8.7: ESTIMATION AND SAMPLE SIZE DETERMINATION FOR FINITE POPULATIONS

Estimating the Mean

In section 7.3 the **finite population correction (fpc) factor** is used to reduce the standard error by a value equal to $\sqrt{(N - n)/(N - 1)}$. When developing confidence interval estimates for population parameters, the fpc factor is used when samples are selected without replacement. Thus, the $(1 - \alpha) \times 100\%$ confidence interval estimate for the mean is calculated as in Equation (8.12).

CONFIDENCE INTERVAL FOR A MEAN (σ UNKNOWN) FOR A FINITE POPULATION

$$\overline{X} \pm t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}} \qquad (8.12)$$

To illustrate the finite population correction factor, refer to the confidence interval estimate for the mean developed for Saxon Plumbing Company in section 8.2. Suppose that in this month there are 5,000 sales invoices. Using $\overline{X}$ = $110.27, S = $28.95, N = 5,000, n = 100, and with 95% confidence, t_{99} = 1.9842. From equation (8.12)

$$110.27 \pm (1.9842) \frac{28.95}{\sqrt{100}} \sqrt{\frac{5{,}000 - 100}{5{,}000 - 1}}$$

$$= 110.27 \pm 5.744(0.99)$$

$$110.27 \pm 5.69$$

$$\$104.58 \leq \mu \leq \$115.96$$

Figure 8.29 illustrates the PHStat output for this example.

	A	B
1	Confidence Interval Estimation for the Mean	
2		
3	Sample Standard Deviation	28.95
4	Sample Mean	110.27
5	Sample Size	100
6	Confidence Level	95%
7	Standard Error of the Mean	2.895
8	Degrees of Freedom	99
9	t Value	1.984217306
10	Interval Half Width	5.744309101
11	Interval Lower Limit	104.53
12	Interval Upper Limit	116.01
13		
14		
15	Finite Populations	
16	Population Size	5000
17	FPC Factor	0.990048503
18	Interval Half Width	5.687144629
19	Interval Lower Limit	104.58
20	Interval Upper Limit	115.96

FIGURE 8.29
Confidence interval estimate for the mean sales invoice amount with the finite population correction for Saxon Plumbing Company

In this case, because the sample is a very small fraction of the population, the correction factor has a minimal effect on the width of the confidence interval. To examine the effect of the correction factor when the sample size is more than 5% of the population size, see Example 8.10.

Example 8.10

ESTIMATING THE MEAN FORCE
FOR INSULATORS

In Example 8.3, a sample of 30 insulators was selected. Suppose a population of 300 insulators were produced by the company. Set up a 95% confidence interval estimate of the population mean.

SOLUTION Using the finite population correction factor, with $\overline{X}$ = 1,723.4 pounds, S = 89.55, n = 30, N = 300, and t_{29} = 2.0452 (for 95% confidence):

$$\overline{X} \pm t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

$$= 1,723.4 \pm (2.0452) \frac{89.55}{\sqrt{30}} \sqrt{\frac{300-30}{300-1}}$$

$$= 1,723.4 \pm 33.44(0.9503)$$

$$= 1,723.4 \pm 31.776$$

$$1,691.62 \le \mu \le 1,755.18$$

Here, because 10% of the population is to be sampled, the fpc factor has a small effect on the confidence interval estimate.

Estimating the Proportion

In sampling without replacement, the $(1 - \alpha) \times 100\%$ confidence interval estimate of the proportion is defined in Equation (8.13).

CONFIDENCE INTERVAL ESTIMATE FOR THE PROPORTION USING THE FINITE POPULATION CORRECTION FACTOR

$$p_s \pm Z \sqrt{\frac{p_s(1-p_s)}{n}} \sqrt{\frac{N-n}{N-1}} \qquad (8.13)$$

To illustrate the use of the finite population correction factor when developing a confidence interval estimate of the population proportion, consider again the estimate developed for Saxon Home Improvement Company in section 8.3. For these data, N = 5,000, n = 100, p_s = 10/100 = 0.10, and with 95% confidence, Z = 1.96. Using Equation (8.13),

$$p_s \pm Z \sqrt{\frac{p_s(1-p_s)}{n}} \sqrt{\frac{N-n}{N-1}}$$

$$= 0.10 \pm (1.96) \sqrt{\frac{(0.10)(0.90)}{100}} \sqrt{\frac{5,000-100}{5,000-1}}$$

$$= 0.10 \pm (1.96)(0.03)(0.99)$$

$$= 0.10 \pm .0582$$

$$0.0418 \le p \le 0.1582$$

In this case, because the sample is a very small fraction of the population, the fpc factor has virtually no effect on the confidence interval estimate.

Determining the Sample Size

Just as the fpc factor is used to develop confidence interval estimates, it also is used to determine sample size when sampling without replacement. For example, in estimating the mean, the sampling error is

$$e = Z \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}$$

and in estimating the proportion, the sampling error is

$$e = Z \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}}$$

To determine the sample size in estimating the mean or the proportion from Equations (8.4) and (8.5),

$$n_0 = \frac{Z^2\sigma^2}{e^2} \text{ and } n_0 = \frac{Z^2 p(1-p)}{e^2}$$

where n_0 is the sample size without considering the finite population correction factor. Applying the fpc factor results in the actual sample size n, computed as in Equation (8.14).

