11 General continuous random variables

11.1 Overview

Continuous random variables such as height have values that depend upon the precision of the devices used to measure them. Modern technology has allowed standardised measurements of length to a reasonable level of precision; however, the first 'standardised' units of length used over five thousand years ago were a bit unreliable. Length was described in terms of cubits, where one cubit was equal to the distance between the elbow and the tips of the finger. Obviously, with so much variation between human beings, an 'official cubit' had to be decided upon, but even then there were problems. Cubit rods found by archaeologists have varied from 47.2 to 52.5 cm in length, with each cubit broken up into 5 to 7 'palms' and a palm broken up into anywhere between 4 and 7 'fingers'. So, depending on the cubit rod used, the same person could be described as being 3 cubits, 2 palms and 1 finger tall; 3 cubits, 5 palms and 5 fingers tall; or somewhere in between.

Today, when people measure their height they can be confident that they will get the same value each time when rounded to the nearest whole centimetre. The size of this centimetre is standardised as one-hundredth of a metre. The metre itself is defined as being the distance travelled by light in a vacuum in $\frac{1}{299\,792\,458}$ of a second.

LEARNING SEQUENCE

Fully worked solutions for this chapter are available in the Resources section of your eBookPLUS at www.jacplus.com.au.

11.2 Continuous random variables and the probability density function

In Chapter 10, we dealt with discrete random variables, that is, data which is finite or countable. The number of white cars in a car park, the number of students in a class and the number of lollies in a jar are all examples of discrete random variables. A **continuous random variable**, however, can assume any value within a given range.

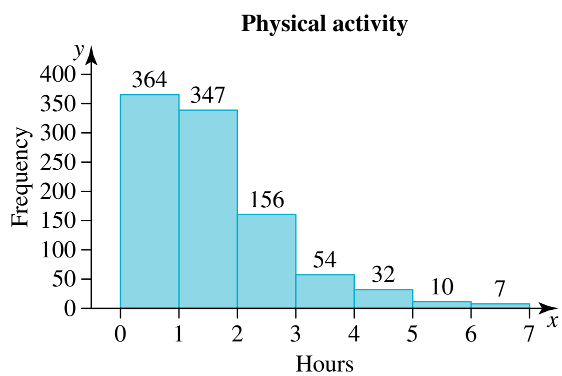
11.2.1 Discrete and continuous variables

For example, imagine that you are measuring the length of a piece of string. Using a metre ruler, you find that its length is somewhere between 9 and 10 centimetres. But is there any way of determining an *exact* value for its length?

This is the fundamental difference between discrete and continuous variables: although you can count the number of white cars in a car park and get an exact number, there is no such thing as an exact value for a continuous variable. You can use more and more precise instruments to measure that string length. For example, you might get 9.5 centimetres using a school ruler, 9.492 centimetres using engineering callipers, or even 9.492 000 000 862 centimetres using a high precision laser. But this is only narrowing the range of values within which a theoretical 'exact' value would lie — determining if it lies between 9 and 10, between 9.4 and 9.5, between 9.49 and 9.50 centimetres and so on. As a result, a continuous random variable is described in terms of the **interval** in which its value lies.

Examples of continuous random variables include time, height and weight — all quantities that must be measured rather than counted.

Consider an Australian health study that was conducted. The study targeted young people aged 5 to 17 years old. They were asked to estimate the average number of hours of physical activity they participated in each week. The results of this study are shown in the following histogram.



The frequencies of individual activity times cannot be determined due to the fact that the times have been grouped into class intervals. This limits the information we are able to extract from the histogram. For example, to determine the number of young people who did less than 2 hours of activity, we simply add the frequencies of the $0 < x < 1$ and $1 < x < 2$ class intervals; that is, $364 + 347 = 711$ people. However, we would not be able to determine the number of people who did less than 1.5 hours of physical activity as this value lies within a class interval rather than being an end point.

Since the value that a continuous random variable can assume is always measured in some way, exact values are unable to be obtained. As there are an infinite number of values that the variable can have within that interval, the probability of a continuous random variable assuming an exact value is zero.

11.2.2 Probability and relative frequency

Because histograms are drawn such that the intervals for each class are of uniform width, the probability of a variable lying within a particular interval is also equal to the fraction of the histogram area that the column for that interval occupies. In other words, the probability that x lies within a particular interval on a histogram is approximately equal to the **relative frequency** of the interval:

$$\mathrm{P}(a < x < b) = \frac{f(a < x < b)}{\sum f}$$

So, the probability that a randomly selected person in the study did between 1 and 2 hours of activity can be determined:

$$\begin{aligned}\mathrm{P}(1 < x < 2) &= \frac{f(1 < x < 2)}{\sum f}\\ &= \frac{347}{364 + 347 + 156 + 54 + 32 + 10 + 7}\\ &= \frac{347}{970}\\ &= 0.358\end{aligned}$$

