

## ESTIMATING WITH CONFIDENCE

Statistical **inference** provides methods for drawing conclusions about a population from sample data.

### SAT Math Scores in California Example 10.2

Suppose you want to estimate the mean SAT Math score for the more than 350,000 high school seniors in California. Only about 49% of California students take the SAT. These self-selected seniors are planning to attend college and so are not representative of all California seniors. You know better than to make inferences about the population based on any sample data. At considerable effort and expense, you give the test to a simple random sample (SRS) of 500 California high school seniors. The mean for your sample is  $\bar{x} = 461$ . What can you say about the mean score  $\mu$  in the population of all 350,000 seniors?

The law of large numbers tells us that the sample mean  $\bar{x}$  from a large SRS will be close to the unknown population mean  $\mu$ . Because  $\bar{x} = 461$ , we guess that  $\mu$  is "somewhere around 461." To make "somewhere around 461" more precise, we ask: *How would the sample mean  $\bar{x}$  vary if we took many samples of 500 seniors from this same population?*

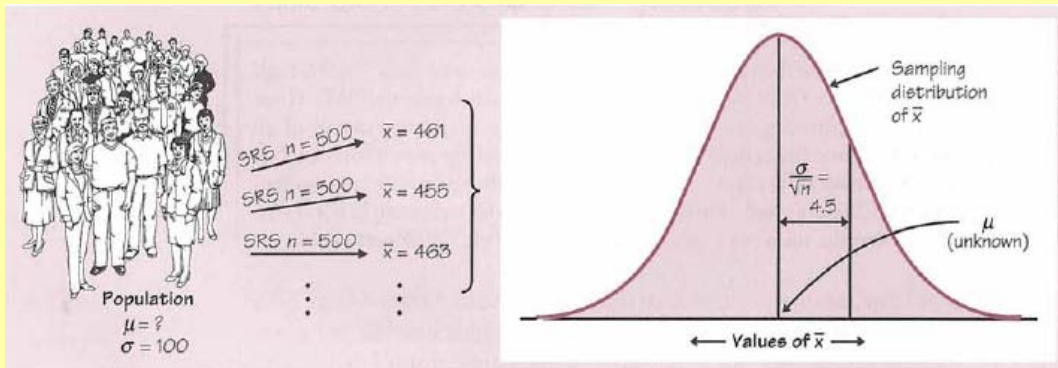
Recall the essential facts about the sampling distribution of  $\bar{x}$ :

- The central limit theorem tells us that the mean  $\bar{x}$  of 500 scores has a distribution that is close to normal.
- The mean of this normal sampling distribution is the same as the unknown mean  $\mu$  of the entire population.
- The standard deviation of  $\bar{x}$  for an SRS of 500 students is  $\frac{\sigma}{\sqrt{500}}$  where  $\sigma$  is the standard deviation of individual SAT Math scores among all California high school seniors.
- Let us suppose that we know that the standard deviation of SAT Math scores in the population of all California seniors is  $\sigma = 100$ . The standard deviation of  $\bar{x}$  is then

$$\frac{\sigma}{\sqrt{n}} = \frac{100}{\sqrt{500}} = 4.5$$

- (It is usually not realistic to assume we know  $\sigma$ . We will see in the next chapter how to proceed when  $\sigma$  is not known. For now, we are more interested in statistical reasoning than in details of realistic methods.)

