8 Interval Estimation

CHAPTER CONTENTS

LEARNING OBJECTIVES

After reading this chapter and doing the exercises, you should be able to:

1 Explain the purpose of an interval estimate of a population parameter.

2 Explain the terms margin of error, confidence interval, confidence level and confidence coefficient.

3 Construct confidence intervals for a population mean:

   3.1 When the population standard deviation is known, using the normal distribution.

   3.2 When the population standard deviation is unknown, using the t distribution.

4 Construct large-sample confidence intervals for a population proportion.

5 Calculate the sample size required to construct a confidence interval with a given margin of error for a population mean, when the population standard deviation is known.

6 Calculate the sample size required to construct a confidence interval with a given margin of error for a population proportion.

In Chapter 7, we stated that a point estimator is a sample statistic used to estimate a population parameter. For example, the sample mean is a point estimator of the population mean, and the sample proportion is a point estimator of the population proportion. Because a point estimator cannot be expected to provide the exact value of the population parameter, an **interval estimate** is often computed, by adding and subtracting a **margin of error**.

STATISTICS IN PRACTICE

How accurate are opinion polls and market research surveys?

IPSOS and ICM Research are two large, global market research and opinion polling companies. IPSOS has its global headquarters in Paris, ICM Research is based in London.

In July 2012, IPSOS released a report covering opinion surveys in 24 countries across the globe, entitled *IPSOS Global @dvisory: The Economic Pulse of the World*. About eighteen and a half thousand interviews were conducted over the 24 countries. Interviewees were asked to assess the economic situation in their home country: as either 'Very good', or 'Somewhat good', or 'Somewhat bad' or 'Very bad'.

In Spain, 78 per cent of respondents described the economic situation as very bad. This compared with 25 per cent of respondents in Great Britain, 15 per cent in South Africa, only 3 per cent in Germany and just 2 per cent in Sweden.

But how accurate are estimates like these based on sample evidence?

The issue of survey accuracy and margin of error features on the ICM website, www.icmresearch.co.uk. On one page, there is an interactive 'ready-reckoner' that will calculate the margin of error for any given percentage result, like those above, and for any given sample size. For example, in respect of the 78 per cent of 1012 respondents in Spain who considered the economic situation to be very bad, the ICM ready-reckoner calculates the 'accuracy at 95 per cent confidence level' to be plus or minus 2.6 percentage points. In other words, this implies we can be 95 per cent confident that the percentage of all adults in Spain who thought the economic situation was very bad was between 75.4 per cent and 80.6 per cent. By comparison, for the corresponding figure of 15 per cent in South Africa, where the sample size was 506, ICM's ready-reckoner gives the margin of error at the 95 per cent confidence level as plus or minus 3.1 per cent.

In this chapter, you will learn the basis for these margins of error, the confidence level of 95 per cent associated with them, and the calculations that underlie the ICM's ready-reckoner.

The purpose of an interval estimate is to provide information about how close the point estimate might be to the value of the population parameter. In relatively simple cases, the general form of an interval estimate is:

$$\text{Point estimate} \pm \text{Margin of error}$$

In this chapter we show how to compute interval estimates of a population mean μ and a population proportion π. The interval estimates have the same general form:

$$\text{Population mean: } \bar{x} \pm \text{Margin of error}$$
$$\text{Population proportion: } p \pm \text{Margin of error}$$

The sampling distributions of $\bar{X}$ and P play key roles in computing these interval estimates.

8.1 POPULATION MEAN: σ KNOWN

To construct an interval estimate of a population mean, either the population standard deviation σ or the sample standard deviation s must be used to compute the margin of error. Although σ is rarely known exactly, historical data sometimes permit us to obtain a good estimate of the population standard deviation prior to sampling. In such cases, the population standard deviation can be considered known for practical purposes. We refer to such cases as the σ **known** case. In this section we show how a simple random sample can be used to construct an interval estimate of a population mean for the σ known case.

Consider the monthly customer service survey conducted by CJW Limited, who has a website for taking customer orders and providing follow-up service. The company prides itself on providing easy online ordering, timely delivery and prompt response to customer enquiries. Good customer service is critical to the company's ongoing success.

CJW's quality assurance team uses a customer service survey to measure satisfaction with its website and online customer service. Each month, the team sends a questionnaire to a random sample of customers who placed an order or requested service during the previous month. The questionnaire asks customers to rate their satisfaction with such things as ease of placing orders, timely delivery, accurate order filling and technical advice. The team summarizes each customer's questionnaire by computing an overall satisfaction score x that ranges from 0 (worst possible score) to 100 (best possible score). A sample mean customer satisfaction score is then computed.

The sample mean satisfaction score provides a point estimate of the mean satisfaction score μ for the population of all CJW customers. With this regular measure of customer service, CJW can promptly take corrective action if a low satisfaction score results. The company conducted this satisfaction survey for a number of months, and consistently obtained an estimate near 12 for the standard deviation of satisfaction scores. Based on these historical data, CJW now assumes a known value of $\sigma = 12$ for the population standard deviation. The historical data also indicate that the population of satisfaction scores follows an approximately normal distribution.

CJW

During the most recent month, the quality assurance team surveyed 100 customers ($n = 100$) and obtained a sample mean satisfaction score of $\bar{x} = 72$. This provides a point estimate of the population mean satisfaction score μ. We show how to compute the margin of error for this estimate and construct an interval estimate of the population mean.

Margin of error and the interval estimate

In Chapter 7 we showed that the sampling distribution of the sample mean $\bar{X}$ can be used to compute the probability that $\bar{X}$ will be within a given distance of μ. In the CJW example, the historical data show that the population of satisfaction scores is normally distributed with a standard deviation of $\sigma = 12$. So, using what we learned in Chapter 7, we can conclude that the sampling distribution of $\bar{X}$ follows a normal distribution with a standard error of:

$$\sigma_{\bar{X}} = \sigma/\sqrt{n} = 12/\sqrt{100} = 1.2$$

This sampling distribution is shown in Figure 8.1.*

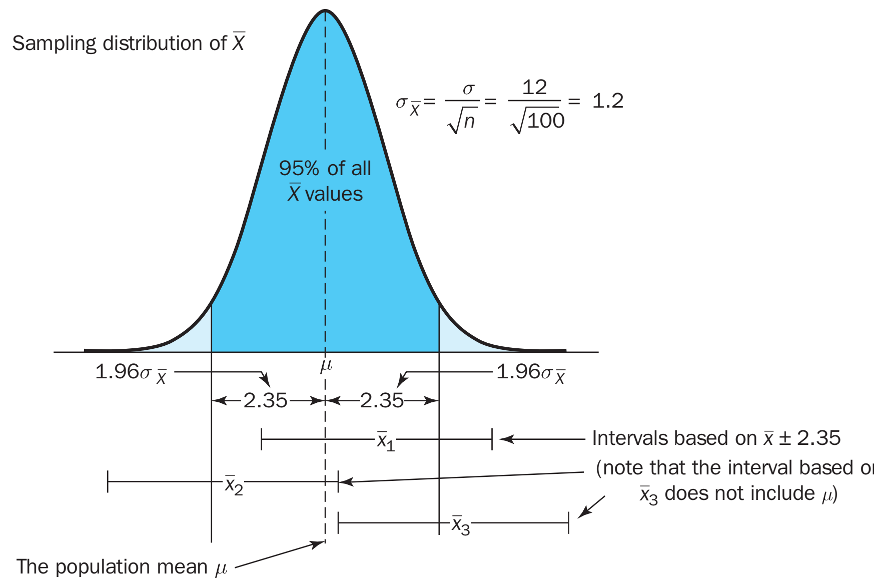
Using the table of cumulative probabilities for the standard normal distribution, we find that 95 per cent of the values of any normally distributed random variable are within ± 1.96 standard deviations of the mean. So, 95 per cent of the $\bar{X}$ values must be within $\pm 1.96\sigma_{\bar{X}}$ of the mean μ. In the CJW example, we know that the sampling distribution of $\bar{X}$ is normal with a standard error of $\sigma_{\bar{X}} = 1.2$. Because $\pm 1.96\sigma_{\bar{X}} = \pm 1.96(1.2) = \pm 2.35$, we conclude that 95 per cent of all $\bar{X}$ values obtained using a sample size of $n = 100$ will be within ± 2.35 units of the population mean μ. See Figure 8.1.

We said above that the general form of an interval estimate of the population mean μ is $\bar{x} \pm$ margin of error. For the CJW example, suppose we set the margin of error equal to 2.35 and compute the interval estimate of μ using $\bar{x} \pm 2.35$. To provide an interpretation for this interval estimate, let us consider the values of $\bar{x} \pm 2.35$ that could be obtained if we took three different simple random samples, each consisting of 100 CJW customers.

The first sample mean might turn out to have the value shown as $\bar{x}_1$ in Figure 8.1. In this case, the interval formed by subtracting 2.35 from $\bar{x}_1$ and adding 2.35 to $\bar{x}_1$ includes the population mean μ.

*The population of satisfaction scores has a normal distribution, so we can conclude that the sampling distribution of $\bar{X}$ is a normal distribution. If the population did not have a normal distribution, we could rely on the central limit theorem, and the sample size of $n = 100$, to conclude that the sampling distribution of $\bar{X}$ is approximately normal. In either case, the sampling distribution would appear as shown in Figure 8.1.

FIGURE 8.1
Sampling distribution of the sample mean satisfaction score from simple random samples of 100 customers, also showing the location of sample means that are within ± 2.35 units of μ, and intervals calculated from selected sample means at locations $\bar{x}_1$, $\bar{x}_2$ and $\bar{x}_3$.



Now consider what happens if the second sample mean turns out to have the value shown as $\bar{x}_2$ in Figure 8.1. Although $\bar{x}_2$ differs from $\bar{x}_1$, we see that the interval formed by $\bar{x}_2 \pm 2.35$ also includes the population mean μ. However, consider what happens if the third sample mean turns out to have the value shown as $\bar{x}_3$ in Figure 8.1. In this case, because $\bar{x}_3$ falls in the upper tail of the sampling distribution and is further than 2.35 units from μ, the interval $\bar{x}_3 \pm 2.35$ does not include the population mean μ.

Any sample mean that is within the darkly shaded region of Figure 8.1 will provide an interval estimate that contains the population mean μ. Because 95 per cent of all possible sample means are in the darkly shaded region, 95 per cent of all intervals formed by subtracting 2.35 from $\bar{x}$ and adding 2.35 to $\bar{x}$ will include the population mean μ.

The general form of an interval estimate of a population mean for the σ known case is:

Interval estimate of a population mean: σ known

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad \textbf{(8.1)}$$

where $(1 - \alpha)$ is the confidence coefficient and $z_{\alpha/2}$ is the z value providing an area $\alpha/2$ in the upper tail of the standard normal probability distribution.

Let us use expression (8.1) to construct a 95 per cent confidence interval for the CJW problem. For a 95 per cent confidence interval, the confidence coefficient is $(1 - \alpha) = 0.95$ and so $\alpha = 0.05$. As we saw above, an area of $\alpha/2 = 0.05/2 = 0.025$ in the upper tail gives $z_{0.025} = 1.96$. With the CJW sample mean $\bar{x} = 72$, $\sigma = 12$ and a sample size $n = 100$, we obtain:

$$72 \pm 1.96\frac{12}{\sqrt{100}} = 72 \pm 2.35$$

The specific interval estimate of μ based on the data from the most recent month is $72 - 2.35 = 69.65$, to $72 + 2.35 = 74.35$. Because 95 per cent of all the intervals constructed using $\bar{x} \pm 2.35$ will contain the population mean, we say that we are 95 per cent confident that the interval 69.65 to 74.35 includes the population mean μ. We say that this interval has been established at the 95 per cent **confidence level**. The value 0.95 is referred to as the **confidence coefficient**, and the interval 69.65 to 74.35 is called the 95 per cent **confidence interval**.

TABLE 8.1 Values of $z_{\alpha/2}$ for the most commonly used confidence levels

Confidence level	α	$\alpha/2$	$z_{\alpha/2}$
90%	0.10	0.05	1.645
95%	0.05	0.025	1.960
99%	0.01	0.005	2.576

Although a 95 per cent confidence level is frequently used, other confidence levels such as 90 per cent and 99 per cent may be considered. Values of $z_{\alpha/2}$ for the most commonly used confidence levels are shown in Table 8.1. Using these values and expression (8.1), the 90 per cent confidence interval for the CJW problem is:

$$72 \pm 1.645 \frac{12}{\sqrt{100}} = 72 \pm 1.97$$

At 90 per cent confidence, the margin of error is 1.97 and the confidence interval is $72 - 1.97 = 70.03$, to $72 + 1.97 = 73.97$. Similarly, the 99 per cent confidence interval is:

$$72 \pm 2.576 \frac{12}{\sqrt{100}} = 72 \pm 3.09$$

At 99 per cent confidence, the margin of error is 3.09 and the confidence interval is $72 - 3.09 = 68.93$, to $72 + 3.09 = 75.09$.

Comparing the results for the 90 per cent, 95 per cent and 99 per cent confidence levels, we see that, in order to have a higher degree of confidence, the margin of error and consequently the width of the confidence interval must be larger.

Practical advice

If the population follows a normal distribution, the confidence interval provided by expression (8.1) is exact. Therefore, if expression (8.1) were used repeatedly to generate 95 per cent confidence intervals, 95 per cent of the intervals generated (in the long run) would contain the population mean. If the population does not follow a normal distribution, the confidence interval provided by expression (8.1) will be approximate. In this case, the quality of the approximation depends on both the distribution of the population and the sample size.

In most applications, a sample size of $n \geq 30$ is adequate when using expression (8.1) to construct an interval estimate of a population mean. If the population is not normally distributed but is roughly symmetrical, sample sizes as small as 15 can be expected to provide good approximate confidence intervals. With smaller sample sizes, expression (8.1) should be used only if the analyst believes, or is willing to assume, that the population distribution is at least approximately normal.

