Lesson 8
Confidence Intervals for Proportions

Outline of the Lesson

As the tax-filing deadline approached in early April of 2011, a national polling organization conducted a survey. They asked the question, "Do you think the amount of taxes you pay is fair?" News media reported the results as, "According to a recent survey, 54% of Americans age 18 and above believe that the amount of taxes they pay is fair, with a margin of error ±4.2%." This is an example of ***inferential statistics***. As we have discussed in the previous two lessons, inferential statistics has two important features:

- Information is obtained from a *sample*.
- The information from the sample is used to draw a conclusion (an *inference*) about the entire *population* from which the sample was drawn.

In this example, the polling organization certainly did not poll every person age 18 and above in the entire country, but the results of the poll were reported as a fact about the entire country. The pollsters made an *inference* about the entire country, based on the results of the poll.

This course covers two major types of inferential statistics, *confidence intervals* and *hypothesis tests*. In the previous two lessons we studied the use of hypothesis tests to draw conclusions about population proportions. In this lesson we learn about confidence intervals, again in the context of population proportions. As we did in the previous two lessons, we will take advantage of the connection between proportions and probabilities. For the poll described above, there are two ways to interpret the 54% figure reported in the news media:

- The proportion of Americans, age 18 or above, who believe the amount of tax they pay is fair, is 54%.
- The probability that a randomly selected American, age 18 or above, would believe the amount of tax they pay is fair, is 0.54.

Building on this connection between proportions and probabilities, we will use a strategy similar to that taken in Lessons 6 and 7. We begin our explanation by thinking about probabilities for random events such as coin tosses and rolls of dice. Then we apply what we learn to the polling process.

Confidence interval vs. Hypothesis test

Confidence intervals, like hypothesis tests, are used to make a statement about a population proportion, with the following difference:

- We use a hypothesis test to *test a claim* about a population proportion. If we reject the null hypothesis, we wind up making a statement about what we believe the population proportion *is not*.
- In contrast, we will use a confidence interval to indicate what we believe the population proportion *is*. Put another way, a confidence interval is used to *estimate* the population proportion.

For example, in Section 7.2 of Lesson 7 we considered the proportion of smokers in a certain town. Based on a hypothetical sample in which 215 out of 1100 people were smokers, we came to the following conclusion:

> At significance level 0.05, we conclude that the proportion of smokers in this town is not 16.8%.

Using this same sample, the method of confidence intervals developed in this lesson would allow us to state our conclusions in a positive manner, perhaps something like this:

> According to a recent survey, the proportion of smokers in that town is 19.55%, with a margin of error of ±2.34%.

Observe the confidence interval conclusion that the proportion *is* 19.55%, as contrasted with the hypothesis test conclusion that it *is not* 16.8%.

Comment: We will explain the phrase “margin of error” later in this lesson.

<u>8.1 – Estimating Probabilities</u>

When we toss a coin the probability of getting *heads* is ½, or 0.5, or 50%. This does not mean, however, that if you toss the coin twice, you must get exactly one head and one tail. It does not even mean that you will get 500 heads out of 1000 tosses. What it does mean is this:

> *If the coin is tossed a large number of times, the proportion of heads will be approximately 0.5 = 50%. In addition, the more times the coin is tossed the closer to 50% you can expect the proportion to be.*

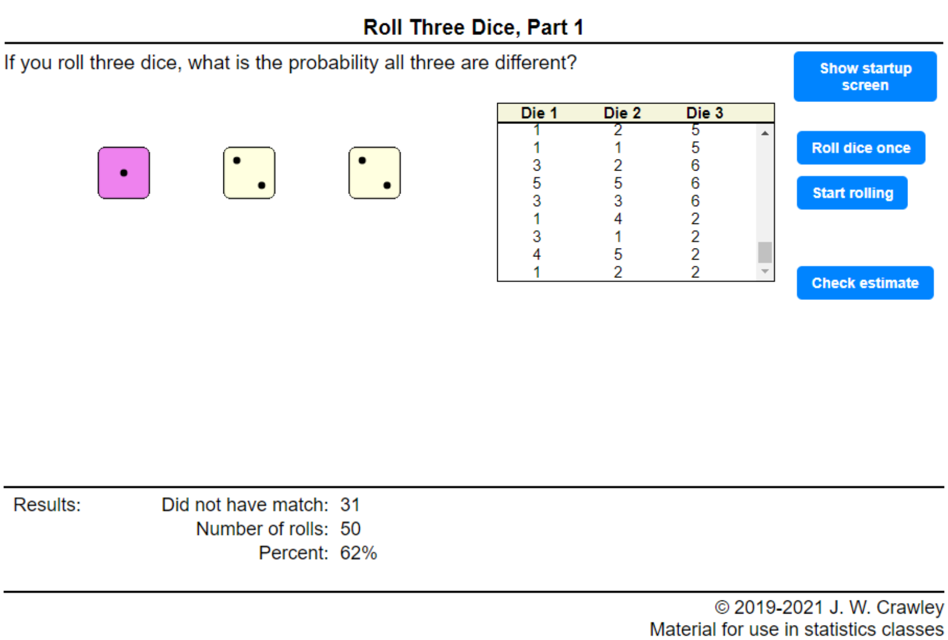
Here is an important consequence of this long-term behavior: *if we were unable to calculate that the theoretical probability is 0.5, we could use the long-term proportion to estimate that probability*. The more times we toss the coin, the closer we can expect our proportion to be to the actual probability of 0.5. This fact is crucial to understanding how confidence intervals can be used to estimate proportions.

An experiment with three dice

Consider the following: roll three dice, and determine the probability that all three are different. Although it is not impossible to calculate the theoretical probability, that calculation is not the subject matter of this course. Our approach is to estimate the probability by experimentation. We can do this one of two ways: roll some actual physical dice, or use the applet "Roll Three Dice, Part 1" which is supplied with this material. Your instructor may ask you to do both – use actual dice, and then use the applet at this link:

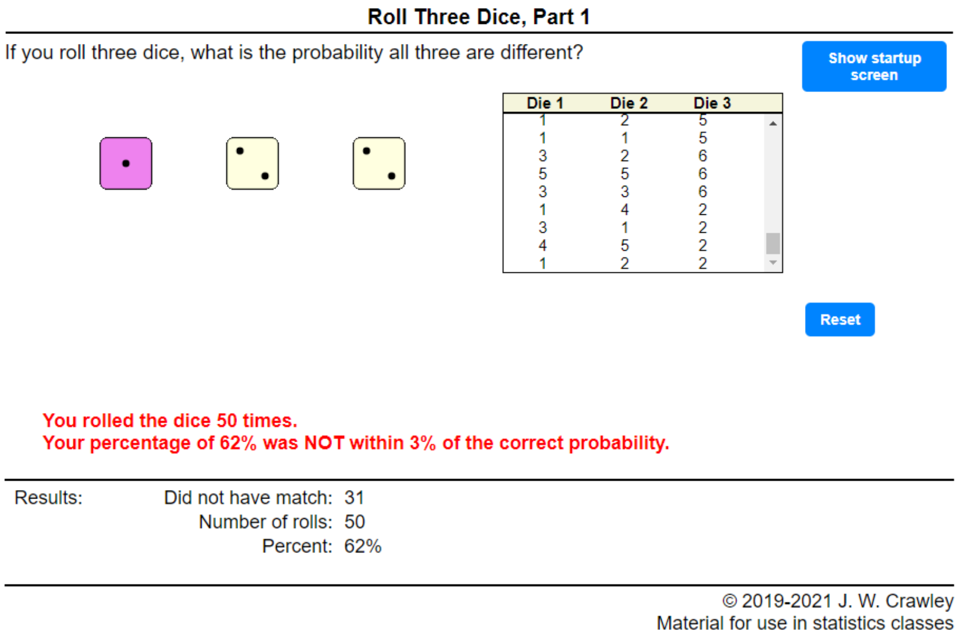
Roll 3 dice, part 1

Here is what you should do. Roll the dice approximately 50 times, keeping track of two things: how many times you rolled the dice, and how many times all three were different. You can either press the "Roll dice once" repeatedly, or use the "Start rolling" and "Stop rolling" to carry out the experiment more quickly. Here are the results obtained by the author:

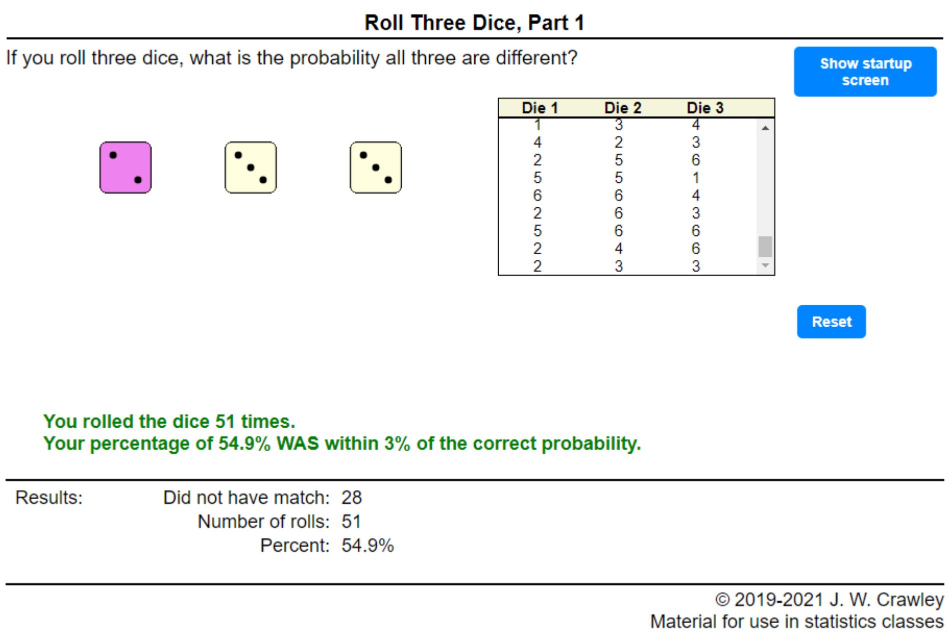


In this experiment, for 31 of the 50 rolls all three dice were different, for a proportion of 31/50 = 0.62 = 62%. Based on this experiment, we would estimate the probability that all three dice will be different as 0.62 or 62%.