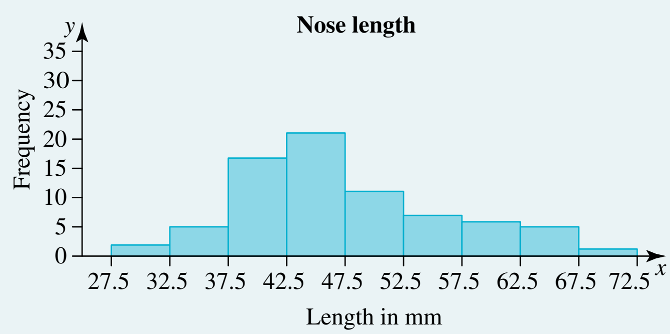
Similarly, we could find the probability that a person did more than 4 hours of activity per week by adding the interval frequencies for all intervals higher than 4:

$$\begin{aligned}\mathrm{P}(x > 4) &= \frac{f(4 < x < 5) + f(5 < x < 6) + f(6 < x < 7)}{\sum f}\\ &= \frac{32 + 10 + 7}{970}\\ &= \frac{49}{970}\\ &= 0.051\end{aligned}$$

WORKED EXAMPLE 1

In a study, the nose lengths, X millimetres, of 75 adults were measured. The results of the study are shown in the table and histogram.

Nose length (mm)	Frequency
$27.5 < X \le 32.5$	2
$32.5 < X \le 37.5$	5
$37.5 < X \le 42.5$	17
$42.5 < X \le 47.5$	21
$47.5 < X \le 52.5$	11
$52.5 < X \le 57.5$	7
$57.5 < X \le 62.5$	6
$62.5 < X \le 67.5$	5
$67.5 < X \le 72.5$	1



Determine the probability of someone in the study having a nose length that is:

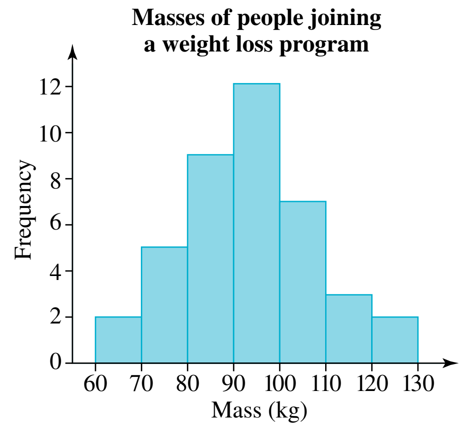
a. **between 42.5 mm and 47.5 mm**

b. **less than or equal to 47.5 mm.**

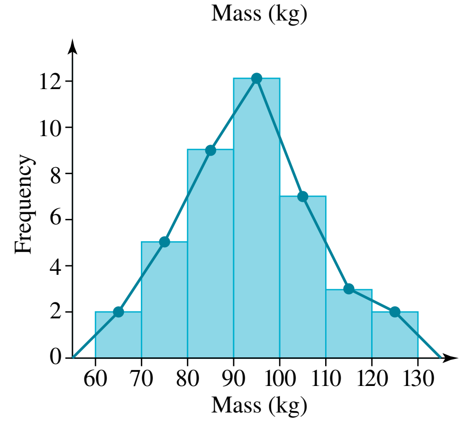
THINK	WRITE
a. 1. Use the table to determine the frequency of the interval and the sum of the frequencies for the study.	**a.** $f(42.5 < x < 47.5) = 21$ $\sum f = 2 + 5 + 17 + 21 + 11 + 7 + 6 + 5 + 1$ $= 75$
2. Substitute values into the equation and evaluate.	$\text{P}(42.5 < x < 47.5) = \dfrac{f(42.5 < x < 47.5)}{\sum f}$ $= \dfrac{21}{75}$ $= 0.28$
3. Answer the question.	The probability of a person having a nose length between 42.5 and 47.5 mm is 0.28.
b. 1. Determine the total of the frequencies for all of the intervals such that $x \le 47.5$.	**b.** $f(x \le 47.5) = 2 + 5 + 17 + 21 = 45$
2. Substitute values into the equation and evaluate.	$\text{P}(x \le 47.5) = \dfrac{f(x \le 47.5)}{\sum f}$ $= \dfrac{45}{75}$ $= 0.6$
3. Answer the question.	The probability of a person having a nose length less than or equal to 47.5 mm is 0.6.

11.2.3 Modelling continuous random variables

The histogram shown displays the masses (in kg) of members joining the Toowong branch of a popular weight loss program.

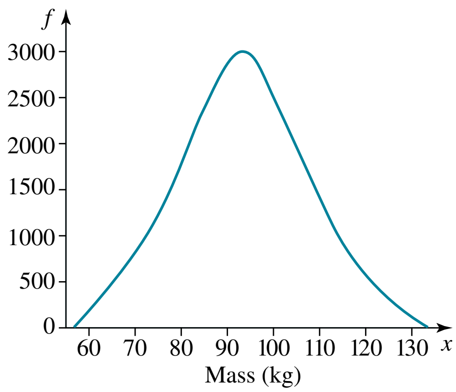


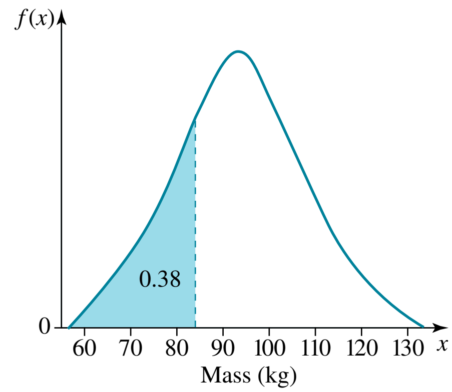
A **frequency polygon** is drawn by joining the midpoints at the top of each bar representing an interval.



