% confidence interval was 2 *SE*. What if we wanted to be more confident? To be more confident, we'll need to capture p more often, and to do that we'll need to make the interval wider. For example, if we want to be 99.7% confident, the margin of error will have to be 3 *SE*.

Figure 18.3

The horizontal green line shows the true percentage of all sea fans that are infected. Most of the 20 simulated samples produced confidence intervals that captured the true value, but a few missed.
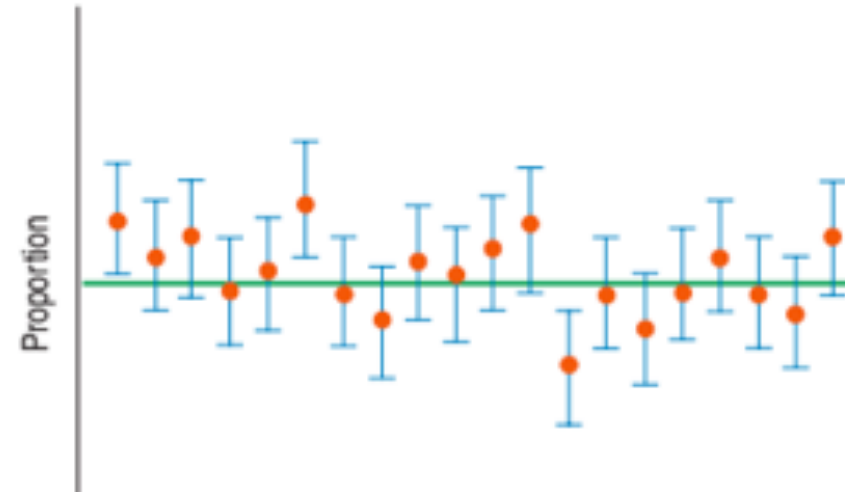
In April and May 2011, the Yale Project on Climate Change Communication and the George Mason University Center for Climate Change Communication interviewed 1010 U.S. adults about American's global warming beliefs and attitudes.[5]

QUESTION: It is standard among pollsters to use a 95% confidence level unless otherwise stated. Given that, what do these researchers mean by their confidence interval in this context?

For Example POLLS AND MARGIN OF ERROR

In April and May 2011, the Yale Project on Climate Change Communication and the George Mason University Center for Climate Change Communication interviewed 1010 U.S. adults about American's global warming beliefs and attitudes.[5]

QUESTION: It is standard among pollsters to use a 95% confidence level unless otherwise stated. Given that, what do these researchers mean by their confidence interval in this context?

ANSWER: *If this polling were done repeatedly, 95% of all random samples would yield confidence intervals that contain the true proportion of all U.S. adults who believe that there's a lot of disagreement among scientists about global warming.*

The more confident we want to be, the larger the margin of error must be. We can be 100% confident that the proportion of infected sea fans is between 0% and 100%, but this isn't likely to be very useful. On the other hand, we could give a confidence interval from 51.8% to 52.0%, but we can't be very confident about a precise statement like this. Every confidence interval is a balance between certainty and precision.

The tension between certainty and precision is always there. Fortunately, in most cases we can be both sufficiently certain and sufficiently precise to make useful statements. There is no simple answer to the conflict. You must choose a confidence level yourself. The data can't do it for you. The choice of confidence level is somewhat arbitrary. The most commonly chosen confidence levels are 90%, 95%, and 99%, but any percentage can be used. (In practice, though, using something like 92.9% or 97.2% is likely to make people think you're up to something.)

$$CI = \hat{p} \pm z^* (SE)$$

more confident

more margin of error

Example FINDING THE MARGIN OF ERROR (TAKE 1)

RECAP: An April 2011 Yale/George Mason poll of 1010 U.S. adults asking questions about current topics reported a margin of error of 3%. It is a convention among pollsters to use a 95% confidence level and to report the "worst case" margin of error, based on $p = 0.5$.

QUESTION: How did the researchers calculate their margin of error?

$$P = 0.5 \qquad q = 0.5$$

(true p)

$$SD = \sqrt{\frac{.5 \cdot .5}{1010}} \approx 0.0157$$

$$0.0157 \cdot 1.96 = 0.030$$

Example FINDING THE MARGIN OF ERROR (TAKE 1)

RECAP: An April 2011 Yale/George Mason poll of 1010 U.S. adults asking questions about current topics reported a margin of error of 3%. It is a convention among pollsters to use a 95% confidence level and to report the "worst case" margin of error, based on $p = 0.5$.