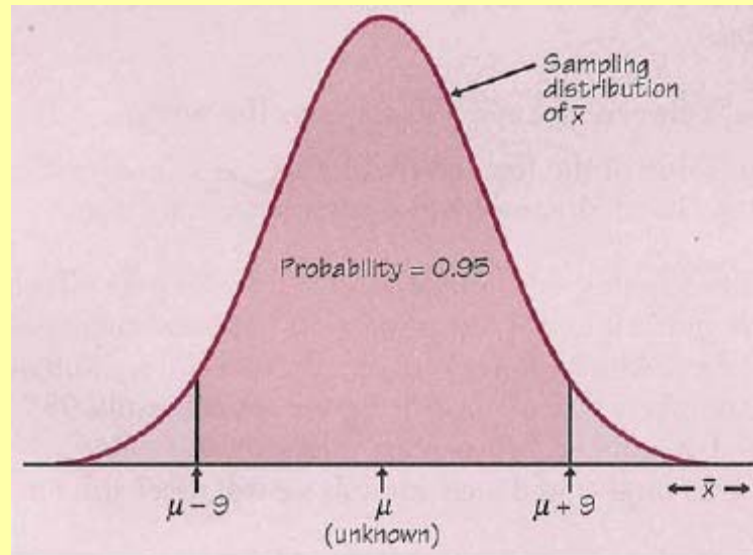
- If we choose many samples of size 500 and find the mean SAT Math score for each sample, we might get mean  $\bar{x} = 461$  from the first sample,  $\bar{x} = 455$  from the second,  $\bar{x} = 463$  from the third sample, and so on. If we collect all these sample means and display their distribution, we get the normal distribution with mean equal to the unknown  $\mu$  and standard deviation 4.5. Inference about the unknown  $\mu$ , starts from this sampling distribution. Figure 10.1 displays the distribution. The different values of  $\bar{x}$  appear along the axis in the figure, and the normal curve shows how probable these values are.



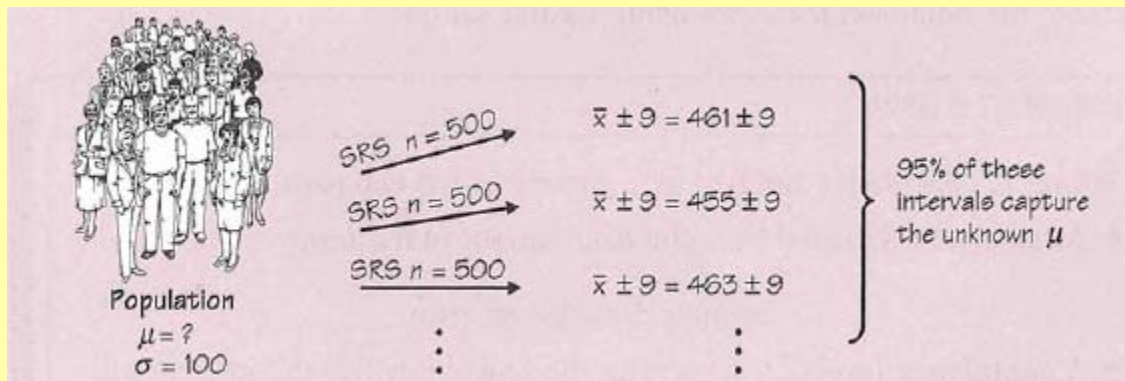
The sampling distribution of the mean score  $\bar{x}$  of SRS of 500 California seniors on the SAT Math quantitative test.

## STATISTICAL CONFIDENCE

- The 68-95-99.7 rule says that in 95% of all samples, the mean score  $\bar{x}$  for the sample will be within two standard deviations of the population mean score  $\mu$ . So the mean of 500 SAT Math scores will be within 9 points of  $\mu$  in 95% of all samples.
- Whenever  $\bar{x}$  is within 9 points of the unknown  $\mu$ ,  $\mu$  is within 9 points of the observed  $\bar{x}$ . This happens in 95% of all samples.
- So in 95% of all samples, the unknown  $\mu$  lies between  $\bar{x} - 9$  and  $\bar{x} + 9$ .



**In 95% of all samples,  $\bar{x}$  lies within  $\pm 9$  of the unknown population mean  $\mu$ .  
So  $\mu$  also lies within  $\pm 9$  of  $\bar{x}$  in those samples.**



**To say that  $\bar{x} \pm 9$  is a 95% confidence interval for the population mean  $\mu$  is to say that in repeated samples, 95% of these intervals capture  $\mu$ .**

**This conclusion just restates a fact about the sampling distribution of  $\bar{x}$ . The language of statistical inference uses this fact about what would happen in the long run to express our confidence in the results of any one sample.**

