16

The normal distribution

Objectives

- ▶ To introduce the **standard normal distribution**.
- ▶ To introduce the family of normal distributions as transformations of the standard normal distribution.
- ▶ To investigate the effect that changing the values of the parameters defining the normal distribution has on the graph of the probability density function.
- ▶ To recognise the **mean**, **median**, **variance** and **standard deviation** of a normal distribution.
- ▶ To use technology to determine **probabilities for intervals** in the solution of problems where the normal distribution is appropriate.

The most useful continuous distribution, and one that occurs frequently, is the normal distribution. The probability density functions of normal random variables are symmetric, single-peaked, bell-shaped curves.

Data sets occurring in nature will often have such a bell-shaped distribution, as measurements on many random variables are closely approximated by a normal probability distribution.

Variables such as height, weight, IQ and the volume of milk in a milk carton are all examples of normally distributed random variables.

As well as helping us to understand better the behaviour of many real-world variables, the normal distribution also underpins the development of statistical estimation, which is the topic of Chapter 17.

ISBN 978-1-009-11049-5

Cambridge University Press

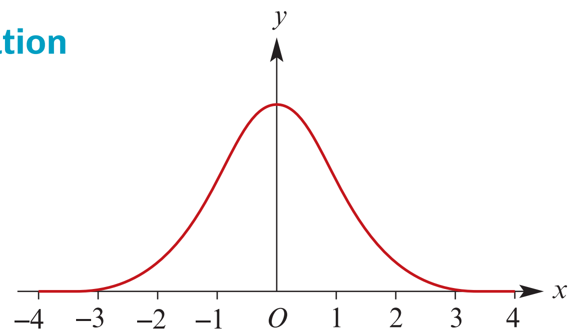
16A The normal distribution

The standard normal distribution

The simplest form of the normal distribution is a random variable with probability density function f given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

The domain of f is $\mathbb{R}$.



Because it is the simplest form of the normal distribution, it is given a special name: the **standard normal distribution**. The graph of the standard normal distribution is as shown.

The graph of the standard normal probability density function f is symmetric about $x = 0$, since $f(-x) = f(x)$. That is, the function f is even.

The line $y = 0$ is an asymptote: as $x \to \pm\infty$, $y \to 0$. Almost all of the area under the probability density function lies between $x = -3$ and $x = 3$.

The mean and standard deviation of the standard normal distribution

It can be seen from the graph that the mean and median of this distribution are the same, and are equal to 0. While the probability density function for the standard normal distribution cannot be integrated exactly, the value of the mean can be verified by observing the symmetry of the two integrals formed below. One is just the negative of the other.

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x)\, dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{1}{2}x^2}\, dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\int_{0}^{\infty} x e^{-\frac{1}{2}x^2}\, dx + \int_{-\infty}^{0} x e^{-\frac{1}{2}x^2}\, dx \right)
\end{aligned}$$

Thus the mean, $\mathrm{E}(X)$, of the standard normal distribution is 0.

What can be said about the standard deviation of this distribution? It can be shown that

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}x^2}\, dx = 1$$

Therefore

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = 1 - 0 = 1 \quad \text{and} \quad \mathrm{sd}(X) = \sqrt{\mathrm{Var}(X)} = 1$$

Standard normal distribution

A random variable with the standard normal distribution has probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

The standard distribution has mean $\mu = 0$ and standard deviation $\sigma = 1$.

Henceforth, we will denote the random variable of the standard normal distribution by Z.

ISBN 978-1-009-11049-5 Cambridge University Press

The general normal distribution

The normal distribution does not apply just to the special circumstances where the mean is 0 and the standard deviation is 1.

Transformations of the standard normal distribution

The graph of the probability density function for a normal distribution with mean μ and standard deviation σ may be obtained from the graph of the probability density function for the standard normal distribution by the transformation with rule:

$$(x, y) \to \left(\sigma x + \mu, \frac{y}{\sigma}\right)$$

This is a dilation of factor σ from the y-axis and a dilation of factor $\frac{1}{\sigma}$ from the x-axis, followed by a translation of μ units in the positive direction of the x-axis, for $\mu > 0$.
(In Section 15D, this was discussed for probability density functions in general.)

Conversely, the transformation which maps the graph of a normal distribution with mean μ and standard deviation σ to the graph of the standard normal distribution is given by

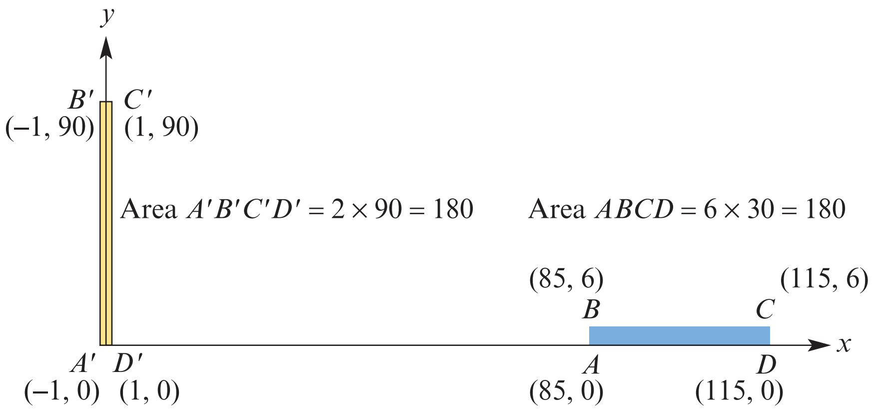
$$(x, y) \to \left(\frac{x - \mu}{\sigma}, \sigma y\right)$$

This is a translation of μ units in the negative direction of the x-axis, followed by a dilation of factor $\frac{1}{\sigma}$ from the y-axis and a dilation of factor σ from the x-axis.

For example, if $\mu = 100$ and $\sigma = 15$, then this transformation is

$$(x, y) \to \left(\frac{x - 100}{15}, 15y\right)$$

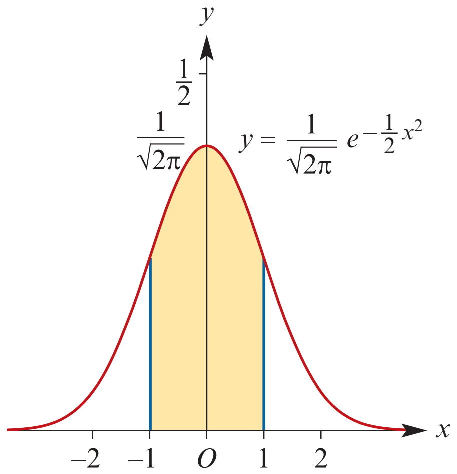
This transformation is area-preserving. In the following diagram, the rectangle $ABCD$ is mapped to $A'B'C'D'$. Both rectangles have an area of 180 square units.

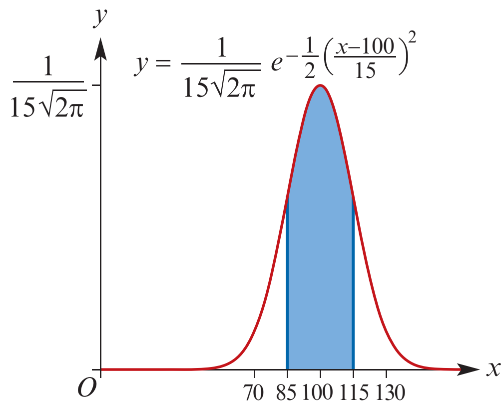


This property enables the probabilities of any normal distribution to be determined from the probabilities of the standard normal distribution.

ISBN 978-1-009-11049-5

Cambridge University Press





The shaded regions are of equal area.

This leads to the general rule for the family of normal probability distributions.

The rule for the general normal distribution

If X is a **normally distributed random variable** with mean μ and standard deviation σ, then the probability density function of X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

and

$$\Pr(X \le a) = \Pr\left(Z \le \frac{a-\mu}{\sigma}\right)$$

where Z is the random variable of the standard normal distribution.

The general form of the normal density function involves two parameters, μ and σ, which are the mean (μ) and the standard deviation (σ) of that particular distribution.

When a random variable has a distribution described by a normal density function, the random variable is said to have a **normal distribution**.

As with all probability density functions, the normal density function has the fundamental properties that:

- probability corresponds to an area under the curve
- the total area under the curve is 1.

However, it has some additional special properties.

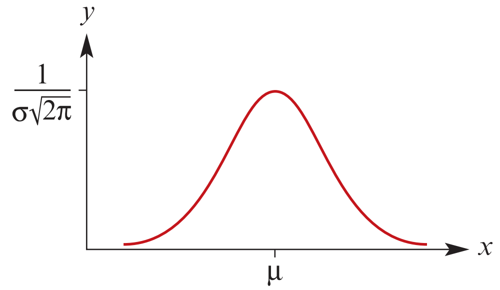
The graph of a normal density function is symmetric and bell-shaped:

- its **centre** is determined by the **mean** of the distribution
- its **width** is determined by the **standard deviation** of the distribution.

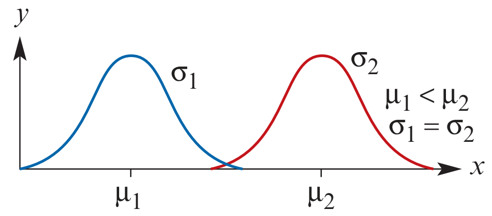
ISBN 978-1-009-11049-5 Cambridge University Press

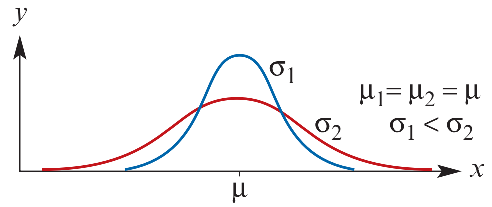
The graph of $y = f(x)$ is shown on the right.