When we click the "Check estimate" button in the applet, the applet lets us know how good the estimate was. If the estimate was "close to" the actual probability, we get a message printed in green, otherwise a message printed in red. Any estimate within 3% of the actual probability is considered close in this applet. For our estimate of 62%, here is the result:



When we roll the dice 50 times, as you can see, our estimate may not be very close to the actual probability. However, sometimes it is; here is another experiment using the applet again:



Exercise 1: Use the applet several times, each time rolling the dice about 50 times. Record your results here: the estimate and the answer to the question, "Was the estimate close – within 3% of the actual probability?" The first two are filled in based on the results shown above.

Percent (estimate)	**Close?**
62%	no
54.9%	yes

Margin of error (closeness, correctness, and interval estimates)

In the original example about taxes, the polling organization reported the population proportion as 54%, "with a margin of error $\pm$4.2%." This raises the question:

What exactly does the phrase "margin of error ±4.2%" mean? More generally, what is a *margin of error*?

We are now in a position to answer this question, by thinking about what happened with the dice.

Before we talk about the dice, however, think again about tossing a coin. If you toss a coin 1000 times, do you expect to get *exactly* 500 heads? Not really – but you do expect that the percentage of heads will be close to 50%. This illustrates the following important point: *When we obtain an estimate of a probability by experimentation, we do not expect the estimate to be absolutely correct*.

Because of this fact, estimates obtained by experimentation are often given as ***interval estimates*** rather than just ***point estimates*****.** Let's explain the terminology just a bit. Using the applet, our first estimate for the probability was 62.00%. This is called a *point estimate*, because it consists of a single point on the number line. Now let's use the point estimate to create an interval estimate, as follows:

For the point estimate of 62.00%, we are really saying, "We think the actual probability is close to 62.00%. By 'close' we mean within 3%. That is, we think the probability is between 59.00% and 65.00%.

An interval on the number line consists of numbers between two given numbers. The phrase, "between 59.00% and 65.00%" describes an interval, so it is called an *interval estimate*. This estimate is sometimes written in mathematical interval notation: (59.00%, 65.00%).

What does this have to do with "margin of error"? The answer is simple – when statisticians use the terminology margin of error, they are describing what they mean by "close." Here are three different ways to describe the same estimate:

- 62.00%, with a margin of error of ±3%.
- Between 59.00% and 65.00%.
- In the interval (59.00%, 65.00%).

To calculate the interval estimate, we subtract the margin of error from the point estimate, and add that same margin of error to the point estimate.

Now to a crucial question – should the estimate earn a message printed in green, or one printed in red? There are two equivalent ways to answer this question. To earn a green message, the point estimate must be close (within the margin of error ±3%). This is the same as saying that the interval estimate must be correct, it must contain the actual probability.

The first point estimate, 62.00%, earned a red message. The actual probability *is not* close to (within ±3% of) this estimate. The actual probability *is not* in the interval between 59.00% and 65.00%.

On the other hand, the second point estimate, 54.9%, earned a green message. The actual probability *is* close to (within ±3% of) this estimate. The actual probability *is* in the interval between 51.9% and 57.9%.

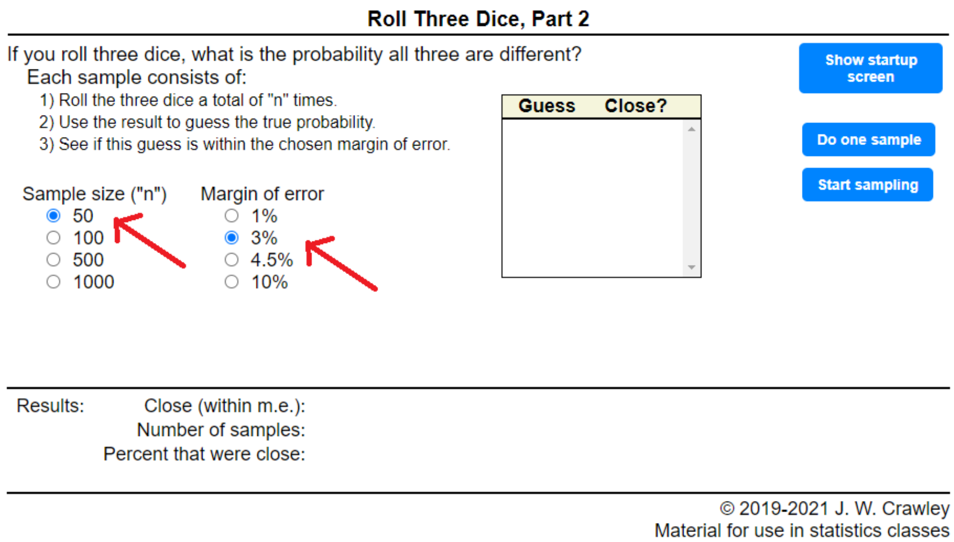
> *You have heard it said that close only counts in horseshoes. However, close also counts in estimating probabilities using experimentation (long-term proportions). An estimate is considered correct if the corresponding interval estimate contains the actual probability. This means that the point estimate is close (within the margin of error).*
>
> *When the interval estimate is correct, instead of saying that the point estimate is close, we sometimes say that the point estimate is "correct within the margin of error." But that's just a fancy way of saying that it is close.*

Confidence level (do you think your estimate is correct??)

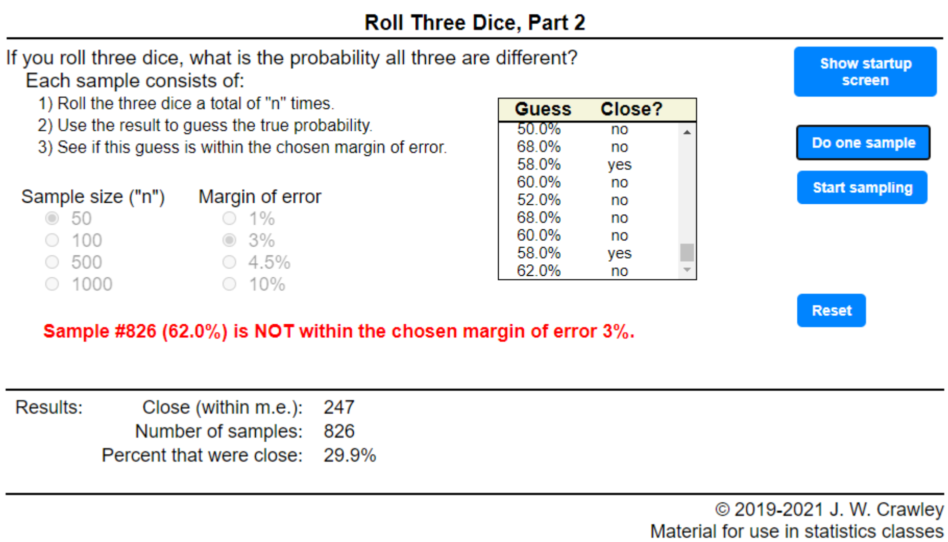
When we roll the dice 50 times, sometimes the point estimate is close – that is, the interval estimate is correct – and sometimes it is not. If you were asked to perform the experiment one more time, again rolling the dice 50 times, how confident would you be that the resulting point estimate would be close? To get a preliminary idea, you might look back at the table above where you recorded your own results. Are you almost certain to be correct? Is it more likely correct than not? Is it about a 50-50 proposition? To get a better idea, you might decide to use the applet several more times. To make that easier for you, we have created a second applet that automatically runs the first applet multiple times.

Roll 3 dice, part 2

As you can see highlighted with the arrows in the screen shot below, this applet is initially set up to repeat exactly the experiment you did – 50 rolls of the dice, and a margin of error of 3%. But you can change those numbers if you wish.



