

## 7.4: Sample Size Considerations

### Learning Objectives

- To learn how to apply formulas for estimating the size sample that will be needed in order to construct a confidence interval for a population mean or proportion that meets given criteria.

Sampling is typically done with a set of clear objectives in mind. For example, an economist might wish to estimate the mean yearly income of workers in a particular industry at 90% confidence and to within \$500. Since sampling costs time, effort, and money, it would be useful to be able to estimate the smallest size sample that is likely to meet these criteria.

### Estimating $\mu$

The confidence interval formulas for estimating a population mean  $\mu$  have the form  $\bar{x} \pm E$ . When the population standard deviation  $\sigma$  is known,

$$E = \frac{z_{\alpha/2}\sigma}{\sqrt{n}}$$

The number  $z_{\alpha/2}$  is determined by the desired level of confidence. To say that we wish to estimate the mean to within a certain number of units means that we want the margin of error  $E$  to be no larger than that number. Thus we obtain the minimum sample size needed by solving the displayed equation for  $n$ .

### Minimum Sample Size for Estimating a Population Mean

The estimated minimum sample size  $n$  needed to estimate a population mean  $\mu$  to within  $E$  units at  $100(1 - \alpha)\%$  confidence is

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} \text{ (rounded up)} \quad (7.4.1)$$

To apply Equation 7.4.1, we must have prior knowledge of the population in order to have an estimate of its standard deviation  $\sigma$ . In all the examples and exercises the population standard deviation will be given.

#### ✓ Example 7.4.1

Find the minimum sample size necessary to construct a 99% confidence interval for  $\mu$  with a margin of error  $E = 0.2$ . Assume that the population standard deviation is  $\sigma = 1.3$ .

#### Solution

Confidence level 99% means that  $\alpha = 1 - 0.99 = 0.01$  so  $\alpha/2 = 0.005$ . From the last line of Figure 7.1.6 we obtain  $z_{0.005} = 2.576$ . Thus

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} = \frac{(2.576)^2 (1.3)^2}{(0.2)^2} = 280.361536$$

which we round up to 281, since it is impossible to take a fractional observation.

#### ✓ Example 7.4.2

An economist wishes to estimate, with a 95% confidence interval, the yearly income of welders with at least five years experience to within \$1,000. He estimates that the range of incomes is no more than \$24,000, so using the Empirical Rule he estimates the population standard deviation to be about one-sixth as much, or about \$4,000. Find the estimated minimum sample size required.

#### Solution

Confidence level 95% means that  $\alpha = 1 - 0.95 = 0.05$  so  $\alpha/2 = 0.025$ . From the last line of Figure 7.1.6 we obtain  $z_{0.025} = 1.960$ .

To say that the estimate is to be “to within \$1,000” means that  $E = 1000$ . Thus

$$n = \frac{(z_{\alpha/2})^2 \sigma^2}{E^2} = \frac{(1.960)^2 (4000)^2}{(1000)^2} = 61.4656$$

which we round up to 62.

## Estimating $p$

The confidence interval formula for estimating a population proportion  $p$  is  $\hat{p} \pm E$ , where

$$E = z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

The number  $z_{\alpha/2}$  is determined by the desired level of confidence. To say that we wish to estimate the population proportion to within a certain number of percentage points means that we want the margin of error  $E$  to be no larger than that number (expressed as a proportion). Thus we obtain the minimum sample size needed by solving the displayed equation for  $n$ .

### Minimum Sample Size for Estimating a Population Proportion

The estimated minimum sample size  $n$  needed to estimate a population proportion  $p$  to within  $E$  at  $100(1 - \alpha)\%$  confidence is

$$n = \frac{(z_{\alpha/2})^2 \hat{p}(1 - \hat{p})}{E^2} \text{ (rounded up)}$$

There is a dilemma here: the formula for estimating how large a sample to take contains the number  $\hat{p}$ , which we know only after we have taken the sample. There are two ways out of this dilemma. Typically the researcher will have some idea as to the value of the population proportion  $p$ , hence of what the sample proportion  $\hat{p}$  is likely to be. For example, if last month 37% of all voters thought that state taxes are too high, then it is likely that the proportion with that opinion this month will not be dramatically different, and we would use the value 0.37 for  $\hat{p}$  in the formula.

The second approach to resolving the dilemma is simply to replace  $\hat{p}$  in the formula by 0.5. This is because if  $\hat{p}$  is large then  $1 - \hat{p}$  is small, and vice versa, which limits their product to a maximum value of 0.25, which occurs when  $\hat{p} = 0.5$ . This is called the **most conservative estimate**, since it gives the largest possible estimate of  $n$ .

### ✓ Example 7.4.3

Find the necessary minimum sample size to construct a 98% confidence interval for  $p$  with a margin of error  $E = 0.05$ ,

- assuming that no prior knowledge about  $p$  is available; and
- assuming that prior studies suggest that  $p$  is about 0.1.

#### Solution

Confidence level 98% means that  $\alpha = 1 - 0.98 = 0.02$  so  $\alpha/2 = 0.01$ . From the last line of Figure 7.1.6 we obtain  $z_{0.01} = 2.326$ .

- Since there is no prior knowledge of  $p$  we make the most conservative estimate that  $\hat{p} = 0.5$ . Then

$$n = \frac{(z_{\alpha/2})^2 \hat{p}(1 - \hat{p})}{E^2} = \frac{(2.326)^2 (0.5)(1 - 0.5)}{0.05^2} = 541.0276$$

which we round up to 542.

- Since  $p \approx 0.1$  we estimate  $\hat{p}$  by 0.1, and obtain

$$n = \frac{(z_{\alpha/2})^2 \hat{p}(1 - \hat{p})}{E^2} = \frac{(2.326)^2 (0.1)(1 - 0.1)}{0.05^2} = 194.769936$$

which we round up to 195.

#### ✓ Example 7.4.4

A dermatologist wishes to estimate the proportion of young adults who apply sunscreen regularly before going out in the sun in the summer. Find the minimum sample size required to estimate the proportion to within three percentage points, at 90% confidence.

##### Solution

Confidence level 90% means that  $\alpha = 1 - 0.90 = 0.10$  so  $\alpha/2 = 0.05$ . From the last line of Figure 7.1.6 we obtain  $z_{0.05} = 1.645$ .

Since there is no prior knowledge of  $p$  we make the most conservative estimate that  $\hat{p} = 0.5$ . To estimate “to within three percentage points” means that  $E = 0.03$ . Then

$$n = \frac{(z_{\alpha/2})^2 \hat{p}(1 - \hat{p})}{E^2} = \frac{(1.645)^2 (0.5)(1 - 0.5)}{0.03^2} = 751.6736111$$

which we round up to 752.

#### Key Takeaway

- If the population standard deviation  $\sigma$  is known or can be estimated, then the minimum sample size needed to obtain a confidence interval for the population mean with a given maximum error of the estimate and a given level of confidence can be estimated.
- The minimum sample size needed to obtain a confidence interval for a population proportion with a given maximum error of the estimate and a given level of confidence can always be estimated. If there is prior knowledge of the population proportion  $p$  then the estimate can be sharpened.

This page titled [7.4: Sample Size Considerations](#) is shared under a [CC BY-NC-SA 3.0](#) license and was authored, remixed, and/or curated by [Anonymous](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.