The graph is symmetric about the line $x = \mu$, and has a maximum value of $\frac{1}{\sigma\sqrt{2\pi}}$, which occurs when $x = \mu$.



Thus the *location* of the curve is determined by the value of μ, and the *steepness* of the curve by the value of σ.





Irrespective of the values of the mean and standard deviation of a particular normal density function, the area under the curve within a given number of standard deviations from the mean is always the same.

Example 1

On the same set of axes, sketch the graphs of the probability density functions of the standard normal distribution and the normal distribution with:

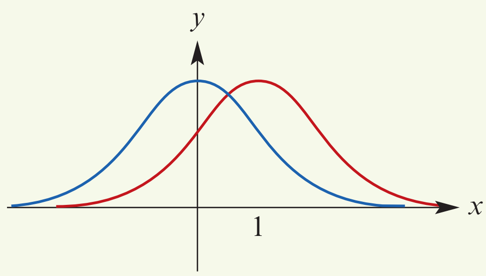
a mean 1 and standard deviation 1

b mean 1 and standard deviation 2.

(A calculator can be used to help.)

Solution

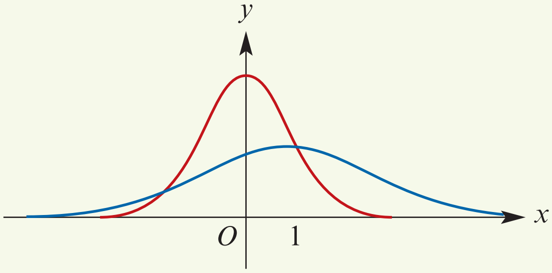
a The graph has been translated 1 unit in the positive direction of the x-axis.



The rules of the two density functions are $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ and $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2}$.

ISBN 978-1-009-11049-5 Cambridge University Press

b The graph has been dilated from the y-axis by factor 2 and from the x-axis by factor $\frac{1}{2}$, and then translated 1 unit in the positive direction of the x-axis.



The rules of the two density functions are $y = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ and $y = \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1}{2}\right)^2}$.

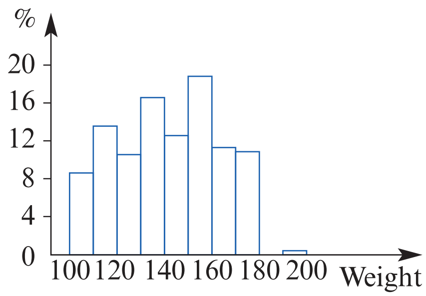
Exercise 16A

Example 1

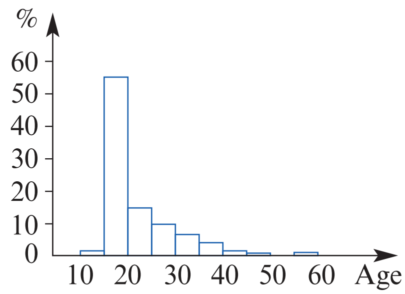
1 Both the random variables X_1 and X_2 are normally distributed, with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively. If $\mu_1 < \mu_2$ and $\sigma_1 < \sigma_2$, sketch both distributions on the same diagram.

2 Which of the following data distributions are approximately normally distributed?

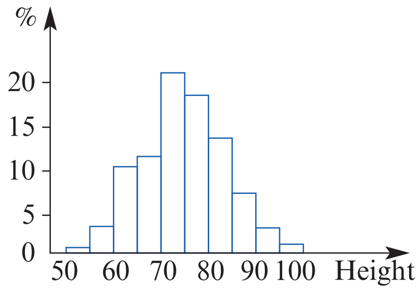
a



b



c



3 Consider the normal probability density function:

$$f(x) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-2}{3}\right)^2}, \quad x \in \mathbb{R}$$

a Use your calculator to find $\int_{-\infty}^{\infty} f(x)\,dx$.

b **i** Express $E(X)$ as an integral.

ii Use your calculator to evaluate the integral found in **i**.

c **i** Write down an expression for $E(X^2)$. **ii** What is the value of $E(X^2)$?

iii What is the value of σ?

ISBN 978-1-009-11049-5 Cambridge University Press

4 Consider the normal probability density function:

$$f(x) = \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x+4}{5}\right)^2}, \quad x \in \mathbb{R}$$

a Use your calculator to find $\int_{-\infty}^{\infty} f(x)\, dx$.

b **i** Express $\mathrm{E}(X)$ as an integral.

ii Use your calculator to evaluate the integral found in **i**.

c **i** Write down an expression for $\mathrm{E}(X^2)$.

ii What is the value of $\mathrm{E}(X^2)$?

iii What is the value of σ?

5 The probability density function of a normal random variable X is given by

$$f(x) = \frac{1}{10\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-3}{10}\right)^2}$$

a Write down the mean and the standard deviation of X.

b Sketch the graph of $y = f(x)$.

6 The probability density function of a normal random variable X is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x+3)^2}$$

a Write down the mean and the standard deviation of X.

b Sketch the graph of $y = f(x)$.

7 The probability density function of a normal random variable X is given by

$$f(x) = \frac{1}{\sqrt{18\pi}} e^{-\frac{1}{2}\left(\frac{x}{3}\right)^2}$$

a Write down the mean and the standard deviation of X.

b Sketch the graph of $y = f(x)$.

8 Describe the sequence of transformations which takes the graph of the probability density function of the standard normal distribution to the graph of the probability density function of the normal distribution with:

a $\mu = 3$ and $\sigma = 2$ **b** $\mu = 3$ and $\sigma = \frac{1}{2}$ **c** $\mu = -3$ and $\sigma = 2$

9 Describe the sequence of transformations which takes the graph of the probability density function of the normal distribution with the given mean and standard deviation to the graph of the probability density function of the standard normal distribution:

a $\mu = 3$ and $\sigma = 2$ **b** $\mu = 3$ and $\sigma = \frac{1}{2}$ **c** $\mu = -3$ and $\sigma = 2$

ISBN 978-1-009-11049-5

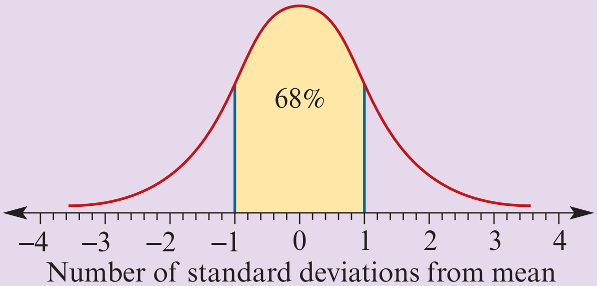
16B Standardisation and the 68–95–99.7% rule

For a set of data values that are normally distributed, approximately 68% of the values will lie within one standard deviation of the mean, approximately 95% of the values will lie within two standard deviations of the mean, and almost all (99.7%) within three standard deviations. This gives rise to what is known as the **68–95–99.7% rule**.

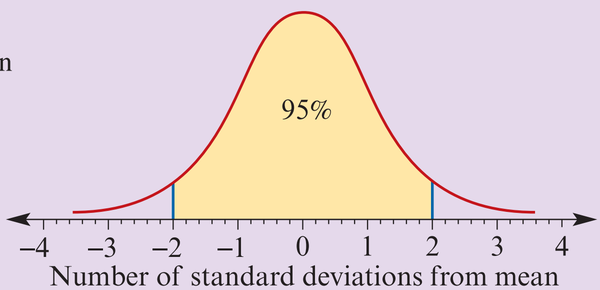
The 68–95–99.7% rule

For a normally distributed random variable, approximately:

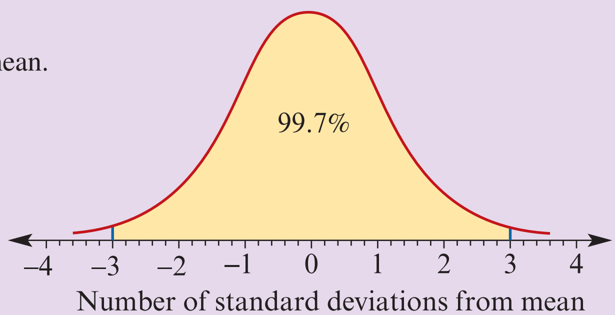
- 68% of the values lie within one standard deviation of the mean



- 95% of the values lie within two standard deviations of the mean



- 99.7% of the values lie within three standard deviations of the mean.



If we know that a random variable is approximately normally distributed, and we know its mean and standard deviation, then we can use the 68–95–99.7% rule to quickly make some important statements about the way in which the data values are distributed.

ISBN 978-1-009-11049-5 Cambridge University Press

Example 2

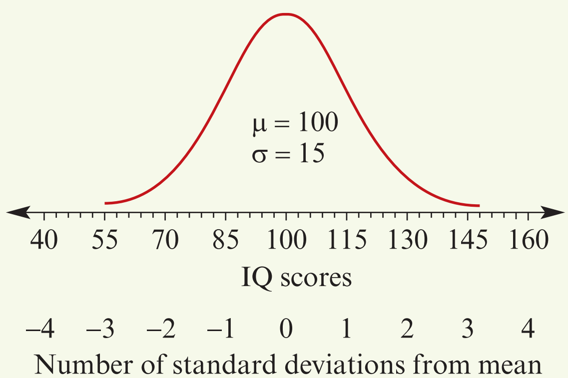
Experience has shown that the scores obtained on a commonly used IQ test can be assumed to be normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$.

Approximately what percentage of the distribution lies within one, two or three standard deviations of the mean?