QUESTION: How did the researchers calculate their margin of error?

ANSWER: Assuming $p = 0.5$, for random samples of $n = 1010$,

$$SD(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.5)(0.5)}{1010}} = 0.0157.$$

For a 95% confidence level, $ME = 2(0.0157) = 0.031$, so their margin of error is just a bit over 3%.
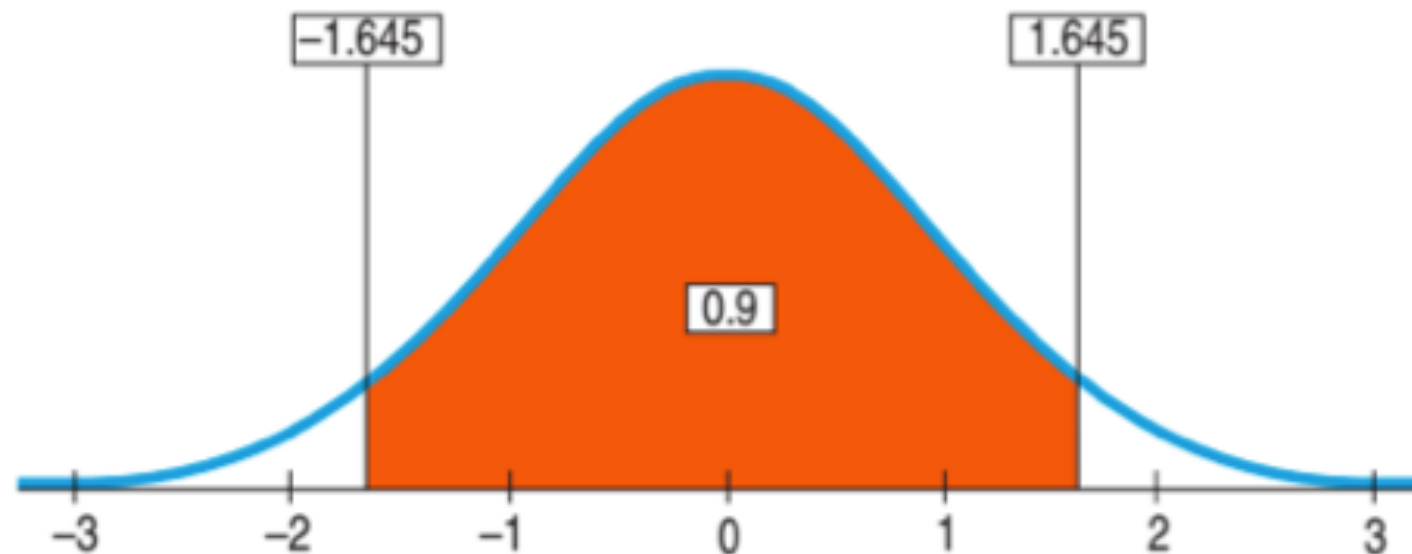
Critical Values

In our sea fans example we used $2SE$ to give us a 95% confidence interval. To change the confidence level, we'd need to change the *number* of SEs so that the size of the margin of error corresponds to the new level. This number of SEs is called the **critical value.** Here it's based on the Normal model, so we denote it z^*. For any confidence level, we can find the corresponding critical value from a computer, a calculator, or a Normal probability table, such as Table Z.

For a 95% confidence interval, you'll find the precise critical value is $z^* = 1.96$. That is, 95% of a Normal model is found within ± 1.96 standard deviations of the mean. We've been using $z^* = 2$ from the 68–95–99.7 Rule because it's easy to remember.

For Example FINDING THE MARGIN OF ERROR (TAKE 2)

RECAP: In April 2011, a Yale/George Mason poll of 1010 U.S. adults found that 40% of the respondents believed that scientists disagreed about whether global warming exists. They reported a 95% confidence interval with a margin of error of 3%.

QUESTIONS: Using the critical value of z and the standard error based on the observed proportion, what would be the margin of error for a 90% confidence interval? What's good and bad about this change?

