12.

Sample means

- Sampling
- Distribution of sample means
- Increasing the number of samples
- Increasing the sample size
- The central limit theorem
- Sample size
- Samples with means that are unusually high or unusually low
- Inferring population parameters from sample statistics
- So do most 95% confidence intervals really contain the population mean? Let’s check
- Choosing the sample size
- Miscellaneous exercise twelve

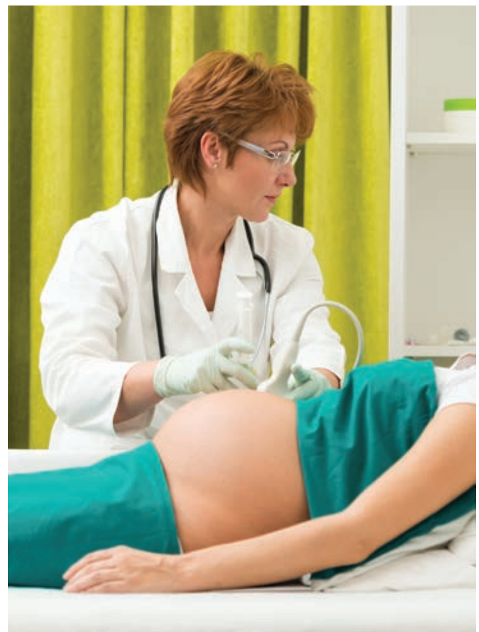
Note

For the purposes of this chapter it is anticipated that, from your study of *Mathematics Methods*, you are now familiar with the idea of probability distributions and the **binomial**, **uniform** and **normal distributions** in particular.

Sampling

In some cases it is simply too expensive, too inconvenient or just unwise to obtain information about a particular set by testing all of the elements that make up the set.

For example, if a doctor wishes to check on the progress of a human pregnancy they may carry out an amniocentesis on the mother. This process involves the doctor in taking, and having analysed, a sample of the amniotic fluid (the fluid surrounding the baby) from the mother. In this case, a sample of the amniotic fluid should be taken, not all of it!



iStock.com/Topalov

Shutterstock.com/Shulevskyy Volodymyr

If a commercial cherry grower wants to test how nice his cherries are he might taste a few – he would not eat them all!

Judgements can then be made about the **population** as a whole based on what we find from the sample. In this chapter we will consider making judgements about what the **mean** of a population might be (be it mean height, mean weight, mean sugar content or whatever) from the mean of that quantity found in a sample that we take.

If you are also studying *Mathematics Methods* Unit Four at this time it is likely that you are considering, or are soon to consider, the idea of predicting a *population proportion* based on a sample proportion. This prediction of a population characteristic follows very similar ideas whether we are predicting a population proportion, as in *Mathematics Methods*, or predicting a population mean, as in this unit of *Mathematics Specialist*.

Numerical characteristics about an entire population, for example, the proportion of Australians who are left-handed, or the mean length of the crocodiles in an area, are called **population parameters**. Predicting a value for a population parameter, based on the equivalent **sample statistic**, is often why we collect data from a sample.

The bigger our sample the more confident we can be that information from the sample reflects the same information about the population.

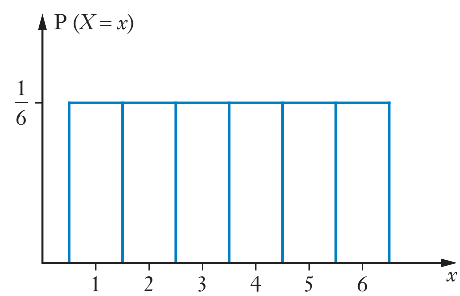
We might consider taking a number of samples to assess whether our one sample is typical of others and to 'get a feel for' the variability that may exist between samples.

If we wanted information about Australian school children we might use a sample involving, say, 500 Australian school children, and use the information from this sample to suggest information about the entire population of Australian school children.

If our sample of 500 Australian school children is a **random sample** then it is constructed in such a way that each Australian school child has an equal chance of being in the sample of 500.

Distribution of sample means

Suppose we want to investigate the mean value obtained when a normal six-sided die is rolled many times.



We know from the symmetry of the situation that each of the numbers

$$1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6$$

has an equal chance of occurring on each roll.

The mean or expected value, i.e. the long term average, of this uniform distribution is 3.5 (and the standard deviation is $\sqrt{\frac{35}{12}}$ or, to two decimal places, 1.71).

Suppose we were to investigate this situation by sampling just 4 rolls of a normal die.

Question: Would the mean of the 4 rolls necessarily be 3.5?

Answer: Whilst the mean could be 3.5 it does not have to be 3.5.
i.e. the mean is *not necessarily* equal to 3.5.

One such sample of 4 rolls of a normal die gave

$$4, \quad 6, \quad 3, \quad 2. \qquad \text{Mean} = 3.75$$

Suppose we again roll the die 4 times.

Question: Would the mean of the 4 rolls again be 3.75?

Answer: Not necessarily.

The next sample of 4 rolls of a normal die gave

$$6, \quad 1, \quad 2, \quad 1. \qquad \text{Mean} = 2.5$$

If we were to continue rolling the die 4 times and were to calculate the mean score each time, we would expect some variation to occur in these mean scores, i.e., we would expect a **distribution of sample means**, the first two of which were 3.75 and 2.5.

 ISBN 9780170395274

Suppose we continue the process to obtain 30 samples of 4 rolls, i.e. 30 samples of 'sample size' 4:

4	6	3	2	mean 3.75	6	6	4	1	mean 4.25
6	1	2	1	2.5	3	1	1	4	2.25
1	5	4	5	3.75	4	3	5	1	3.25
2	5	1	2	2.5	6	1	1	4	3
6	5	3	6	5	5	4	3	4	4
4	1	3	2	2.5	1	5	3	5	3.5
6	2	6	6	5	1	1	3	3	2
6	6	2	4	4.5	3	5	6	6	5
2	2	1	4	2.25	6	6	2	1	3.75
4	4	1	2	2.75	2	6	3	4	3.75
3	5	5	4	4.25	6	1	4	2	3.25
1	2	1	3	1.75	5	5	4	3	4.25
2	6	5	4	4.25	1	6	1	5	3.25
1	2	3	4	2.5	5	4	4	2	3.75
2	2	2	2	2	5	4	1	2	3

For these 30 samples of 4 rolls, our **distribution of sample means** is as follows:

Mean	1.75	2	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	5
Frequency	1	2	2	4	1	2	3	1	5	1	4	1	0	3

Mean of sample means (2 decimal places): 3.38

Standard deviation of sample means (2 decimal places): $\sigma_n = 0.93$, $\sigma_{n-1} = 0.95$.

Grouping the data:

Mean ($\bar{x}$)	$1 \le \bar{x} < 2$	$2 \le \bar{x} < 3$	$3 \le \bar{x} < 4$	$4 \le \bar{x} < 5$	$5 \le \bar{x} < 6$
Frequency	1	9	11	6	3

Note:
- Whilst dice-rolling involves a uniform distribution of outcomes the distribution of sample means does not display this same uniformity. Most of our means are close to the population mean of 3.5.
- The standard deviation of a sampling distribution of mean values is sometimes referred to as the **standard error** of the mean.

Increasing the number of samples

Suppose we increase the number of samples of 4 rolls of the die from 30 to, say 100, or even 200, to get more sample means.

How do you think the mean and the standard deviation of the sample means would compare to the
3.38 (mean)
and 0.93 (standard deviation)
values obtained above for 30 samples?