**Our sample of 500 California seniors gave  $\bar{x} = 461$ . We say that we are 95% confident that the unknown mean SAT Math score for all California high school seniors lies between**

$$\bar{x} - 9 = 461 - 9 = 452$$

**and**

$$\bar{x} + 9 = 461 + 9 = 470$$

**Be sure you understand the grounds for our confidence. There are only two possibilities:**

- 1. The interval between 452 and 470 contains the true  $\mu$**
- 2. Our SRS was one of the few samples for which  $\bar{x}$  is not within 9 points of the true  $\mu$ . Only 5% of all samples give such inaccurate results.**
- 3. We cannot know whether our sample is one of the 95% for which the interval  $\pm 9$  catches  $\mu$ , or if it is one of the unlucky 5%. The statement that we are 95% confident that the unknown  $\mu$  lies between 452 and 470 is shorthand for saying, "We got these numbers by a method that gives correct results 95% of the time."**
- 4. The interval of numbers between the values  $\bar{x} \pm 9$  is called a 95% confidence interval for  $\mu$ . Like most confidence intervals we will meet, this one has the form**

**estimate  $\pm$  margin of error**

- 5. The estimate ( $\bar{x}$  in this case) is our guess for the value of the unknown parameter. The margin of  $\pm 9$  shows how accurate we believe our guess is, based on the variability of the estimate. This is a 95% confidence interval because it catches the unknown  $\mu$  in 95% of all possible samples.**

## CONFIDENCE INTERVALS

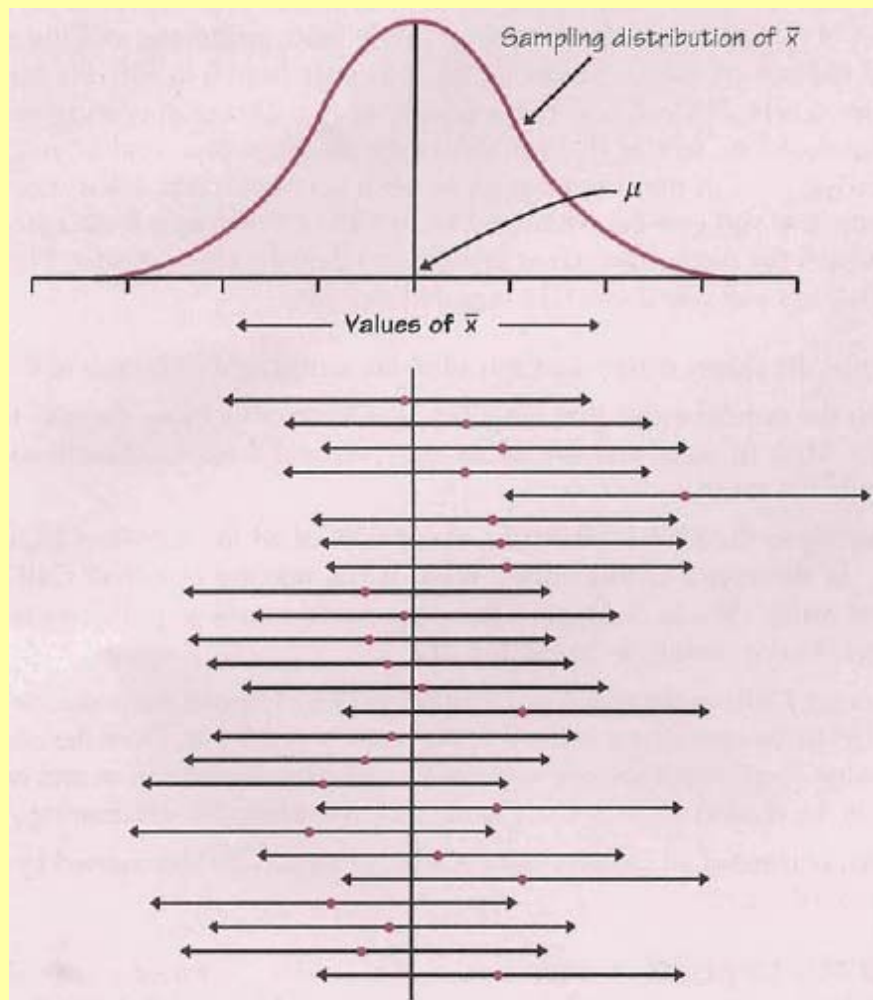
**A level C confidence interval for a parameter has two parts:**

- **An interval calculated from the data, usually of the form**

$$\text{estimate} \pm \text{margin of error}$$

- **A confidence level C, which gives the probability that the interval will capture the true parameter value in repeated samples.**

- The figure below shows the result of drawing many SRSs from the same population and calculating a 95% confidence interval from each sample. The center of each interval is at  $\bar{x}$  and therefore varies from sample to sample.
- The sampling distribution of  $\bar{x}$  appears at the top of the figure to show the long-term pattern of this variation. The 95% confidence intervals from 25 SRSs appear below. The center  $\bar{x}$  of each interval is marked by a dot. The arrows on either side of the dot span the confidence interval.
- All except one of these 25 intervals cover the true value of  $\mu$ . In a very large number of samples, 95% of the confidence intervals would contain  $\mu$ .