Solution

Since the scores are normally distributed with $\mu = 100$ and $\sigma = 15$, the 68–95–99.7% rule means that approximately:

- 68% of the scores will lie between 85 and 115
- 95% of the scores will lie between 70 and 130
- 99.7% of the scores will lie between 55 and 145.



Note: In this example, we are using a continuous distribution to model a discrete situation.

Statements can also be made about the percentage of scores that lie in the tails of the distribution, by using the symmetry of the distribution and noting that the total area under the curve is 100%.

Example 3

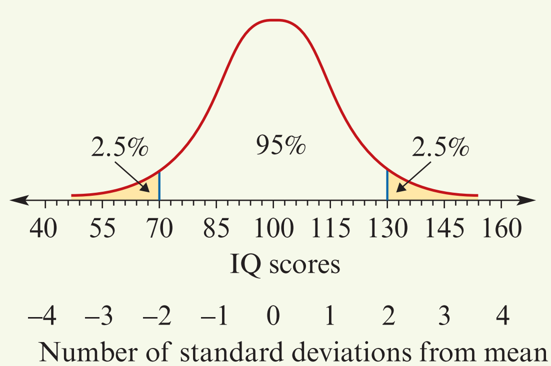
From Example 2, we know that 95% of the scores in the IQ distribution lie between 70 and 130 (that is, within two standard deviations of the mean). What percentage of the scores are *more* than two standard deviations above or below the mean (in this instance, less than 70 or greater than 130)?

Solution

If we focus our attention on the tails of the distribution, we see that 5% of the IQ scores lie outside this region.

Using the symmetry of the distribution, we can say that 2.5% of the scores are below 70, and 2.5% are above 130.

That is, if you obtained a score greater than 130 on this test, you would be in the top 2.5% of the group.



ISBN 978-1-009-11049-5

Standardised values

Clearly, the standard deviation is a natural measuring stick for normally distributed data. For example, a person who obtained a score of 112 on an IQ test with a mean of $\mu = 100$ and a standard deviation of $\sigma = 15$ is less than one standard deviation from the mean. Their score is typical of the group as a whole, as it lies well within the middle 68% of scores. In contrast, a person who scored 133 has done exceptionally well; their score is more than two standard deviations from the mean and this puts them in the top 2.5%.

Because of the additional insight provided, it is usual to convert normally distributed data to a new set of units which shows the number of standard deviations each data value lies from the mean of the distribution. These new values are called **standardised values** or ***z*-values**. To standardise a data value x, we first subtract the mean μ of the normal random variable from the value and then divide the result by the standard deviation σ. That is,

$$\text{standardised value} = \frac{\text{data value} - \text{mean of the normal curve}}{\text{standard deviation of the normal curve}}$$

or symbolically,

$$z = \frac{x - \mu}{\sigma}$$

Standardised values can be positive or negative:

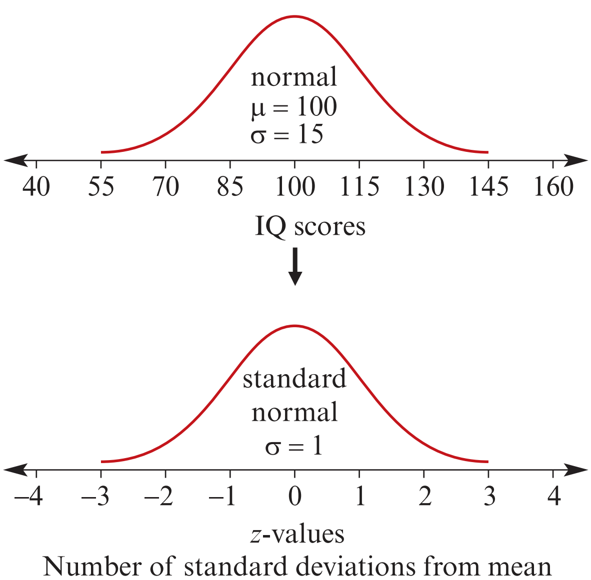
- A **positive *z*-value** indicates that the data value it represents lies **above** the mean.
- A **negative *z*-value** indicates that the data value lies **below** the mean.

For example, an IQ score of 90 lies *below* the mean and has a standardised value of

$$z = \frac{x - \mu}{\sigma} = \frac{90 - 100}{15} = \frac{-10}{15} \approx -0.67$$

There are as many different normal curves as there are values of μ and σ. But if the measurement scale is changed to 'standard deviations from the mean' or z-values, all normal curves reduce to the same normal curve with mean $\mu = 0$ and standard deviation $\sigma = 1$.

The figures on the right show how standardising IQ scores transforms a normal distribution with mean $\mu = 100$ and standard deviation $\sigma = 15$ into the standard normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$.



ISBN 978-1-009-11049-5

Cambridge University Press

Exercise 16B

Example 2

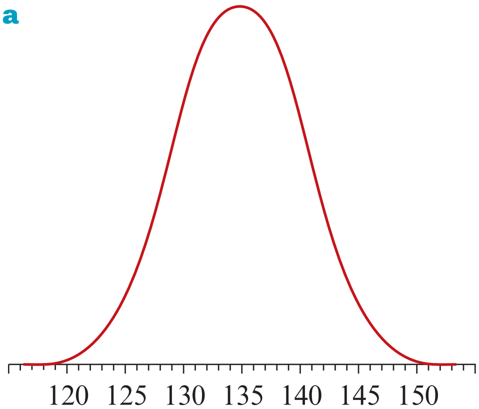
1 The scores obtained on an IQ test can be assumed to be normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$. What percentage of scores lie:

a above 115 **b** below 85

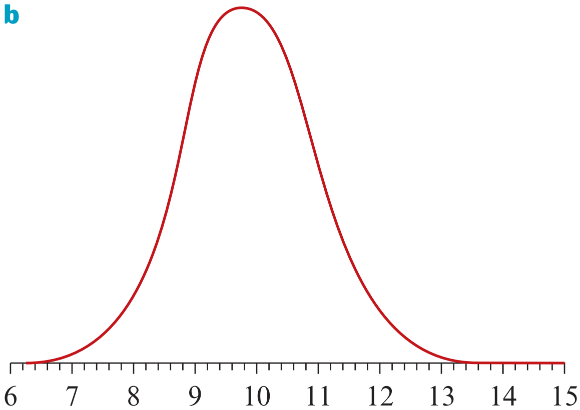
c above 130 **d** below 70?

2 State the values of the mean and standard deviation of the normal distributions shown:

a



b



Example 3

3 The heights of young women are normally distributed with mean $\mu = 160$ cm and standard deviation $\sigma = 8$ cm. What percentage of the women would you expect to have heights:

a between 152 cm and 168 cm

b greater than 168 cm

c less than 136 cm?

4 Fill in the blanks in the following paragraph.

The age at marriage of males in the US in the 1980s was approximately normally distributed with a mean of $\mu = 27.3$ years and a standard deviation of $\sigma = 3.1$ years. From this data, we can conclude that in the 1980s about 95% of males married between the ages of ______ and ______.

5 Fill in the blanks in the following statement of the 68–95–99.7% rule.

For any normal distribution, about:

- 68% of the values lie within ______ standard deviation of the mean
- ______% of the values lie within two standard deviations of the mean
- ______% of the values lie within ______ standard deviations of the mean.

6 If you are told that in Australian adults, nostril width is approximately normally distributed with a mean of $\mu = 2.3$ cm and a standard deviation of $\sigma = 0.3$ cm, find the percentage of people with nostril widths less than 1.7 cm.

ISBN 978-1-009-11049-5

Cambridge University Press

7 The distribution of IQ scores for the inmates of a certain prison is approximately normal with mean $\mu = 85$ and standard deviation $\sigma = 15$.

a What percentage of the prison population have an IQ of 100 or higher?

b If someone with an IQ of 70 or less can be classified as having special needs, what percentage of the prison population could be classified as having special needs?

8 The distribution of the heights of navy officers was found to be normal with a mean of $\mu = 175$ cm and a standard deviation of $\sigma = 5$ cm. Determine:

a the percentage of navy officers with heights between 170 cm and 180 cm

b the percentage of navy officers with heights greater than 180 cm

c the approximate percentage of navy officers with heights greater than 185 cm.

9 The distribution of blood pressures (systolic) among women of similar ages is normal with a mean of 120 (mm of mercury) and a standard deviation of 10 (mm of mercury). Determine the percentage of women with a systolic blood pressure:

a between 100 and 140

b greater than 130

c greater than 120

d between 90 and 150.

10 The heights of women are normally distributed with mean $\mu = 160$ cm and standard deviation $\sigma = 8$ cm. What is the standardised value for the height of a woman who is:

a 160 cm tall **b** 150 cm tall **c** 172 cm tall?

11 The length of pregnancy for a human is approximately normally distributed with a mean of $\mu = 270$ days and a standard deviation of $\sigma = 10$ days. How many standard deviations away from the mean is a pregnancy of length:

a 256 days **b** 281 days **c** 305 days?

12 Michael scores 85 on the mathematics section of a scholastic aptitude test, the results of which are known to be normally distributed with a mean of 78 and a standard deviation of 5. Cheryl sits for a different mathematics ability test and scores 27. The scores from this test are normally distributed with a mean of 18 and a standard deviation of 6. Assuming that both tests measure the same kind of ability, who has the better score?

13 The following table gives a student's results in Biology and History. For each subject, the table gives the student's mark (x) and also the mean (μ) and standard deviation (σ) for the class.

	Mark (x)	Mean (μ)	Standard deviation (σ)	Standardised mark (z)
Biology	77	68.5	4.9	
History	79	75.3	4.1	

Complete the table by calculating the student's standardised mark for each subject, and use this to determine in which subject the student did best *relative* to her peers.