ISBN 9780170395274

The following statistics were obtained for 100 and 200 samples of 4 rolls:

For 100 samples of four rolls:	Mean of sample means	3.523
	Standard deviation of sample means	$\sigma_{n-1} = 0.847$, $\sigma_n = 0.843$
For 200 samples of four rolls:	Mean of sample means	3.518
	Standard deviation of sample means	$\sigma_{n-1} = 0.836$, $\sigma_n = 0.834$

Do these figures support what you thought would be the case?

Increasing the sample size

Suppose we increase the sample size from samples involving 4 rolls of the die to samples involving 10 rolls of the die, or perhaps 35 rolls of the die or even 100 rolls of the die.

How would you expect this increased sample size to affect the mean and the standard deviation of the sample means?

The sample means for 30 samples each involving 10 rolls (i.e. sample size 10) were:

4.1	4.1	3	3.7	2.9	4	3.8	3.9	2.6	3.8
3.9	4.2	3.8	3.6	3.8	2.9	2.9	3.5	3.5	4.4
3.9	3.4	3.7	3.8	3.2	4.1	3.2	3.4	3.2	2.7

Mean of sample means: 3.567
Standard deviation of sample means: $\sigma_{n-1} = 0.477$, $\sigma_n = 0.469$

The sample means for 30 samples each involving 35 rolls (i.e. sample size 35) were:

4.057	3.371	3.114	3.571	3.228	3.429	4.086	3.227	3.457	3.4
3.757	3.2	3.457	3.029	3.171	3.657	2.886	3.286	3.086	3.057
3.771	3.629	3.457	3.6	3.343	3.086	3.743	3.029	3.571	3.629

Mean of sample means: 3.413
Standard deviation of sample means: $\sigma_{n-1} = 0.303$, $\sigma_n = 0.298$

The sample means for 30 samples each involving 100 rolls (i.e. sample size 100) were:

3.31	3.45	3.66	3.54	3.43	3.71	3.32	3.48	3.36	3.42
3.25	3.52	3.35	3.44	3.53	3.54	3.5	3.53	3.34	3.64
3.57	3.29	3.41	3.89	3.54	3.23	3.47	3.64	3.2	3.68

Mean of sample means: 3.475
Standard deviation of sample means: $\sigma_{n-1} = 0.157$, $\sigma_n = 0.155$

Notice that the mean of the sample means is quite close to 3.5, the mean of the uniform distribution from which the samples were drawn. Also note that in each case the standard deviation of the sample means is less than 1.71, the standard deviation of the uniform distribution from which the samples were drawn, and that as the sample size increased the standard deviation decreased. The greater the sample size, the less variation there is in the sample means.

ISBN 9780170395274

The previous paragraph made two important claims, based on the results of the previous pages:

- The sample means are quite close to the mean of the distribution from which the samples were drawn, in this case 3.5.
- The greater the sample size, the less variation there is in the sample means, i.e. as the sample size increased the standard deviation of the sample means decreased.

Let us take this second statement even further and suggest that for a sample size of n the standard deviation of the sample means is close to $\frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation of the population.

In this dice rolling situation $\sigma = 1.7078$, correct to 4 decimal places, so this further suggestion would mean that:

For a sample size of 4 the standard deviation would be approximately 0.85.
For a sample size of 10 the standard deviation would be approximately 0.54.
For a sample size of 35 the standard deviation would be approximately 0.29.
For a sample size of 100 the standard deviation would be approximately 0.17.

Compare these theoretical figures for the standard deviations with those actually obtained as shown on the previous page.

Okay, now it's time that you checked the above claims based on data you obtain, rather than data presented to you.

In a moment you will roll four dice (or one die four times), find the mean of the four scores obtained, and note the result. Repeating this a further 99 times will give you 100 mean values, each from samples of size 4.

You can then check if your results agree with the above statements, i.e. is the mean of your 100 sample means close to 3.5, the mean of the distribution from which the samples are drawn, and is your standard deviation of the 100 means close to $\frac{1.7078}{\sqrt{4}}$?

Then you can check for sample sizes of 20, 50, …

But first let us consider how best to proceed with this data collection.

One way would be to physically roll four dice, note the four scores, find the mean, record the result and then repeat the process until you have 100 sample means recorded.

Alternatively you could **simulate** such an activity using the random number generator of some calculators, spreadsheets or internet sites, as the next page explains.

iStock.com/ThomasVogel

Some calculators can randomly generate a list consisting of a specified number of integers in a given range, as the first line of the display shown on the right suggests. The second line then calculates the mean of these four numbers.

Alternatively generate four numbers and find their mean 'in one go', as the third line suggests.

```
randList(4,1,6)
                                {5,5,4,1}
sum(ans) / 4
                                     3.75
sum(randList(4,1,6)) / 4
                                      2.5
```

Twenty mean values generated in this way are shown below:

4	4	3	2.25	3.75	2.75	4.75	3.5	3.25	2.25
3.25	3	4.25	3.5	3.25	3.75	4.75	3	2.5	3.25

These same twenty mean values presented as a frequency table:

Value	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75
Frequency	2	1	1	3	4	2	2	2	1	0	2

Analysing these values with a calculator gives $\bar{x} = 3.4$, $\sigma_n = 0.7$, $\sigma_{n-1} = 0.718$.

Alternatively, using a computer spreadsheet:

RANDBETWEEN(1,6)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
	5	5	6	5	6	2	6	4	2	4	3	5	3	3	6	4	3	2	4	5
	3	2	1	4	1	6	5	6	4	4	3	3	4	5	5	1	5	4	3	6
	3	3	3	5	5	6	5	3	1	4	5	2	6	3	6	1	3	1	2	4
	1	3	1	3	2	1	4	2	5	4	4	3	1	5	2	4	3	3	3	5
Mean	3	3.25	2.75	4.25	3.5	3.75	5	3.75	3	4	3.75	3.25	3.5	4	4.75	2.5	3.5	2.5	3	5
Mean	2.25	2.5	2.75	3	3.25	3.5	3.75	4	4.25	4.5	4.75	5								
Count	0	2	1	3	2	3	3	2	1	0	1	2								
Mean of sample means						3.6														
Standard deviation σ					n	0.73	n–1	0.75												

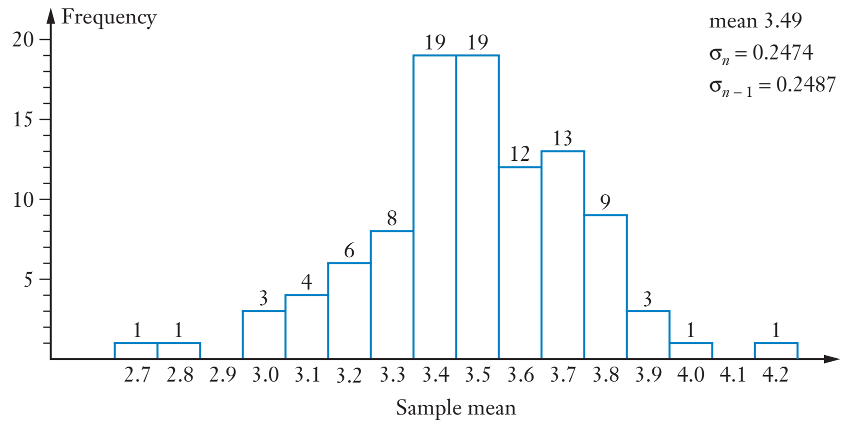
Calculators and spreadsheets can be instructed to display and analyse lists of random numbers, and their mean values, in a number of ways. There are also interactive websites that will display sample means. The reader is encouraged to investigate the capability of their own calculator, spreadsheets and the internet in this regard.