EXERCISES

Methods

1. A simple random sample of 40 items results in a sample mean of 25. The population standard deviation is $\sigma = 5$.

   a. What is the value of the standard error of the mean, $\sigma_{\bar{X}}$?

   b. At 95 per cent confidence, what is the margin of error for estimating the population mean?

2. A simple random sample of 50 items from a population with $\sigma = 6$ results in a sample mean of 32.

a. Construct a 90 per cent confidence interval for the population mean.
b. Construct a 95 per cent confidence interval for the population mean.
c. Construct a 99 per cent confidence interval for the population mean.

COMPLETE SOLUTIONS

3. A simple random sample of 60 items results in a sample mean of 80. The population standard deviation is $\sigma = 15$.

a. Compute the 95 per cent confidence interval for the population mean.
b. Assume that the same sample mean was obtained from a sample of 120 items. Construct a 95 per cent confidence interval for the population mean.
c. What is the effect of a larger sample size on the interval estimate?

4. A 95 per cent confidence interval for a population mean was reported to be 152 to 160. If $\sigma = 15$, what sample size was used in this study?

Applications

5. In an effort to estimate the mean amount spent per customer for dinner at a Johannesburg restaurant, data were collected for a sample of 49 customers. Assume a population standard deviation of 40 South African rand (ZAR).

a. At 95 per cent confidence, what is the margin of error?
b. If the sample mean is ZAR186, what is the 95 per cent confidence interval for the population mean?

6. A survey of small businesses with websites found that the average amount spent on a site was €11 500 per year. Given a sample of 60 businesses and a population standard deviation of $\sigma =$ €4000, what is the margin of error in estimating the population mean spend per year? Use 95 per cent confidence.

COMPLETE SOLUTIONS

7. A survey of 750 university students found they were paying on average €108 per week in accommodation costs. Assume the population standard deviation for weekly accommodation costs is €22.

a. Construct a 90 per cent confidence interval estimate of the population mean.
b. Construct a 95 per cent confidence interval estimate of the population mean.
c. Construct a 99 per cent confidence interval estimate of the population mean.
d. Discuss what happens to the width of the confidence interval as the confidence level is increased. Does this result seem reasonable? Explain.

8.2 POPULATION MEAN: σ UNKNOWN

If a good estimate of the population standard deviation σ cannot be obtained prior to sampling, we must use the sample standard deviation s to estimate σ. This is the **σ unknown** case. When s is used to estimate σ, the margin of error and the interval estimate for the population mean are based on a probability distribution known as the ***t* distribution**. Although the mathematical development of the t distribution is based on the assumption of a normal distribution for the population from which we are sampling, research shows that the t distribution can be successfully applied in many situations where the population deviates from normal. Later in this section we provide guidelines for using the t distribution if the population is not normally distributed.

The t distribution is a family of similar probability distributions, with a specific t distribution depending on a parameter known as the **degrees of freedom**.

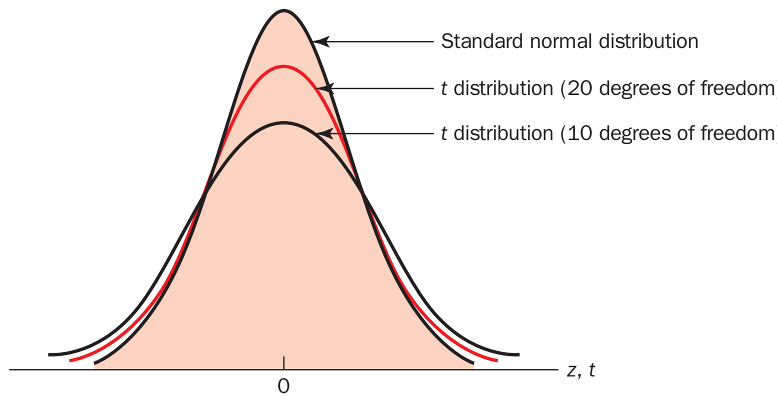


FIGURE 8.2
Comparison of the standard normal distribution with t distributions having 10 and 20 degrees of freedom

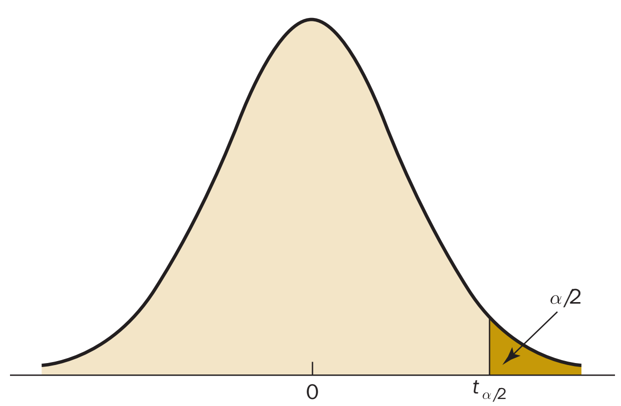
The t distribution with one degree of freedom is unique, as is the t distribution with two degrees of freedom with three degrees of freedom and so on. As the number of degrees of freedom increases, the difference between the t distribution and the standard normal distribution becomes smaller and smaller. Figure 8.2 shows t distributions with ten and 20 degrees of freedom and their relationship to the standard normal probability distribution. Note that, the higher the degrees of freedom, the lower is the variability, and the greater the resemblance to the standard normal distribution. Note also that the mean of the t distribution is zero.

We place a subscript on t to indicate the area in the upper tail of the t distribution. For example, just as we used $z_{0.025}$ to indicate the z value providing a 0.025 area in the upper tail of a standard normal distribution, we will use $t_{0.025}$ to indicate a 0.025 area in the upper tail of a t distribution. So, in general, the notation $t_{\alpha/2}$ will represent a t value with an area of $\alpha/2$ in the upper tail of the t distribution. See Figure 8.3.

Table 2 of Appendix B is a table for the t distribution. Each row in the table corresponds to a separate t distribution with the degrees of freedom shown. For example, for a t distribution with ten degrees of freedom, $t_{0.025} = 2.228$. Similarly, for a t distribution with 20 degrees of freedom, $t_{0.025} = 2.086$. As the degrees of freedom continue to increase, $t_{0.025}$ approaches $z_{0.025} = 1.96$. The standard normal distribution z values can be found in the infinite degrees of freedom row (labelled ∞) of the t distribution table. If the degrees of freedom exceed 100, the infinite degrees of freedom row can be used to approximate the actual t value. In other words, for more than 100 degrees of freedom, the standard normal z value provides a good approximation to the t value.

William Sealy Gosset, writing under the name 'Student', was the originator of the t distribution. Gosset, an Oxford graduate in mathematics, worked for the Guinness Brewery in Dublin, Ireland. The distribution is sometimes referred to as 'Student's t distribution'.