This is a relatively small set of data, as the study only involved 40 people. However, if we increased the size of the study so that the members of more branches of the weight loss program were included and the class intervals were made smaller, the frequency polygon would take on the appearance of a smooth curve.

As the shape of a curve may be modelled by a mathematical formula, it should therefore be possible to develop a function, $f(x)$, that will allow us to approximate relative frequency values — and therefore probability values — for intervals regardless of the original interval boundaries that generated the function curve.





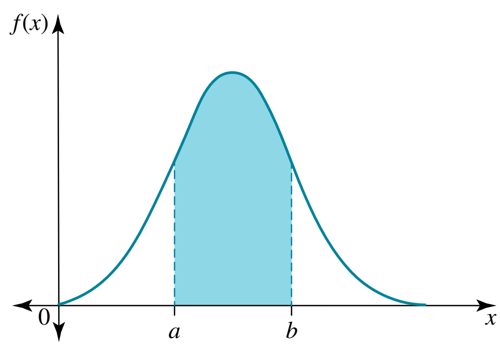
For our weight loss program curve, 38% of the area under the curve is bounded by the x-axis and the line $x = 84$. This represents the relative frequency of members having a mass of 84 kg or less. This also means that the probability of a randomly chosen person in the program having a mass of 84 kg or less is approximately 0.38; that is, $\text{P}(x < 84) = 0.38$.

The absolute area bounded by the curve and the x-axis can have any value greater than 0. However, if the curve is scaled down in such a way that the total area under the curve is equal to 1, it is referred to as the **probability density curve** (or **pdf**) for the continuous random variable.

11.2.4 The probability density curve

As you will recall from Chapter 7, the area under a curve $f(x)$ for any interval may be determined by evaluating the integral of the function over that interval. Thus, the probability that a continuous random variable X lies between the values a and b is given by:

$$\text{P}(a < X < b) = \int_a^b f(x)\,dx$$



Using this relationship, we can clarify why probabilities are assessed for continuous random variables lying within intervals rather than at exact values.

Consider a continuous variable X such that $X = x$ where $x \in R$.

$$\begin{aligned}
\text{P}(X = x) &= \int_x^x f(x)\,dx \\
&= [F(x)]_x^x \\
&= F(x) - F(x) \\
\text{P}(X = x) &= 0
\end{aligned}$$

Thus, $\mathrm{P}(X = x) = 0$ when $x \in [a, b]$.

There are two critical conditions that constrain a probability density function.

First, as probability cannot be a negative number, the probability density function must be greater than or equal to zero over the interval being considered; that is,

$$f(x) \geq 0 \text{ for all } x \in [a, b].$$

Secondly, the sum of probabilities for all possible values of x must be equal to 1:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$

(This is analogous to the condition for the discrete probability distribution that $\sum_{\text{all } x} \mathrm{P}(X = x) = 1$.)

In the event that an interval is bounded by a and b such that $\int_a^b f(x)\, dx = 1$, then the function must be 0 everywhere else outside this interval.

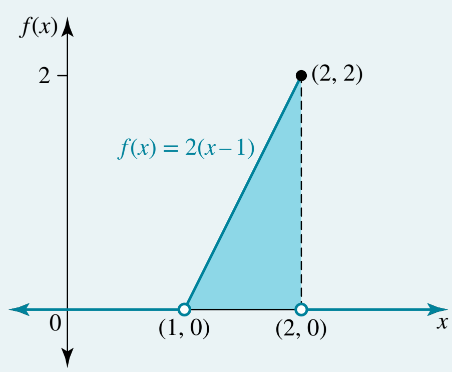
WORKED EXAMPLE 2

Sketch the graph of each of the following functions and state whether each function is a probability density function.

a. $f(x) = \begin{cases} 2(x-1), & 1 \leq x \leq 2 \\ 0, & \text{elsewhere} \end{cases}$

b. $f(x) = \begin{cases} 0.5, & 2 \leq x \leq 4 \\ 0, & \text{elsewhere} \end{cases}$

c. $f(x) = \begin{cases} 2e^{-x} & 0 \leq x \leq 2 \\ 0, & \text{elsewhere} \end{cases}$

THINK	WRITE
a. 1. Sketch the graph of $f(x) = 2(x-1)$ over the domain $1 \leq x \leq 2$, giving an x-intercept of 1 and an end point of $(2, 2)$. Make sure to include the horizontal lines for $y = 0$ either side of this graph. *Note:* This function is known as a triangular probability function because of its shape.	a. 
2. Inspect the graph to determine if the function is always positive or zero, that is, $f(x) \geq 0$ for all $x \in [a, b]$.	Yes, $f(x) \geq 0$ for all x-values.
3. Calculate the area of the shaded region to determine if $\int_1^2 2(x-1)\, dx = 1$.	Method 1: Using the area of triangles Area of shaded region $= \frac{1}{2} \times \text{base} \times \text{height}$ $= \frac{1}{2} \times 1 \times 2$ $= 1$