Now carry out the activity mentioned on the previous page:

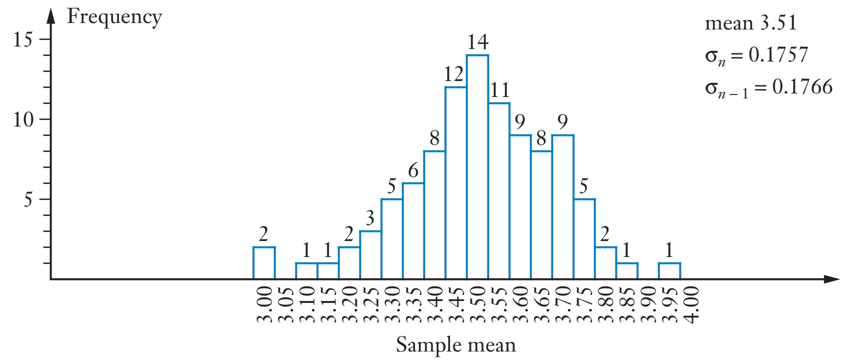
- Find the mean score obtained for rolling four normal dice (simulated), and repeat this activity to give 100 such sample means (i.e. obtain 100 sample means for sample size 4), recording the sample means in a frequency table. Analyse your results and see if the sample means have a mean close to 3.5 and a standard deviation close to $\frac{1.7078}{\sqrt{n}}$, in this case for $n = 4$.
- Repeat the activity but now obtaining 100 sample means for sample size 20.
- Repeat the activity but now obtaining 100 sample means for sample size 50.

ISBN 9780170395274

The graph below shows the distribution of sample means for one particular set of 100 samples of sample size 50. (The graph involves grouped data but the summary statistics shown have been calculated from the original 100 sample means.)



The graph below shows the distribution of sample means for one particular set of 100 samples of sample size 100. (The graph involves grouped data but the summary statistics shown have been calculated from the original 100 sample means.)



The central limit theorem

According to the central limit theorem, as the sample size, n, increases:

- **the distribution of sample means approaches a normal distribution**
- **the mean of the sample means, $\bar{x}$, approaches the population mean**
- **the standard deviation of the sample means approaches $\frac{\sigma}{\sqrt{n}}$ where σ is the standard deviation of the population**

We have seen that the above statements seem to be the case when the samples are drawn from a uniform distribution. Will they also be the case if the samples are drawn from other distributions, for example, the normal distribution itself?

Some calculators and spreadsheets can generate random numbers from normal distributions. Investigate sample means from such a distribution. Do the three bold statements above still seem to be the case?

Sample size

The central limit theorem applies irrespective of the type of distribution the samples are drawn from. However, note the following points regarding the sample size needed for the 'approximation to normal' to be valid.

- If the population from which our samples are drawn is not normally distributed then, as an approximate 'rule of thumb', we assume that for $\boldsymbol{n \geq 30}$ the sample means can be well-modelled by a normal distribution.
- If the population from which our samples are drawn is normally distributed we can assume that the sample means will be normally distributed for all n.

EXAMPLE 1

Note: For the repeated rolling of a normal die we expect a long term mean of 3.5 and standard deviation 1.71.

A computer is used to simulate the rolling of a normal fair six-sided die 100 times and to calculate and record the mean of these 100 scores.

The computer carries out this process 250 times.

a How would you expect the 250 mean scores to be distributed?

b How would you expect the 250 mean scores to be distributed if instead the repeated simulation involved 200 rolls of the die?

Solution

We are sampling from a population with mean 3.5 and standard deviation 1.71.

a The sample size is 100.

From the central limit theorem the sample means will be approximately normally distributed with a mean of 3.5 and a standard deviation of $\frac{1.71}{\sqrt{100}} = 0.171$.

b The sample size is 200.

From the central limit theorem the sample means will be approximately normally distributed with a mean of 3.5 and a standard deviation of $\frac{1.71}{\sqrt{200}} = 0.121$.

Knowing that the sample means **form a distribution that is approximately normal** allows us to use our understanding of the normal distribution:

- to determine the likelihood of the mean of a particular sample lying in a given interval

and

- to thereby gain some indication of whether a particular sample mean seems unexpectedly high or unexpectedly low.

ISBN 9780170395274

EXAMPLE 2

A random variable X is known to have a mean of 65
and a standard deviation of 8.

A random sample involves 100 measurements of X.

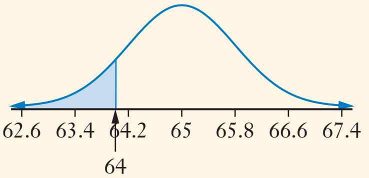
What is the probability that the mean of such a sample is less than 64?

Solution

From the central limit theorem we can assume that the distribution of sample means will be approximately normal with mean 65 and standard deviation $\frac{8}{\sqrt{100}}$, i.e. the sample means are normally distributed with a mean of 65 and a standard deviation of 0.8.

Using a calculator, if $Y \sim N(65, 0.8^2)$

$P(Y < 64) = 0.1056$



The probability that the mean of such a sample is less than 64 is approximately 0.11.

EXAMPLE 3

A random variable X is normally distributed with a mean of 27.6
and a standard deviation of 3.6.

A random sample involves 16 measurements of X.

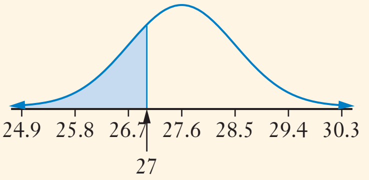
What is the probability that the mean of such a sample is less than 27?

Solution

Whilst the sample size is small the sample means will be normally distributed because the population from which the sample is drawn is normally distributed. Thus the sample means will be normally distributed with mean 27.6 and standard deviation $\frac{3.6}{\sqrt{16}}$, i.e., $N(27.6, 0.9^2)$.

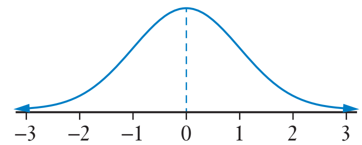
Using a calculator, if $Y \sim N(27.6, 0.9^2)$

$P(Y < 27) = 0.2525$



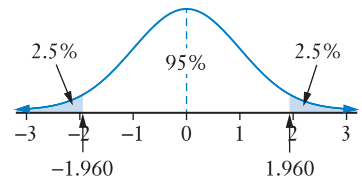
The probability that the mean of such a sample is less than 27 is approximately 0.25.

Samples with means that are unusually high or unusually low



From your studies of *Mathematics Methods* Unit Four you should be familiar with the **standard normal distribution** being a normal distribution with mean 0 and standard deviation 1, i.e. $Z \sim \mathrm{N}(0, 1^2)$, and with its associated 'z scores'.

$$\begin{aligned} \text{Solving} \quad \mathrm{P}(Z < k) &= 0.975 \\ \text{gives} \quad k &= 1.960 \end{aligned}$$



Thus for a normal distribution 95% of the distribution lies within 1.96 standard deviations of the mean.

With the sample means belonging to a normal distribution, if a sample mean is found to lie further than 1.96 standard deviations from the mean of this distribution we may wonder if the sample is unusual in some way. (In such a case a possible course of action might be to carry out further sampling.)