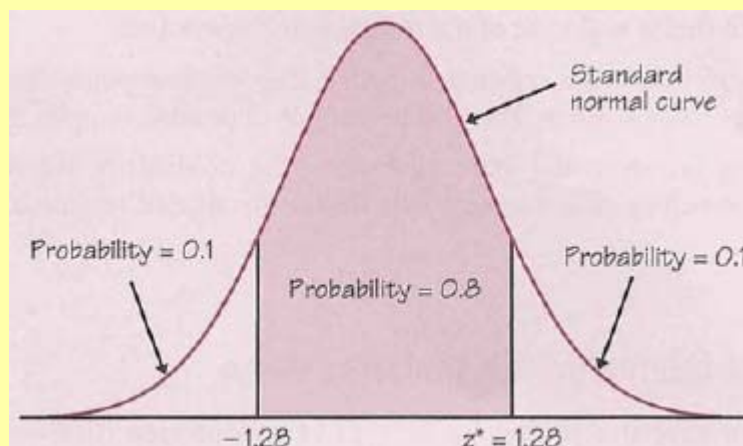


## CONFIDENCE INTERVAL FOR A POPULATION MEAN

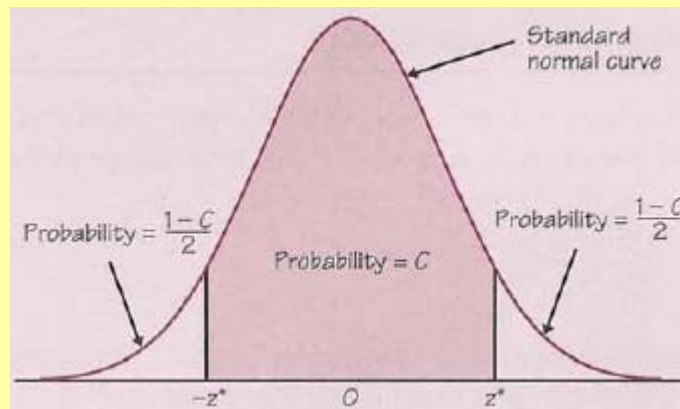
**How to construct a level  $C$  confidence interval for the mean  $\mu$  of a population when the data come from an SRS of size  $n$ .**

- The construction of the interval depends on the fact that the sampling distribution of the sample mean  $\bar{x}$  is at least approximately normal.
- This distribution is exactly normal if the population distribution is normal. When the population distribution is not normal, the central limit theorem tells us that the sampling distribution of  $\bar{x}$  will be approximately normal if  $n$  is sufficiently large.
- Be sure to check that these conditions are satisfied before you construct a confidence interval.
  1. The data must come from an SRS from the population of interest.
  2. The sampling distribution of  $\bar{x}$  is approximately normal.
- To construct a level  $C$  confidence interval, we want to catch the central probability  $C$  under a normal curve.
- To do that, we must go out  $z^*$  standard deviations on either side of the mean.
- Since any normal distribution can be standardized, we can get the value  $z^*$  from the standard normal table. Here is an example of how to find  $z^*$ .

**To find an 80% confidence interval, we must catch the central 80% of the normal sampling distribution of  $\bar{x}$ . So  $z^*$  is the point with area 0.1 to its right (and 0.9 to its left) under the standard normal curve.**



In general, the central probability  $C$  under a standard normal curve lies between  $-z^*$  and  $z^*$ . Because  $z^*$  has area  $\frac{(1-C)}{2}$  to its right under the curve, we call it the upper  $\frac{(1-C)}{2}$  critical value.