ISBN 978-1-009-11049-5 Cambridge University Press

14 Three students took different tests in French, English and Mathematics:

Student	Subject	Mark (x)	Mean (μ)	Standard deviation (σ)	Standardised mark (z)
Mary	French	19	15	4	
	English	42	35	8	
	Mathematics	20	20	5	
Steve	French	21	23	4	
	English	39	42	3	
	Mathematics	23	18	4	
Sue	French	15	15	5	
	English	42	35	10	
	Mathematics	19	20	5	

a Determine the standardised mark for each student on each test.

b Who is the best student in:

i French **ii** English **iii** Mathematics?

c Who is the best student overall? Give reasons for your answer.

16C Determining normal probabilities

A CAS calculator can be used to determine areas under normal curves, allowing us to find probabilities for ranges of values other than one, two or three standard deviations from the mean. The following example is for the standard normal distribution, but the same procedures can be used for any normal distribution by entering the appropriate values for μ and σ.

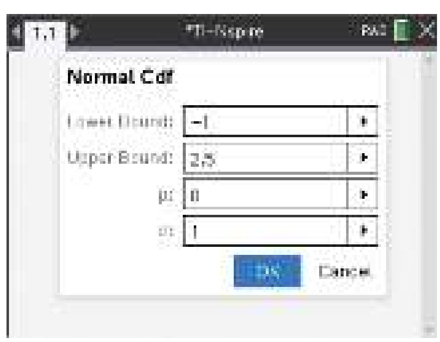
Example 4

Suppose that Z is a standard normal random variable (that is, it has mean $\mu = 0$ and standard deviation $\sigma = 1$). Find:

a $\Pr(-1 < Z < 2.5)$ **b** $\Pr(Z > 1)$

Using the TI-Nspire

a Use (menu) > **Probability** > **Distributions** > **Normal Cdf** and complete as shown. (Use (tab) or ▼ to move between cells.)

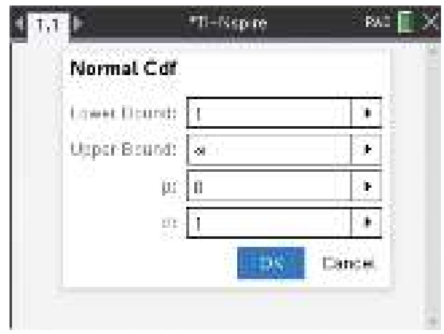


ISBN 978-1-009-11049-5

Cambridge University Press

The answer is:
$\Pr(-1 < Z < 2.5) = 0.8351$

b Use (menu) > **Probability** > **Distributions** > **Normal Cdf** and complete as shown.
(The symbol ∞ can be found using (π▸) or (ctrl)(📖).)



The answer is:
$\Pr(Z > 1) = 0.1587$

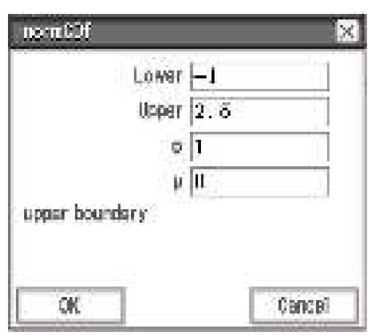
Note: You can enter the commands and parameters directly if preferred. The commands are not case sensitive.

Using the Casio ClassPad

Method 1

a ■ In Main $\sqrt{\alpha}$, go to **Interactive** > **Distribution** > **Continuous** > **normCDf**.

■ Enter the lower and upper bounds and tap OK.



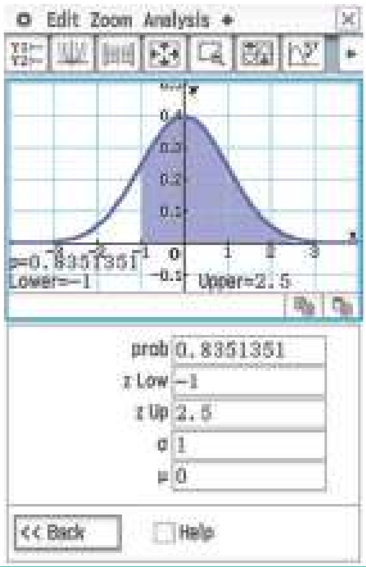
ISBN 978-1-009-11049-5

Cambridge University Press

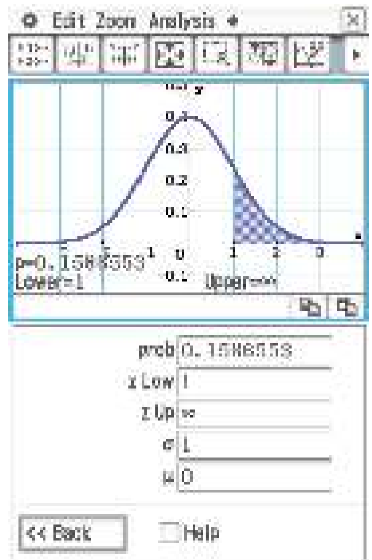
Method 2

- In **Statistics**, go to **Calc** > **Distribution** and select **Normal CD**. Tap `Next`.
- Enter values for the lower and upper bounds. Tap `Next` to view the answer.
- Select the graph icon to view the graph with the answer.

a



b



The calculator can also be used to determine **percentiles** of any normal distribution.

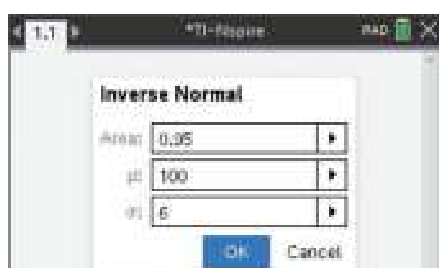
Example 5

Suppose X is normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 6$.

Find k such that $\Pr(X \leq k) = 0.95$.

Using the TI-Nspire

Use `menu` > **Probability** > **Distributions** > **Inverse Normal** and complete as shown.



The value of k is 109.869.

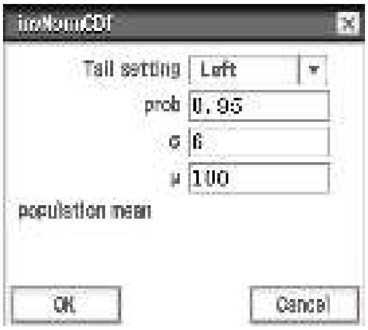
Note: You can enter the command and parameters directly if preferred. The command is not case sensitive.

ISBN 978-1-009-11049-5

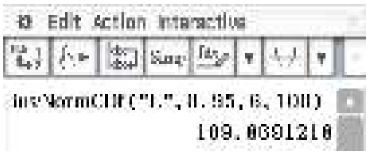
Using the Casio ClassPad

Method 1

- In Main $\sqrt{\alpha}$, go to **Interactive** > **Distribution** > **Inverse** > **InvNormCDf**.
- Set the 'Tail setting' as 'Left'.
- Enter the probability, 0.95, to the left of the required value k.
- Enter the standard deviation σ and the mean μ.
- Tap OK.

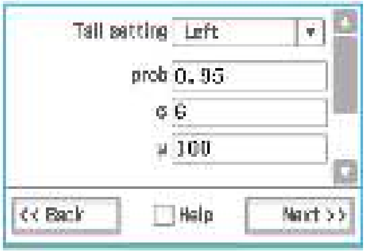


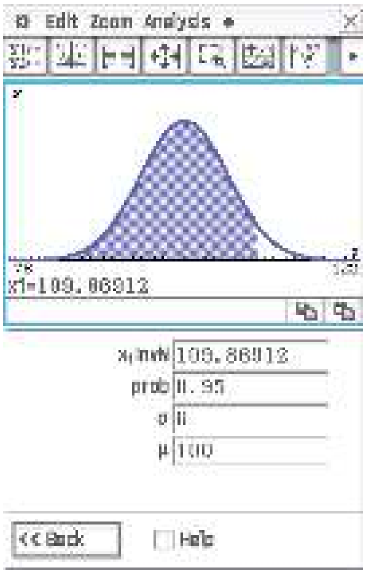
Note: The tail setting is 'Left' to indicate that we seek the value k such that 95% of the area lies to the left of k for this normal distribution.



Method 2

- In the **Statistics** application, go to **Calc** > **Inv. Distribution**.
- Select **Inverse Normal CD** and tap Next.
- Set the 'Tail setting' as 'Left'.
- Enter the probability, 0.95, to the left.
- Enter the standard deviation σ and the mean μ.
- Tap Next to view the answer.
- Select the graph icon to view the graph with the answer.





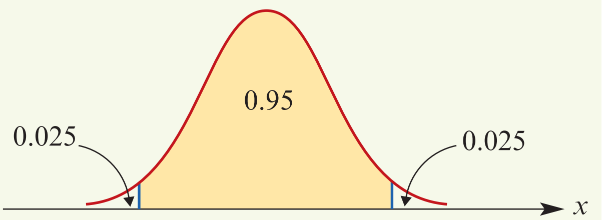
Example 6

Suppose X is normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 6$.

Find c_1 and c_2 such that $\Pr(c_1 < X < c_2) = 0.95$.

Solution

Examining the normal curve, we see that there are (infinitely) many intervals which enclose an area of 0.95. By convention, we choose the interval which leaves equal areas in each tail.



To find c_1 using the inverse-normal facility of your calculator, enter 0.025 as the area. To find c_2, enter 0.975.