When a sample mean is outside of this 95% interval we say that this particular sample mean is *significantly different* from the expected mean *at the 5% level*.

We say there is: **a significant difference at the 5% level.**

I.e., the fact that the mean of our sample is at one of the extremes, where we would expect only 5% of the sample means to lie, is significantly unusual.

EXAMPLE 4

A company manufacturing breakfast cereal claims that the weight of cereal in packets of a particular brand of their cereal, each claiming to contain 500 grams of the cereal, is normally distributed with mean 504.3 grams and standard deviation 2.1 grams.

An independent survey samples such packets of cereal to check this claim.

In each of the following cases, determine whether the sample mean is significantly different to the expected sample mean at the 5% level.

a sample size 30, sample mean 504.9 g,

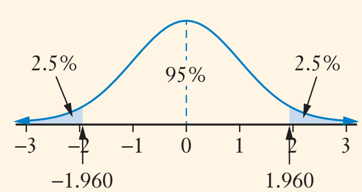
b sample size 100, sample mean 504.9 g.

Solution

a We would expect the sample means to be normally distributed with mean 504.3 g and standard deviation $\frac{2.1}{\sqrt{30}}$ g. Hence 95% of the sample means will lie between

$$504.3 - 1.96 \times \frac{2.1}{\sqrt{30}} \quad \text{and} \quad 504.3 + 1.96 \times \frac{2.1}{\sqrt{30}}$$

i.e. between 503.55 and 505.05.



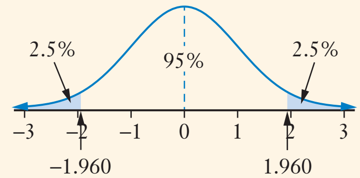
With our sample mean of 504.9 g lying within this interval there is not a significant difference at the 5% level.

ISBN 9780170395274

b We would expect the sample means to be normally distributed with mean 504.3 g and standard deviation $\dfrac{2.1}{\sqrt{100}}$ g. Hence 95% of the sample means will lie between

$$504.3 - 1.96 \times \frac{2.1}{\sqrt{100}} \quad \text{and} \quad 504.3 + 1.96 \times \frac{2.1}{\sqrt{100}}$$

i.e. between 503.89 and 504.71.



With our sample mean of 504.9 g lying outside of this interval there is a significant difference at the 5% level.

- Alternatively, rather than setting up the 95% interval, the above example could be answered by comparing P(mean ≥ 504.9 g) to 2.5% or by seeing how many standard deviations 504.9 g is from 504.3 g.
- The values for the central 95% interval can be determined directly from some calculators. Explore the ability of your calculator in this regard.

Exercise 12A

1 Note: For the repeated rolling of a normal die we expect a long-term mean of 3.5 and standard deviation 1.71.

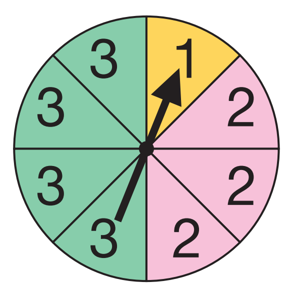
A computer is used to simulate the rolling of a normal fair six-sided die 50 times and to calculate and record the mean of these 50 scores.

The computer carries out this task 200 times.

How would you expect the 200 mean scores to be distributed?

How would the distribution of means differ if instead the repeated simulation involved 150 rolls of the die each time?

2 If we define the random variable X as the number obtained from the spinner on the right, the probability distribution for X is as shown below.



x	1	2	3
$P(X = x)$	0.125	0.375	0.5

Use your calculator to confirm that the mean and standard deviation of this distribution are respectively 2.375 and 0.696 (to three decimal places).

The spinner is spun 60 times and the mean of the 60 numbers obtained is determined.

If this process were repeated a large number of times how would the mean values be distributed?

How would the distribution compare with the previous case if instead the process involved repeatedly calculating the mean of 100 spins?

ISBN 9780170395274

3 If we define the random variable X as the number obtained when two normal dice are rolled and the numbers on the uppermost faces are added together, X has the probability distribution shown below:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Use your calculator to confirm that the mean and standard deviation of this distribution are respectively 7 and 2.415 (to three decimal places).

A computer is used to simulate the rolling of two normal dice 36 times, each time recording the sum of the two numbers obtained, and then calculating the mean of the 36 numbers. The computer carries this simulation of 36 rolls of the two dice 100 times.

How would you expect the 100 mean scores to be distributed?

How would the distribution of the 100 means compare with the previous case if instead each simulation involved 120 rolls of the two dice rather than 36?

iStock.com/tatniz

4 A random variable X is known to have a mean of 2145
and a standard deviation of 132.

A random sample involves 50 measurements of X.

What is the probability that the mean of such a sample is less than 2175?

5 A random variable X is known to have a mean of 16.8
and a standard deviation of 3.2.

A random sample involves 64 measurements of X.

What is the probability that the mean of such a sample is more than 17.5?

6 A random variable X is known to have a mean of 145
and a standard deviation of 20.

A random sample involves 100 measurements of X.

What is the probability that the mean of such a sample lies between 144 and 150?

7 A random variable, X, is such that $X \sim N(5, 1)$.

The random variable Y is the mean of 25 random measurements of X.

a What will be the distribution of Y?

b Find $P(Y \geq 5.5)$.

8 A random variable, X, is such that $X \sim Bin(50, 0.6)$.

The random variable Y is the mean of 50 random measurements of X.

What will be the distribution of Y?

Note: Bin (n, p) has mean np and standard deviation $\sqrt{npq}$ where $q = (1 - p)$.

ISBN 9780170395274

9 For a particular brand of breakfast cereal the weight of cereal in packets that claim to contain 500 grams is actually normally distributed with a mean of 508 g and standard deviation 3 g.

a What percentage of packets will contain less than 500 g?

b What percentage of packets will contain more than 510 g?

c A random sample of 10 packets is selected and the weight of cereal in each packet is measured. What is the probability that the mean of these ten weights is 508 grams when rounded to the nearest gram?

10 A company claims that the weights of its 1 kg bags of tomatoes are normally distributed with mean 1.015 kg and standard deviation 0.006 kg.

Assuming that the company's claim is true:

a What is the probability of a randomly-chosen bag containing less than 1 kg?

b What is the probability of a randomly-chosen bag containing more than 1.02 kg?

c If five bags are randomly selected what is the probability that the mean weight of the five bags will be between 1.01 kg and 1.02 kg?

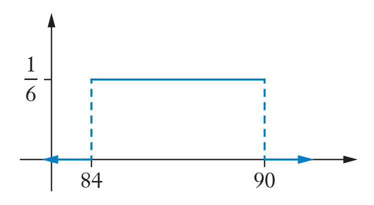
11 Scientists feel confident that the population of adult male lizards of a particular species have lengths that are normally distributed with a mean of 17.4 cm and standard deviation 2.1 cm.

A sample of ten adult male lizards of this species are caught and their lengths measured.

The mean length of this sample of ten is found to be 19.4 cm.

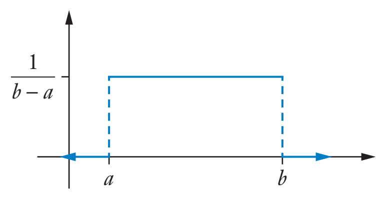
Comment upon this result.