Method 2: Using calculus

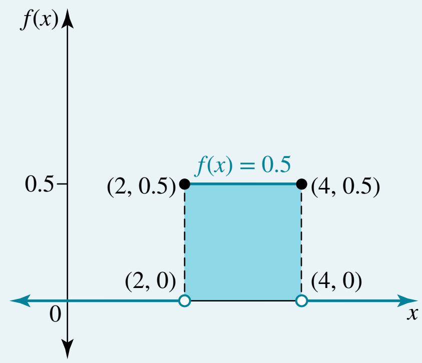
$$\text{Area of shaded region} = \int_2^1 2(x-1)\,dx$$
$$= \int_2^1 (2x-2)dx$$
$$= [x^2 - 2x]_1^2$$
$$= (2^2 - 2(2)) - (1^2 - 2(1))$$
$$= 0 - 1 + 2$$
$$= 1$$

4. Interpret the results.

$f(x) \geq 0$ for all values, and the area under the curve equals 1. Therefore, this is a probability density function.

b. 1. Sketch the graph of $f(x) = 0.5$ for $2 \leq x \leq 4$. This gives a horizontal line with end points of $(2, 0.5)$ and $(4, 0.5)$. Make sure to include the horizontal lines for $y = 0$ on either side of this graph.
Note: This function is known as a uniform or rectangular probability density function because of its rectangular shape.

b.



2. Inspect the graph to determine if the function is always positive or zero, that is, $f(x) \geq 0$ for all $x \in [a, b]$.

Yes, $f(x) \geq 0$ for all x-values.

3. Calculate the area of the shaded region to determine if $\int_2^4 0.5\,dx = 1$.

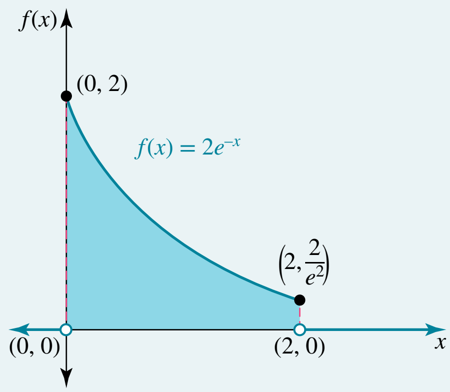
Again, it is not necessary to use calculus to determine the area.
Method 1:
$$\text{Area of shaded region} = \text{length} \times \text{width}$$
$$= 2 \times 0.5$$
$$= 1$$
Method 2:
$$\text{Area of shaded region} = \int_4^2 0.5\,dx$$
$$= [0.5x]_2^4$$
$$= 0.5(4) - 0.5(2)$$
$$= 2 - 1$$
$$= 1$$

4. Interpret the results.

$f(x) \geq 0$ for all values, and the area under the curve equals 1. Therefore, this is a probability density function.

c. 1. Sketch the graph of $f(x) = 2e^{-x}$ for $0 \le x \le 2$. The end points will be $(0, 2)$ and $\left(2, e^{-2}\right)$. Make sure to include the horizontal lines for $y = 0$ on either side of this graph.

c.



2. Inspect the graph to determine if the function is always positive of zero, that is, $f(x) \ge 0$ for all $x \in [a, b]$.

Yes, $f(x) \ge 0$ for all x-values.

3. Calculate the area of the shaded region to determine if $\int_0^2 2e^{-x}dx = 1$.

$$\begin{aligned}\int_0^2 2e^{-x}dx &= 2\int_0^2 e^{-x}dx \\ &= 2\left[-e^{-x}\right]_0^2 \\ &= 2\left(-e^{-2} + e^0\right) \\ &= 2\left(-e^{-2} + 1\right) \\ &= 1.7293\end{aligned}$$

4. Interpret the results.

$f(x) \ge 0$ for all values. However, the area under the curve does not equal 1. Therefore, this is not a probability density function.

WORKED EXAMPLE 3

Given that the following functions are probability density functions, determine the value of a in each function.

a. $f(x) = \begin{cases} a(x-1)^2, & 0 \le x \le 4 \\ 0, & \text{elsewhere} \end{cases}$

b. $f(x) = \begin{cases} ae^{-4x}, & x > 0 \\ 0, & \text{elsewhere} \end{cases}$

THINK

WRITE

a. 1. As the function has already been defined as a probability density function, this means that the area under the graph is definitely 1.

a.
$$\int_0^4 f(x)dx = 1$$
$$\int_0^4 a(x-1)^2dx = 1$$

2. Remove a from the integral, as it is a constant.

$$a\int_0^4 (x-1)^2dx = 1$$

3. Antidifferentiate and substitute in the terminals.

$$a\int_0^4 (x-1)^2dx = 1$$

$$a\left[\frac{(x-1)^3}{3}\right]_0^4 = 1$$

$$a\left[\frac{3^3}{3} - \frac{(-1)^3}{3}\right] = 1$$

4. Solve for a.

$$a\left(9+\frac{1}{3}\right) = 1$$

$$a \times \frac{28}{3} = 1$$

$$a = \frac{3}{28}$$

b. 1. As the function has already been defined as a probability density function, this means that the area under the graph is definitely 1.

