Role of the Population Size *N* Formulas 7-2 and 7-3 are remarkable because they show that the sample size does not depend on the size (N) of the population; the sample size depends on the desired confidence level, the desired margin of error, and sometimes the known estimate of $\hat{p}$. (See Exercise 39 for dealing with cases in which a relatively large sample is selected without replacement from a finite population, so the sample size n does depend on the population size N.)

Example 4 What Percentage of Adults Buy Clothing Online?

Gap, Banana Republic, J. Crew, Yahoo, and America Online are just a few of the many companies interested in knowing the percentage of adults who buy clothing online. How many adults must be surveyed in order to be 95% confident that the sample percentage is in error by no more than three percentage points?

a. Use this recent result from the Census Bureau: 66% of adults buy clothing online.

b. Assume that we have no prior information suggesting a possible value of the proportion.

Solution

a. The prior study suggests that $\hat{p} = 0.66$, so $\hat{q} = 0.34$ (found from $\hat{q} = 1 - 0.66$). With a 95% confidence level, we have $\alpha = 0.05$, so $z_{\alpha/2} = 1.96$. Also, the margin of error is $E = 0.03$ (the decimal equivalent of "three percentage points"). Because we have an estimated value of $\hat{p}$, we use Formula 7-2 as follows:

$$n = \frac{[z_{\alpha/2}]^2 \hat{p}\hat{q}}{E^2} = \frac{[1.96]^2(0.66)(0.34)}{0.03^2}$$
$$= 957.839 = 958 \qquad \text{(rounded up)}$$

We must obtain a simple random sample that includes at least 958 adults.

b. As in part (a), we again use $z_{\alpha/2} = 1.96$ and $E = 0.03$, but with no prior knowledge of $\hat{p}$ (or $\hat{q}$), we use Formula 7-3 as follows:

$$n = \frac{[z_{\alpha/2}]^2 \cdot 0.25}{E^2} = \frac{[1.96]^2 \cdot 0.25}{0.03^2}$$
$$= 1067.11 = 1068 \qquad \text{(rounded up)}$$

Interpretation

To be 95% confident that our sample percentage is within three percentage points of the true percentage for all adults, we should obtain a simple random sample of 1068 adults. By comparing this result to the sample size of 958 found in part (a), we can see that if we have no knowledge of a prior study, a larger sample is required to achieve the same results as when the value of $\hat{p}$ can be estimated.

CAUTION Try to avoid these two common errors when calculating sample size:

1. Don't make the mistake of using $E = 3$ as the margin of error corresponding to "three percentage points." If the margin of error is three percentage points, use $E = 0.03$.
2. Be sure to substitute the critical z score for $z_{\alpha/2}$. For example, if you are working with 95% confidence, be sure to replace $z_{\alpha/2}$ with 1.96. Don't make the mistake of replacing $z_{\alpha/2}$ with 0.95 or 0.05.