12 A continuous random variable X is uniformly distributed on the interval $84 \rightarrow 90$.



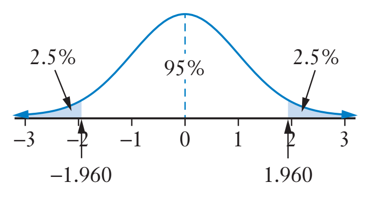
a Find $P(88 < X < 90)$.

b Use the fact that a continuous random variable, uniformly distributed between the values a and b, has a standard deviation of $\frac{b-a}{\sqrt{12}}$ to determine the standard deviation of X.



c The random variable Z is the mean of 48 randomly chosen values of X. Find $P(Z > 87.5)$.

13 For a large sample of given size the population of sample means is normally distributed. If we take a sample of this size and find that the mean of the sample lies further than 1.96 standard deviations from the mean of this normal distribution we say that there is *a significant difference at the 5% level*.



A continuous random variable X has mean 513 and standard deviation 26. Values of X are randomly sampled. For each of the following determine, with reasoning, whether there is *a significant difference at the 5% level*.

a First sample taken: Sample size 64, sample mean 505.

b Second sample taken: Sample size 100, sample mean 510.

Population parameters

Inferring population parameters from sample statistics

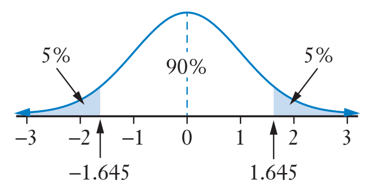
If we wish to use a sample to estimate the mean of the population from which the sample was drawn we could simply assume the population mean to be the same as the sample mean. However, we do have a problem with doing this because we would expect there to be some variation in sample means. If we used a different sample, drawn from the same population, we would have a different sample mean and therefore a different estimate of the population mean. However, from the central limit theorem, we know that sample means from large samples will be normally distributed, or approximately normally distributed. Hence we can use our understanding of the normal distribution to give a range of values that we can, with a particular level of confidence, expect the population mean to lie within.

First let us revisit the **standard normal distribution**, i.e. $Z \sim N(0, 1^2)$, with its associated 'z scores' and establish some important numbers relating to the 90%, 95% and 99% confidence levels (one of which we used when considering samples with means that were unusually high or unusually low).

For a *90% confidence interval*

Solving $P(Z < k) = 0.95$

gives $k = 1.6449$

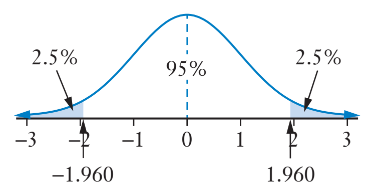


90% of the scores from a normal distribution lie within 1.645 standard deviations of the mean.

For a *95% confidence interval*

Solving $P(Z < k) = 0.975$

gives $k = 1.9600$

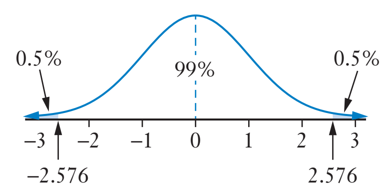


95% of the scores from a normal distribution lie within 1.960 standard deviations of the mean.

For a *99% confidence interval*

Solving $P(Z < k) = 0.995$

gives $k = 2.5758$



99% of the scores from a normal distribution lie within 2.576 standard deviations of the mean.

Shutterstock.com/ Denis Cristo

ISBN 9780170395274

From the central limit theorem we know that the mean of our single sample comes from a distribution of sample means that approximate a normal distribution with a mean equal to the mean of the population. Hence we can be 90% confident that our sample mean is within 1.645 standard deviations of the population mean, (1.645 being the critical score for the 90% confidence interval). Therefore:

We can be 90% confident that the population mean is within 1.645 standard deviations of the sample mean.

(If A is within 1.645 units of some fixed value B then B is within 1.645 units of A.)

Similarly:

We can be 95% confident that the population mean is within 1.960 standard deviations of the sample mean.

and

We can be 99% confident that the population mean is within 2.576 standard deviations of the sample mean.

If σ is the population standard deviation then the standard deviation referred to in each of the three previous italicised statements is $\frac{\sigma}{\sqrt{n}}$.

If we do not know the population standard deviation we can use the standard deviation of our sample as an estimate of the population standard deviation. Remember that calculators usually give two standard deviation values, σ_n and σ_{n-1}, or σ_x and s_x, the second of these being a little larger than the first. It is this second one, σ_{n-1} or s_x (called the *sample standard deviation*) that is used as an estimate for the population standard deviation as its slightly larger value allows for the fact that the variation in the sample slightly underestimates the variation in the population as a whole.

Thus, to infer the mean, μ, and standard deviation, σ, of a population from the mean, $\bar{x}$, and sample standard deviation, s_x, of a sample:

- assume that the standard deviation of the population, σ, is equal to s_x (i.e σ_{n-1}), the sample standard deviation.
- use the fact that our sample mean comes from a normal distribution with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$ to say that μ lies in the interval

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \quad \leq \quad \mu \quad \leq \quad \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

where k is the appropriate number of standard deviations for the required confidence interval,

i.e., to 3 decimal places, $k = 1.645$ for a 90% confidence interval
$k = 1.960$ for a 95% confidence interval
and $k = 2.576$ for a 99% confidence interval.

EXAMPLE 5

A random sample of 50 'top-grade' ripe plums of a particular species has a mean weight of 125 grams and a sample standard deviation of 12 grams.

Determine the 95% confidence interval for the mean weight of the population.

Solution

Not knowing the standard deviation of the population we use the sample standard deviation as an estimate of the population standard deviation, i.e. 12 grams.

Thus if μ is the population mean, for the 95% confidence interval we use

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \quad \leq \quad \mu \quad \leq \quad \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

with $\bar{x} = 125$ grams, $k = 1.960$, $\sigma = 12$ grams and $n = 50$.

Thus $$121.67 \text{ g} \quad \leq \quad \mu \quad \leq \quad 128.33 \text{ g}$$

The 95% confidence interval for the population mean is 121.7 g to 128.3 g.

Note

- It could be argued that any rounding should be such that it does not decrease the interval. The final statement of the answer in the previous example would then be stated as

  $$121.6 \text{ g} \leq \mu \leq 128.4 \text{ g}.$$

  The answers in this text will ignore this technicality and will give answers rounded in the usual way.

- Some calculators can determine confidence intervals, as the display below suggests.

C-Level	0.95
σ	12
$\bar{x}$	125
n	50

→

Lower	121.67383
Upper	128.32617
$\bar{x}$	125
n	50

- The 95% (or 90% or 99%) confidence interval tells us that if we were to construct confidence intervals in this way then we would expect 95% (or 90% or 99%) of them to contain the population mean.

- If we are told the standard deviation of the population we would use that, not the sample standard deviation, in our calculation of confidence intervals for the mean of the population.

ISBN 9780170395274

EXAMPLE 6

A sample of thirty-six 20 kg bags of cement has a mean weight of 20.095 kg and sample standard deviation 42 grams. Find the 99% confidence interval for the mean weight of the population.

Solution

Not knowing the standard deviation of the population we use the sample standard deviation as an estimate of the population standard deviation, i.e. we use 42 grams.

Thus if μ is the population mean, for the 99% confidence interval we use

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \quad \le \quad \mu \quad \le \quad \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

with $\bar{x} = 20.095$ kg, $k = 2.576$, $\sigma = 0.042$ kg and $n = 36$.