b. $$\int_0^{\infty} f(x)dx = 1$$

$$\int_0^{\infty} ae^{-4x}dx = 1$$

2. Remove a from the integral, as it is a constant.

$$a\int_0^{\infty} e^{-4x}dx = 1$$

3. To evaluate an integral containing infinity as one of the terminals, we determine the appropriate limit.

$$a \times \lim_{k\to\infty} \int_0^k e^{-4x}dx = 1$$

4. Antidifferentiate and substitute in the terminals.

$$a \times \lim_{k\to\infty} \int_k^0 e^{-4x}dx = 1$$

$$a \times \lim_{k\to\infty} \left[-\frac{1}{4}e^{-4x}\right]_0^k = 1$$

$$a \times \lim_{k\to\infty} \left(-\frac{e^{-4k}}{4} + \frac{1}{4}\right) = 1$$

$$a \times \lim_{k\to\infty} \left(-\frac{e^{-4k}}{4} + \frac{1}{4}\right) = 1$$

$$a \times \lim_{k\to\infty} \left(-\frac{1}{4e^{4k}} + \frac{1}{4}\right) = 1$$

5. Solve for a. Remember that a number divided by an extremely large number is effectively 0, so $\lim\limits_{k\to\infty}\left(\frac{1}{e^{4k}}\right) = 0$.

$$a\left(0+\frac{1}{4}\right) = 1$$

$$\frac{a}{4} = 1$$

$$a = 4$$

on Resources

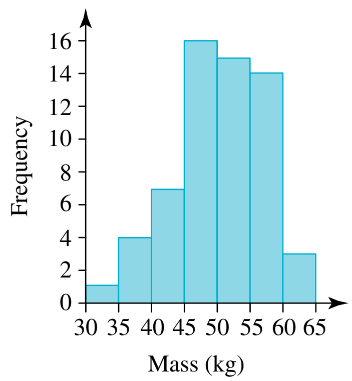
Interactivity: Continuous random variables (int-6433)

Exercise 11.2 Continuous random variables and the probability density function

Technology free

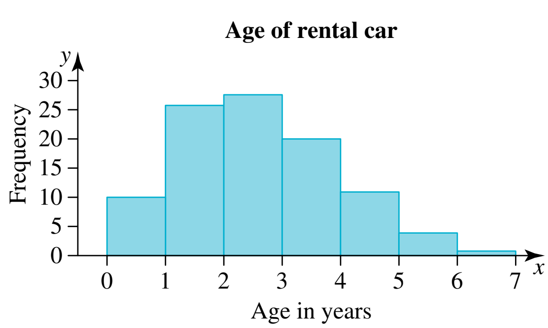
1. Which of the following are continuous random variables?
   a. The population of your town or city
   b. The types of motorbike in a parking lot
   c. The heights of people in an identification line-up
   d. The masses of babies in a group
   e. The languages spoken at home by students in your class
   f. The time spent watching TV
   g. The number of children in the families in your suburb
   h. The air pressure in your car's tyres
   i. The number of puppies in a litter
   j. The types of radio program listened to by teenagers
   k. The times for swimming 50 metres
   l. The quantity of fish caught in a net
   m. The number of CDs you own
   n. The types of shops in a shopping centre
   o. The football competition ladder at the end of each round
   p. The lifetimes of torch batteries
   q. The number of people attending a rock concert
   r. Exam grades
   s. The types of magazine sold at a news agency
   t. Hotel accommodation ratings

2. **WE1** The frequency histogram shows the distribution of masses (in kilograms) of 60 students in Year 7 at Northwood State High School.
   Determine the probability that a random student has a mass:
   a. between 40 and 60 kilograms
   b. less than 45 kilograms
   c. greater than 55 kilograms.



3. A small car-hire firm keeps note of the age and kilometres covered by each of the cars in their fleet. Generally, cars are no longer used once they have either covered 350 000 kilometres or are more than 5 years old. The following information describes the ages of the cars in their current fleet.

Age	Frequency
$0 < x \le 1$	10
$1 < x \le 2$	26
$2 < x \le 3$	28
$3 < x \le 4$	20
$4 < x \le 5$	11
$5 < x \le 6$	4
$6 < x \le 7$	1



a. Determine:

i. $P(X \le 2)$

ii. $P(X > 4)$

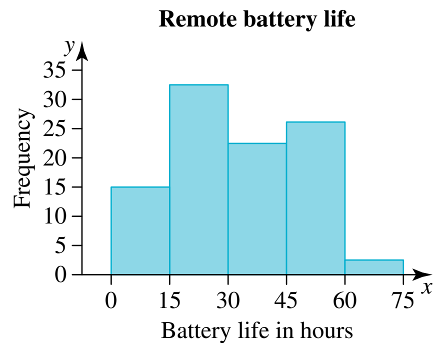
b. Determine:

i. $P(1 < X \le 4)$

ii. $P(X > 1 | X \le 4)$

4. The battery life for batteries in television remote controls was investigated in a study. The results are shown in the table and histogram.