With $n = 1010$ and $\hat{p} = 0.40$, $SE(\hat{p}) = \sqrt{\dfrac{\hat{p}\hat{q}}{n}} = \sqrt{\dfrac{(0.40)(0.60)}{1010}} = 0.0154$

For a 90% confidence level, $z^* = 1.645$, so $ME = 1.645(0.0154) = 0.0254$.

$$\hat{P} = 0.40$$

$$\hat{q} = 0.60$$

$$SE(\hat{p}) = \sqrt{\dfrac{(0.40)(0.60)}{1010}} = 0.0154$$

$$90\% \; CI \quad z^* \; 1.645 \cdot (0.0154) = 0.0254$$

$$95\% \; CI \quad z^* \quad 1.96 \cdot (0.0154) = 0.0301$$

For Example FINDING THE MARGIN OF ERROR (TAKE 2)

RECAP: In April 2011, a Yale/George Mason poll of 1010 U.S. adults found that 40% of the respondents believed that scientists disagreed about whether global warming exists. They reported a 95% confidence interval with a margin of error of 3%.

QUESTIONS: Using the critical value of z and the standard error based on the observed proportion, what would be the margin of error for a 90% confidence interval? What's good and bad about this change?

$$\text{With } n = 1010 \text{ and } \hat{p} = 0.40, \; SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(0.40)(0.60)}{1010}} = 0.0154$$

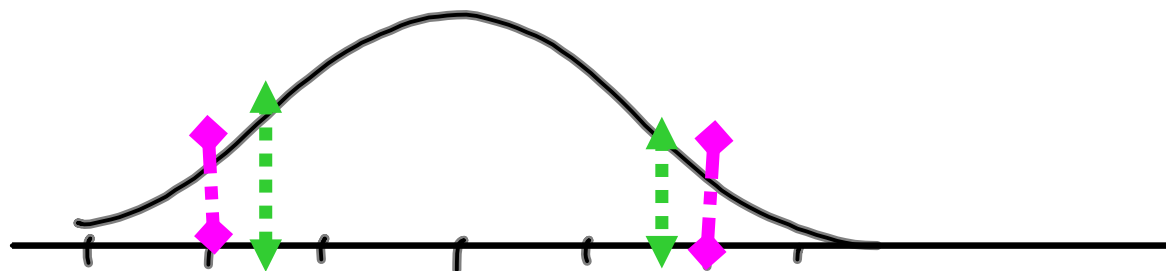For a 90% confidence level, $z* = 1.645$, so $ME = 1.645(0.0154) = 0.0254$.

ANSWER: Now the margin of error is only about 2.5%, producing a narrower interval. What's good about the change is that we now have a smaller interval, but what's bad is that we are less certain that the interval actually contains the true proportion of adults who think that scientists disagree about global warming.

Assumptions and Conditions

We've just made some pretty sweeping statements about sea fans. Those statements were possible because we used a Normal model for the sampling distribution. But is that model appropriate?

We've also said that the same basic ideas will work for other statistics. One difference for those statistics is that some will have sampling distribution models that are different than the Normal. But the background theory (which we won't bother with) is so similar for all of those models that they share the same basic assumptions and conditions about independence and sample size. Then for each statistic, we usually tack on a special-case assumption or two. We'll deal with those as we get to them in later chapters.

We saw the assumptions and conditions for using the Normal to model the sampling distribution for a proportion in the last chapter. Because they are so crucial to making sure our confidence interval is useful, we'll repeat them here.

Independence Assumption

Independence Assumption: The data values must be independent. To think about whether this assumption is plausible, we often look for reasons to suspect that it fails. We wonder whether there is any reason to believe that the data values somehow affect each other. (For example, might the disease in sea fans be contagious?) Whether you decide that the **Independence Assumption** is plausible depends on your knowledge of the situation. It's not one you can check by looking at the data.

However, now that we have data, there are two conditions that we can check:

Randomization Condition: Were the data sampled at random or generated from a properly randomized experiment? Proper randomization can help ensure independence.

10% Condition: If you sample more than 10% of a population, the formula for the standard error won't be quite right. There is a special formula (found in advanced books) that corrects for this, but it isn't a common problem unless your population is small.