> **SAMPLE SIZE DETERMINATION USING THE FINITE POPULATION CORRECTION FACTOR**
>
> $$n = \frac{n_0 N}{n_0 + (N-1)}$$ (8.14)

In determining the sample size for Saxon Home Improvement Company, a sample size of 97 was needed (rounded up from 96.04) for the mean and a sample of 100 (rounded up from 99.96) was needed for the proportion. Using the fpc factor in Equation (8.14) for the mean, with $N = 5,000$, $e = \$5$, $S = \$25$, and $Z = 1.96$ (for 95% confidence), leads to

$$n = \frac{(96.04)(5,000)}{96.04 + (5,000 - 1)} = 94.24$$

Thus, $n = 95$.

Using the fpc factor in Equation (8.14) for the proportion, with $N = 5,000$, $e = 0.07$, $p = 0.15$, and $Z = 1.96$ (for 95% confidence),

$$n = \frac{(99.96)(5,000)}{99.96 + (5,000 - 1)} = 98.02$$

Thus, $n = 99$.

To satisfy *both requirements simultaneously with one sample*, the larger sample size of 99 is needed. PHStat output is displayed in Figure 8.30.

	A	B
1	Sample Size for Estimating the Mean	
2		
3	Population Standard Deviation	25
4	Sampling Error	5
5	Confidence Level	95%
6	Z Value	-1.95996108
7	Calculated Sample Size	96.03618611
8	Sample Size Needed	97
9		
10		
11	Finite Populations	
12	Population Size	5000
13	Sample Size with FPC	94.24485185
14	Sample Size Needed	95

FIGURE 8.30
Sample size for estimating the mean sales invoice amount with the finite population correction for the Saxon Home Improvement Company

PROBLEMS FOR SECTION 8.7

Learning the Basics

8.88 If $\bar{X} = 75$, $S = 24$, $n = 36$, and $N = 200$, set up a 95% confidence interval estimate of the population mean μ if sampling is done *without* replacement.

8.89 Consider a population of 1,000 where the standard deviation is assumed equal to 20. What sample size would be required if sampling is done *without* replacement if you desire 95% confidence and a sampling error of ±5?

Applying the Concepts

8.90 The quality control manager at a lightbulb factory needs to estimate the mean life of a large shipment of lightbulbs. The process standard deviation is known to be 100 hours. Assume that the shipment contains a total of 2,000 lightbulbs and that sampling is done *without* replacement.
  a. Set up a 95% confidence interval estimate of the population mean life of lightbulbs in this shipment if a random sample of 50 lightbulbs selected from the shipment indicates a sample average life of 350 hours.
  b. Determine the sample size needed to estimate the average life to within ±20 hours with 95% confidence.
  c. What are your answers to (a) and (b) if the shipment contains 1,000 lightbulbs?

8.91 A survey is planned to determine the mean annual family medical expenses of employees of a large company. The management of the company wishes to be 95% confident that the sample average is correct to within ±$50 of the true average annual family medical expenses. A pilot study indicates that the standard deviation is estimated as $400. How large a sample size is necessary if the company has 3,000 employees and if sampling is done *without* replacement?

8.92 The manager of a bank that has 1,000 depositors in a small city wants to determine the proportion of its depositors with more than one account at the bank.
  a. Set up a 90% confidence interval estimate of the population proportion of the bank's depositors who have more than one account at the bank if a random sample of 100 depositors is selected *without* replacement and 30 state that they have more than one account at the bank.
  b. A bank manager wants to be 90% confident of being correct to within ±0.05 of the true population proportion of depositors who have more than one account at the bank. What sample size is needed if sampling is done *without* replacement?
  c. What are your answers to (a) and (b) if the bank has 2,000 depositors?

8.93 An automobile dealer wants to estimate the proportion of customers who still own the cars they purchased 5 years earlier. Sales records indicate that the population of owners is 4,000.
  a. Set up a 95% confidence interval estimate of the population proportion of all customers who still own their cars 5 years after they were purchased if a random sample of 200 customers selected *without* replacement from the automobile dealer's records indicate that 82 still own cars that were purchased 5 years earlier.
  b. What sample size is necessary to estimate the true proportion to within ±0.025 with 95% confidence?
  c. What are your answers to (a) and (b) if the population consists of 6,000 owners?

8.94 The inspection division of the Lee County Weights and Measures Department is interested in estimating the actual amount of soft drink that is placed in 2-liter bottles at the local bottling plant of a large nationally known soft-drink company. The population consists of 2,000 bottles. The bottling plant has informed the inspection division that the standard deviation for 2-liter bottles is 0.05 liter.
  a. Set up a 95% confidence interval estimate of the population mean amount of soft drink per bottle if a random sample of one hundred 2-liter bottles obtained *without* replacement from this bottling plant indicates a sample average of 1.99 liters.
  b. Determine the sample size necessary to estimate the population mean amount to within ±0.01 liter with 95% confidence.
  c. What are your answers to (a) and (b) if the population consists of 1,000 bottles?

8.95 A stationery store wants to estimate the mean retail value of greeting cards that it has in its 300-card inventory.
  a. Set up a 95% confidence interval estimate of the population mean value of all greeting cards that are in its inventory if a random sample of 20 greeting cards selected *without* replacement indicates an average value of $1.67 and a standard deviation of $0.32.
  b. What is your answer to (a) if the store has 500 greeting cards in its inventory?