Hours of life	Frequency
$0 < x \le 15$	15
$15 < x \le 30$	33
$30 < x \le 45$	23
$45 < x \le 60$	26
$60 < x \le 75$	3



a. How many remote control batteries were included in the study?

b. What is the probability that a battery will last more than 45 hours?

c. What is the probability that a battery will last between 15 and 60 hours?

d. A new battery producer is advocating that their batteries have a long life of 60 + hours. If it is known that this is just advertising hype because these batteries are no different from the batteries in the study, what is the probability that these new batteries will have a life of 60 + hours?

5. **WE2** Sketch each of the following functions and determine whether each one is a probability density function.

a. $f(x) = \begin{cases} \frac{1}{4}e^{2x}, & 0 \le x \le \log_e(3) \\ 0, & \text{elsewhere} \end{cases}$

b. $f(x) = \begin{cases} 0.25, & -2 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}$

6. Sketch each of the following functions and determine whether each one is a probability density function.

a. $f(x) = \begin{cases} \frac{1}{2}\cos(x), & -\frac{\pi}{2} \le x \le \frac{\pi}{2} \\ 0, & \text{elsewhere} \end{cases}$

b. $f(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & \frac{1}{2} \le x \le 4 \\ 0, & \text{elsewhere} \end{cases}$

7. **WE3** Given that the following function is a probability density function, determine the value of n.

$$f(x) = \begin{cases} n(x^3 - 1), & 1 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

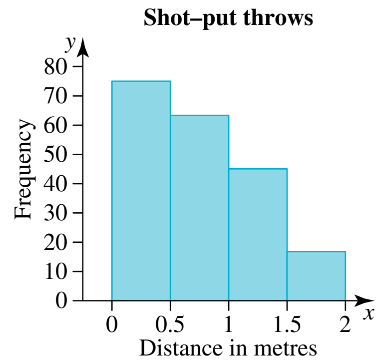
8. Given that the following function is a probability density function, determine the value of a.

$$f(x) = \begin{cases} -ax, & -2 \le x \le 0 \\ 2ax, & 0 < x \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

9. A number of experienced shot-putters were asked to aim for a line 10 metres away.

After each of them put their shot, its distance from the 10-metre line was measured. All of the shots were on or between the 8- and 10-metre lines. The results of the measurements are shown, where X is the distance in metres from the 10-metre line.

Metres	Frequency
$0 < x \le 0.5$	75
$0.5 < x \le 1$	63
$1 < x \le 1.5$	45
$1.5 < x \le 2$	17



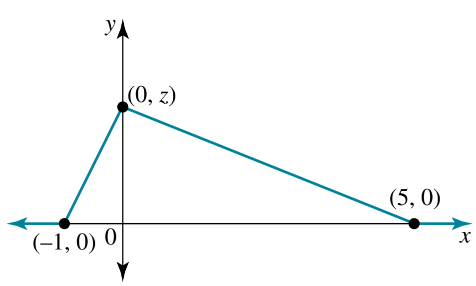
a. How many shot-put throws were measured?
b. Calculate:
   i. $P(X > 0.5)$
   ii. $P(1 < X \le 2)$
c. A guest shot-putter is visiting the athletics club where the measurements are being conducted. His shot-putting ability is equivalent to the abilities of the club members. Determine the probability that he puts the shot between 50 cm and 1 m of the 10-metre line if it is known that he put the shot within 1 metre of the 10-metre line.

10. The rectangular function, f, is defined by the rule

$$f(x) = \begin{cases} c, & 0.25 < x < 1.65 \\ 0, & \text{elsewhere} \end{cases}.$$

Determine the value of the constant c, given that f is a probability density function

11. The graph of a function, f, is shown. If f is known to be a probability density function, show that the value of z is $\frac{1}{3}$.



Technology active

12. Determine the value of the constant m in each of the following if each function is a probability density function.

a. $f(x) = \begin{cases} m(6-2x), & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}$

b. $f(x) = \begin{cases} me^{-2x}, & x \ge 0 \\ 0, & \text{elsewhere} \end{cases}$

c. $f(x) = \begin{cases} me^{2x}, & 0 \le x \le \log_e(3) \\ 0, & \text{elsewhere} \end{cases}$

13. Let X be a continuous random variable with the probability density function

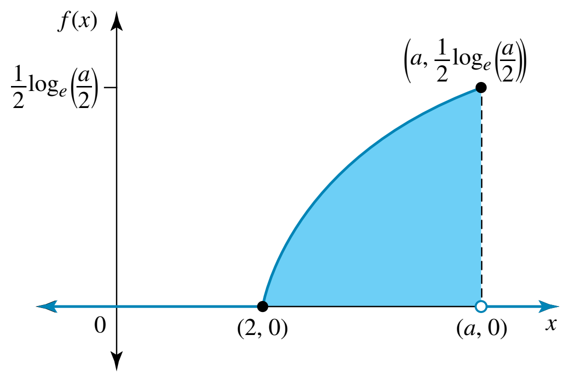
$$f(x) = \begin{cases} x^2 + 2kx + 1, & 0 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}.$$

Show that the value of k is $-\dfrac{11}{9}$.