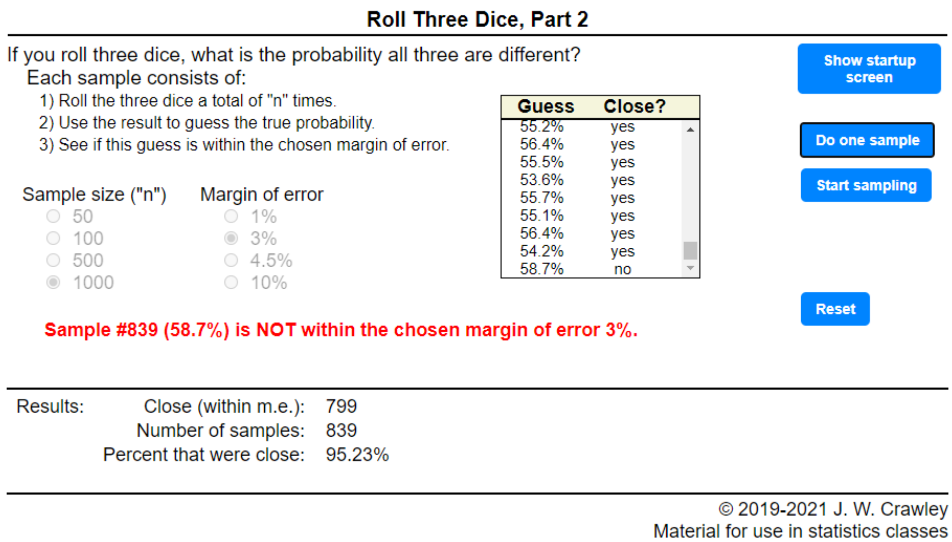
When we ran the applet with these values, here is what we saw after a large number of samples (we did 826 samples, where each sample consisted of 50 rolls of the dice):



It happened that our final sample's proportion (62.0%) was not close, it was not within the margin of error. Put another way, our final interval estimate of (59.0%, 65.0%) was incorrect. In fact, only about 30% of the samples yielded results that were close. In answer to the question, "How confident are you that the next sample will yield a correct answer?" we would say, "Not very confident – it looks like only about 30% of the samples yield correct answers."

Exercise 2. Use the applet to run between 800 and 1000 samples. Record the results here.
Are your results fairly consistent with ours?

Based on these results, it looks like rolling the dice 50 times, and using a margin of error of 3%, isn't a very good strategy, if we would like to get a correct answer most of the time. The proportion obtained in 50 rolls is not very likely to be within 3% of the actual probability. If we think about it, this really shouldn't come as a shock. We might expect to have to roll the dice more than 50 times to begin to get an accurate estimate. The applet allows us to do this. For example, here is what we obtained when we increased the sample size from 50 rolls to 1000 rolls, and ran 839 samples:



It happened that our final sample was not close (correct within the margin of error), but over 95% of the samples were. If we did one more sample, we could be pretty confident that it would be correct. Our ***confidence level*** would be about 95%. We are 95% confident of getting a correct answer because it appears that about 95% of the samples do give a correct answer.

Note: Just as each sample yields only an approximate indication of the probability for the dice, this process yields only an approximate indication of how confident we should be (the confidence level). But since we did a large number of samples, namely 839, it is a pretty good indication.

Exercise 3. Use the applet to run between 800 and 1000 samples, with the sample size set to 1000 and the margin of error still 3%. Record the results here. Are your results fairly consistent with ours?

There is obviously a connection between sample size and confidence level. This should not come as a surprise. If you think about your experience tossing a coin, you should realize that if you just toss a coin a few times the proportion of heads may not be very close to 50%. When you do lots of tosses the proportion is more likely to be close to the 50% theoretical probability. The more we toss the coin, the more confident we are that our proportion will be close to the actual probability. Paraphrased, "increasing the sample size increases the confidence level."

There is also obviously a connection between margin of error and confidence level. The margin of error indicates how close the estimate has to be to be considered correct. If our margin of error is reduced from 3% to 1%, we are less confident that our sample will satisfy the closeness criteria. Conversely, increasing the margin of error to 10% raises our confidence level. In short, "increasing the margin of error increases the confidence level."

Exercise 4: Use the applet several times, experimenting with changes in the sample size and the margin of error. Each time do at least 800 samples. Record your results here. The first two are filled in based on the results shown above.

n	m.e.	# samples	% correct (approximate confidence level)
50	3%	826	29.9%
1000	3%	839	95.23%

Do the results support the discussion about the relationships among sample size, margin of error, and confidence level?

Terminology: We have been discussing interval estimates, and the confidence level associated with those estimates. Because there is always an associated confidence level, these interval estimates are generally called **confidence intervals.**

8.2 – Estimating Population Proportions

We are now ready to apply what we have learned to the process of sampling from a large population. Although there are a few subtle differences, with proper sampling methods these differences are very minor and do not affect our final results.

The key to our discussion is the connection between proportion and probability. For example, consider the question which began this lesson, “Do you think the amount of taxes you pay is fair?” For that question, the population the pollsters were interested in consisted of all Americans age 18 and above. The pollsters wanted to estimate the proportion of that population that would answer “yes” to that question. The basis for what they did is the fact that these two numbers are the same:

- The proportion of Americans, age 18 or above, who would answer “yes.”
- The probability that a randomly selected American, age 18 or above, would answer “yes.”

As a result, the pollsters need to estimate a probability, just as we estimated a probability for the roll of three dice.

How should they proceed? Just as we did for the dice, the method consists of estimating the probability experimentally. To get an estimate for the dice-rolling problem, we repeated the following process 50 times: roll three dice and record the result. To get an estimate for the tax question, the pollsters could repeat this process 50 times: select an American age 18 or above, ask the question, and record the answer. For our process of rolling the dice to have any meaning, the rolling of the dice had to be random. Similarly, for the process of selecting individuals for the survey to have any meaning, the selection must be random.

To illustrate the similarity to the dice rolling situation, we have three applets which are quite similar to the applets you used to investigate dice rolling. In the next subsections we will discuss these applets and analyze what they show us about sampling.

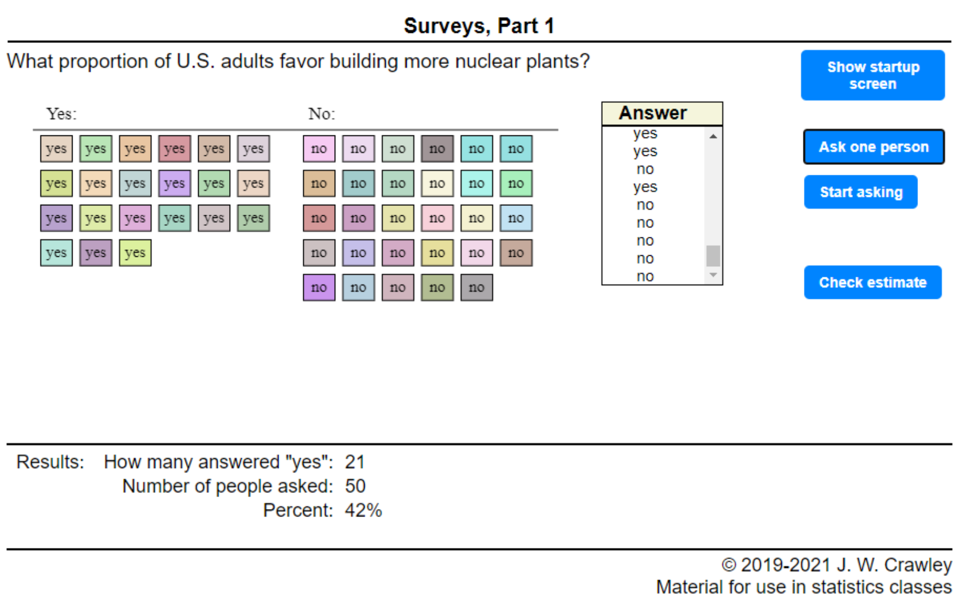
Reminder: Statisticians frequently use the variable p to stand for a population proportion, and the variable $\hat{p}$ to stand for a sample proportion. In terms of these variables, we are trying to estimate p, and we will use $\hat{p}$ as our point estimate.

Surveys, Part 1 – taking a sample to estimate a population proportion *p*

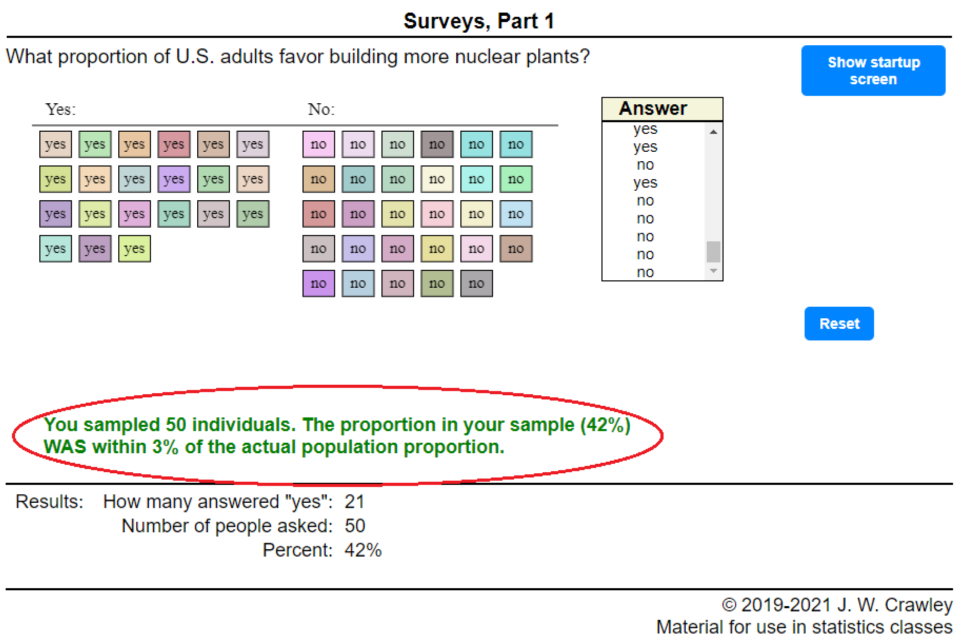
The applet at the following link deals with asking the question, "Should additional nuclear plants be built?"