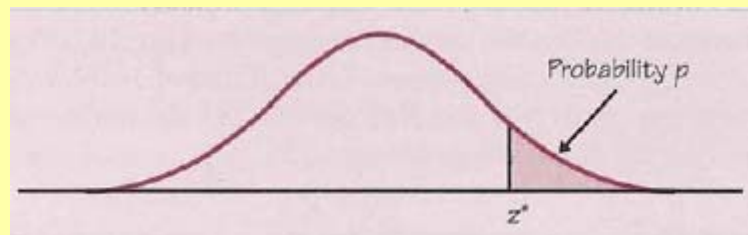


Confidence Level	Tail Area	$Z^*$
90%	0.05	1.645
95%	0.025	1.960
99%	0.005	2.576

Notice that for 95% confidence we use  $z^* = 1.960$ . This is more exact than the approximate value  $z^* = 2$  given by the 68-95-99.7 rule.

### CRITICAL VALUES

The number  $z^*$  with probability  $p$  lying to its right under the standard normal curve is called the upper  $p$  critical value of the standard normal distribution.



Important concepts to keep in mind when thinking on level C confidence intervals:

- Any normal curve has probability C between the point  $z^*$  standard deviations below its mean and the point  $z^*$  standard deviations above its mean.
- The standard deviation of the sampling distribution of  $\bar{x}$  is  $\frac{\sigma}{\sqrt{n}}$ , and its mean is the population mean  $\mu$ . So there is probability C that the observed sample mean  $\bar{x}$  takes a value between

$$\mu - z^* \frac{\sigma}{\sqrt{n}} \quad \text{and} \quad \mu + z^* \frac{\sigma}{\sqrt{n}}$$

- Whenever this happens, the population mean,  $\mu$  is contained between

$$\bar{x} - z^* \frac{\sigma}{\sqrt{n}} \quad \text{and} \quad \bar{x} + z^* \frac{\sigma}{\sqrt{n}}$$

- That is our confidence interval. The estimate of the unknown  $\mu$  is  $\bar{x}$  and the margin of error is  $z^* \frac{\sigma}{\sqrt{n}}$ .

#### CONFIDENCE INTERVAL FOR A POPULATION MEAN

Choose an SRS of size n from a population having unknown mean  $\mu$  and known standard deviation  $\sigma$ . A level C confidence interval for  $\mu$  is

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$$

Here  $z^*$  is the value with area C between  $-z^*$  and  $z^*$  under the standard normal curve. This interval is exact when the population distribution is normal and is approximately correct for large n in other cases.

## VIDEO SCREEN TENSION

### Example 10.5

A manufacturer of high-resolution video terminals must control the tension on the mesh of fine wires that lies behind the surface of the viewing screen. Too much tension will tear the mesh and too little will allow wrinkles. The tension is measured by an electrical device with output readings in millivolts (mV). Some variation is inherent in the production process. Careful study has shown that when the process is operating properly, the standard deviation of the tension readings is  $\sigma = 43$  mV. Here are the tension readings from an SRS of 20 screens from a single day's production.

269.5	297.0	269.6	283.3	304.8	280.4	233.5	257.4	317.5	327.4
264.7	307.7	310.0	343.3	328.1	342.6	338.8	340.1	374.6	336.1

Construct a 90% confidence interval for the mean tension  $\mu$  of all the screens produced on this day.

***Step 1: Identify the population of interest and the parameter you want to draw conclusions about.***

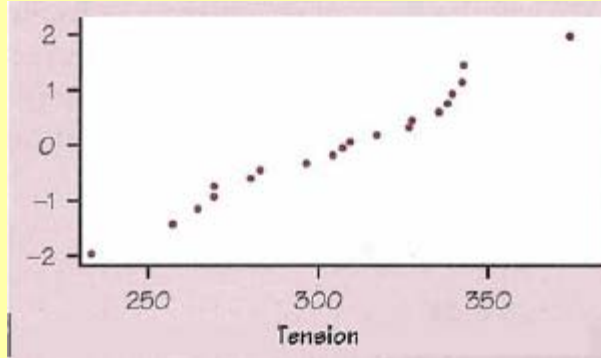
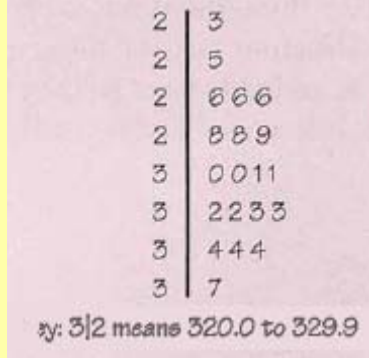
The population of interest is all of the video terminals produced on the day in question. We want to estimate  $\mu$ , the mean tension for all of these screens.