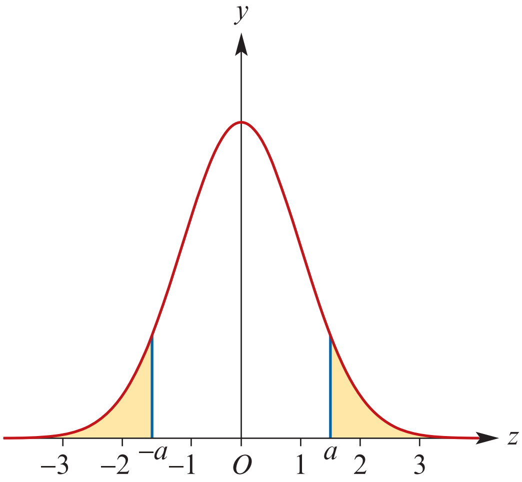
This will give the answer $c_1 = 88.240$ and $c_2 = 111.760$.

Symmetry properties

Probabilities associated with a normal distribution can often be determined by using its symmetry properties.

Here we work with the standard normal distribution, as it is easiest to use the symmetry properties in this situation:

- $\Pr(Z > a) = 1 - \Pr(Z \le a)$
- $\Pr(Z < -a) = \Pr(Z > a)$
- $\Pr(-a < Z < a) = 1 - 2\Pr(Z \ge a)$
  $= 1 - 2\Pr(Z \le -a)$



Exercise 16C

Example 4

1 Suppose Z is a standard normal random variable (that is, it has mean $\mu = 0$ and standard deviation $\sigma = 1$). Find the following probabilities, drawing an appropriate diagram in each case:

a $\Pr(Z < 2)$ **b** $\Pr(Z < 2.5)$ **c** $\Pr(Z \le 2.5)$ **d** $\Pr(Z < 2.53)$
e $\Pr(Z \ge 2)$ **f** $\Pr(Z > 1.5)$ **g** $\Pr(Z \ge 0.34)$ **h** $\Pr(Z > 1.01)$

2 Suppose Z is a standard normal random variable. Find the following probabilities, drawing an appropriate diagram in each case:

a $\Pr(Z > -2)$ **b** $\Pr(Z > -0.5)$ **c** $\Pr(Z > -2.5)$ **d** $\Pr(Z \ge -1.283)$
e $\Pr(Z < -2)$ **f** $\Pr(Z < -2.33)$ **g** $\Pr(Z \le -1.8)$ **h** $\Pr(Z \le -0.95)$

ISBN 978-1-009-11049-5 Cambridge University Press

3 Suppose Z is a standard normal random variable. Find the following probabilities, drawing an appropriate diagram in each case:

a $\Pr(-1 < Z < 1)$ **b** $\Pr(-2 < Z < 2)$ **c** $\Pr(-3 < Z < 3)$

How do these results compare with the 68–95–99.7% rule discussed in Section 16B?

4 Suppose Z is a standard normal random variable. Find the following probabilities, drawing an appropriate diagram in each case:

a $\Pr(2 < Z < 3)$ **b** $\Pr(-1.5 < Z < 2.5)$

c $\Pr(-2 < Z < -1.5)$ **d** $\Pr(-1.4 < Z < -0.8)$

Example 5 **5** Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \leq c) = 0.9$.

6 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \leq c) = 0.75$.

7 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \leq c) = 0.975$.

8 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \geq c) = 0.95$.

9 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \geq c) = 0.8$.

10 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \leq c) = 0.10$.

11 Suppose Z is a standard normal random variable. Draw an appropriate diagram and then find the value c such that $\Pr(Z \leq c) = 0.025$.

12 Let X be a normal random variable with mean $\mu = 100$ and standard deviation $\sigma = 6$. Find:

a $\Pr(X < 110)$ **b** $\Pr(X < 105)$ **c** $\Pr(X > 110)$ **d** $\Pr(105 < X < 110)$

13 Let X be a normal random variable with mean $\mu = 40$ and standard deviation $\sigma = 5$. Find:

a $\Pr(X < 48)$ **b** $\Pr(X < 36)$ **c** $\Pr(X > 32)$ **d** $\Pr(32 < X < 36)$

14 Let X be a normal random variable with mean $\mu = 6$ and standard deviation $\sigma = 2$.

a Find c such that $\Pr(X < c) = 0.95$.

b Find k such that $\Pr(X < k) = 0.90$.

15 Let X be a normal random variable with mean $\mu = 10$ and standard deviation $\sigma = 3$.

a Find c such that $\Pr(X < c) = 0.50$.

b Find k such that $\Pr(X < k) = 0.975$.

ISBN 978-1-009-11049-5 Cambridge University Press

16 The 68–95–99.7% rule tells us approximately the percentage of a normal distribution which lies within one, two or three standard deviations of the mean. If Z is the standard normal random variable find, correct to two decimal places:

a a such that $\Pr(-a < Z < a) = 0.68$ **b** b such that $\Pr(-b < Z < b) = 0.95$

c c such that $\Pr(-c < Z < c) = 0.997$

17 Given that X is a normally distributed random variable with a mean of 22 and a standard deviation of 7, find:

a $\Pr(X < 26)$ **b** $\Pr(25 < X < 27)$

c $\Pr(X < 26 \mid 25 < X < 27)$ **d** c such that $\Pr(X < c) = 0.95$

e k such that $\Pr(X > k) = 0.9$ **f** c_1 and c_2 such that $\Pr(c_1 < X < c_2) = 0.95$

18 Let X be a normal random variable with mean $\mu = 10$ and standard deviation $\sigma = 0.5$. Find:

a $\Pr(X < 11)$ **b** $\Pr(X < 11 \mid X < 13)$

c c such that $\Pr(X < c) = 0.95$ **d** k such that $\Pr(X < k) = 0.2$

e c_1 and c_2 such that $\Pr(c_1 < X < c_2) = 0.95$

16D Solving problems using the normal distribution

The normal distribution can be used to solve many practical problems.

Example 7

The time taken to complete a logical reasoning task is normally distributed with a mean of 55 seconds and a standard deviation of 8 seconds.

a Find the probability, correct to four decimal places, that a randomly chosen person will take less than 50 seconds to complete the task.

b Find the probability, correct to four decimal places, that a randomly chosen person will take less than 50 seconds to complete the task, if it is known that this person took less than 60 seconds to complete the task.

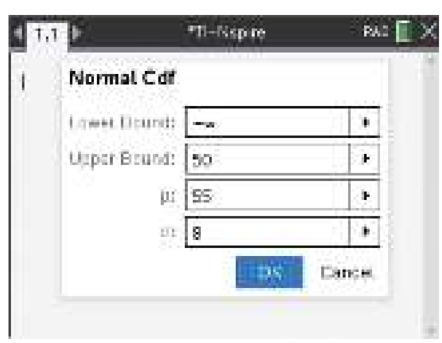
Using the TI-Nspire

a Method 1

Use (menu) > **Probability** > **Distributions** > **Normal Cdf** and complete as shown.

The answer is:

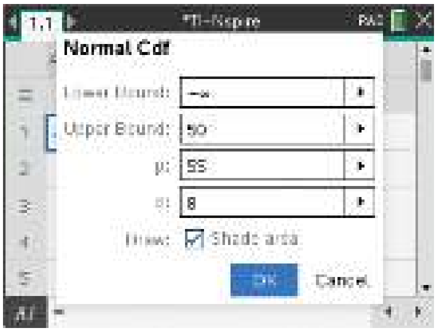
$\Pr(X < 50) = 0.2660$

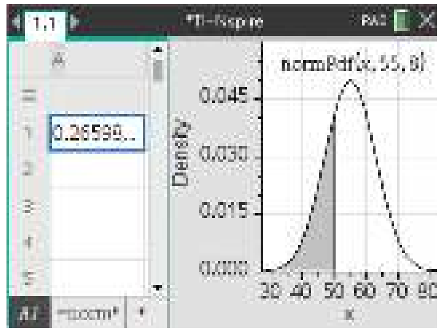


ISBN 978-1-009-11049-5

Method 2

You can also solve this problem in a **Lists & Spreadsheet** page and plot the graph. Use (menu) > **Statistics** > **Distributions** > **Normal Cdf** and complete as shown below.





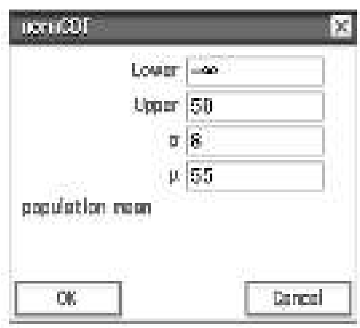
b $\Pr(X < 50 \mid X < 60) = \dfrac{\Pr(X < 50 \cap X < 60)}{\Pr(X < 60)}$

$$= \frac{\Pr(X < 50)}{\Pr(X < 60)} = \frac{0.2660}{0.7340} = 0.3624$$

Using the Casio ClassPad

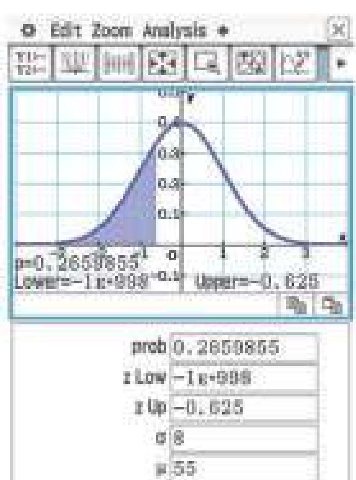
a Method 1

- In $\sqrt[\text{Main}]{\alpha}$, go to **Interactive** > **Distribution** > **Continuous** > **normCDf**.
- Enter values for the lower and upper bounds, the standard deviation and the mean. Tap OK.