FIGURE 8.3
t distribution with $\alpha/2$ area of probability in the upper tail



Margin of error and the interval estimate

In Section 8.1 we showed that an interval estimate of a population mean for the σ known case is:

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

To compute an interval estimate of μ for the σ unknown case, the sample standard deviation s is used to estimate σ and $z_{\alpha/2}$ is replaced by the t distribution value $t_{\alpha/2}$. The margin of error is then $\pm\ t_{\alpha/2}s/\sqrt{n}$, and the general expression for an interval estimate of a population mean when σ is unknown is:

Interval estimate of a population mean: σ unknown

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \quad \textbf{(8.2)}$$

where s is the sample standard deviation, $(1 - \alpha)$ is the confidence coefficient, and $t_{\alpha/2}$ is the t value providing an area of $\alpha/2$ in the upper tail of the t distribution with $n - 1$ degrees of freedom*.

Consider a study designed to estimate the mean credit card debt for a defined population of households. A sample of $n = 85$ households provided the credit card balances in the file 'Balance' on the online platform. The first few rows of this data set are shown in the EXCEL screenshot in Figure 8.4 below. For this situation, no previous estimate of the population standard deviation σ is available. As a consequence, the sample data must be used to estimate both the population mean and the population standard deviation.

FIGURE 8.4
First few data rows and summary statistics for credit card balances

	A	B	C	D
1	Balance			
2	9619			
3	5364		mean =	5900
4	8348		standard deviation =	3058
5	7348			
6	381			
7	2998			
8	1686			
9	1962			
10	4920			

BALANCE

*The reason the number of degrees of freedom associated with the t value in expression (8.2) is $n - 1$ concerns the use of s as an estimate of the population standard deviation. The expression for the sample standard deviation is $s = \sqrt{\Sigma(x_i - \bar{x})^2/(n-1)}$. Degrees of freedom refers to the number of independent pieces of information that go into the computation of $\Sigma(x_i - \bar{x})^2$. The n pieces of information involved in computing $\Sigma(x_i - \bar{x})^2$ are as follows: $x_1 - \bar{x},\ x_2 - \bar{x}, \cdots, x_n - \bar{x}$. In Section 3.2 we indicated that $\Sigma(x_i - \bar{x}) = 0$. Hence, only $n - 1$ of the $x_i - \bar{x}$ values are independent; that is, if we know $n - 1$ of the values, the remaining value can be determined exactly by using the condition that $\Sigma(x_i - \bar{x}) = 0$. So $n - 1$ is the number of degrees of freedom associated with $\Sigma(x_i - \bar{x})^2$ and hence the number of degrees of freedom for the t distribution in expression (8.2).

Using the data in the 'Balance' file, we compute the sample mean $\overline{x} = 5900$ (€) and the sample standard deviation $s = 3058$ (€).

With 95 per cent confidence and $n - 1 = 84$ degrees of freedom, Table 2 in Appendix B gives $t_{0.025} = 1.989$. We can now use expression (8.2) to compute an interval estimate of the population mean:

$$5900 \pm 1.989\left(\frac{3058}{\sqrt{85}}\right) = 5900 \pm 660$$

The point estimate of the population mean is €5900, the margin of error is €660, and the 95 per cent confidence interval is 5900 − 660 = €5240 to 5900 + 660 = €6560. We are 95 per cent confident that the population mean credit card balance for all households in the defined population is between €5240 and €6560.

The procedures used by MINITAB, EXCEL and IBM SPSS to construct confidence intervals for a population mean are described in the software guides on the online platform.

Practical advice

If the population follows a normal distribution, the confidence interval provided by expression (8.2) is exact and can be used for any sample size. If the population does not follow a normal distribution, the confidence interval provided by expression (8.2) will be approximate. In this case, the quality of the approximation depends on both the distribution of the population and the sample size.

In most applications, a sample size of $n \geq 30$ is adequate when using expression (8.2) to construct an interval estimate of a population mean. However, if the population distribution is highly skewed or contains outliers, the sample size should be 50 or more. If the population is not normally distributed but is roughly symmetrical, sample sizes as small as 15 can be expected to provide good approximate confidence intervals. With smaller sample sizes, expression (8.2) should only be used if the analyst is confident that the population distribution is at least approximately normal.

Using a small sample

In the following example we construct an interval estimate for a population mean when the sample size is small. An understanding of the distribution of the population becomes a factor in deciding whether the interval estimation procedure provides acceptable results.

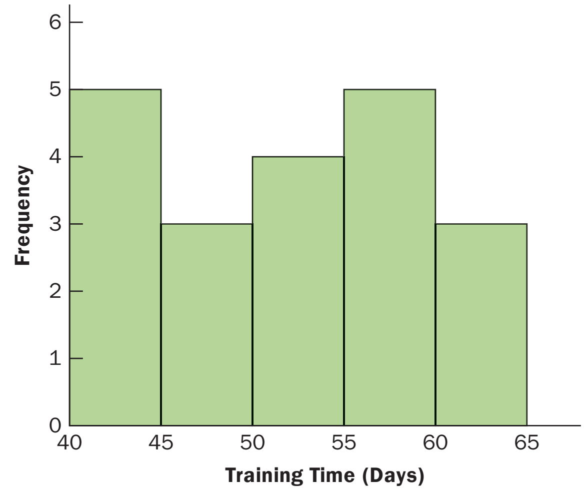
Scheer Industries is considering a new computer-assisted program to train maintenance employees to do machine repairs. To fully evaluate the program, the director of manufacturing requested an estimate of the population mean time required for maintenance employees to complete the training.

A sample of 20 employees is selected, with each employee in the sample completing the training program. Data on the training time in days for the 20 employees are shown in Table 8.2. A histogram of the sample data appears in Figure 8.5. What can we say about the distribution of the population based on this histogram? First, the sample data do not support with certainty the conclusion that the distribution of the population is normal, but we do not see any evidence of skewness or outliers. Therefore, using the guidelines in the previous subsection, we conclude that an interval estimate based on the t distribution appears acceptable for the sample of 20 employees.

TABLE 8.2 Training time in days for a sample of 20 Scheer Industries employees

52	59	54	42
44	50	42	48
55	54	60	55
44	62	62	57
45	46	43	56

FIGURE 8.5
Histogram of training times for the Scheer Industries sample



We compute the sample mean and sample standard deviation as follows:

$$\bar{x} = \frac{\Sigma x_i}{n} = \frac{1030}{20} = 51.5 \text{ days}$$

$$s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{889}{20-1}} = 6.84 \text{ days}$$

For a 95 per cent confidence interval, we use Table 2 from Appendix B and $n - 1 = 19$ degrees of freedom to obtain $t_{0.025} = 2.093$. Expression (8.2) provides the interval estimate of the population mean:

$$51.5 \pm 2.093\left(\frac{6.84}{\sqrt{20}}\right) = 51.5 \pm 3.2$$

The point estimate of the population mean is 51.5 days. The margin of error is 3.2 days and the 95 per cent confidence interval is $51.5 - 3.2 = 48.3$ days to $51.5 + 3.2 = 54.7$ days.