Sample Size Assumption

The model we use for inference for proportions is based on the Central Limit Theorem. We need to know whether the sample is large enough to make the sampling model for the sample proportions approximately Normal. It turns out that we need more data as the proportion gets closer and closer to either extreme (0 or 1). That's why we check the:

Success/Failure Condition: We must expect at least 10 "successes" and at least 10 "failures." Recall that by tradition we arbitrarily label one alternative (usually the outcome being counted) as a "success" even if it's something bad (like getting a disease). The other alternative is, of course, then a "failure."

One-Proportion z-Interval

When the conditions are met, we are ready to find a level C confidence interval for the population proportion, p. The confidence interval is $\hat{p} \pm z^* \times SE(\hat{p})$ where the standard deviation of the proportion is estimated by $SE(\hat{p}) = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$ and the critical value, z^*, specifies the number of SEs needed for C% of random samples to yield confidence intervals that capture the true parameter value.

In October 2010, the Gallup Poll[6] asked 510 randomly sampled adults the question "Generally speaking, do you believe the death penalty is applied fairly or unfairly in this country today?" Of these, 58% answered "Fairly," 36% said "Unfairly," and 7% said they didn't know. (Percentages add up to 101% due to rounding).

Fairly
Unfairly
Don't Know
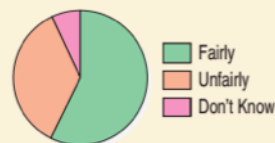
Who	Adults in the United States
What	Response to a question about the death penalty
When	October 2010
Where	United States
How	510 adults were randomly sampled and asked by the Gallup Poll
Why	Public opinion research

Question: From this survey, what can we conclude about the opinions of *all* adults?

To answer this question, we'll build a confidence interval for the proportion of all U.S. adults who believe the death penalty is applied fairly. There are four steps to building a confidence interval for proportions: Plan, Model, Mechanics, and Conclusion.

Fair ···· 0.58   ····→ assume this is $\hat{p}$

Unfair ···· 0.36

$\boxed{0.58}$

(Don't know ··· 0.07 ;)

$\hat{p} = 0.58 \qquad \hat{q} = 0.42$

① Conditions less than 10%

② SRS

③
$\begin{pmatrix} 510 \cdot .58 = 296 \\ 510 \cdot .42 = 212 \\ \text{more than 10} \end{pmatrix}$

$SE_{(\hat{p})} = \sqrt{\dfrac{.58 \cdot .42}{510}} \approx 0.02\overset{2}{18}$

CI 95%   1.96

$0.58 \pm 1.96 \,(0.022)$

$(0.537, \; 0.623)$

I want to find an interval that is likely, with 95% confidence, to contain the true proportion, p, of U.S. adults who think the death penalty is applied fairly. I have a random sample of 510 U.S. adults.

✓ **Randomization Condition:** Gallup drew a random sample from all U.S. adults. I can be confident that the respondents are independent.

✓ **10% Condition:** The sample is certainly less than 10% of the population.

✓ **Success/Failure Condition:**
$n\hat{p} = 510(58\%) = 296 \geq 10$ and
$n\hat{q} = 510(42\%) = 214 \geq 10$,
so the sample appears to be large enough to use the Normal model.

The conditions are satisfied, so I can use a Normal model to find a **one-proportion z-interval**.

$n = 510$, $\hat{p} = 0.58$, so

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(0.58)(0.42)}{510}} = 0.022.$$

Because the sampling model is Normal, for a 95% confidence interval, the critical value $z^* = 1.96$.

The margin of error is

$$ME = z^* \times SE(\hat{p}) = 1.96(0.022) = 0.043.$$

So the 95% confidence interval is

$$0.58 \pm 0.043 \text{ or } (0.537, 0.623).$$

$n = 510, \hat{p} = 0.58$, so

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{(0.58)(0.42)}{510}} = 0.022.$$

Because the sampling model is Normal, for a 95% confidence interval, the critical value $z^* = 1.96$.

The margin of error is

$$ME = z^* \times SE(\hat{p}) = 1.96(0.022) = 0.043.$$

So the 95% confidence interval is

$$0.58 \pm 0.043 \text{ or } (0.537, 0.623).$$

I am 95% confident that between 53.7% and 62.3% of all U.S. adults think that the death penalty is applied fairly.