14. X is a continuous random variable such that

$$f(x) = \begin{cases} \frac{1}{2}\log_e\left(\dfrac{x}{2}\right), & 2 \le x \le a \\ 0, & \text{elsewhere} \end{cases}$$

and $\int_2^a f(x)\ dx = 1$. The graph of this function is shown.



a. Differentiate $x\log_e\left(\dfrac{x}{2}\right)$ and hence determine an antiderivative of $\log_e\left(\dfrac{x}{2}\right)$.

b. Using the answer from part **a**. determine the value of the constant a.

11.3 Cumulative distribution functions

11.3.1 Hybrid probability density functions

Given that the probability value of a continuous random variable X lying in the interval between a and b is described by

$$P(a < x < b) = \int_a^b f(x)\ dx$$

and that the probability of X having an exact value is 0, that is,

$$P(X = a) = P(X = b) = 0,$$

it follows that

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$$

and

$$P(a \le X \le b) = P(a \le X \le c) + P(c < X \le b), \text{ where } a \le c \le b.$$

These properties are particularly useful when the probability density function is a hybrid function and the required probability encompasses two functions.

WORKED EXAMPLE 4

A continuous random variable, Y, has a probability density function, f, defined by

$$f(y) = \begin{cases} -ay, & -3 \le y \le 0 \\ ay, & 0 < y \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

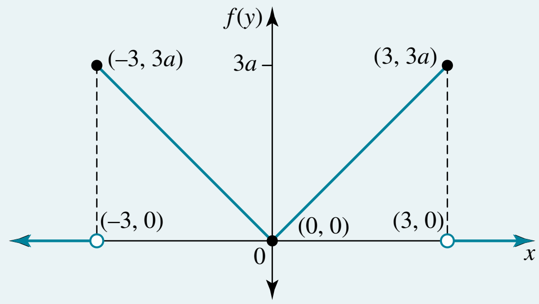
where a is a positive constant.

a. **Sketch the graph of f.**

b. **Determine the value of the constant, a.**

c. **Determine $P(1 \le Y \le 3)$.**

d. **Determine $P(Y < 2 | Y > -1)$**

THINK	WRITE
a. The hybrid function contains three sections. The first graph, $f(y) = -ay$, is a straight line with end points of $(0, 0)$ and $(-3, 3a)$. The second graph is also a straight line and has end points of $(0, 0)$ and $(3, 3a)$. Don't forget to include the $f(y) = 0$ lines for $x > 3$ and $x < -3$.	a. $f(-3) = 3a$ and $f(3) = 3a$ 
b. Use the fact that $\int_{-3}^{3} f(y)dy = 1$ to solve for a.	b. $\int_{-3}^{3} f(y)dy = 1$ Using the area of a triangle, we find: $\frac{1}{2} \times 3 \times 3a + \frac{1}{2} \times 3 \times 3a = 1$ $\frac{9a}{2} + \frac{9a}{2} = 1$ $9a = 1$ $a = \frac{1}{9}$
c. Identify the part of the function that the required y-values sit within: the values $1 \le Y \le 3$ are within the region where $f(y) = \frac{1}{9}y$.	c. $P(1 \le Y \le 3) = \int_{1}^{3} f(y)dy$ $= \int_{1}^{3} \left(\frac{1}{9}y\right) dy$ $= \left[\frac{1}{18}y^2\right]_1^3$ $= \frac{1}{18}(3)^2 - \frac{1}{18}(1)^2$ $= \frac{8}{18}$ $= \frac{4}{9}$ *Note:* The method of finding the area of a trapezium could also be used.

d. 1. State the rule for the conditional probability.

d. $P(Y < 2|Y > -1) = \dfrac{P(Y < 2, Y > -1)}{P(Y > -1)}$

$= \dfrac{P(-1 < Y < 2)}{P(Y > -1)}$

2. Determine $P(-1 < Y < 2)$. As the interval is across two functions, the interval needs to be split.

$P(-1 < Y < 2) = P(-1 < Y < 0) + P(0 \le Y < 2)$

3. To calculate the probabilities we need to determine the areas under the curve.

$= \int_{-1}^{0} -\frac{1}{9} ydy + \int_{0}^{2} \frac{1}{9} ydy$

$= -\int_{-1}^{0} \frac{1}{9} ydy + \int_{0}^{2} \frac{1}{9} ydy$

4. Antidifferentiate and evaluate after substituting the terminals.

$= -\left[\frac{1}{18}y^2\right]_{1}^{0} + \left[\frac{1}{18}y^2\right]_{0}^{2}$

$= -\left(\frac{1}{18}(0)^2 - \frac{1}{18}(-1)^2\right) + \frac{1}{18}(2)^2 - \frac{1}{18}(0)^2$

$= \frac{1}{18} + \frac{4}{18}$

$= \frac{5}{18}$

5. Determine $P(Y > -1)$. As the interval is across two functions, the interval needs to be split.

$P(Y > -1) = P(-1 < Y < 0) + P(0 \le Y \le 3)$

6. To determine the probabilities we need to determine the areas under the curve. As $P(0 \le Y \le 3)$ covers exactly half the area under the curve, $P(0 \le Y \le 3) = \frac{1}{2}$. (The entire area under the curve is always 1 for a probability density function.)