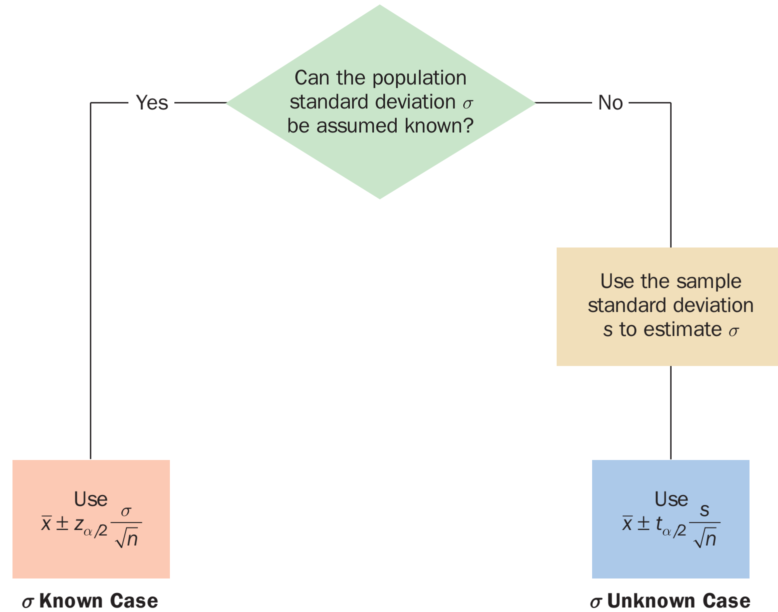
Using a histogram of the sample data to learn about the distribution of a population is rarely conclusive, but in many cases it provides the only information available. The histogram, along with judgement on the part of the analyst, can often be used to decide if expression (8.2) can be used to construct the interval estimate.

Summary of interval estimation procedures

We provided two approaches to computing the margin of error and constructing an interval estimate of a population mean. For the σ known case, σ and the standard normal distribution are used in expression (8.1). For the σ unknown case, the sample standard deviation s and the t distribution are used in expression (8.2).

A summary of the interval estimation procedures for the two cases is shown in Figure 8.6. In most applications, a sample size of $n \geq 30$ is adequate. If the population has a normal or approximately normal distribution, however, smaller sample sizes may be used. For the σ unknown case a sample size of $n \geq 50$ is recommended if the population distribution is believed to be highly skewed or has outliers.

FIGURE 8.6
Summary of interval estimation procedures for a population mean



EXERCISES

Methods

8. For a t distribution with 16 degrees of freedom, find the area, or probability, in each region.

- **a.** To the right of 2.120.
- **b.** To the left of 1.337.
- **c.** To the left of -1.746.
- **d.** To the right of 2.583.
- **e.** Between -2.120 and 2.120.
- **f.** Between -1.746 and 1.746.

9. Find the t value(s) for each of the following cases.

- **a.** Upper-tail area of 0.025 with 12 degrees of freedom.
- **b.** Lower-tail area of 0.05 with 50 degrees of freedom.
- **c.** Upper-tail area of 0.01 with 30 degrees of freedom.
- **d.** Where 90 per cent of the area falls between these two t values with 25 degrees of freedom.
- **e.** Where 95 per cent of the area falls between these two t values with 45 degrees of freedom.

COMPLETE SOLUTIONS

10. The following sample data are from a normal population: 10, 8, 12, 15, 13, 11, 6, 5.

- **a.** What is the point estimate of the population mean?
- **b.** What is the point estimate of the population standard deviation?
- **c.** With 95 per cent confidence, what is the margin of error for the estimation of the population mean?
- **d.** What is the 95 per cent confidence interval for the population mean?

11. A simple random sample with $n = 54$ provided a sample mean of 22.5 and a sample standard deviation of 4.4.

a. Construct a 90 per cent confidence interval for the population mean.
b. Construct a 95 per cent confidence interval for the population mean.
c. Construct a 99 per cent confidence interval for the population mean.
d. What happens to the margin of error and the confidence interval as the confidence level is increased?

Applications

12. Sales personnel for Emirates Distributors submit weekly reports listing the customer contacts made during the week. A sample of 65 weekly reports showed a sample mean of 19.5 customer contacts per week. The sample standard deviation was 5.2. Provide 90 per cent and 95 per cent confidence intervals for the population mean number of weekly customer contacts for the sales personnel.

13. Consumption of alcoholic beverages by young women of drinking age is of concern in the UK and some other European countries. Annual consumption data (in litres) are shown below for a sample of 20 European young women.

ALCOHOL

266	82	199	174	97
170	222	115	130	169
164	102	113	171	0
93	0	93	110	130

Assuming the population is distributed roughly symmetrically, construct a 95 per cent confidence interval for the mean annual consumption of alcoholic beverages by young European women.

14. The International Air Transport Association (IATA) surveys business travellers to develop quality ratings for international airports. The maximum possible rating is 10. Suppose a simple random sample of business travellers is selected and each traveller is asked to provide a rating for Kuwait International Airport. The ratings obtained from the sample of 50 business travellers follow. Construct a 95 per cent confidence interval estimate of the population mean rating for Kuwait International.

IATA

2	1	8	7	3	1	8	1	7	9	2	9	10	9	7	8	9
1	0	3	0	1	6	2	3	1	6	8	7	7	7	7	7	1
2	5	2	1	2	2	0	2	2	7	0	8	7	0	2	8	

15. Suppose a survey of 40 first-time home buyers finds that the mean of annual household income is €40 000 and the sample standard deviation is €15 300.

COMPLETE SOLUTIONS

a. At 95 per cent confidence, what is the margin of error for estimating the population mean household income?
b. What is the 95 per cent confidence interval for the population mean annual household income for first-time home buyers?

16. A sample of 30 fast-food restaurants including McDonald's and Burger King were visited. During each visit, the customer went to the drive-through and ordered a basic meal such as a burger, fries and drink. The time between pulling up to the order kiosk and receiving the filled order was recorded. The times in minutes for the 30 visits are as follows:

FASTFOOD

0.9	1.0	1.2	2.2	1.9	3.6	2.8	5.2	1.8	2.1	6.8	1.3	3.0	4.5	2.8
2.3	2.7	5.7	4.8	3.5	2.6	3.3	5.0	4.0	7.2	9.1	2.8	3.6	7.3	9.0

a. Provide a point estimate of the population mean drive-through time at fast-food restaurants.
b. At 95 per cent confidence, what is the margin of error?
c. What is the 95 per cent confidence interval estimate of the population mean?
d. Discuss skewness that may be present in this population. What suggestion would you make for a repeat of this study?

ACTTEMPS

17. A survey by Accountemps asked a sample of 200 executives to provide data on the number of minutes per day office workers waste trying to locate mislabelled, misfiled or misplaced items. Data consistent with this survey are contained in the data set 'ActTemps'.

a. Use 'ActTemps' to develop a point estimate of the number of minutes per day office workers waste trying to locate mislabelled, misfiled or misplaced items.
b. What is the sample standard deviation?
c. What is the 95 per cent confidence interval for the mean number of minutes wasted per day?