Method 2

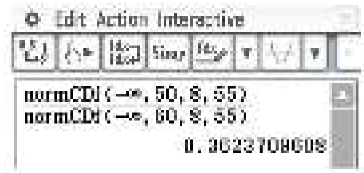
- In (Statistics icon), go to **Calc** > **Distribution** and select **Normal CD**. Tap Next.
- Enter the lower and upper bounds, the standard deviation and the mean.
- Tap Next to view the answer.
- Select (graph icon) to view the graph with the answer.



ISBN 978-1-009-11049-5 Cambridge University Press

b ■ In Main $\sqrt{\alpha}$, select the fraction template.

■ Enter as shown and tap EXE.



When the mean and standard deviation of a normal distribution are unknown, it is sometimes necessary to transform to the standard normal distribution. This is demonstrated in the following example.

Example 8

Limits of acceptability imposed on the lengths of a certain batch of metal rods are 1.925 cm and 2.075 cm. It is observed that, on average, 5% are rejected as undersized and 5% are rejected as oversized.

Assuming that the lengths are normally distributed, find the mean and standard deviation of the distribution.

Solution

It is given that $\Pr(X > 2.075) = 0.05$ and $\Pr(X < 1.925) = 0.05$.

Symmetry tells us that the mean is equal to

$$\mu = \frac{2.075 + 1.925}{2} = 2$$

Transforming to the standard normal gives

$$\Pr\left(Z > \frac{2.075 - \mu}{\sigma}\right) = 0.05 \quad \text{and} \quad \Pr\left(Z < \frac{1.925 - \mu}{\sigma}\right) = 0.05$$

The first equality can be rewritten as

$$\Pr\left(Z < \frac{2.075 - \mu}{\sigma}\right) = 0.95$$

Use the inverse-normal facility of your calculator to obtain

$$\frac{2.075 - \mu}{\sigma} = 1.6448\ldots \qquad \text{and} \qquad \frac{1.925 - \mu}{\sigma} = -1.6448\ldots$$

These equations confirm that $\mu = 2$.

Substitute $\mu = 2$ into the first equation and solve for σ:

$$\frac{2.075 - 2}{\sigma} = 1.6448\ldots$$

$$\therefore \qquad \sigma = 0.045596\ldots$$

Thus $\sigma = 0.0456$, correct to four decimal places.

ISBN 978-1-009-11049-5 Cambridge University Press

Skill-sheet

Exercise 16D

1 Suppose that IQ scores are normally distributed with mean $\mu = 100$ and standard deviation $\sigma = 15$.

Example 7

a What is the probability that a person chosen at random has an IQ:

i greater than 110

ii less than 75

iii greater than 130, given that they have an IQ greater than 110?

b To be allowed to join an elite club, a potential member must have an IQ in the top 5% of the population. What IQ score would be necessary to join this club?

2 The heights of women are normally distributed with a mean of $\mu = 160$ cm and a standard deviation of $\sigma = 8$ cm.

a What is the probability that a woman chosen at random would be:

i taller than 155 cm

ii shorter than 170 cm

iii taller than 170 cm, given that her height is between 168 cm and 174 cm?

b What height would put a woman among the tallest 10% of the population?

c What height would put a woman among the shortest 20% of the population?

3 The results of a mathematics exam are normally distributed with mean $\mu = 50$ and standard deviation $\sigma = 7$.

a What is the probability that a student chosen at random has an exam mark:

i greater than 60

ii less than 75

iii greater than 60, given that they passed? (Assume a pass mark of 50.)

b The top 15% of the class are to be awarded a distinction. What mark would be required to gain a distinction in this exam?

4 The lengths of a species of fish are normally distributed with a mean length of 40 cm and a standard deviation of 4 cm. Find the percentage of these fish having lengths:

a greater than 45 cm

b between 35.5 cm and 45.5 cm.

Example 8

5 The weights of cats are normally distributed. It is known that 10% of cats weigh more than 1.8 kg, and 15% of cats weigh less than 1.35 kg. Find the mean and the standard deviation of this distribution.

6 The marks of a large number of students in a statistics examination are normally distributed with a mean of 48 marks and a standard deviation of 15 marks.

a If the pass mark is 53, find the percentage of students who passed the examination.

b If 8% of students gained an A on the examination by scoring a mark of at least c, find the value of c.

ISBN 978-1-009-11049-5

7 The height of a certain population of adult males is normally distributed with mean 176 cm and standard deviation 7 cm.

a Find the probability that the height of a randomly selected male will exceed 190 cm.

b If two males are selected at random, find the probability that both of their heights will exceed 190 cm.

c Suppose 10 males are selected at random. Find the probability that at least two will have heights that exceed 190 cm.

8 **a** Machine A is packaging bags of mints with a mean weight of 300 grams. The bags are considered underweight if they weigh less than 295 grams. It is observed that, on average, 5% of bags are rejected as underweight. Assuming that the weights of the bags are normally distributed, find the standard deviation of the distribution.

b In the same factory, machine B is packaging bags of liquorice. The bags from this machine are considered underweight if they weigh less than 340 grams. It is observed that, on average, 2% of bags from machine B are rejected as underweight. Assuming that the weights are normally distributed with a standard deviation of 5 grams, find the mean of the distribution.

9 The volume of soft drink in a 1-litre bottle is normally distributed. The soft drink company needs to calibrate its filling machine. They don't want to put too much soft drink into each bottle, as it adds to their expense. However, they know they will be fined if more than 2% of bottles are more than 2 millilitres under volume. The standard deviation of the volume dispensed by the filling machine is 2.5 millilitres. What should they choose as the target volume (i.e. the mean of the distribution)? Give your answer to the nearest millilitre.

10 The weights of pumpkins sold to a greengrocer are normally distributed with a mean of 1.2 kg and a standard deviation of 0.4 kg. The pumpkins are sold in three sizes:

Small: under 0.8 kg   Medium: from 0.8 kg to 1.8 kg   Large: over 1.8 kg

a Find the proportions of pumpkins in each of the three sizes.

b The prices of the pumpkins are $2.80 for a small, $3.50 for a medium, and $5.00 for a large. Find the expected cost for 100 pumpkins chosen at random from the greengrocer's supply.

11 Potatoes are delivered to a chip factory in semitrailer loads. A sample of 1 kg of the potatoes is chosen from each load and tested for starch content. From past experience it is known that the starch content is normally distributed with a standard deviation of 2.1.

a For a semitrailer load of potatoes with a mean starch content of 22.0:

i What is the probability that the test reading is 19.5 or less?

ii What reading will be exceeded with a probability of 0.98?

b If the starch content is greater than 22.0, the potatoes cannot be used for chips, and so the semitrailer load is rejected. What is the probability that a load with a mean starch content of 18.0 will be rejected?

ISBN 978-1-009-11049-5 Cambridge University Press

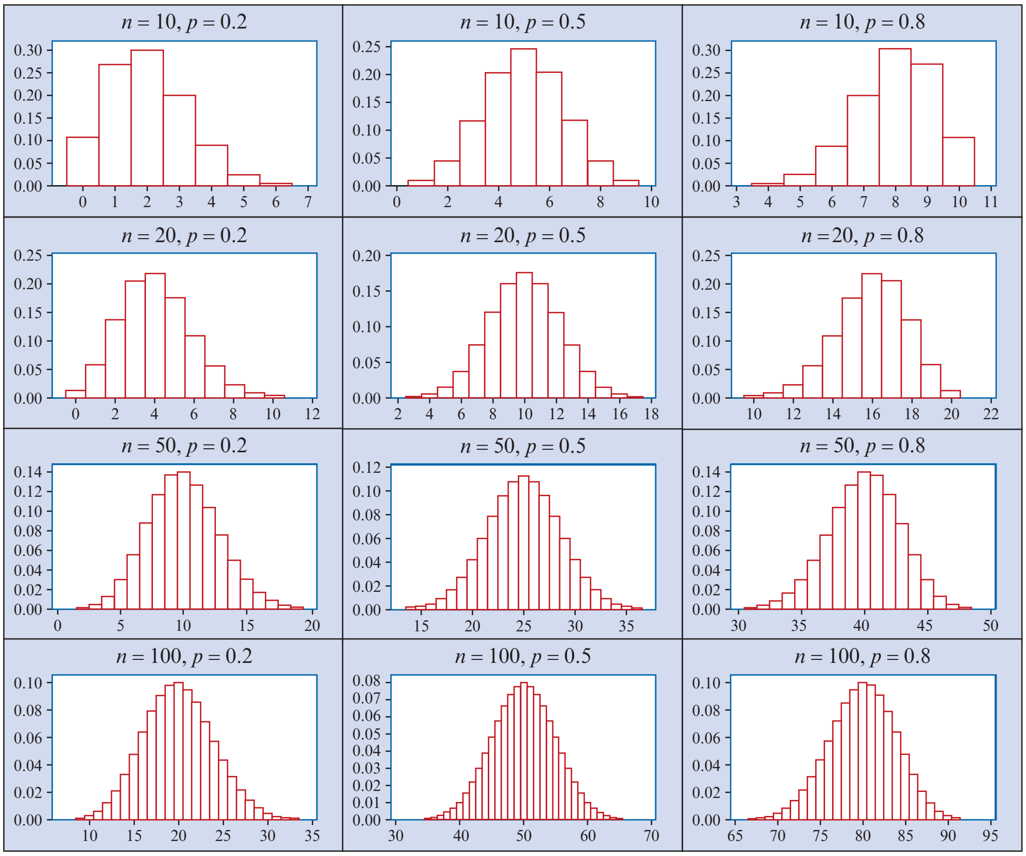
12 The amount of a certain chemical in a type A cell is normally distributed with a mean of 10 and a standard deviation of 1. The amount in a type B cell is normally distributed with a mean of 14 and a standard deviation of 2. To determine whether a cell is type A or type B, the amount of chemical in the cell is measured. The cell is classified as type A if the amount is less than a specified value c, and as type B otherwise.

a If $c = 12$, calculate the probability that a type A cell will be misclassified, and the probability that a type B cell will be misclassified.

b Find the value of c for which the two probabilities of misclassification are equal.