Surveys, part 1

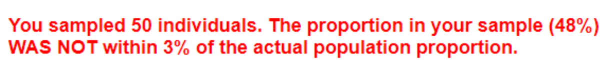
Each time the *Ask one person* button is pressed, the person's answer is shown. You can use that button repeatedly, or the *Start asking* and *Stop asking* buttons, to obtain a sample of whatever size you wish. The following screen shot shows the author's results after 50 persons have been asked the question.



After 50 people have been asked, 21 of those 50 people (42%) have answered "yes." We could stop now, and use that percentage to estimate the proportion p of the ***entire population*** that would answer yes. If we do, the estimate might be close or it might not. Just as in the dice applet, this applet will tell us whether or not the estimate is close, as illustrated below (see the text in the red circle).



Here is the result for another run of the applet, which again asks 50 randomly chosen individuals:



Some observations are in order:

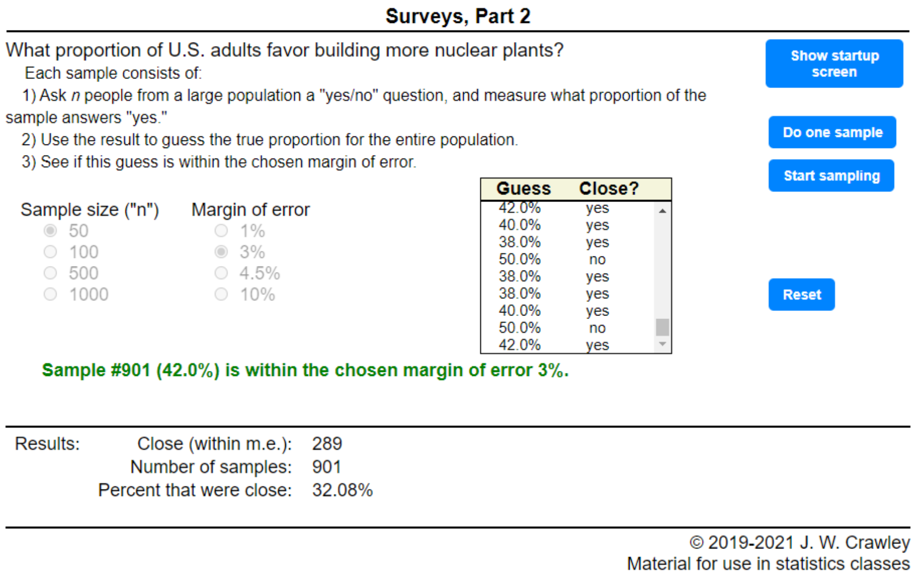
- Just as happened for rolling dice, sometimes the estimate is close and sometimes it is not.
- This applet, like the first dice-rolling applet, uses a margin of error of ±3% as its measure of closeness.
- The point estimates from the two different surveys are 42% for the first, 48% for the second.
- The interval estimates (found by adding and subtracting the margin of error) are (39%, 45%) for the first, and (45%, 51%) for the second.
- The second interval estimate is incorrect – it does *not* contain the actual probability. The first interval estimate is correct.
- Remember that these interval estimates are called *confidence intervals*, because there is always a confidence level lurking in the background – more on this later.

Surveys, Part 2 – confidence level

To get a handle on the relationship between sample size and confidence level (the likelihood that the estimate will be close), you could repeat the previous process indefinitely. However, just as we did for dice rolling, we provide a second applet which carries out that process a large number of times automatically.

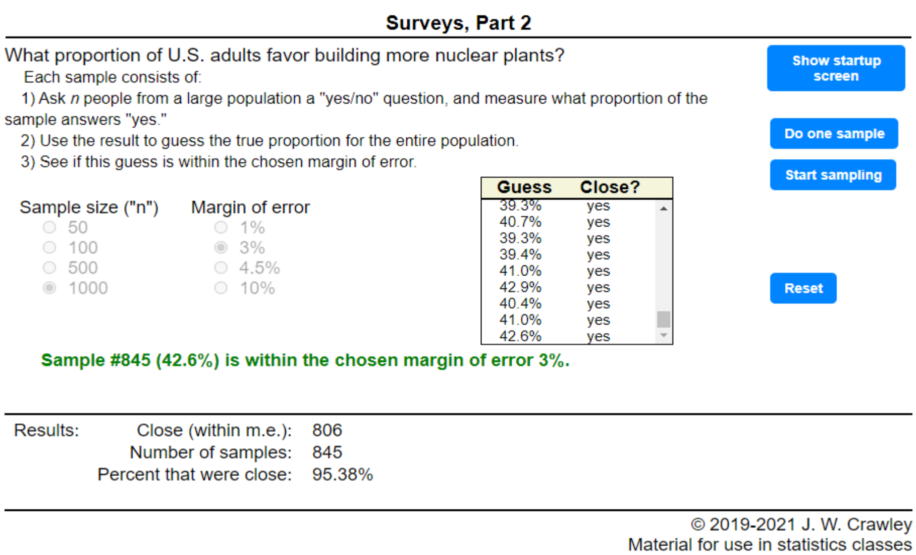
Surveys, part 2

Just as in the dice applet, the default sample size and margin of error are 50 and ±3%, respectively, but you can change those values. The goal is the same – take samples of a given size, and see if the estimate they provide is close to the actual proportion *p*. As for the dice, we let the applet do hundreds of samples (minimum of 800) in order to get a good feel for how confident we should be that the sample yields a close estimate. Here is the result for one run; out of 901 samples, 289 – that is, 32.08% of them – were close (within the chosen margin of error).



Exercise 5. Use the applet to run between 800 and 1000 samples, with n = 50 and m.e. = 3%. Are your results fairly consistent with ours?

Just as for rolling dice, it appears that using sample size 50 and margin of error 3%, isn't a very good strategy, if we would like to get a correct answer most of the time. The confidence level seems to be well below 50%. Just as for the dice, we can improve the results by using a larger sample. Here is an example:



Just as for rolling dice, increasing the sample size increases the confidence level. And, just as for rolling dice, increasing the margin of error increases the confidence level. And, just as happened for the dice, we obtain an apparent confidence level close to 95% by using *n* = 1000 and margin of error = 3%.

Exercise 6: Use the applet several times, experimenting with changes in the sample size and the margin of error. Each time do between 800 and 1000 samples. Record your results here. The first two are filled in based on the results shown above. *Use n = 1000 and m.e. = 3% for several of the runs.*

n	**m.e.**	**# samples**	**% correct (approximate confidence level)**
50	3%	901	32.08%
1000	3%	845	95.38%

Do the results support the discussion about the relationships among sample size, margin of error, and confidence level? Do the results seem to indicate that *n* = 1000 and m.e. = 3% gives a confidence level close to 95%?

8.3 – 95% Confidence Intervals

In this section we will develop a general strategy for calculating confidence intervals for which the confidence level is 95%. This means that we are using a method which gives correct results for 95% of the possible samples. Such an interval is called a ***95% confidence interval***.

Here is a brief summary of ideas related to the example reported at the beginning of the lesson: "According to a recent survey, 54% of Americans age 18 and above believe that the amount of taxes they pay is fair, with a margin of error ±4.2%."

- Researchers wanted to make a statement about the *population* consisting of *all* Americans age 18 and above.
- To do this, they chose a *sample* of people from that population.
- They measured the proportion, $\hat{p}$, for this sample. This proportion is a *statistic*, because it is a fact about the sample. The value of $\hat{p}$ was reported as 54%.
- They used this statistic to make a statement about *p*, the corresponding proportion for the entire population. Because *p* is a fact about the population, it is called a *parameter*.
- The statement consisted of two parts:
  - A *point estimate* – the estimate that *p* is 54%.
  - An *interval estimate*, created by combining the point estimate with a margin of error.

  In this case, they estimated that *p* is 54%, with a margin of error ±4.2%. In other words, they estimated that *p* is somewhere in the interval

  $(54\% - 4.2\%, 54\% + 4.2\%)$, which is equal to $(49.8\%, 58.2\%)$.
- Although it was not reported by the media, the interval is a *95% confidence interval*. The researchers used a method which gets the right answer 95% of the time for the size of sample they used, so the *confidence level* for this interval estimate is 95%.
- The mathematics of what they did is valid, but it depends on obtaining a *simple random sample*. To the extent that the sample was not truly random, the results become less reliable. That is, the results may turn out to be incorrect more than 5% of the time if the sample was not randomly chosen.

The only missing piece in this description is the question of how the researchers arrived at a margin of error of $\pm 4.2\%$. In this section we will develop a strategy for calculating the margin of error, based on the sample size and the sample proportion, which will give us a 95% confidence level. In the next section we will consider calculations for other desired confidence levels.