Thus $20.077 \text{ kg} \le \mu \le 20.113 \text{ kg}$

The 99% confidence interval for the population mean μ is

$$20.077 \text{ kg} \le \mu \le 20.113 \text{ kg}.$$

zInterval 0.042,20.095,36,0.99

"Title"	"z Interval"
"CLower"	20.076969
"CUpper"	20.113031
"$\bar{x}$"	20.095
"ME"	0.01803081
"n"	36
"σ"	0.042

EXAMPLE 7

The lengths of ten adult moths of a certain species were determined and the mean length was found to be 5.2 cm, sample standard deviation 0.3 cm. Given that the lengths of the adults of this species of moth are normally distributed determine the 95% confidence interval for μ, the mean length of this population of adult moths, and explain what your answer means.

Solution

Though our sample size of just ten is small the sample means will be normally distributed because the population from which the sample is taken is normally distributed.

Not knowing the standard deviation of the population we will use the sample standard deviation, i.e. 0.3 cm.

Thus if μ cm is the population mean, for the 95% confidence interval we use

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \quad \le \quad \mu \quad \le \quad \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

with $\bar{x} = 5.2$ cm, $k = 1.96$, $\sigma = 0.3$ cm and $n = 10$.

Thus $5.014 \text{ cm} \le \mu \le 5.386 \text{ cm}$

The 95% confidence interval for the population mean length is

5.01 cm to 5.39 cm.

zInterval 0.3,5.2,10,0.95

"Title"	"z Interval"
"CLower"	5.0140615
"CUpper"	5.3859385
"$\bar{x}$"	5.2
"ME"	0.18593851
"n"	10
"σ"	0.3

We can be 95% confident that the mean lengths of adult moths of this species lies between 5.01 cm and 5.39 cm (because 95% of the 95% confidence intervals constructed in this way would contain the population mean).

ISBN 9780170395274

Exercise 12B

1 If we use a given sample of size $n > 30$, mean $\bar{x}$ and sample standard deviation s, to determine confidence intervals for the population mean, which has the smaller width – the 90% confidence interval or the 95% confidence interval?

2 If we use a given sample of size $n > 30$, mean $\bar{x}$ and sample standard deviation s, to determine confidence intervals for the population mean, which has the smaller width – the 95% confidence interval or the 99% confidence interval?

3 Two different sized samples, taken from the same population, have the same mean value. The population standard deviation is known.

If the sample mean and population standard deviation are used to determine the 95% confidence interval for the mean of the population, which sample will give the narrower confidence interval – the bigger size sample or the smaller size sample?

4 Find the 90% confidence interval for the mean of the population given that a sample of size 100 taken from this population had a mean of 573 cm and a sample standard deviation of 48 cm. (Give interval boundaries to the nearest centimetre.)

5 Find the 95% confidence interval for the mean of the population given that a sample of size 50 taken from this population had a mean of 26.14 kg and a sample standard deviation of 3.67 kg. (Give interval boundaries to the nearest 0.01 kg.)

6 Find the 99% confidence interval for the mean of the population given that a sample of size 80 taken from this population had a mean of 17.2 cm and a sample variance of 5.76 cm.

7 The lengths of ten 12-month-old baby girls were recorded and the mean length was found to be 74.6 cm, sample standard deviation 1.4 cm. Given that the lengths of 12-month-old baby girls are normally distributed determine the 95% confidence interval for μ, the mean length of 12-month-old baby girls. Explain what your confidence interval means.

8 A random sample of 40 three-month-old seedlings of a particular plant type had a mean height of 17.8 cm, sample standard deviation 2.4 cm.

Determine a 90% confidence interval for the mean height of three-month-old seedlings of this plant type and explain what this confidence interval means.

Shutterstock.com/Sergio33

9 A random sample of 200 birds of a particular species had a mean wing span of length 18.3 cm, sample standard deviation 2.7 cm.

Use this information to determine a 95% confidence interval for the mean wing span for the entire population of birds of this species.

Had the same mean and sample standard deviation come from a random sample of 300 birds of this species instead, what would the 95% confidence interval be now?

ISBN 9780170395274

So do most 95% confidence intervals really contain the population mean? Let's check

The following statement was made a few pages earlier:

> *The 95% (or 90% or 99%) confidence interval tells us that if we were to construct confidence intervals in this way then we would expect 95% (or 90% or 99%) of them to contain the population mean.*

Let's check this using a situation for which we know what the population mean is.

If we could roll a die an infinite number of times, and find the mean score obtained we know that in theory the mean would be 3.5 and standard deviation 1.7078. Suppose we did not know these figures. Would the 95% confidence intervals, constructed from samples that we take, contain the population mean as we claim?

Using a computer or calculator spreadsheet simulate 100 such rolls and determine the mean and the standard deviation.

The display on the right shows a sample mean for such a simulation as 3.25 and sample standard deviation, s_x, as 1.7019.

Thus to estimate the population mean, μ, using the 95% confidence interval, we use

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

with $\bar{x} = 3.25$, $k = 1.96$, $\sigma = 1.7019$ and $n = 100$.

Thus $$2.916 \leq \mu \leq 3.584$$

We can be 95% confident that the population mean lies between 2.916 and 3.584, i.e. in the interval 3.25 ± 0.334.

	A	B	C
1	2		
2	5		
3	3		
4	1		
5	1		
6	6		

=int(rand()•6+1)

$\bar{x} = 3.25$
$\Sigma x = 325$
$\Sigma x^2 = 1343$
$\sigma_x = 1.6933694$
$s_x = 1.7019003$
n = 100

C-Level	0.95
σ	1.7019
$\bar{x}$	3.25
n	100

Lower	2.9164337
Upper	3.5835663
$\bar{x}$	3.25
n	100

The known population mean does indeed lie in the 95% confidence interval obtained from this sample.

Note: The population mean of 3.5 and standard deviation of 1.7078 gives a 95% interval of 3.5 ± 0.3347, but we would not normally know these figures.

Use a spreadsheet to produce 50 such samples of 100 rolls of a die and see how many of the 50 produce 95% confidence intervals that do contain 3.5, our known population mean.

ISBN 9780170395274

Choosing the sample size

The fact that, with a particular level of confidence, we can say the population mean, μ, lies in the interval

$$\bar{x} - k\frac{\sigma}{\sqrt{n}} \quad \le \quad \mu \quad \le \quad \bar{x} + k\frac{\sigma}{\sqrt{n}}$$

can be used to determine the required sample size, n, needed to give us a confidence interval of a desired width for the particular confidence level.

EXAMPLE 8

A random sample is to be taken from a random variable with population standard deviation of 5 units.

If we want to be 95% confident that the mean of the sample is within 1.7 units of the population mean how large should the sample be?

Solution

Writing the sample mean as $\bar{x}$ and the population mean as μ we have, for the 95% confidence interval

$$\bar{x} - 1.96 \times \frac{\sigma}{\sqrt{n}} \quad \le \quad \mu \quad \le \quad \bar{x} + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

Thus we require $\quad 1.96 \times \frac{5}{\sqrt{n}} = 1.7$

Solving gives $\quad n = 33.2$

We need to be *within* 1.7 units so we need to round n up. Hence the sample size needs to be 34.

EXAMPLE 9

An earlier sample of 1 litre bottles of a particular type of fruit drink suggests that the amount of fruit juice in each litre of the fruit drink is distributed with a mean of 153 millilitres and standard deviation 6.4 millilitres.

A new sample is to be carried out for which we want to be 99% confident that the mean of the sample is within 2 millilitres of the population mean. How large should the sample be?