16E The normal approximation to the binomial distribution

We saw in Chapter 14 that the shape of the binomial distribution depends on n and p. The following plots show the binomial distribution for $n = 10, 20, 50, 100$ and $p = 0.2, 0.5, 0.8$.



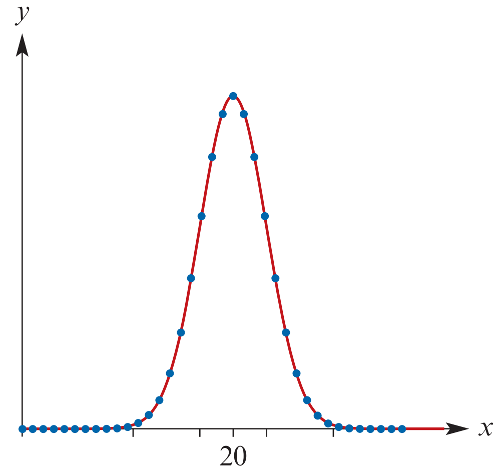
We can see that, if n is small and p is close to 0 or 1, these distributions are skewed. Otherwise, they look remarkably symmetric. In fact, if n is large enough and p is not too close to 0 or 1, the binomial distribution is approximately normal. Moreover, the mean and standard deviation of this normal distribution agree with those of the binomial distribution.

ISBN 978-1-009-11049-5 Cambridge University Press

In the figure opposite, the binomial distribution with $n = 40$ and $p = 0.5$ is plotted (the blue points). This distribution has mean $\mu = 20$ and standard deviation $\sigma = \sqrt{10}$.

On the same axes, the probability density function of the normal distribution with mean $\mu = 20$ and standard deviation $\sigma = \sqrt{10}$ is drawn (the red curve).

We will see that this approximation has important uses in statistics.



When is it appropriate to use the normal approximation?

If n is large enough, the skew of the binomial distribution is not too great. In this case, the normal distribution can be used as a reasonable approximation to the binomial distribution. The approximation is generally better for larger n and when p is not too close to 0 or 1.

If n is sufficiently large, the binomial random variable X will be approximately normally distributed, with a mean of $\mu = np$ and a standard deviation of $\sigma = \sqrt{np(1-p)}$.

One rule of thumb is that:

Both np and $n(1-p)$ must be greater than 5 for a satisfactory approximation.

In the example shown in the figure above, we have $np = 20$ and $n(1-p) = 20$. There are ways of improving this approximation but we will not go into that here.

Example 9

A sample of 1000 people from a certain city were asked to indicate whether or not they were in favour of the construction of a new freeway. It is known that 30% of people in this city are in favour of the new freeway. Find the approximate probability that between 270 and 330 people in the sample were in favour of the new freeway.

Solution

Let X be the number of people in the sample who are in favour of the freeway. Then we can assume that X is a binomial random variable with $n = 1000$ and $p = 0.3$.

Therefore

$$\begin{aligned}\mu &= np \\ &= 1000 \times 0.3 \\ &= 300\end{aligned} \qquad \text{and} \qquad \begin{aligned}\sigma &= \sqrt{np(1-p)} \\ &= \sqrt{1000 \times 0.3 \times 0.7} \\ &= \sqrt{210}\end{aligned}$$

ISBN 978-1-009-11049-5 Cambridge University Press

Thus

$$\begin{aligned}\Pr(270 < X < 330) &\approx \Pr\left(\frac{270-300}{\sqrt{210}} < Z < \frac{330-300}{\sqrt{210}}\right)\\ &\approx \Pr(-2.070 < Z < 2.070)\\ &\approx 0.9616\end{aligned}$$

Note: When we calculate this probability directly using the binomial distribution, we find that $\Pr(270 \le X \le 330) = 0.9648$ and $\Pr(270 < X < 330) = 0.9583$.

Exercise 16E

In each of the following questions, use the normal approximation to the binomial distribution.

Example 9

1 A die is rolled 100 times. What is the probability that more than 10 sixes will be observed?

2 If 50% of the voting population in a particular state favour candidate A, what is the approximate probability that more than 156 in a sample of 300 will favour that candidate.

3 A sample of 100 people is drawn from a city in which it is known that 10% of the population is over 65 years of age. Find the approximate probability that the sample contains:

- **a** at least 15 people who are over 65 years of age
- **b** no more than 8 people over 65 years of age.

4 A manufacturing process produces on average 40 defective items per 1000. What is the approximate probability that a random sample of size 400 contains:

- **a** at least 10 and no more than 20 defective items
- **b** 25 or more defective items?

5 A survey of the entire population in a particular city found that 40% of people regularly participate in sport. What is the approximate probability that fewer than 38% of a random sample of 200 people regularly participate in sport?

6 An examination consists of 25 multiple-choice questions. Each question has four possible answers. At least 10 correct answers are required to pass the examination. Suppose that a student guesses the answer to each question.

- **a** What is the approximate probability that the student will pass the examination?
- **b** What is the approximate probability that the student guesses between 12 and 14 answers correctly?

ISBN 978-1-009-11049-5

Chapter summary

Assign-ment

Nrich

- A special continuous random variable X, called a **normal random variable**, has a probability density function given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

where μ and σ are the mean and standard deviation of X.
- In the special case that $\mu = 0$ and $\sigma = 1$, this probability density function defines the **standard normal distribution**. A random variable with this distribution is usually denoted by Z.
- The graph of a normal density function is a symmetric, bell-shaped curve; its centre is determined by the mean, μ, and its width by the standard deviation, σ.
- The **68–95–99.7% rule** states that, for any normal distribution:
  - approximately 68% of the values lie within one standard deviation of the mean
  - approximately 95% of the values lie within two standard deviations of the mean
  - approximately 99.7% of the values lie within three standard deviations of the mean.
- If X is a normally distributed random variable with mean μ and standard deviation σ, then to **standardise** a value x of X we subtract the mean and divide by the standard deviation:

$$z = \frac{x - \mu}{\sigma}$$

The standardised value z indicates the number of standard deviations that the value x lies above or below the mean.
- A calculator can be used to evaluate the cumulative distribution function of a normal random variable – that is, to find the area under the normal curve up to a specified value.
- The inverse-normal facility of a calculator can be used to find the value of a normal random variable corresponding to a specified area under the normal curve.

Technology-free questions

1 Given that $\Pr(Z \le a) = p$ for the standard normal random variable Z, find in terms of p:

a $\Pr(Z > a)$ **b** $\Pr(Z < -a)$ **c** $\Pr(-a \le Z \le a)$

2 Let X be a normal random variable with mean 4 and standard deviation 1. Let Z be the standard normal random variable.

a If $\Pr(X < 3) = \Pr(Z < a)$, then $a =$ ____.

b If $\Pr(X > 5) = \Pr(Z > b)$, then $b =$ ____.

c $\Pr(X > 4) =$ ____

3 A normal random variable X has mean 8 and standard deviation 3. Give the rule for a transformation that maps the graph of the density function of X to the graph of the density function for the standard normal distribution.

ISBN 978-1-009-11049-5 Cambridge University Press

4 Let X be a normal random variable with mean μ and standard deviation σ. If $\mu < a < b$ with $\Pr(X < b) = p$ and $\Pr(X < a) = q$, find:

a $\Pr(X < a \mid X < b)$ **b** $\Pr(X < 2\mu - a)$ **c** $\Pr(X > b \mid X > a)$

5 Let X be a normal random variable with mean 4 and standard deviation 2. Write each of the following probabilities in terms of Z:

a $\Pr(X < 5)$ **b** $\Pr(X < 3)$ **c** $\Pr(X > 5)$

d $\Pr(3 < X < 5)$ **e** $\Pr(3 < X < 6)$

In Questions 6 to 8, you will use the following:

$\Pr(Z<1) = 0.84$ $\Pr(Z<2) = 0.98$ $\Pr(Z<0.5) = 0.69$

6 A machine produces metal rods with mean diameter 2.5 mm and standard deviation 0.05 mm. Let X be the random variable of the normal distribution. Find:

a $\Pr(X < 2.55)$ **b** $\Pr(X < 2.5)$

c $\Pr(X < 2.45)$ **d** $\Pr(2.45 < X < 2.55)$

7 Nuts are packed in tins such that the mean weight of the tins is 500 g and the standard deviation is 5 g. The weights are normally distributed with random variable W. Find:

a $\Pr(W > 505)$ **b** $\Pr(500 < W < 505)$

c $\Pr(W > 505 \mid W > 500)$ **d** $\Pr(W > 510)$

8 A random variable X has a normal distribution with mean 6 and standard deviation 1. Find:

a $\Pr(X < 6.5)$ **b** $\Pr(6 < X < 6.5)$

c $\Pr(6.5 < X < 7)$ **d** $\Pr(5 < X < 7)$

9 Suppose that three tests were given in your mathematics course. The class means and standard deviations, together with your scores, are listed in the table.

	μ	σ	Your score
Test A	50	11	62
Test B	47	17	64
Test C	63	8	73

On which test did you do best and on which did you do worst?

10 Let X be a normally distributed random variable with mean 10 and variance 4, and let Z be a random variable with the standard normal distribution.