***Step 2: Choose the appropriate inference procedure.***

Verify the conditions for using the selected procedure. Since we know  $\sigma$ , we should use the confidence interval for a population mean that was just introduced to estimate  $\mu$ . Now we must check that the two required conditions-

- (1) SRS from the population of interest and
- (2) sampling distribution of  $\bar{x}$  approximately normal-are met.

- The data come from an SRS of 20 screens from the population of **all screens** produced that day.
- Is the sampling distribution off approximately normal? Past experience suggests that the tension readings of screens produced on a single day follow a normal distribution quite closely. If the population distribution is normal, the sampling distribution of  $\bar{x}$  will be normal. Let's examine the sample data.



A stemplot of the tension readings shows no outliers or strong skewness.

The normal probability plot tells us that the sample data are approximately normally distributed. These data give us no reason to doubt the normality of the population from which they came.

**Step 3: If the conditions are met, carry out the inference procedure.**

You can check that the mean tension reading for the 20 screens in our sample is  $\bar{x} = 306.3$  mV.

The confidence interval formula is  $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$ . For a 90% confidence level, the critical value is  $z^* = 1.645$ . So the 90% confidence interval for  $\mu$  is

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 306.3 \pm 1.645 \frac{43}{\sqrt{20}} = 306.3 \pm 15.8 = (290.5, 322.1)$$

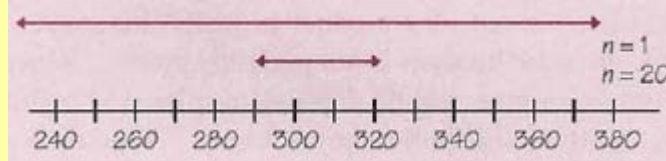
**Step 4: Interpret your results in the context of the problem.**

We are 90% confident that the true mean tension in the entire batch of video terminals produced that day is between 290.5 and 322.1 mV.

Suppose that a single computer screen had a tension reading of 306.3 mV, the same value as the sample mean in Example 10.5. Repeating the calculation with  $n = 1$  shows that the 90% confidence interval based on a single measurement is

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 306.3 \pm 1.645 \frac{43}{\sqrt{1}} = 306.3 \pm 70.7 = (235.6, 377.0)$$

The mean of twenty measurements gives a smaller margin of error and therefore a shorter interval than a single measurement.



Larger samples give shorter intervals.

## HOW CONFIDENCE INTERVALS BEHAVE

The confidence interval  $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$  for the mean of a normal population illustrates several important properties that are shared by all confidence intervals in common use. The user chooses the confidence level, and the margin of error follows from this choice.

- We would like high confidence and also a small margin of error.
- High confidence says that our method almost always gives correct answers.
- A small margin of error says that we have pinned down the parameter quite precisely.
- The margin of error is  $z^* \frac{\sigma}{\sqrt{n}}$
- This expression has  $z^*$  and  $\sigma$  in the numerator and  $\sqrt{n}$  in the denominator.
- So the margin of error gets smaller when
  - a. Getting a smaller  $z^*$ . Smaller  $z^*$  is the same as smaller confidence level  $C$ .
  - b. Increasing the sample size  $n$ .
  - c. Getting a smaller  $\sigma$ .
- There is a trade-off between the confidence level and the margin of error.
- To obtain a smaller margin of error from the same data, you must be willing to accept lower confidence.
- The standard deviation  $\sigma$  measures the variation in the population. You can think of the variation among individuals in the population as noise that obscures the average value  $\mu$ . It is easier to pin down  $\mu$  when  $\sigma$  is small.
- Increasing the sample size  $n$  reduces the margin of error for any fixed confidence level. Because  $n$  appears under a square root sign, we must take four times as many observations in order to cut the margin of error in half.