Solution

Writing the sample mean as $\bar{x}$ and the population mean as μ we have, for the 99% confidence interval

$$\bar{x} - 2.576 \times \frac{\sigma}{\sqrt{n}} \quad \le \quad \mu \quad \le \quad \bar{x} + 2.576 \times \frac{\sigma}{\sqrt{n}}$$

Thus we require $\quad 2.576 \times \frac{6.4}{\sqrt{n}} = 2$

Solving gives $\quad n = 67.95$

We want to be *within* 2 millilitres so we need to round n up. Hence the sample size needs to be 68.

ISBN 9780170395274

Exercise 12C

1 A random sample is to be taken from a random variable with population standard deviation of 8 units.

If we want to be 95% confident that the mean of the sample is within 1.5 units of the population mean how large should the sample be?

2 A random sample is to be taken from a random variable with population standard deviation of 18.7 units.

If we want to be 99% confident that the mean of the sample is within 2.5 units of the population mean how large should the sample be?

3 A random sample is to be taken from a random variable with a normally distributed population, standard deviation 7.3 units.

If we want to be 95% confident that the mean of the sample is within 3 units of the population mean how large should the sample be?

4 A random sample of 30 measurements of a random variable is taken in order to have some idea of the standard deviation of the variable.

The sample standard deviation of this sample was found to be 8.4 units.

If we assume that this standard deviation is also the standard deviation of the random variable, find the sample size needed for a second sample if we want to be 90% confident that the mean of this second sample is within 2 units of the population mean.

5 A previous sample involving measuring the lengths of 30 adult snakes of a particular species found that the lengths were distributed with a mean of 28.4 cm and sample standard deviation 3.6 cm.

A new sample is to be taken for which the investigators want to be 95% confident that the mean length of this second sample is within 0.5 cm of the population mean.

Assuming the standard deviation of the lengths of the population is the same as the sample standard deviation of the lengths in the first sample, what should be the size of this second sample?

6 A drink dispenser is programmed to dispense quantities with a mean volume of 250 mL, standard deviation 3 mL.

With age, dispensers of this type are known to start dispensing amounts that no longer have a mean volume of 250 mL, though the standard deviation tends to remain at 3 mL.

A fresh sample of dispensed amounts is to be checked to allow an estimate to be made of the mean volume dispensed.

If we want to be 95% confident that the mean volume being dispensed is within 1 mL of our sample mean, how many 'dispenses' should we include in our sample?

iStock.com/4kodiak

Miscellaneous exercise twelve

This miscellaneous exercise may include questions involving the work of this chapter, the work of any previous chapters in this unit, and the ideas mentioned in the Preliminary work section at the beginning of this unit.

1 Find an expression for $\frac{dC}{dx}$ for each of the following cost functions.

a $C(x) = 300 + 7x$

b $C(x) = 500 + 4x + x^2$

c $C(x) = \frac{x^3}{12} - 12x^2 + 800x + 1000$

d $C(x) = 60 + 4\sqrt{x} - \frac{1000}{x}$

2 Find an expression for $\frac{dy}{dx}$ as a function of x given that $x^3 + 2x^2y + y^3 = 10$.

3 If $y^4 = x^4 - 4xy - 7$ find

a an expression for $\frac{dy}{dx}$,

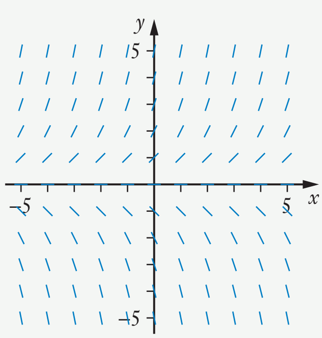
b the gradient at the point (2, 1).

4 The three slope fields shown below have their associated differential equations included in the following list. Choose the correct equation for each slope field.

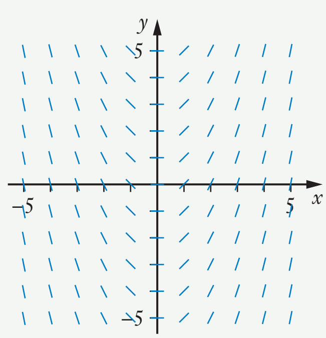
$\frac{dy}{dx} = 2y$ $\qquad$ $\frac{dy}{dx} = y$ $\qquad$ $\frac{dy}{dx} = x$

$\frac{dy}{dx} = \frac{x-3}{y-2}$ $\qquad$ $\frac{dy}{dx} = (x-2)(y-3)$ $\qquad$ $\frac{dy}{dx} = (x-3)(y-2)$

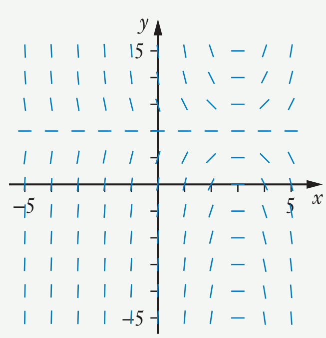
Slope field A



Slope field B



Slope field C



5 Clearly showing your use of the technique of expressing an algebraic fraction as partial fractions, determine:

$$\int \frac{3x^3 + 6x^2 - 4x - 8}{(x+1)(x^2-2)}\,dx.$$

6 Find an expression in terms of k (not left as a definite integral) for the *exact* area between the x-axis and the curve $y = 3\sin^2 x \cos x$ from $x = 0$ to $x = k$ where

a $0 < k < \frac{\pi}{2}$

b $\frac{\pi}{2} < k < \pi$.

ISBN 9780170395274

7 A particle moves such that its velocity, v m/s, is a function of its displacement from some fixed point O, x metres, according to the rule $v = 3 + 0.1x$, $x > 0$.

When timing commenced, i.e. when $t = 0$, $x = 20$.

Find **a** the acceleration of the particle when $x = 2$,

b x when $t = 5$.

8 If $\frac{dA}{dt} = 6e^{2t}$ and $A = 4$ when $t = 0$, find **a** A in terms of t,

b the exact value of A when $t = 0.5$.

c Use $\frac{dA}{dt}$ to find the approximate change in A when t changes from 0 to 0.01.

9 A savings account is opened with an initial deposit of \$10 000.

The account attracts interest at a rate of 10% per annum compounded continuously.

Thus if the value of the account is \$$P$ then $\frac{dP}{dt} = 0.1P$.

Assuming no further deposits are made,

a how long will it take for the initial value of the account to double?

b how long will it take for the value of the account to become \$40 000?

c how long will it take for the value of the account to become \$80 000?

10 A sample of 64 observations is randomly selected from a population and is found to have a mean of 53.24 and a sample standard deviation of 5.12.

a Find the range of values, with 53.24 at its centre, in which we can have 90% confidence that the mean of the population will lie (i.e. find the 90% confidence interval).

b Find the 95% confidence interval for the mean of the population.

c Find the 99% confidence interval for the mean of the population.

11 The mean lifetime of a random sample of 40 'triple A' batteries of a particular brand is found to be 223 hours, sample standard deviation 18 hours.

Find the 95% confidence interval for the mean lifetime of all 'triple A' batteries of this brand giving the interval boundaries to the nearest 0.1 hour.

What does this confidence interval tell us?

12 **a** The displacement, x metres, of an object from an origin O is given by

$$x = A\cos kt,$$

where A and k are constants.