> **Note:** From using our applets, we have some empirical evidence that for a sample of size 1000, using a margin of error of $\pm 3\%$ yields a confidence level close to 95%. You may want to experiment with the following "Surveys, Part 3" applet for further investigation of this phenomenon.
>
> Surveys, part 3
>
> This applet is identical to the second survey applet, with only one difference. The second applet always samples from a population with population proportion *p* = 0.40; the third uses a different population proportion every time the *Reset* button is chosen.
>
> Our purpose in this section, however, is to develop the calculations that will allow us to achieve 95% confidence for *any size* sample we have.

The sampling distribution

In Lesson 6, we learned that the logic of hypothesis testing for proportions is based on the so-called *sampling distribution of the sample proportions*. It turns out that our calculations for confidence intervals are based on that same concept. We will therefore quickly review a few key ideas from that lesson, and apply them to our current situation.

If the population proportion is *p*, and we sample 1000 randomly chosen individuals, we expect the proportion in the sample to be close to *p* – maybe not exactly *p*, but pretty close. If we repeat the random sampling process over and over, sometimes it will be very close, sometimes not so close. In addition, it will be very close more often than it will be not so close.

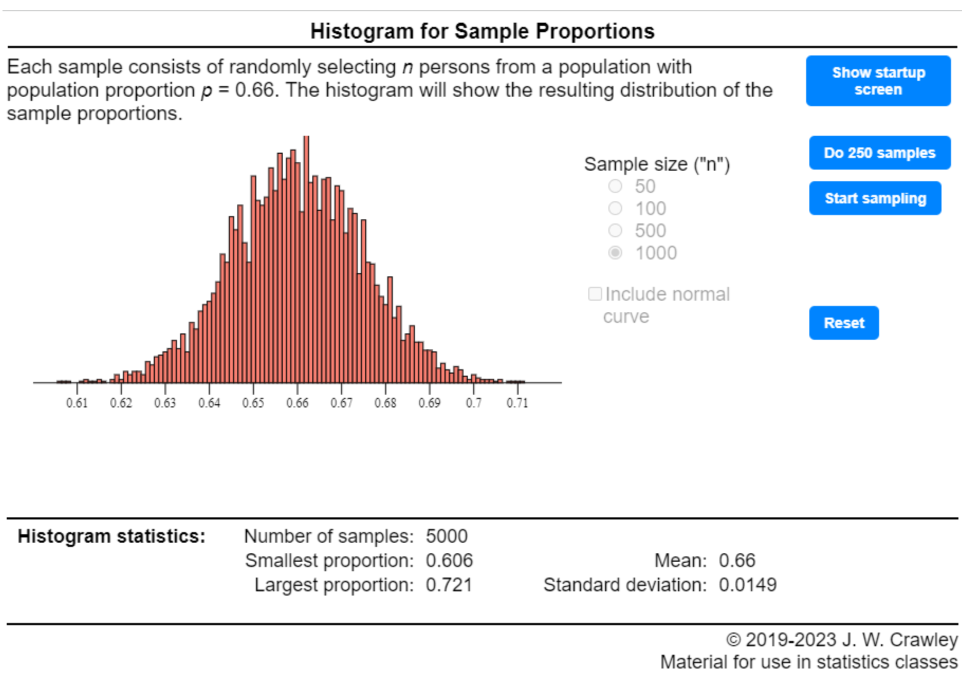
Reminder: The proportion we measure in the sample is called *p*-hat, written $\hat{p}$. Remember that *p*, without the "hat," indicates the proportion in the entire population.

The link below runs an applet that illustrates this. We have used the same applet[1] earlier, in Lesson 7. In the applet you select a sample size *n*, then click on the "Do 250 samples" button to take 250 samples from a population with a particular population proportion *p*. For each sample it calculates the sample proportion, $\hat{p}$, creating a histogram of all these $\hat{p}$ values. You can use the controls to add more samples to the graph, start over with the same sample size, or change the sample size while starting over.

Histogram of p-hat values

NOTE: When you click on the link, the app will choose a population proportion; for the author's output shown below, the choice was 0.66, but for you it will be different. It will also be different every time you restart the app by clicking on the link. Running the app multiple times will allow you to see that the results are similar, no matter what the population proportion might be.

Here is the result of a run by the author, in which he generated 5000 samples from a population whose population proportion *p* is 0.66.



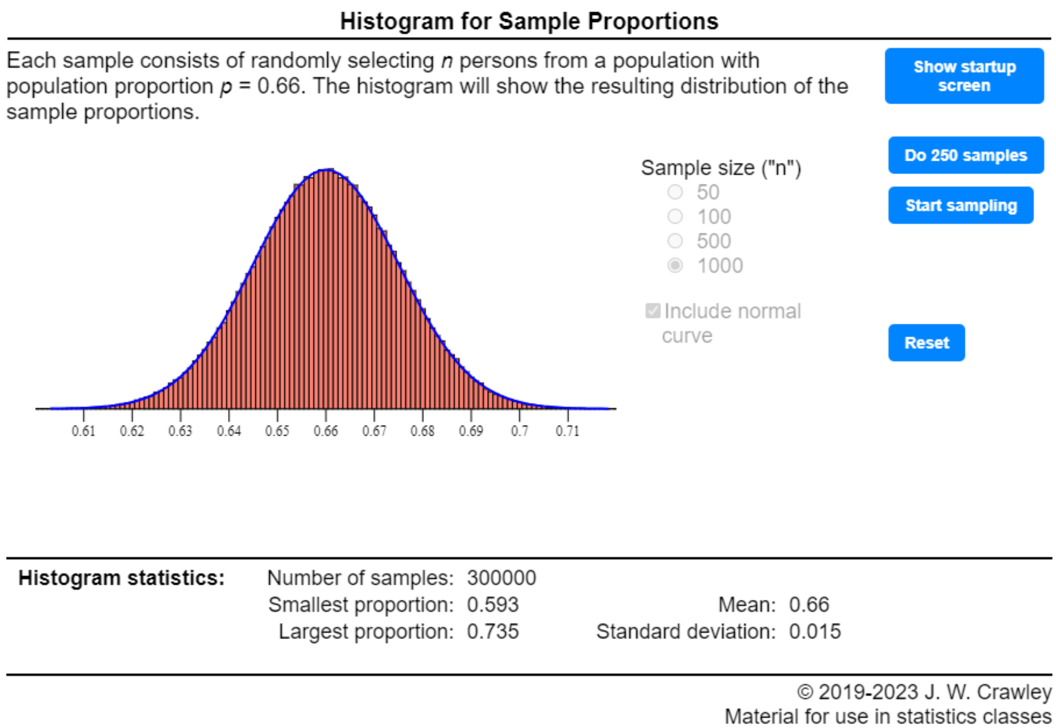
[1] As in earlier lessons, we actually have two versions of the applet. In the first, the scale for the histograms is based on the range of values in the histogram, so the graphs for *n* = 50, 100, and so on, look quite similar. In the second, at the following link, all the graphs are drawn on the same scale. Using the second applet will emphasize how much closer the various sample proportions are to the mean when the sample size is larger. Here is the link:

Histogram of p-hat values

As we learned in Lesson 6, statisticians call the histograms you have viewed, or more precisely the theoretical histogram containing the $\hat{p}$ values for *all possible* samples of the chosen size, the **sampling distribution of the sample proportions.** We also learned the following facts about the sampling distribution:

- It is mound-shaped; indeed, if n is large enough, it is approximately normal.
- The mean is equal to the population proportion p
- The standard deviation is given by the formula $\sqrt{\frac{p(1-p)}{n}}$. As we noted in Lesson 6, for sampling distributions the standard deviation is referred to as the standard error.

For example, if p is 0.66 and n is 1000, the sampling distribution will be approximately normal, with mean = 0.66 and standard deviation (standard error) = $\sqrt{\frac{0.66(1-0.66)}{1000}}$ = 0.0150 = 1.5%. Note that the theoretical results for the sampling distribution of all possible sample proportions are quite consistent with our results for 5000 randomly generated sample proportions (mound shaped, mean 0.66, standard deviation 0.0149).. Here is another run of the applet in which the author chose the option to overlay a normal curve with the same mean and standard deviation, then generated 300,000 samples.



The author then changed the sample size to 50 and generated 50,000 samples, with these results:

Histogram statistics:	Number of samples: 50000		
	Smallest proportion: 0.36	Mean:	0.6604
	Largest proportion: 0.9	Standard deviation:	0.0669

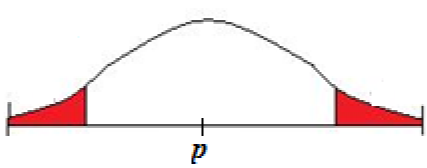
Once again, the mean (0.6604) is quite close to the population proportion (0.66), but the data is much more spread out, as reflected in the much smaller "smallest proportion", the much larger "largest proportion", and the much larger standard deviation. For n = 50, the standard error for the sampling distribution is $\sqrt{\frac{0.66(1-0.66)}{50}}$ = 0.0670, and again our histogram for 50,000 samples closely matches that value.

Exercise 7. Use the applet to experiment. In particular, for each of the sample sizes (50, 100, 500, 1000) use the start sampling / stop sampling buttons to generate about 100,000 samples with an overlaid normal curve. For each sample size, record your answers to these questions:

- **a.** Does the histogram match the normal curve fairly closely?
- **b.** What is the mean for the histogram?
- **c.** What is the standard deviation?
- **d.** Calculate the mean and standard error for the sampling distribution of sample proportions. Do your answers in (b) and (c) match these values fairly closely?