8.3 DETERMINING THE SAMPLE SIZE

We commented earlier on the role of the sample size in providing good approximate confidence intervals when the population is not normally distributed. In this section, we focus on another aspect of the sample size issue. We describe how to choose a sample size large enough to provide a desired margin of error. To understand how this process is done, we return to the σ known case presented in Section 8.1. Using expression (8.1), the interval estimate is $\bar{x} \pm z_{\alpha/2}\sigma/\sqrt{n}$. We see that $z_{\alpha/2}$, the population standard deviation σ, and the sample size n combine to determine the margin of error. Once we select a confidence coefficient $1 - \alpha$, $z_{\alpha/2}$ can be determined. Then, if we have a value for σ, we can determine the sample size n needed to provide any desired margin of error. Let $E =$ the desired margin of error.

$$E = z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Solving for $\sqrt{n}$, we have:

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

Squaring both sides of this equation, we obtain the following expression for the sample size.

Sample size for an interval estimate of a population mean

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2} \qquad \textbf{(8.3)}$$

This sample size provides the desired margin of error at the chosen confidence level.

In equation (8.3), E is the acceptable margin of error, and the value of $z_{\alpha/2}$ follows directly from the confidence level to be used. Although user preference must be considered, 95 per cent confidence is the most frequently chosen value ($z_{0.025} = 1.96$). Equation (8.3) can be used to provide a good sample size recommendation. However, the analyst should use judgement in deciding whether the recommendation given by equation (8.3) needs adjustment.

Use of equation (8.3) requires a value for the population standard deviation σ. However, even if σ is unknown, we can use equation (8.3) provided we have a preliminary or *planning value* for σ. In practice, one of the following procedures can be chosen:

1. Use an estimate of the population standard deviation computed from data of previous studies as the planning value for σ.
2. Use a pilot study to select a preliminary sample. The sample standard deviation from the preliminary sample can be used as the planning value for σ.
3. Use judgement or a 'best guess' for the value of σ. For example, we might begin by estimating the largest and smallest data values in the population. The difference between the largest and smallest values provides an estimate of the range for the data. The range divided by four is often suggested as a rough approximation of the standard deviation and hence an acceptable planning value for σ.

Consider the following example. A travel organization would like to conduct a study to estimate the population mean daily rental cost for a family car in Ireland. The director specifies that the population mean daily rental cost be estimated with a margin of error of €2 and a 95 per cent level of confidence. A previous study some years before had found a mean cost of approximately €80 per day for renting a family car, with a standard deviation of about €10.

The director specified a desired margin of error of $E = 2$, and the 95 per cent level of confidence indicates $z_{0.025} = 1.96$. We only need a planning value for the population standard deviation σ to compute the required sample size. Using €10 (from the previous study) as the planning value for σ, we obtain:

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2} = \frac{(1.96)^2(10)^2}{(2)^2} = 96.04$$

The sample size for the new study needs to be at least 96.04 family car rentals in order to satisfy the director's €2 margin-of-error requirement. In cases where the computed n is not an integer, we usually round up to the next integer value, in this case 97. Here, the sample size might be rounded for convenience to 100.

EXERCISES

Methods

COMPLETE SOLUTIONS

18. How large a sample should be selected to provide a 95 per cent confidence interval with a margin of error of 10? Assume that the population standard deviation is 40.

19. The range for a set of data is estimated to be 36.

a. What is the planning value for the population standard deviation?
b. At 95 per cent confidence, how large a sample would provide a margin of error of 3?
c. At 95 per cent confidence, how large a sample would provide a margin of error of 2?

Applications

20. Refer to the Scheer Industries example in Section 8.2. Use 6.82 days as a planning value for the population standard deviation.

a. Assuming 95 per cent confidence, what sample size would be required to obtain a margin of error of 1.5 days?
b. If the precision statement was made with 90 per cent confidence, what sample size would be required to obtain a margin of error of two days?

21. Suppose you are interested in estimating the average cost of staying for one night in a double room in a three-star hotel in France (outside Paris). Using €30.00 as the planning value for the population standard deviation, what sample size is recommended for each of the following cases? Use €3 as the desired margin of error.

a. A 90 per cent confidence interval estimate of the population mean cost.
b. A 95 per cent confidence interval estimate of the population mean cost.
c. A 99 per cent confidence interval estimate of the population mean cost.
d. When the desired margin of error is fixed, what happens to the sample size as the confidence level is increased? Would you recommend a 99 per cent confidence level be used? Discuss.

22. Suppose the price/earnings (P/E) ratios for stocks listed on a European Stock Exchange have a mean value of 35 and a standard deviation of 18. We want to estimate the population mean P/E ratio for all stocks listed on the exchange. How many stocks should be included in the sample if we want a margin of error of 3? Use 95 per cent confidence.

23. Fuel consumption tests are conducted for a particular model of car. If a 98 per cent confidence interval with a margin of error of 0.2 litres per 100km is desired, how many cars should be used in the test? Assume that preliminary tests indicate the standard deviation is 0.5 litres per 100km.

COMPLETE SOLUTIONS

24. In developing patient appointment schedules, a medical centre wants to estimate the mean time that a staff member spends with each patient. How large a sample should be taken if the desired margin of error is two minutes at a 95 per cent level of confidence? How large a sample should be taken for a 99 per cent level of confidence? Use a planning value for the population standard deviation of eight minutes.

8.4 POPULATION PROPORTION

We said earlier that the general form of an interval estimate of a population proportion π is: $p \pm$ margin of error. The sampling distribution of the sample proportion of plays a key role in computing the margin of error for this interval estimate.

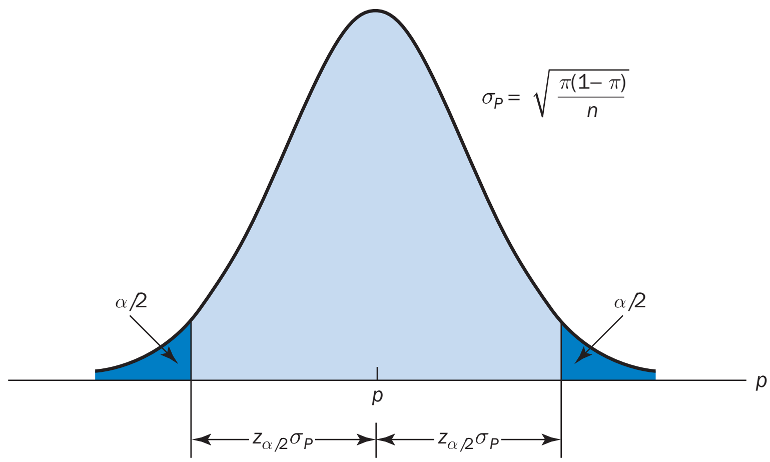
In Chapter 7 we said that the sampling distribution of the sample proportion P can be approximated by a normal distribution whenever $n\pi \geq 5$ and $n(1-\pi) \geq 5$. Figure 8.7 shows the normal approximation of the sampling distribution of P. The mean of the sampling distribution of P is the population proportion π, and the standard error of P is:

$$\sigma_P = \sqrt{\frac{\pi(1-\pi)}{n}} \tag{8.4}$$

If we choose $z_{\alpha/2}\sigma_P$ as the margin of error in an interval estimate of a population proportion, we know that $100(1-\alpha)$ per cent of the intervals generated will contain the true population proportion. But σ_P cannot be used directly in the computation of the margin of error because π will not be known; π is what we are trying to estimate. So, p is substituted for π and the margin of error for an interval estimate of a population proportion is given by:

$$\text{Margin of error} = z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \tag{8.5}$$

FIGURE 8.7
Normal approximation of the sampling distribution of P



The general expression for an interval estimate of a population proportion is:

Interval estimate of a population proportion

$$p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \qquad \textbf{(8.6)}$$

where $1 - \alpha$ is the confidence coefficient and $z_{\alpha/2}$ is the z value providing an area of $\alpha/2$ in the upper tail of the standard normal distribution.

