

7.6 (CD-ROM TOPIC) SAMPLING FROM FINITE POPULATIONS

The Central Limit Theorem and the standard errors of the mean and of the proportion are based on samples selected with replacement. However, in virtually all survey research, you sample without replacement from populations that are of a finite size N . In these cases, particularly when the sample size n is more than 5% of the population size N (i.e., $n/N > 0.05$), you use a **finite population correction factor (fpc)** defined in Equation (7.9), to calculate the standard error of the mean and the standard error of the proportion.

FINITE POPULATION CORRECTION FACTOR

$$fpc = \sqrt{\frac{N-n}{N-1}} \quad (7.9)$$

where

$$\begin{aligned} n &= \text{sample size} \\ N &= \text{population size} \end{aligned}$$

Therefore, when referring to means,

STANDARD ERROR OF THE MEAN FOR FINITE POPULATIONS

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \quad (7.10)$$

When referring to proportions,

STANDARD ERROR OF THE PROPORTION FOR FINITE POPULATIONS

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \sqrt{\frac{N-n}{N-1}} \quad (7.11)$$

Examining the formula for the finite population correction factor [Equation (7.9)], you see that the numerator is always smaller than the denominator, since n is greater than 1 for all practical cases. Therefore, the correction factor is less than 1. Because this finite population correction factor is multiplied by the standard error, the standard error becomes smaller when corrected. Therefore, you get more precise estimates when you use the finite population correction factor.

The application of the finite population correction factor is illustrated using two examples previously discussed in this chapter.

EXAMPLE 7.7

USING THE FINITE POPULATION CORRECTION FACTOR WITH THE MEAN

In the cereal-filling example in section 7.1 on page 234, you selected a sample of 25 cereal boxes from a filling process with $\mu = 368$ grams. Suppose that 2,000 boxes (i.e., the population) are filled on this particular day. Using the finite population correction factor, what is the probability that the sample mean is below 365 grams?

SOLUTION Using the finite population correction factor, $\sigma = 15$, $n = 25$, and $N = 2,000$, so that

$$\begin{aligned}\sigma_{\bar{X}} &= \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \\ &= \frac{15}{\sqrt{25}} \sqrt{\frac{2,000-25}{2,000-1}} \\ &= 3\sqrt{0.988} = 2.982\end{aligned}$$

The probability that the sample mean is below 365 is computed as follows.

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{-3}{2.982} = -1.01$$

From Table E.2, the area below 365 grams is 0.1562.

The finite population correction factor has a very small effect on the standard error of the mean and the subsequent area under the normal curve because the sample size is only 1.25% of the population size (i.e., $n/N = 25/2000 = 0.0125$).

EXAMPLE 7.8

USING THE FINITE POPULATION CORRECTION FACTOR WITH THE PROPORTION

In the example concerning multiple banking accounts on page 240, suppose there are a total of 1,000 different depositors at the bank. Using the finite population correction factor, what is the probability that the sample proportion of depositors having multiple bank accounts is less than 0.30?

SOLUTION Using the finite population correction factor with the previous sample of $n = 200$ results in the following.

$$\begin{aligned}\sigma_p &= \sqrt{\frac{\pi(1-\pi)}{n}} \sqrt{\frac{N-n}{N-1}} \\ &= \sqrt{\frac{(0.40)(0.60)}{200}} \sqrt{\frac{1,000-200}{1,000-1}} \\ &= \sqrt{\frac{0.24}{200}} \sqrt{\frac{800}{999}} = \sqrt{0.0012} \sqrt{0.801} \\ &= (0.0346)(0.895) = 0.031\end{aligned}$$

With the standard error of the sample proportion = 0.031 from Equation (7.11),

$$\begin{aligned}Z &= \frac{0.30 - 0.40}{0.031} \\ &= -3.23\end{aligned}$$

From Table E.2, the appropriate area below $p = 0.30$ is 0.00062. In this example, the use of the finite population correction factor has a moderate effect on the standard error of the proportion and on the area under the normal curve because the sample size is 20% of the population (i.e., $n/N = 200/1,000 = 0.20$).

PROBLEMS FOR SECTION 7.6

Learning the Basics

7.67 Given that $N = 80$ and $n = 10$ and the sample is selected without replacement, determine the finite population correction factor.

7.68 Which of the following finite population factors will have a greater effect in reducing the standard error—one based on a sample of $n = 100$ selected without replacement from a population of $N = 400$ or one based on a sample of $n = 200$ selected without replacement from a population of $N = 900$? Explain.

7.69 Given that $N = 60$ and $n = 20$ and the sample is selected *with* replacement, should you use the finite population correction factor? Explain.

Applying the Concepts

7.70 The diameter of Ping-Pong balls manufactured at a large factory is approximately normally distributed with a mean of 1.30 inches and a standard deviation of 0.04 inch. If a random sample of 16 Ping-Pong balls is selected from a population of 200 Ping-Pong balls without replacement, what is the probability that the sample mean is between 1.31 and 1.33 inches?

7.71 The amount of time a bank teller spends with each customer has a population mean $\mu = 3.10$ minutes and stan-

dard deviation $\sigma = 0.40$ minute. If a random sample of 16 customers is selected without replacement from a population of 500 customers,

- what is the probability that the mean time spent per customer is at least 3 minutes?
- there is an 85% chance that the sample mean is less than how many minutes?

7.72 Historically, 10% of a large shipment of machine parts are defective. If a random sample of 400 parts is selected without replacement from a shipment that included 5,000 machine parts, what is the probability that the sample will have

- between 9% and 10% defective parts?
- less than 8% defective parts?

7.73 Historically, 93% of the deliveries of an overnight mail service arrive before 10:30 the following morning. If a random sample of 500 deliveries is selected without replacement from a population that consisted of 10,000 deliveries, what is the probability that the sample will have

- between 93% and 95% of the deliveries arriving before 10:30 the following morning?
- more than 95% of the deliveries arriving before 10:30 the following morning?