## CHANGING THE CONFIDENCE LEVEL

### Example 10.6

Suppose that the manufacturer in Example 10.5 wants 99% confidence rather than 90%. Table C gives the critical value for 99% confidence as  $z^* = 2.575$ . The 99% confidence interval for  $\mu$  based on an SRS of 20 video monitors with mean  $\bar{x} = 306.3$  is

$$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}} = 306.3 \pm 2.575 \frac{43}{\sqrt{20}} = 306.3 \pm 24.8 = (281.5, 331.1)$$

Demanding 99% confidence instead of 90% confidence has increased the margin of error from 15.8 to 24.8.



## CHOOSING THE SAMPLE SIZE

### Example 10.7

Company management wants a report of the mean screen tension for the day's production accurate to within  $\pm 5$  mV with 95% confidence. How large a sample of video monitors must be measured to comply with this request?

For 95% confidence, Table C gives  $z^* = 1.96$ . We know that  $\sigma = 43$ . Set the margin of error to be at most 5:

$$m \leq 5$$

$$z^* \frac{\sigma}{\sqrt{n}} \leq 5$$

$$1.96 \frac{43}{\sqrt{n}} \leq 5$$
$$\sqrt{n} \geq \frac{(1.96)(43)}{5}$$

$$\sqrt{n} \geq 16.856$$

$$n \geq 284.125 \text{ so take } n = 285$$

### **SAMPLE SIZE FOR DESIRED MARGIN OF ERROR**

**To determine the sample size  $n$  that will yield a confidence interval for a population mean with a specified margin of error  $m$ , set the expression for the margin of error to be less than or equal to  $m$  and solve for  $n$ :**

$$z^* \frac{\sigma}{\sqrt{n}} \leq m$$

### **Some cautions**

Any formula for inference is correct only in specific circumstances. If statistical procedures carried warning labels like those on drugs, most inference methods would have long labels

indeed. Our handy formula  $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$  for estimating a normal mean comes with the

following list of **warning Labels**

- CI are completely safe when the data is SRS. The margin of error in a confidence interval covers only random sampling errors.
- A CI margin of error is obtained from the sampling distribution and indicates how much error can be expected because of chance variation in randomized data production.
- Practical difficulties, such as undercoverage and nonresponse in a sample survey, can cause additional errors that may be larger than the random sampling error.
- Using CI is not in great danger if data can plausibly be thought of as observations taken at random from a population.
- The formula for CI is not correct if the data was collected with sampling designs more complex than an SRS.
- Multistage or stratified samples are not present in this inference procedure.
- There is no correct method for inference from data collected haphazardly with bias of unknown size.
- Because  $\bar{x}$  is strongly influenced by a few extreme observations, outliers can have a large effect on the confidence interval. Try to correct them or justify their removal before computing the interval. If the outlier cannot be removed, ask your statistical consultant about procedures that are not sensitive to outliers.
- If the sample size is too small and the population is not normal, the true confidence level will be different from the value C used in computing the interval. Examine your data carefully for skewness and other signs of non-normality. The interval relies only on the distribution of  $\bar{x}$ , which even for quite small sample sizes is much closer to normal than the individual observations. When  $n \geq 15$ , the confidence interval is not greatly disturbed by nonnormal populations unless extreme outliers or quite strong skewness are present.
- You must know the standard deviation  $\sigma$  of the population. This unrealistic requirement renders the interval  $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$  of little use in **statistical practice**.
- We will learn in the next chapter what to do when  $\mu$  is unknown. However, if the sample is large, the sample standard deviation  $s$  will be close to the unknown  $\sigma$ . Then  $\bar{x} \pm z^* \frac{s}{\sqrt{n}}$  is an approximate confidence interval for  $\mu$ .

Every inference procedure that we will meet has its own list of warnings. It is easy to state (from the mathematics of probability) conditions under which a method of inference is exactly correct. These conditions are never fully met in practice.

For example, no population is exactly normal. Deciding when a statistical procedure should be used in practice often requires judgment assisted by exploratory analysis of the data.

#### What you can say

- We are 95% confident that the mean SAT Math score for all California high school seniors lies between 452 and 470.
- That is, these numbers were calculated by a method that gives correct results in 95% of all possible samples.

#### What you can not say

We cannot say that the probability is 95% that the true mean falls between 452 and 470.

- No randomness remains after we draw one particular sample and get from it one particular interval.
- The true mean either is or is not between 452 and 470.
- The probability calculations of standard statistical inference describe how often the method gives correct answers.