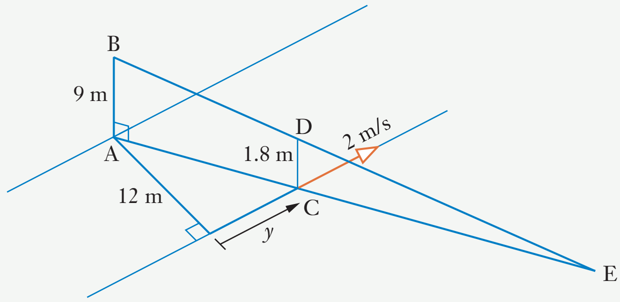
Prove that the object is moving with SHM and that it is initially at an extreme position.

b The depth of water in a harbour, above and below the mean depth, is an example of simple harmonic motion. In a particular harbour the low tide depth of 3 metres is recorded at 7 a.m. one morning and the next high tide is expected to record a depth of 15 metres at 1.20 p.m. later that same day.

A particular container ship requires a depth of at least 5 metres for safe entry into the harbour, for unloading at the dock side and for leaving. Determine, to the nearest 5 minutes, the times between 7 a.m. and 9 p.m. that day, between which it is safe for the ship to engage in these activities.

13 If $\frac{dy}{dx} = xe^x$ use the incremental formula to determine the approximate change in y when x changes from 1.25 to 1.26.

14 The diagram, not drawn to scale, shows a lamppost, AB, of height 9 metres, situated on one side of a straight road.



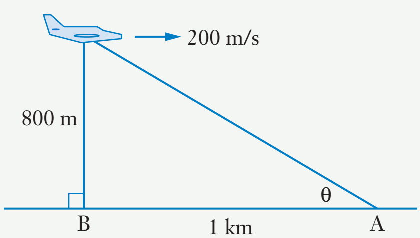
The road is of width 12 metres and line CD (see diagram) represents a person of height 1.8 metres walking at 2 metres/second along the other side of the road from the lamp. The light at B causes the person to have a shadow shown as CE.

Find the rate at which the length of the shadow is changing at the instant when AC is of length 20 metres.

15 An aircraft is following a flightpath that will take it directly over a searchlight at A (see diagram).

The speed of the aircraft is 200 m/s and it maintains a steady altitude of 800 metres.

At what rate must the searchlight be rotating at the instant when the aircraft is 1 km from A, measured horizontally, if it is to keep the aircraft in the light beam?



iStock.com/Starkblast

 ISBN 9780170395274

16 A processing company processes a particular ore in one-tonne 'batches'. Previous statistics from the company indicate that the amounts of a particular element extracted from the one tonne batches are normally distributed with a mean of 78 kg and standard deviation 12 kg.

a What is the probability of more than 100 kg of the particular element being extracted from a randomly selected one tonne batch?

b What is the probability of between 70 kg and 90 kg of the particular element being extracted from a randomly selected one tonne batch?

c If 4 separate batches are sampled and the amount of the element extracted from each is noted, what is the probability that the mean of these four amounts will be less than 75 kg?

17 The shape of the interior of a particular flower vase is the same as that formed by rotating about the y-axis the area in the first quadrant that is enclosed between the x-axis, the y-axis, $y = x^2 - 4$ and $y = h$ ($h > 0$). With one centimetre to one unit on each axis determine the value of h, to the nearest centimetre, if the capacity of the vase is 2.5 L.

If water were poured into this vase at a constant rate of 5 cm^3/second, find the rate at which the water level is rising when the depth of water in the vase is 6 cm.

18 At 1.30 p.m. a police doctor arrives at a house where a suspicious death has occurred. Upon arrival the doctor takes the temperature of the corpse and finds it to be 28.6°C. By 2.30 p.m. the temperature of the corpse is 27.7°C.

The house is air-conditioned so the room temperature may be assumed to have been 22°C all day.

According to Newton's law of cooling, if the temperature of an object exceeds that of its surroundings by T°C then the rate of change of T is proportional to T itself. If normal body temperature is 37°C estimate the time at which the suspicious death occurred.

19 **a** Let us suppose that the standard deviation of h, the time, in hours, of 'non-stop operation before requiring recharging', of a particular electrical component made by company A is 80 hours.

The engineers from a space program wish to estimate the mean value of h by sampling a number of these components and running them non-stop until recharging is required, before deciding whether to use them in their space vehicle.

The engineers want to be 95% confident that the mean value of h for their sample is within 20 hours of the population mean.

How large should their sample be?

b A second company, company B, also makes the component and they claim that for their components, h is normally distributed with a mean of 1850 hours and standard deviation 65 hours.

The space program engineers sample 100 of these company B components and find that for their sample the sample standard deviation of h is indeed very close to 65 hours but the mean for the sample is 1800 hours.

Comment on this result including in your comment mention of the 95% confidence interval.

20 Without the assistance of your calculator, determine $\int \frac{2x^3 + 3x^2 - 24x - 29}{x^3 - 7x - 6}\,dx$.

ISBN 9780170395274

21 In the planning for a fish farming business, mathematical modelling is carried out to assess N, the likely number of fish, over a particular minimum size, in each breeding pond, t months after the initial 250 fish are placed in the pond.

It is felt that, with no removal of fish, the rate of change of N will be such that

$$\frac{dN}{dt} = \frac{2N}{5} - \frac{N^2}{15\,625} \quad \text{(i.e. a logistic model),} \qquad \text{for } 0 < N < 6250.$$

a Clearly showing your use of the technique of separating the variables, and partial fractions, find, in the form $\frac{k}{1+ce^{-at}}$, a formula for N in terms of t.

b Show that your part **a** answer does give the limiting value for N as 6250.

c According to this model what, to the nearest ten, will be the value of N when

i $t = 6$? **ii** $t = 12$?

22 In chapter 9, one example determined the antiderivative of $\sin^5 x$ as:

$$-\cos x + \frac{2}{3}\cos^3 x - \frac{1}{5}\cos^5 x + c.$$

However, as the displays below suggest, one calculator claims that the answer is

$$\cos x\left(\frac{-\sin^4 x}{5} - \frac{4\sin^2 x}{15} - \frac{8}{15}\right)$$

and another claims the answer is

$$\frac{-(150\cos x + 3\cos 5x - 25\cos 3x)}{240}.$$

$$\int \left((\sin(x))^5\right) dx$$
$$\left(\frac{-(\sin(x))^4}{5} - \frac{4\cdot(\sin(x))^2}{15} - \frac{8}{15}\right)\cdot\cos(x)$$

$$\int_{\square}^{\square} (\sin(x))^5\, dx$$
$$\frac{-(150\cdot\cos(x) + 3\cdot\cos(5\cdot x) - 25\cdot\cos(3\cdot x))}{240}$$

Show that, with the exception of the fact that the calculator displays do not show the constant, the three expressions are the same.

23 The parts of this question give answers that involve inverse trigonometrical functions, e.g. $\sin^{-1} x$, also written as $\arcsin x$. (This inverse function should not be confused with $\frac{1}{\sin x}$, which is better written as $(\sin x)^{-1}$, or $\operatorname{cosec} x$.)

Determine the following indefinite integrals using the suggested substitution.

a $\int \frac{1}{\sqrt{1-x^2}}\,dx, \quad x = \sin u$

b $\int \frac{1}{\sqrt{25-x^2}}\,dx, \quad x = 5\sin u$

c $\int \frac{1}{\sqrt{9-4x^2}}\,dx, \quad x = \frac{3}{2}\sin u$

d $\int \sqrt{1-x^2}\,dx, \quad x = \sin u$

e $\int \sqrt{4-x^2}\,dx, \quad x = 2\sin u$

f $\int \sqrt{4-x^2}\,dx, \quad x = 2\cos u$

ISBN 9780170395274