TI Tips FINDING CONFIDENCE INTERVALS

```
EDIT CALC TESTS
7↑ZInterval…
8:TInterval…
9:2-SampZInt…
0:2-SampTInt…
A:1-PropZInt…
B:2-PropZInt…
C↓χ²-Test…
```

It will come as no surprise that your TI can calculate a confidence interval for a population proportion. Remember the sea fans? Of 104 sea fans, 54 were diseased. To find the resulting confidence interval, we first take a look at a whole new menu.

- Under STAT go to the TESTS menu. Quite a list! Commands are found here for the inference procedures you will learn through the coming chapters.

```
1-PropZInt
 x:54
 n:104
 C-Level:.95
 Calculate
```

- We're using a Normal model to find a confidence interval for a proportion based on one sample. Scroll down the list and select 1-PropZInt.
- Enter the number of successes observed and the sample size.
- Specify a confidence level and then Calculate.

```
1-PropZInt
 (.42321,.61525)
 p̂=.5192307692
 n=104
■
```

And there it is! Note that the TI calculates the sample proportion for you, but the important result is the interval itself, 42% to 62%. The calculator did the easy part—just Show. Tell is harder. It's your job to interpret that interval correctly.

```
ERR:DOMAIN
1▉Quit
```

    Beware: You may run into a problem. When you enter the value of x, you need a *count*, not a percentage. Suppose the marine scientists had reported that 52% of the 104 sea fans were infected. You can enter x:.52*104, and the calculator will evaluate that as 54.08. Wrong. Unless you fix that result, you'll get an error message. Think about it—the number of infected sea fans must have been a whole number, evidently 54. When the scientists reported the results, they rounded off the actual percentage $(54 \div 104 = 51.923\%)$ to 52%. Simply change the value of x to 54 and you should be able to Calculate the correct interval.

Choosing Your Sample Size

The question of how large a sample to take is an important step in planning any study. We weren't ready to make that calculation when we first looked at study design in Chapter 11, but now we can—and we always should.

Suppose a candidate is planning a poll and wants to estimate voter support within 3% with 95% confidence. How large a sample does she need?

Let's look at the margin of error:

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.03 = 1.96 \sqrt{\frac{\hat{p}\hat{q}}{n}}.$$

We want to find n, the sample size. To find n we need a value for $\hat{p}$. We don't know $\hat{p}$ because we don't have a sample yet, but we can probably guess a value. The worst case— the value that makes $\hat{p}\hat{q}$ (and therefore n) largest—is 0.50, so if we use that value for $\hat{p}$, we'll certainly be safe. Our candidate probably expects to be near 50% anyway.

Our equation, then, is

$$0.03 = 1.96 \sqrt{\frac{(0.5)(0.5)}{n}}.$$

To solve for n, we first multiply both sides of the equation by $\sqrt{n}$ and then divide by 0.03:

$$0.03\sqrt{n} = 1.96\sqrt{(0.5)(0.5)}$$

$$\sqrt{n} = \frac{1.96\sqrt{(0.5)(0.5)}}{0.03} \approx 32.67$$

Notice that evaluating this expression tells us the *square root* of the sample size. We need to square that result to find *n:*

$$n \approx (32.67)^2 \approx 1067.1$$

To be safe, we round up and conclude that we need at least 1068 respondents to keep the margin of error as small as 3% with a confidence level of 95%.

For Example CHOOSING A SAMPLE SIZE

RECAP: The Yale/George Mason poll that estimated that 40% of all voters believed that scientists disagree about whether global warming exists had a margin of error of $\pm 3\%$. Suppose an environmental group planning a follow-up survey of voters' opinions on global warming wants to determine a 95% confidence interval with a margin of error of no more than $\pm 2\%$.

QUESTION: How large a sample do they need? (You could take $p = 0.5$, but we have data that indicate $p = 0.40$, so we can use that.)

For Example CHOOSING A SAMPLE SIZE

RECAP: The Yale/George Mason poll that estimated that 40% of all voters believed that scientists disagree about whether global warming exists had a margin of error of $\pm 3\%$. Suppose an environmental group planning a follow-up survey of voters' opinions on global warming wants to determine a 95% confidence interval with a margin of error of no more than $\pm 2\%$.