Using the sampling distribution to build a confidence interval

The diagram below illustrates the general situation. The mean for the *sampling distribution of sample proportions* is the population proportion *p*. Each value in the histogram represents a possible sample proportion. We have shaded the 5% of the data for which the sample proportion is the furthest from the correct value.



Red shading indicates the 5% of data furthest from the mean.

Our strategy will be to build our confidence interval to achieve the following:

- Those samples for which the sample proportion lies in the red-shaded area of the histogram give an incorrect answer; and
- Those lying in the non-shaded area give a correct answer.

In this way, 95% of the possible samples will give a correct answer, and 5% an incorrect answer – that is, our confidence level will be 95%.

From the empirical rule, for any mound-shaped distribution, we know that approximately the middle 95% of the data lies within 2 standard deviations of the mean. So a "ball-park" strategy would be to use 2 standard deviations (2 standard errors) as our margin of error. Any sample within 2 standard deviations will yield a correct answer, and this is approximately 95% of the samples. However, we can refine this strategy just a bit.

Provided the sample is large enough[2], the sampling distribution of the sample proportions is not just mound shaped but indeed approximately normal. This means that we can use properties of a normal distribution to calculate a good value for the margin of error. To get started, let's do a very quick review of some ideas first discussed in Lesson 4.

Recall that a *z*-score is simply a measure of the distance from a data item to the mean, measured in terms of standard deviations. In terms of *z*-scores, the empirical rule states that for any mound-shaped distribution:

- Approximately 95% of the data has a *z*-score between –2 and +2.

In Lesson 4 we learned that we can be more precise when we have a normal distribution. We can use Table A to find a *z*-score with the property that 2.5% of the area lies to the left of that *z*-score. Here is the pertinent part of Table A:

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
⋮										
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294

We see that the *z*-score –1.96 has 2.5% of the area to the left of that *z*-score. By symmetry, 2.5% of the area is to the right of $z = 1.96$. Therefore, for a normal distribution we can state that:

- 95% of the data has a *z*-score between –1.96 and +1.96

Based on this fact, we will calculate the margin of error as 1.96 standard deviations, that is 1.96 standard errors = $1.96 \cdot \text{se} = 1.96\sqrt{\frac{p(1-p)}{n}}$. If we do this, our interval estimate will be correct 95% of the time. That is, our interval will be a 95% confidence interval.

There is only one problem with this formula. How can we calculate a margin of error using a formula that contains *p*, the proportion for the entire population? We *don't know* the proportion for the entire population. The whole point of taking a sample is to try to obtain an estimate for the unknown population proportion.

What shall we do? Fortunately, if we just do the best we can, the results will still be good. We don't know *p*, the population proportion, but we do have an estimate for it. Let's just use that estimate, $\hat{p}$,

[2] How large is "large enough"? The general rule is that your sample must contain at least 15 *successes* ("yes" answers to the question) and 15 *failures* ("no" answers). In terms of the variables *n* and $\hat{p}$, we can write two conditions:

$n\hat{p} \geq 15$ that is, #successes is at least 15

$n(1-\hat{p}) \geq 15$ that is, #failures is at least 15 the proportion from the sample. So we will calculate the standard error using the formula $se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, which means we will use $1.96 \cdot se = 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ as the margin of error. The good news, as illustrated by the following exercise, is that this doesn't change the result by very much at all – so we will still be correct about 95% of the time.

> **Exercise 8:** Suppose that for a sample of size 850, your sample proportion is $\hat{p}$ = 54%, but the population proportion is actually p = 56.7%.
>
> a. Calculate the margin of error you *should* use for a 95% confidence interval.
>
> b. Do the same calculation but using the sample proportion instead. Is the answer much different?

A note on standard error calculations: The standard error for the sampling distribution is given by the formula $\sqrt{\frac{p(1-p)}{n}}$. When we are calculating confidence intervals, we do not know the value for p (the proportion in the entire population). We do the best we can – we use $\hat{p}$, the proportion in the sample. Our formula becomes $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. The first formula gives the "exact standard error," the second the "approximate standard error." Textbooks differ in how they refer to these quantities. For our purposes, we will just use the generic term "standard error," realizing that if we do not know p we must instead use $\hat{p}$ in the calculation.

Recall that in Lesson 7, we used p_0, the value for p which the null hypotheses claimed was correct. So we have the following brief summary:

- For hypothesis tests, $se = \sqrt{\frac{p_0(1-p_0)}{n}}$
- For confidence intervals, $se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Calculating a 95% confidence interval

Here is a summary of the process we use to calculate a 95% confidence interval. Our goal is to estimate a population proportion p, by taking a random sample from the population. Let n indicate the size of the sample, and let $\hat{p}$ be the proportion for the sample.

- The point estimate is $\hat{p}$.
- We calculate the (approximate) standard error using $se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.
- We calculate the margin of error using $m.e. = 1.96 \cdot se$
- We can write the interval as $(\hat{p} - m.e., \hat{p} + m.e.)$ which is the same as $(\hat{p} - 1.96\ se, \hat{p}\ + 1.96\ se)$.
- This is a 95% confidence interval. That is, the method gives a correct answer for 95% of the possible samples.

- To use this method, n must be "large enough":

$n\hat{p} \geq 15$ — that is, #successes is at least 15
$n(1-\hat{p}) \geq 15$ — that is, #failures is at least 15

Example: Give a point estimate and a 95% confidence interval for a population proportion, based on a sample in which 753 of the 1077 people in the sample respond yes to the question posed by the researchers.

Solution: The point estimate is the proportion we obtain from the sample, that is, the sample proportion:

$$\hat{p} = \frac{count\ for\ sample}{sample\ size} = \frac{753}{1077} = 0.6992$$

Our best estimate is that, if we asked the same question to the entire population, 69.92% of them would answer yes.

Because this is only an estimate, we also want to calculate the margin of error. This is a two step process. First we calculate the standard error:

$$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.6992(1-0.6992)}{1077}} = 0.0140$$

Then we use the standard error to calculate the margin of error:

$$m.e. = 1.96 \cdot se = 1.96(0.0140) = 0.0274$$

We could report the results as 69.92%, with a margin of error of 2.74%. Or we could calculate the confidence interval by subtracting and adding the margin of error to the sample proportion:

(0.6992 – 0.0274, 0.6992 + 0.0274) = (0.6718, 0.7266)

Writing this as percents, we would write (67.18%, 72.66%). We are 95% confident that between 67.18% and 72.66% of the entire population would answer yes to the question.

Note: To justify the use of the normal distribution, both the number of "successes" and the number of "failures" must be at least 15. The number of successes is the number of yes answers, namely 753. The number of failures is a count of the others in the sample, 1077 – 753 = 324. Both are certainly much larger than 15, so the method is justified.

Exercise 9: Estimate the population proportion (using a point estimate and a 95% confidence interval) for each of the following.

a. In a random sample from the population, 382 of the 779 people answer yes to the question posed.

b. The researchers obtain 519 yes answers and 982 no answers from the sample.

The applet at the following link gives additional practice calculating point estimates and 95% confidence intervals.

95% confidence intervals

8.4 – Interpretation for Confidence Intervals

Stating the results, in the context of the problem

There is another important issue to be addressed: *How do we report the results?* We have already seen one way, using the margin of error terminology:

> According to a recent survey, 54% of Americans age 18 and above believe that the amount of taxes they pay is fair, with a margin of error ±4.2%.

This is the method usually used in the media. One problem with this way of reporting the result is that it does not mention the confidence level. Here is a reporting method that does contain the confidence level:

> We are 95% confident that between 49.8% and 58.2% of all Americans age 18 and above believe that the amount of taxes they pay is fair.

This method gives both the interval and the confidence level. We have included the word "all" to emphasize that this is a statement about the entire population, not just about the people in the sample. In your own work, it is a good idea to include the word "all" to emphasize exactly what the population is. Here is a general template that can be used:

> We are 95% confident that between _______% and _______% of _____describe the population_____ _______describe what a "yes" answer to the question indicates

Example. In Section 7.2 of Lesson 7 we considered the proportion of smokers in a certain town. Based on a hypothetical sample in which 215 out of 1100 people were smokers, we can calculate a 95% confidence interval using these steps:

$\hat{p} = \frac{215}{1100} = 0.1955$

$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.1955(1-0.1955)}{1100}} = 0.0120$

$m.e. = 1.96 \cdot se = 1.96(0.0120) = 0.0235$

(0.1955 – 0.0235, 0.1955 + 0.0235) = (0.1720, 0.2190) = (17.2%, 21.9%)

These results could be reported using "margin of error" terminology as:

> According to a recent survey, the proportion of smokers in that town is 19.55%, with a margin of error of ±2.35%.

To include the confidence level and write the result as an interval we might report the results as

> We are 95% confident that between 17.2% and 21.9% of all the people in that town are smokers.

Interpreting reported population proportions

When an estimate of a population proportion is reported, there are additional things to keep in mind. For example, consider our primary example: "54% of Americans age 18 and above believe that the amount of taxes they pay is fair, with a margin of error ±4.2%." This simple sentence conveys a lot of unstated information.