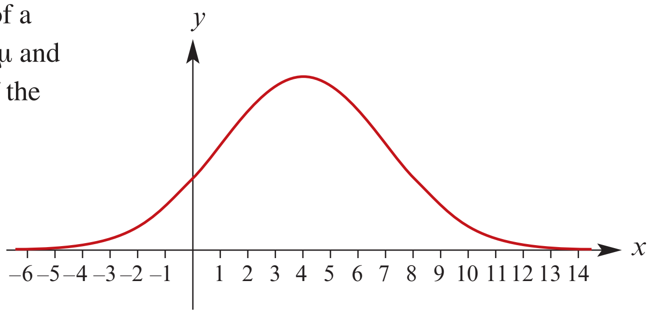
a Find $\Pr(X > 10)$.

b Find b such that $\Pr(X > 13) = \Pr(Z < b)$.

ISBN 978-1-009-11049-5 Cambridge University Press

Multiple-choice questions

1 The diagram shows the graph of a normal distribution with mean μ and standard deviation σ. Which of the following statements is true?



A $\mu = 4$ and $\sigma = 3$
B $\mu = 3$ and $\sigma = 4$
C $\mu = 4$ and $\sigma = 2$
D $\mu = 3$ and $\sigma = 2$
E $\mu = 4$ and $\sigma = 4$

2 If Z is a standard normal random variable and $\Pr(Z < c) = 0.25$, then the value of c is closest to

A 0.6745 **B** −0.6745 **C** 0.3867 **D** 0.5987 **E** −0.5987

3 The random variable X has a normal distribution with mean 12 and variance 9. If Z is a standard normal random variable, then the probability that X is more than 15 is equal to

A $\Pr(Z < 1)$ **B** $\Pr(Z > 1)$ **C** $\Pr\left(Z > \frac{1}{3}\right)$
D $1 - \Pr\left(Z > \frac{1}{3}\right)$ **E** $1 - \Pr(Z > 1)$

4 The actual length of an AFL game is normally distributed with a mean of 102 minutes. If the percentage of games that last more than 110 minutes is approximately 0.38% then the standard deviation of the distribution is

A 1 **B** 2 **C** 3 **D** 4 **E** 5

5 If X is a normally distributed random variable with mean $\mu = 6$ and standard deviation $\sigma = 3$, then the transformation which maps the graph of the density function f of X to the graph of the standard normal distribution is

A $(x, y) \to \left(\frac{x-3}{6}, 6y\right)$ **B** $(x, y) \to \left(\frac{x-6}{3}, \frac{y}{3}\right)$ **C** $(x, y) \to \left(\frac{x-6}{3}, 3y\right)$
D $(x, y) \to (3(x+6), 3y)$ **E** $(x, y) \to \left(3(x+6), \frac{y}{3}\right)$

6 The amount of water that Steve uses to water the garden is normally distributed with a mean of 100 litres and a standard deviation of 14 litres. On 20% of occasions it takes him more than k litres to water the garden. What is the value of k?

A 88.2 **B** 110.7 **C** 120.0 **D** 111.8 **E** 114.0

7 The marks achieved by Angie in Mathematics, Indonesian and Politics, together with the mean and standard deviation for each subject, are given in the following table:

ISBN 978-1-009-11049-5 Cambridge University Press

Subject	Mark	Mean (μ)	Standard deviation (σ)
Mathematics	72	72	5
Indonesian	57	59	2
Politics	68	64	4

Which of the following statements is correct?

A Angie's best subject was Politics, followed by Mathematics and then Indonesian.
B Angie's best subject was Mathematics, followed by Politics and then Indonesian.
C Angie's best subject was Politics, followed by Indonesian and then Mathematics.
D Angie's best subject was Mathematics, followed by Indonesian and then Politics.
E Angie's best subject was Indonesian, followed by Mathematics and then Politics.

8 Suppose that X is normally distributed with mean 11.3 and standard deviation 2.9, and Z is the standard normal random variable. $\Pr(-1 < Z < 2)$ is equal to

A $\Pr(8.4 < X < 17.1)$ **B** $\Pr(-14.2 < X < -5.5)$ **C** $\Pr(-2.9 < X < 5.8)$
D $\Pr(-8.4 < X < 26.7)$ **E** $\Pr(-11.3 < X < 22.6)$

9 The volume of liquid in a 1-litre bottle of soft drink is a normally distributed random variable with a mean of μ litres and a standard deviation of 0.005 litres. To ensure that 99.9% of the bottles contain at least 1 litre of soft drink, the value of μ should be closest to

A 0.995 litres **B** 1.0 litres **C** 1.005 litres **D** 1.015 litres **E** 1.026 litres

10 The gestation period for human pregnancies in a certain country is normally distributed with a mean of 272 days and a standard deviation of σ days. If from a population of 1000 births there were 91 pregnancies of length less than 260 days, then σ is closest to

A 3 **B** 5 **C** 9 **D** 12 **E** 16

Extended-response questions

1 A test devised to measure mathematical aptitude gives scores that are normally distributed with a mean of 50 and a standard deviation of 10. If we wish to categorise the results so that the highest 10% of scores are designated as high aptitude, the next 20% as moderate aptitude, the middle 40% as average, the next 20% as little aptitude and the lowest 10% as no aptitude, then what ranges of scores will be covered by each of these five categories?

2 The amount of anaesthetic required to cause surgical anaesthesia in patients is normally distributed, with a mean of 50 mg and a standard deviation of 10 mg. The lethal dose is also normally distributed, with a mean of 110 mg and a standard deviation of 20 mg. If a dosage that brings 90% of patients to surgical anaesthesia were used, what percentage of patients would be killed by this dose?

ISBN 978-1-009-11049-5 Cambridge University Press

3 Records kept by a manufacturer of car tyres suggest that the distribution of the mileage from their tyres is normal, with mean 60 000 km and standard deviation 5000 km.

a What proportion of the company's tyres last:

i less that 55 000 km

ii more than 50 000 km but less than 74 000 km

iii more than 72 000 km, given that they have already lasted more than 60 000 km?

b The company's advertising manager wishes to claim that '90% of our tyres last longer than c km'. What should c be?

c What is the probability that a customer buys five tyres at the same time and finds that they all last more than 72 000 km?

4 Suppose that L, the useful life (in hours) of a fluorescent tube manufactured by Company A is normally distributed with a mean of 600 and a standard deviation of 4.

a Find correct to four decimal places the probability that a tube lasts longer than 605 hours.

b Find correct to four decimal places the probability that a tube lasts longer than 607 hours, given that it lasts longer than 605 hours.

Company B, also manufactures fluorescent tubes. The useful life of the tubes from this company is a random variable with probability density function

$$b(x) = \begin{cases} \dfrac{1}{27648}(x-588)(612-x)^2 & 588 \le x \le 612 \\ 0 & elsewhere \end{cases}$$

c The fluorescent tubes from Company B are distributed to shops in boxes of 10. Find, correct to three decimal places, the probability that at least three of the tubes in a randomly selected box last longer than 605 hours.

The local lighting store stocks fluorescent globes from both Company A and Company B, in equal quantities. A customer comes into the store and randomly selects a single fluorescent globe.

d What is the probability that the globe they select lasts longer than 605 hours? Give your answer correct to three decimal places.

e Given that the globe selected lasts longer than 605 hours, what is the probability that it was manufactured by Company B? Give your answer correct to three decimal places.

5 In a given manufacturing process, components are rejected if they have a particular dimension greater than 60.4 mm or less than 59.7 mm. It is found that 3% are rejected as being too large and 5% are rejected as being too small. Assume that the dimension is normally distributed.

a Find the mean and standard deviation of the distribution of the dimension, correct to one decimal place.

b Use the result of **a** to find the percentage of rejects if the limits for acceptance are changed to 60.3 mm and 59.6 mm.

ISBN 978-1-009-11049-5 Cambridge University Press

6 The hardness of a metal may be determined by impressing a hardened point into the surface of the metal and then measuring the depth of penetration of the point. Suppose that the hardness of a particular alloy is normally distributed with mean 70 and standard deviation 3.

a If a specimen is acceptable only if its hardness is between 65 and 75, what is the probability that a randomly chosen specimen has an acceptable hardness?

b If the acceptable range of hardness was $(70 - c, 70 + c)$, for what value of c would 95% of all specimens have acceptable hardness?

c If the acceptable range is the same as in **a**, and the hardness of each of 10 randomly selected specimens is independently determined, what is the expected number of acceptable specimens among the 10?

d What is the probability that at most eight out of 10 randomly selected specimens have a hardness less than 73.84?

e The profit on an acceptable specimen is \$20, while unacceptable specimens result in a loss of \$5. If $\$P$ is the profit on a randomly selected specimen, find the mean and variance of P.

7 The weekly error (in seconds) of a brand of watch is known to be normally distributed. Only those watches with an error of less than 5 seconds are acceptable.

a Find the mean and standard deviation of the distribution of error if 3% of watches are rejected for losing time and 3% are rejected for gaining time.

b Determine the probability that fewer than two watches are rejected in a batch of 10 such watches.

8 A brand of detergent is sold in bottles of two sizes: standard and large. For each size, the content (in litres) of a randomly chosen bottle is normally distributed with mean and standard deviation as given in the table:

	Mean	Standard deviation
Standard bottle	0.760	0.008
Large bottle	1.010	0.009

a Find the probability that a randomly chosen standard bottle contains less than 0.75 litres.

b Find the probability that a box of 10 randomly chosen standard bottles contains at least three bottles whose contents are each less than 0.75 litres.

c Using the results

$$\mathrm{E}(aX - bY) = a\mathrm{E}(X) - b\mathrm{E}(Y)$$
$$\mathrm{Var}(aX - bY) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y)$$

find the probability that there is more detergent in four randomly chosen standard bottles than in three randomly chosen large bottles. (Assume that $aX - bY$ is normally distributed.)

ISBN 978-1-009-11049-5 Cambridge University Press