baby is a girl or is not, and there's no probability involved. A population proportion p is like the baby that has been born—the value of p is fixed, so the confidence interval limits either contain p or do not, and that is why it's incorrect to say that there is a 95% chance that p will fall between values such as 0.828 and 0.872.

A confidence level of 95% tells us that the *process* we are using will, in the long run, result in confidence interval limits that contain the true population proportion 95% of the time. Suppose that the proportion of $p = 0.90$ is the true population proportion of all adults who know what Twitter is. Then the confidence interval obtained from the Pew Research Center poll does not contain the population proportion, because $p = 0.90$ is not between 0.828 and 0.872. Figure 7-1 illustrates this, and it also shows that 19 out of 20 (or 95%) different confidence intervals contain the assumed value of $p = 0.90$. Figure 7-1 is trying to tell the story that with a 95% confidence level, we expect about 19 out of 20 confidence intervals (or 95%) to contain the true value of p.

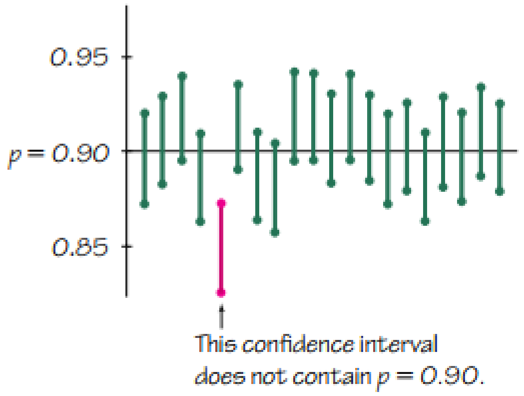

Figure 7-1 Confidence Intervals from 20 Different Samples

CAUTION Confidence intervals can be used informally to compare different data sets, but *the overlapping of confidence intervals should not be used for making formal and final conclusions about equality of proportions.* (See "On Judging the Significance of Differences by Examining the Overlap Between Confidence Intervals," by Schenker and Gentleman, *American Statistician*, Vol. 55, No. 3.)

Using Confidence Intervals for Hypothesis Tests

A confidence interval can be used to *test some claim* made about a population proportion p. Formal methods of hypothesis testing are introduced in Chapter 8, and those methods might require adjustments to confidence intervals that are not described in this chapter.

Some examples and exercises in this chapter require that we address some claim, but in this chapter we do not yet use a formal method of hypothesis testing, so we simply generate a confidence interval and make an *informal judgment* based on the result. For example, if sample results consist of 70 heads in 100 tosses of a coin, the resulting 95% confidence interval of $0.610 < p < 0.790$ can be used to *informally* support a claim that the proportion of heads is *greater than* 50%. (A one-sided claim is a statement that p is *greater than* some value or is *less than* some value, and a formal hypothesis test might require that we construct a one-sided confidence interval, as in Exercise 43, or adjust the confidence level by using 90% instead of 95%.)

BMI a Bad Measure?

The body mass index (BMI) of a person is calculated by finding their weight in kilograms and dividing it by the square of their height in meters. The Centers for Disease Control (CDC) publishes recommendations, including one stating that someone is obese if their BMI is 30.0 or higher. Keith Devlin wrote "Top 10 Reasons Why the BMI is Bogus" for *Your Health*, a podcast by NPR. He argues that, although it is commonly used, the BMI is not a valid measure of fat or obesity. He says that it makes no sense to square a person's height. Also, the BMI makes no adjustment for people with strong bones or good muscle tone. He says that a person with high body fat will have a high BMI, but a person with a high BMI will not necessarily have high body fat. He also says that using the BMI causes us to neglect using more reliable measures of obesity.