- 54% *of the people surveyed* answered yes.
- Although I didn't ask everyone in the entire country, I'm using the result from my survey to *estimate* that for the entire country 54% would also answer yes.
- I *think* the actual value is close to this estimate – no further away than ±4.2%. Put another way, I *think* the actual value is somewhere between 49.8% and 58.2%.

The news media do not always give the margin of error when they report a population proportion. They might simply state something similar to this: A recent survey showed that 43% of all Republicans favor Candidate X. When you hear such a report, you should always keep in mind that there is a margin of error implicit in the results.

There is even more to the story, although the news media almost never reports this part. There is the confidence level. For our example, if you go back to the original report you find a confidence level of 95% reported, although that didn't make it into the news media's version. So there is also this:

- When I say I *think* the actual value is somewhere between 49.8% and 58.2%, I am 95% confident that my result is correct. This is because I used a method that gives a correct answer 95% of the time.
- Since the method gives a correct answer 95% of the time, this means that 5% of the time the method yields an incorrect answer. So 5% of the time the method yields an estimate that is *not* within ±4.2% of the actual population proportion.

To the extent that the sampling method was indeed random, and provided we correctly interpret the results, the methodology is sound. But we should be careful not to hear only the "54%" in the report, but also the rest of the story.

"Plausible" values

In the example, when we indicate the confidence interval as $(49.8\%, 58.2\%)$, we can think of the numbers in the interval as the "most plausible" or "most believable" possible values for the population proportion. That is not to say that the actual proportion *cannot* be outside the interval – after all, this could be one of those intervals which gets an incorrect answer. However, since this is a 95% confidence interval, only 5% of all the possible intervals would give an incorrect answer.

As a result, if we ask ourselves the question, "Based on this interval, is it plausible (believable, feasible) that the population proportion could be 49%?" our answer will be, "No." *Based on the interval*, we believe the value is between 49.8% and 58.2%, and 49% is not between those two numbers.

On the other hand, if we ask, "Based on this interval, is it plausible that the population proportion could be 58%?" the answer will be, "Yes" because 58% is one of the values within the interval.

Example. Researchers select a random sample from a certain population, and ask a yes/no question. In the sample, 513 answer yes and 627 answer no. They report a point estimate and a 95% confidence interval for the proportion of the population that would answer yes. The point estimate is 0.4500 and the confidence interval is $(0.4211, 0.4789)$. Based on these results,

a. Is it plausible that in the entire population 47% of the people would answer yes to the question?
b. Is it plausible that in the entire population 42% of the people would answer yes to the question?
c. Can we conclude that less than half of the population would answer yes to the question?
d. Would it be correct to state that the population proportion is 0.4500?
e. True or false: There is a chance that the researchers have reached an incorrect conclusion, but that chance is not very large.

Solution

a. Because 47% (that is, 0.4700) is one of the numbers in the interval, it is plausible that 47% of the entire population would answer yes.
b. No. Because 42% (that is, 0.4200) is **not** within the interval, it is not a plausible value, based on these results. **Note:** The question of plausibility is based solely on whether the proportion in question is within the interval.
c. Yes – all the plausible proportions in the interval are less than one half.
d. No. It is correct to state that the proportion in the **sample** *is* 0.4500, but that does not imply that the proportion in the entire **population** *is* 0.4500. It would, however, be correct to state that the population proportion *could be* 0.4500, since 0.4500 is one of the plausible values in the interval – but it also *could be* 0.4253, or 0.4788, or in fact any number found within the interval.
e. True. The methodology used to calculate the results yields an interval which does in fact contain the true population proportion 95% of the time (that is the meaning of "95% confidence"). So the results could be incorrect – 5% of the time the methods yield incorrect results. (It is worth noting that generally speaking, even the intervals which "miss" typically do not miss by very much, however.)

Example. A researcher wants to know if college students are in favor of a textbook rental policy. She asks the question to 843 randomly chosen college students and 76% say "yes". Which of the following statements are correct based on this description?

a. The population consists of the 843 randomly chosen college students.
b. Using this sample, we can be very confident that more than half of all college students would say "yes" to this question.
c. The 76% corresponds to the variable $\hat{p}$.
d. The population consists of all college students who favor a textbook rental policy.
e. The population consists of all college students.
f. If another researcher did a similar study with a sample of 843 college students, they would get the same percentage (76%) because the sample is the same size.

Solution

a. Incorrect. It would be correct to state that the **sample** consists of the 843 randomly chosen college students.

b. Correct. Without doing the calculations, we realize that if we calculate a confidence interval all the numbers in the interval will be quite a bit larger than 50%.
c. Correct. The variable $\hat{p}$ refers to the proportion in the sample.
d. Incorrect. See part (e) below.
e. Correct. We have two ways to come to this conclusion. The first sentence of the description indicates the researcher is wonders if *college students* favor the policy, the second sentence indicates that the sample consists of randomly chosen *college students*. In general, if a sample consists of *n* individuals chosen from a certain group, the population is *all* individuals in that same group.
f. Incorrect. If another researcher did a similar study with a sample of 843 college students, we would expect to see similar results (that is, a value reasonably close to 76%), but not necessarily exactly the same percentage.

The following applet provides additional practice in calculating confidence intervals, along with interpreting the results.

Calculations with questions on plausible values

For additional practice in interpreting confidence intervals, you may use these two applets.

Questions on plausible values

Other interpretation questions

8.5 – Other Confidence Levels

It is quite simple to obtain confidence intervals for other confidence levels, if we keep in mind the basis for the 95% confidence interval. In general, the confidence interval will have this form:

$(\hat{p} - m.e., \hat{p} + m.e.)$, where:

- $m.e. = z^* \cdot se$ (for 95% confidence intervals this formula is $m.e. = 1.96 \cdot se$)
- $se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
- z^* is chosen based on the desired confidence level

That is, we form the interval by using the sample proportion as the point estimate, then adding and subtracting a margin of error, $z^* \cdot se$, where z^* is chosen to achieve the desired confidence level. For a normal distribution, 95% of the data lies within 1.96 standard deviations of the mean, so we have been using $z^* = 1.96$.

Now suppose we want a 98% confidence interval. We simply ask this question: "For a normal distribution, for what z^* does 98% of the data lie within z^* standard deviations of the mean?" Put another way, we must fill in the blanks in this statement:

- 98% of the data in a normal distribution has *z*-score between ________ and ________.

We have worked this problem as an exercise back in Lesson 4. There will be 1% of the data to the left of $-z^*$ and 1% to the right of $+z^*$. We look for 1% = 0.0100 in Table A. We do not find this exact value, but the closest we do find is 0.0099, for $z = -2.33$.

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110

So we choose $z^* = 2.33$.

Exercise 10: Give a 94% confidence interval based on a sample for which the sample proportion is 70% and the sample size is 280.

We can use Table A, as illustrated, to determine the appropriate value for z^*. However, for the most common confidence levels there is a shortcut, provided by Table B. This table lists values for z^* to three decimal places, in the bottom row of the table. Here is the pertinent part of Table B:

	Confidence Level							
	80%	90%	95%	96%	98%	99%	99.5%	99.9%
	Right-tail Probability							
df	0.10	0.05	0.025	0.02	0.01	0.005	0.0025	0.0005
1	3.078	6.314	12.706	15.895	31.821	63.657	127.321	636.619
2	1.886	2.920	4.303	4.849	6.965	9.925	14.089	31.599
3	1.638	2.353	3.182	3.482	4.541	5.841	7.453	12.924
⋮								
60	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.460
80	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.416
100	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.390
z*	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.291

Looking at the "Confidence Level" header and the bottom row of the table, we see that z^* should be 1.960 for a 95% confidence interval. For a 98% confidence interval $z^* = 2.326$ (compare this to the rounded value of 2.33 we obtained using Table A).

Example. In a recent survey of 1022 U.S. adults, 500 indicated that they approve of the way the Supreme Court is handling its job. The corresponding 95% confidence interval is (45.86%, 51.99%).

a. Based on the 95% confidence interval, which if any of these conclusions are justified?
   - Less than half the adults in the country approve of the way the Supreme Court is handling its job.
   - More than half the adults in the country approve of the way the Supreme Court is handling its job.

b. Without doing any calculations, would a 99% confidence interval be larger or smaller than this 95% confidence interval?
c. Calculate the 99% confidence interval.
d. Interpret the interval "in the context of the problem."
e. Calculate a 90% confidence interval.

Solution.

a. Since the interval contains plausible values both below and above 50%, neither of the conclusions is justified. The proportion *could be* less than 50%, but it also *could be* greater than 50%.
b. The 99% confidence interval will be larger. There are at least two ways to realize this. First, in order to be more confident in our results, our results will need to include a larger range of plausible values. The smaller the interval, the less confident we become that the true proportion is within the interval. Second, the 99% interval will use a larger $z*$ (2.576) than the $z*$ for the 95% interval (1.96).
c. $\hat{p} = \frac{500}{1022} = 0.4892$

   $$se = \sqrt{\frac{0.4892(1 - 0.4892)}{1022}} = 0.0156$$

   $m.e. = 2.576 * 0.0156 = 0.0402$
   Interval is $(0.4892 - 0.0402, 0.4892 + 0.0402) = (0.4490, 0.5294)$
d. We are 99% confident that for the population consisting of all U.S. adults, between 44.90% and 52.94% approve of the way the Supreme Court is handling its job.
e. The calculations are the same except that we use $z* = 1.645$.
   $m.e. = 1.645 * 0.0156 = 0.0257$
   Interval is $(0.4892 - 0.0257, 0.4892 + 0.0257) = (0.4635, 0.5149)$

Exercise 11: In a recent survey of 1093 people in Pennsylvania, 765 indicated that they favor a tax on companies drilling for natural gas in the state. Give a 99% confidence interval, and interpret the result "in the context of the problem."