Consider the following example. A national survey of 900 women golfers was conducted to learn how women golfers view their treatment at golf courses. (The data are available in the file 'TeeTimes' on the companion online platform.) The survey found that 396 of the women golfers were satisfied with the availability of tee times. So, the point estimate of the proportion of the population of women golfers who are satisfied is $396/900 = 0.44$. Using expression (8.6) and a 95 per cent confidence level,

TEETIMES

$$p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} = 0.44 \pm 1.96\sqrt{\frac{0.44(1-0.44)}{900}} = 0.44 \pm 0.0324$$

The margin of error is 0.0324 and the 95 per cent confidence interval estimate of the population proportion is 0.408 to 0.472. Using percentages, the survey results enable us to state with 95 per cent confidence that between 40.8 per cent and 47.2 per cent of all women golfers are satisfied with the availability of tee times.

Determining the sample size

The rationale for the sample size determination in constructing interval estimates of π is similar to the rationale used in Section 8.3 to determine the sample size for estimating a population mean.

Previously in this section we said that the margin of error associated with an interval estimate of a population proportion is $z_{\alpha/2}\sqrt{p(1-p)/n}$. The margin of error is based on the values of $z_{\alpha/2}$, the sample proportion p, and the sample size n. Larger sample sizes provide a smaller margin of error and better precision. Let E denote the desired margin of error:

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

Solving this equation for n provides a formula for the sample size that will provide a margin of error of size E:

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

However, we cannot use this formula to compute the sample size because p will not be known until after we select the sample. What we need, then, is a planning value for p. Using p^* to denote the planning value for p, the following formula can be used to compute the sample size.

Sample size for an interval estimate of a population proportion

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2} \tag{8.7}$$

In practice, the planning value can be chosen by one of the following procedures:

1. Use the sample proportion from a previous sample of the same or similar units.
2. Use a pilot study to select a preliminary sample. The sample proportion from this sample can be used as the planning value.
3. Use judgement or a 'best guess' for the value of p^*.
4. If none of the preceding alternatives apply, use a planning value of $p^* = 0.50$.

Let us return to the survey of women golfers and assume that the company is interested in conducting a new survey to estimate the current proportion of the population of women golfers who are satisfied with the availability of tee times. How large should the sample be if the survey director wants to estimate the population proportion with a margin of error of 0.025 at 95 per cent confidence? With $E = 0.025$ and $z_{\alpha/2} = 1.96$, we need a planning value p^* to answer the sample size question. Using the previous survey result of $p = 0.44$ as the planning value p^*, equation (8.7) shows that:

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2} = \frac{(1.96)^2(0.44)(1-0.44)}{(0.025)^2} = 1514.5$$

The sample size must be at least 1514.5 women golfers to satisfy the margin of error requirement. Rounding up to the next integer value indicates that a sample of 1515 women golfers is recommended to satisfy the margin of error requirement.

The fourth alternative suggested for selecting a planning value p^* is to use $p^* = 0.50$. This value of p^* is frequently used when no other information is available. To understand why, note that the numerator of equation (8.7) shows that the sample size is proportional to the quantity $p^*(1 - p^*)$. A larger value for this quantity will result in a larger sample size. Table 8.3 gives some possible values of $p^*(1 - p^*)$. Note that the largest value occurs when $p^* = 0.50$. So, in case of any uncertainty about an appropriate planning value, we know that $p^* = 0.50$ will provide the largest sample size recommendation. If the sample proportion turns out to be different from the 0.50 planning value, the margin of error will be smaller than anticipated. In effect, we play it safe by recommending the largest possible sample size.

In the survey of women golfers example, a planning value of $p^* = 0.50$ would have provided the sample size:

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2} = \frac{(1.96)^2(0.5)(1-0.5)}{(0.025)^2} = 1536.6$$

A slightly larger sample size of 1537 women golfers would be recommended.

TABLE 8.3 Some possible values for $p^*(1 - p^*)$

p^*	$p^*(1 - p^*)$	
0.10	$(0.10)(0.90) = 0.09$	
0.30	$(0.30)(0.70) = 0.21$	
0.40	$(0.40)(0.60) = 0.24$	
0.50	$(0.50)(0.50) = 0.25$	← Largest value for $p^*(1 - p^*)$
0.60	$(0.60)(0.40) = 0.24$	
0.70	$(0.70)(0.30) = 0.21$	
0.90	$(0.90)(0.10) = 0.09$	

EXERCISES

Methods

25. A simple random sample of 400 individuals provides 100 Yes responses.

- **a.** What is the point estimate of the proportion of the population that would provide Yes responses?
- **b.** What is your estimate of the standard error of the sample proportion?
- **c.** Compute a 95 per cent confidence interval for the population proportion.

26. A simple random sample of 800 elements generates a sample proportion $p = 0.70$.

- **a.** Provide a 90 per cent confidence interval for the population proportion.
- **b.** Provide a 95 per cent confidence interval for the population proportion.

COMPLETE SOLUTIONS

27. In a survey, the planning value for the population proportion is $p^* = 0.35$. How large a sample should be taken to provide a 95 per cent confidence interval with a margin of error of 0.05?

28. At 95 per cent confidence, how large a sample should be taken to obtain a margin of error of 0.03 for the estimation of a population proportion? Assume that past data are not available for developing a planning value for p.

Applications

29. A survey of 611 office workers investigated telephone answering practices, including how often each office worker was able to answer incoming telephone calls and how often incoming telephone calls went directly to voice mail. A total of 281 office workers indicated that they never need voice mail and are able to take every telephone call.

- **a.** What is the point estimate of the proportion of the population of office workers who are able to take every telephone call?
- **b.** At 90 per cent confidence, what is the margin of error?
- **c.** What is the 90 per cent confidence interval for the proportion of the population of office workers who are able to take every telephone call?

30. The French market research and polling company CSA carried out surveys to investigate job satisfaction among professionally qualified employees of private companies. A total of 629 professionals were involved in the surveys, of whom 195 said that they were dissatisfied with their employer's recognition of their professional experience.

a. What is the point estimate of the proportion of the population of employees who were dissatisfied with their employer's recognition of their professional experience?
b. At 95 per cent confidence, what is the margin of error?
c. What is the 95 per cent confidence interval for the proportion of the population of employees who were dissatisfied with their employer's recognition of their professional experience?

31. In a sample of 162 companies, 104 reported profits that beat prior estimates, 29 matched estimates and 29 fell short of prior estimates.

a. What is the point estimate of the proportion that fell short of estimates?
b. Determine the margin of error and provide a 95 per cent confidence interval for the proportion that fell short of estimates.
c. How large a sample is needed if the desired margin of error is 0.05?