QUESTION: How large a sample do they need? (You could take $p = 0.5$, but we have data that indicate $p = 0.40$, so we can use that.)

ANSWER:

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.02 = 1.96 \sqrt{\frac{(0.40)(0.60)}{n}}$$

$$\sqrt{n} = \frac{1.96\sqrt{(0.40)(0.60)}}{0.02} \approx 48.01$$

$$n = 48.01^2 = 2304.96$$

The environmental group's survey will need at least 2305 respondents.

Unfortunately, bigger samples cost more money and more effort. Because the standard error declines only with the *square root* of the sample size, to cut the standard error (and thus the ME) in half, we must *quadruple* the sample size.

Generally a margin of error of 5% or less is acceptable, but different circumstances call for different standards. For a pilot study, a margin of error of 10% may be fine, so a sample of 100 will do quite well. In a close election, a polling organization might want to get the margin of error down to 2%. Drawing a large sample to get a smaller ME, however, can run into trouble. It takes time to survey 2400 people, and a survey that extends over a week or more may be trying to hit a target that moves during the time of the survey. An important event can change public opinion in the middle of the survey process.

Keep in mind that the sample size for a survey is the number of respondents, not the number of people to whom questionnaires were sent or whose phone numbers were dialed. And also keep in mind that a low response rate turns any study essentially into a voluntary response study, which is of little value for inferring population values. It's almost always better to spend resources on increasing the response rate than on surveying a larger group. A full or nearly full response by a modest-size sample can yield useful results.

Surveys are not the only place where proportions pop up. Banks sample huge mailing lists to estimate what proportion of people will accept a credit card offer. Even pilot studies may mail offers to over 50,000 customers. Most don't respond; that doesn't make the sample smaller—they simply said "No thanks." Those who do respond want the card. To the bank, the response rate[7] is $\hat{p}$. With a typical success rate around 0.5%, the bank needs a very small margin of error—often as low as 0.1%—to make a sound business decision. That calls for a large sample, and the bank must take care in estimating the size needed. For our election poll calculation we used $p = 0.5$, both because it's safe and because we honestly believed p to be near 0.5. If the bank used 0.5, they'd get an absurd answer. Instead, they base their calculation on a proportion closer to the one they expect to find.

For Example SAMPLE SIZE REVISITED

A credit card company is about to send out a mailing to test the market for a new credit card. From that sample, they want to estimate the true proportion of people who will sign up for the card nationwide. A pilot study suggests that about 0.5% of the people receiving the offer will accept it.

QUESTION: To be within a tenth of a percentage point (0.001) of the true rate with 95% confidence, how big does the test mailing have to be?

A credit card company is about to send out a mailing to test the market for a new credit card. From that sample, they want to estimate the true proportion of people who will sign up for the card nation-wide. A pilot study suggests that about 0.5% of the people receiving the offer will accept it.

QUESTION: To be within a tenth of a percentage point (0.001) of the true rate with 95% confidence, how big does the test mailing have to be?

ANSWER: Using the estimate $\hat{p} = 0.5\%$:

$$ME = 0.001 = z^*\sqrt{\frac{\hat{p}\hat{q}}{n}} = 1.96\sqrt{\frac{(0.005)(0.995)}{n}}$$

$$(0.001)^2 = 1.96^2\frac{(0.005)(0.995)}{n} \Rightarrow n = \frac{1.96^2(0.005)(0.995)}{(0.001)^2}$$

$$= 19{,}111.96 \text{ or } 19{,}112$$

That's a lot, but it's actually a reasonable size for a trial mailing such as this. Note, however, that if they had assumed 0.50 for the value of p, they would have found

$$ME = 0.001 = z^*\sqrt{\frac{pq}{n}} = 1.96\sqrt{\frac{(0.5)(0.5)}{n}}$$

$$(0.001)^2 = 1.96^2\frac{(0.5)(0.5)}{n} \Rightarrow n = \frac{1.96^2(0.5)(0.5)}{(0.001)^2} = 960{,}400.$$

Quite a different (and unreasonable) result.

Cars What fraction of cars is made in Japan? The computer output below summarizes the results of a random sample of 50 autos. Explain carefully what it tells you.

z-Interval for proportion
With 90.00% confidence,
$0.29938661 < p(japan) < 0.46984416$