7.3: SAMPLING FROM FINITE POPULATIONS

The central limit theorem and the standard errors of the mean and of the proportion are based on the premise that the samples selected are chosen with replacement. However, in virtually all survey research, sampling is conducted without replacement from populations that are of a finite size N. In these cases, particularly when the sample size n is not small in comparison with the population size N (i.e., more than 5% of the population is sampled) so that $n/N > 0.05$, a **finite population correction factor (fpc)** is used to define both the standard error of the mean and the standard error of the proportion. The finite population correction factor is expressed as

FINITE POPULATION CORRECTION FACTOR

$$fpc = \sqrt{\frac{N-n}{N-1}} \tag{7.9}$$

where

n = sample size

N = population size

Therefore, when dealing with means,

STANDARD ERROR OF THE MEAN FOR FINITE POPULATIONS

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \tag{7.10}$$

When referring to proportions,

STANDARD ERROR OF THE PROPORTION FOR FINITE POPULATIONS

$$\sigma_{p_s} = \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} \tag{7.11}$$

Examining the formula for the finite population correction factor [Equation (7.9)], observe 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 more precise estimates are obtained when the finite population correction factor is used.

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

Example 7.5

USING THE FINITE POPULATION CORRECTION FACTOR WITH THE MEAN

In the cereal-filling example in section 7.1, a sample of 25 cereal boxes was selected from a filling process. Suppose that 2,000 boxes (i.e., the population) are filled on this particular day. Using the finite population correction factor, determine the probability of obtaining a sample whose mean is below 365 grams.

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

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

The probability of obtaining a sample whose mean is between 365 and 368 grams is computed as follows.

$$Z = \frac{\overline{X}-\mu_{\overline{X}}}{\sigma_{\overline{X}}} = \frac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}} = \frac{-3}{2.982} = -1.01$$

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

It is evident in this example that the use of 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 (i.e., $n = 25$ boxes) is only 1.25% of the population size (i.e., $N = 2{,}000$ boxes).

Example 7.6

USING THE FINITE POPULATION CORRECTION FACTOR WITH THE PROPORTION

In the example on page 244 concerning multiple banking accounts, suppose there are a total of 1,000 different depositors at the bank. Using the finite population correction factor, determine the probability of obtaining a sample where the 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.

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

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 $p_s = 0.30$ 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% (i.e., $n/N = 0.20$) of the population.

PROBLEMS FOR SECTION 7.3

Learning the Basics

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

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

7.43 Given that $N = 60$ and $n = 20$ and the sample is obtained *with* replacement, should the finite population correction factor be used? Explain.

Applying the Concepts

7.44 The diameter of Ping-Pong balls manufactured at a large factory is expected to be approximately normally distributed with a mean of 1.30 inches and a standard deviation of 0.04 inch. If many random samples of 16 Ping-Pong balls are selected from a population of 200 Ping-Pong balls *without* replacement, what proportion of the sample means would be between 1.31 and 1.33 inches?

7.45 The amount of time a bank teller spends with each customer has a population mean $\mu = 3.10$ minutes and standard deviation $\sigma = 0.40$ minute. If a random sample of 16 customers is selected *without* replacement from a population of 500 customers,

a. what is the probability that the average time spent per customer will be at least 3 minutes?
b. there is an 85% chance that the sample mean will be below how many minutes?

7.46 Historically, 10% of a large shipment of machine parts are defective. If random samples of 400 parts are selected *without* replacement from a shipment that included 5,000 machine parts, what proportion of the samples will have

a. between 9% and 10% defective parts?
b. less than 8% defective parts?

7.47 Historically, 93% of the deliveries of an overnight mail service arrive before 10:30 the following morning. If random samples of 500 deliveries are selected *without* replacement from a population that consisted of 10,000 deliveries, what proportion of the samples will have

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