32. In early December 2008, the Palestinian Center for Policy and Survey Research carried out an opinion poll among adults in the West Bank and Gaza Strip. Respondents were asked their opinion about the chance of an independent Palestinian state being established alongside Israel in the next five years. Among the 1270 respondents, 34.6 per cent felt there was no chance of this happening.

a. Provide a 95 per cent confidence interval for the population proportion of adults who thought there was no chance of an independent Palestinian state being established alongside Israel in the next five years.
b. Provide a 99 per cent confidence interval for the population proportion of adults who thought there was no chance of an independent Palestinian state being established alongside Israel in the next five years.
c. What happens to the margin of error as the confidence is increased from 95 per cent to 99 per cent?

33. In a survey conducted by ICM Research in the UK, 710 out of 1000 adults interviewed said that, if there were to be a referendum, they would vote for the UK not to join the European currency (the euro). What is the margin of error and what is the interval estimate of the population proportion of British adults who would vote for the UK not to join the European currency? Use 95 per cent confidence.

34. A well-known bank credit card firm wishes to estimate the proportion of credit card holders who carry a non-zero balance at the end of the month and incur an interest charge. Assume that the desired margin of error is 0.03 at 98 per cent confidence.

COMPLETE SOLUTIONS

a. How large a sample should be selected if it is anticipated that roughly 70 per cent of the firm's cardholders carry a non-zero balance at the end of the month?
b. How large a sample should be selected if no planning value for the proportion could be specified?

ONLINE RESOURCES

For the data files, additional online summary, questions, answers and software section go to the accompanying online platform.

SUMMARY

In this chapter we introduced the idea of an interval estimate of a population parameter. A point estimator may or may not provide a good estimate of a population parameter. The use of an interval estimate provides a measure of the precision of an estimate. A common form of interval estimate is a confidence interval.

We presented methods for computing confidence intervals of a population mean and a population proportion. Both are of the form: point estimate $\pm$ margin of error. The confidence interval has a confidence coefficient associated with it.

We presented interval estimates for a population mean for two cases. In the σ known case, historical data or other information is used to make an estimate of σ prior to taking a sample. Analysis of new sample data then proceeds based on the assumption that σ is known. In the σ unknown case, the sample data are used to estimate both the population mean and the population standard deviation. In the σ known case, the interval estimation procedure is based on the assumed value of σ and the use of the standard normal distribution. In the σ unknown case, the interval estimation procedure uses the sample standard deviation *s* and the *t* distribution.

In both cases the quality of the interval estimates obtained depends on the distribution of the population and the sample size. Practical advice about the sample size necessary to obtain good approximations was included in Sections 8.1 and 8.2.

The general form of the interval estimate for a population proportion is *p* $\pm$ margin of error. In practice, the sample sizes used for interval estimates of a population proportion are generally large. Consequently, the interval estimation procedure is based on the standard normal distribution.

We explained how the expression for margin of error can be used to calculate the sample size required to achieve a desired margin of error at a given level of confidence. We did this for two cases: estimating a population mean when the population standard deviation is known, and estimating a population proportion.

KEY TERMS

Confidence coefficient
Confidence interval
Confidence level
Degrees of freedom
Interval estimate
Margin of error
σ known
σ unknown
***t* distribution**

KEY FORMULAE

Interval estimate of a population mean: σ known

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{8.1}$$

Interval estimate of a population mean: σ unknown

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}} \tag{8.2}$$

Sample size for an interval estimate of a population mean

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{E^2} \tag{8.3}$$

Interval estimate of a population proportion

$$p \pm z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \tag{8.6}$$

Sample size for an interval estimate of a population proportion

$$n = \frac{(z_{\alpha/2})^2 p^*(1-p^*)}{E^2} \tag{8.7}$$

CASE PROBLEM 1

International bank

The manager of a city-centre branch of a well-known international bank commissioned a customer satisfaction survey. The survey investigated three areas of customer satisfaction: their experience waiting for service at a till, their experience being served at the till and their experience of self-service facilities at the branch. Within each of these categories, respondents to the survey were asked to give ratings on a number of aspects of the bank's service. These ratings were then summed to give an overall satisfaction rating in each of the three areas of service. The summed ratings are scaled such that they lie between 0 and 100, with 0 representing extreme dissatisfaction and 100 representing extreme satisfaction. The data file for this case study ('IntlBank' on the online platform) contains the 0–100 ratings for the three areas of service, together with particulars of respondents' gender and whether they would recommend the bank to other people (a simple Yes/No response was required to this question). A table containing the first few rows of the data file is shown.

INTNLBANK

Waiting	*Service*	*Self-service*	*Gender*	*Recommend*
55	65	50	male	no
50	80	88	male	no
30	40	44	male	no
65	60	69	male	yes
55	65	63	male	no
40	60	56	male	no
15	65	38	male	yes
45	60	56	male	no
55	65	75	male	no
50	50	69	male	yes

Managerial report

1. Use descriptive statistics to summarize each of the five variables in the data file (the three service ratings, customer gender and customer recommendation).
2. Calculate a 95 per cent confidence interval estimate of the mean service rating for the population of customers of the branch, for each of the three service areas. Provide a managerial interpretation of each interval estimate.
3. Calculate a 95 per cent confidence interval estimate of the proportion of the branch's customers who would recommend the bank, and a 95 per cent confidence interval estimate of

the proportion of the branch's customers who are female. Provide a managerial interpretation of each interval estimate.

4. Suppose the branch manager required an estimate of the percentage of branch customers who would recommend the branch within a margin of error of 3 percentage points. Using 95 per cent confidence, how large should the sample size be?
5. Suppose the branch manager required an estimate of the percentage of branch customers who are female within a margin of error of 5 percentage points. Using 95 per cent confidence, how large should the sample size be?

CASE PROBLEM 2

Consumer Knowhow

Consumer Knowhow is a consumer research organization that conducts surveys designed to evaluate a wide variety of products and services available to consumers. In one particular study, Consumer Knowhow looked at consumer satisfaction with the performance of cars produced by a major European manufacturer. A questionnaire sent to owners of one of the manufacturer's family cars revealed several complaints about early transmission problems.

To learn more about the transmission failures, Consumer Knowhow used a sample of transmission repairs provided by a transmission repair firm located near the manufacturing plant. The data in the file 'Repairs' are the kilometres driven for 50 cars at the time of transmission failure.

REPAIRS

Managerial report

1. Use appropriate descriptive statistics to summarize the transmission failure data.
2. Construct a 95 per cent confidence interval for the mean number of kilometres driven until transmission failure for the population of cars with transmission failure. Provide a managerial interpretation of the interval estimate.
3. Discuss the implication of your statistical findings in relation to the proposition that some owners of the cars experienced early transmission failures.
4. How many repair records should be sampled if the research company wants the population mean number of kilometres driven until transmission failure to be estimated with a margin of error of 5000 kilometres? Use 95 per cent confidence.
5. What other information would you like to gather to evaluate the transmission failure problem more fully?