The following applet provides additional practice in calculating confidence intervals with varying confidence levels.

Confidence interval calculations

For additional practice in interpreting confidence intervals, you may use these two applets.

Questions on plausible values

Other interpretation questions

8.6 – Requirements, and Warnings

In Lesson 7 we discussed the issue of whether a researcher is justified in making a statement about an entire population, based on a survey of only a sample from that population. In this section, we remind you of that discussion, along with some additional requirements developed in the current lesson.

- Mathematically, the method is based on using a normal distribution. To do so, the sample size *n* must be "large enough." In practice, *n* is large enough provided:
  - $n\hat{p} \geq 15$ that is, the number of *successes* is at least 15
  - $n(1-\hat{p}) \geq 15$ that is, the number of *failures* is at least 15

- When the results are reported, the margin of error should be included (either using the phrase "margin of error" or by giving an interval). The researcher (and the media) should not imply that the point estimate is guaranteed to be precisely accurate.

  The sampling process actually yields an *interval* within which we believe the population proportion lies. This is true whether or not the margin of error is included in the report.

- The researcher should also report the confidence level. (The media typically does not, presumably because either they do not understand it or they believe their audience will not understand it.)

  When you see results reported by the media, you should always keep in mind that the method yields correct answers only ____% of the time, filling in the blank with the confidence level. (We believe that 95% is a commonly-used confidence level when the level is not reported, but if the confidence level is not reported it is impossible to be certain what it is.)

- All the previous discussion is about the mathematical process, based on the assumption that the sample was a *simple random sample*. The remaining warnings relate to potential problems in the sampling process itself. Refer back to Lesson 5 for more information; following are just a few reminders of possible problems.

- In surveying people, it can be very difficult to obtain true randomness. For example, some people refuse to answer surveys, and some people don't have phones – just two of the many obstacles to obtaining a truly random sample.

- A related problem occurs when a sample is taken from one group, but the results are reported for another group. (Imagine a presidential election poll which samples from only one state but reports the results as valid for the entire country! This would obviously be problematic.)

- Results can be skewed by the wording of the question. This can be unintentional or, in the case of unethical pollsters, intentional.

Solutions to Exercises

Some of the exercises have no specific solutions, since the results from running the applet will vary.

2: Use the applet to run between 800 and 1000 samples. Record the results here. Are your results fairly consistent with ours?
Here are the results obtained by the author. Yes, they are fairly consistent with the results reported in the discussion just before the exercise. The percent that were close is a little higher, but still in the vicinity of about 30%.

Results:	Close (within m.e.):	291
	Number of samples:	894
	Percent that were close:	32.55%

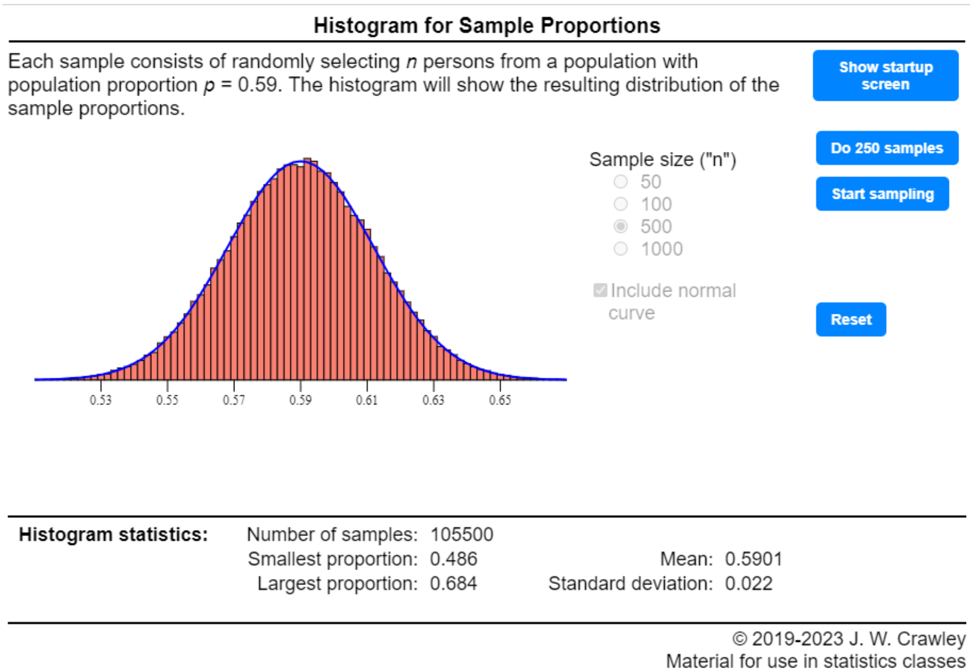
3: Use the applet to run between 800 and 1000 samples, with the sample size set to 1000 and the margin of error still 3%. Record the results here. Are your results fairly consistent with ours?
Here are the results obtained by the author. Yes, they are fairly consistent with the results reported in the discussion just before the exercise. The percent that were close is a little lower, but still pretty close to 95%.

Results:	Close (within m.e.):	782
	Number of samples:	832
	Percent that were close:	93.99%

5: Use the applet to run between 800 and 1000 samples, with n = 50 and m.e. = 3%. Record the results here. Are your results fairly consistent with ours?
Here are the results obtained by the author. Yes, they are fairly consistent with the results reported in the discussion just before the exercise.

Results:	Close (within m.e.):	287
	Number of samples:	852
	Percent that were close:	33.69%

7: Use the applet to experiment. In particular, for each of the sample sizes (50, 100, 500, 1000) use the start sampling / stop sampling buttons to generate about 100,000 samples with an overlaid normal curve.
Here are the results obtained by the author for a run with samples of size 500, with *p* for this run being 0.59.



For each sample size, record your answers to these questions:

a. Does the histogram match the normal curve fairly closely?
Yes

b. What is the mean for the histogram?
0.5901

c. What is the standard deviation?
0.0220

d. Calculate the mean and standard error for the sampling distribution of sample proportions. Do your answers in (b) and (c) match these values fairly closely?
Mean is 0.59 and standard error is 0.0220. Yes.

8: Suppose that for a sample of size 850, your sample proportion is $\hat{p}$ = 54%, but the population proportion is actually p = 56.7%.

a. Calculate the margin of error you *should* use for a 95% confidence interval.

$$1.96\sqrt{\frac{.567(1-.567)}{850}} = 0.0333$$

b. Do the same calculation but using the sample proportion instead. Is the answer much different?

$$1.96\sqrt{\frac{.54(1-.54)}{850}} = 0.0335.$$ The difference is very small, only 0.0002.

9: Estimate the population proportion (using a point estimate and a 95% confidence interval) for each of the following.

a. In a random sample from the population, 382 of the 779 people answer yes to the question posed.

The point estimate is $\hat{p} = \frac{382}{779} = 0.4904$

Here are the calculations for the confidence interval:

$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.4904(1-0.4904)}{779}} = 0.0179$

$m.e. = 1.96 \cdot se = 1.96(0.0179) = 0.0351$

(0.4904 – 0.0351, 0.4904 + 0.0351) = (0.4553, 0.5255) = (45.53%, 52.55%)

b. The researchers obtain 519 yes answers and 982 no answers from the sample.

The sample includes both the yes and the no answers, so $n = 519 + 982 = 1501$. Once we find n, the calculations are the same as for part (a).

$\hat{p} = \frac{519}{1501} = 0.3458$

$se = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.3458(1-0.3458)}{1501}} = 0.0123$

$m.e. = 1.96 \cdot se = 1.96(0.0123) = 0.0241$

(0.3458 – 0.0241, 0.3458 + 0.0241) = (0.3217, 0.3699) = (31.17%, 36.99%)

10: Give a 94% confidence interval based on a sample for which the sample proportion is 70% and the sample size is 280.

$se = \sqrt{\frac{.7(1-.7)}{280}} = 0.0274$

$z^* = 1.88$ (look for 0.0300 in Table A, the closest is 0.0301 for z = –1.88)

$(.7 - 1.88 * 0.0274, .7 + 1.88 * 0.0274) = (.6485, .7515)$

We are 94% confident that the population proportion lies between 64.85% and 75.15%.

11: In a recent survey of 1093 adults in Pennsylvania, 765 indicated that they favor a tax on companies drilling for natural gas in the state. Give a 99% confidence interval, and interpret the result “in the context of the problem.”

$\hat{p} = \frac{765}{1093} = 0.6999$

$se = \sqrt{\frac{0.6999(1-0.6999)}{1093}} = 0.0139$

$z^* = 2.576$ from Table B

m.e. = $z^* * se = 2.576 * 0.0139 = 0.0358$

$(.6999 - 0.0358, .6999 + 0.0358) = (0.6641, 0.7357)$

We are 99% confident that between 66.41% and 73.57% of all the adults in the entire state favors a tax on companies drilling for natural gas in Pennsylvania. It is very clear that far more than half the population of the state is in favor of such a tax.