$= \int_{-1}^{0} -\frac{1}{9} y\,dy + \int_{0}^{3} \frac{1}{9} y\,dy$

$= -\int_{-1}^{0} \frac{1}{9} y\,dy + \frac{1}{2}$

7. Antidifferentiate and evaluate after substituting the terminals.

$= -\left[\frac{1}{18}y^2\right]_{-1}^{0} + \frac{1}{2}$

$= -\left(\frac{1}{18}(0)^2 - \frac{1}{18}(-1)^2\right) + \frac{1}{2}$

$= \frac{1}{18} + \frac{9}{18}$

$= \frac{10}{18}$

$= \frac{5}{9}$

8. Now substitute into the formula to determine $P(Y < 2|Y > -1) = \dfrac{P(-1 < Y < 2)}{P(Y > -1)}$.

$$\begin{aligned} P(Y < 2|Y > -1) &= \frac{P(-1 < Y < 2)}{P(Y > -1)} \\ &= \frac{5}{18} \div \frac{5}{9} \\ &= \frac{5}{18} \times \frac{9}{5} \\ &= \frac{1}{2} \end{aligned}$$

11.3.2 The cumulative distribution function

The **cumulative distribution function** $F(x)$ describes the probability that a continuous random variable X has a value less than or equal to x; that is,

$$F(x) = P(x \le c) = \int_{-\infty}^{c} f(x)\, dx.$$

The cumulative distribution function (or **cdf**) of a continuous random variable has three fundamental properties:

- $F(x)$ must be a non-decreasing function; that is, $F(a) \le F(b)$ when $a < b$.
- As x approaches $-\infty$, $F(x)$ approaches or equals 0; that is, $\lim_{x \to -\infty} F(x) = 0$.
- As x approaches ∞, $F(x)$ approaches or equals 1; that is, $\lim_{x \to \infty} F(x) = 1$.

The probability of a variable falling within a particular interval is easily determined by using the cumulative distribution function.

Given that $P(X \le b) = P(X \le a) + P(a < X \le b)$,

$P(a < X \le b) = P(X \le b) - P(X \le a)$.

Thus, $P(a < X \le b) = F(b) - F(a)$.

Consider the probability distribution function $f(x)$ such that

$$f(x) = \begin{cases} 0.5x & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}.$$

The probabilities for consecutive intervals across the distribution can be tabulated as shown.

$a \le x \le b$	$P(a \le x \le b)$
$0 \le x < 0.25$	0.015 625
$0.25 \le x < 0.5$	0.046 875
$0.5 \le x < 0.75$	0.078 125
$0.75 \le x < 1$	0.109 375
$1 \le x < 1.25$	0.140 625
$1.25 \le x < 1.5$	0.171 875
$1.5 \le x < 1.75$	0.203 125
$1.75 \le x \le 2$	0.234 375

As $\mathrm{P}(a \leq X \leq b) = \mathrm{P}(a \leq X \leq c) + \mathrm{P}(c < X \leq b)$, $a < c < b$,

$$\begin{aligned} \mathrm{P}(X \leq 0.5) &= \mathrm{P}(0 \leq X \leq 0.25) + \mathrm{P}(0.25 < X \leq 0.5) \\ \mathrm{P}(X \leq 0.75) &= \mathrm{P}(X \leq 0.5) + \mathrm{P}(0.5 < X \leq 0.75) \\ \mathrm{P}(X \leq 1) &= \mathrm{P}(X \leq 0.75) + \mathrm{P}(0.75 < X \leq 1) \end{aligned}$$

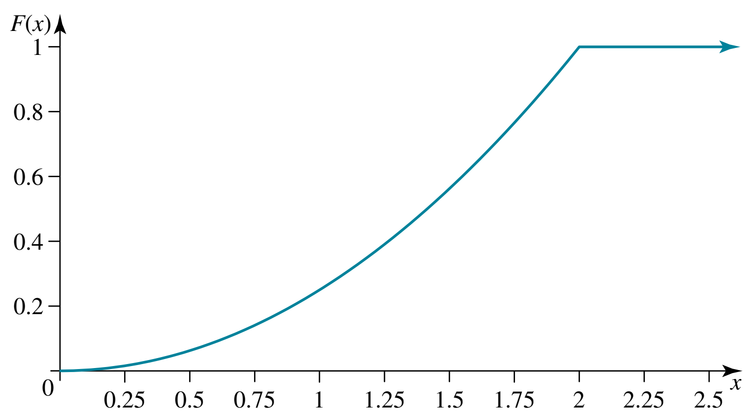
and so on.

We can tabulate the cumulative probability $\mathrm{P}(x \leq c)$ across the interval $0 \leq c \leq 2$:

$x \leq c$	$\mathrm{P}(x \leq c)$
$x \leq 0.25$	0.015 625
$x \leq 0.5$	0.062 5
$x \leq 0.75$	0.140 625
$x \leq 1$	0.25
$x \leq 1.25$	0.390 625
$x \leq 1.5$	0.562 5
$x \leq 1.75$	0.765 625
$x \leq 2$	1

Note that as $\int_{-\infty}^{c} f(x)\, dx = \int_{0}^{c} f(x)\, dx$ for this pdf, $\mathrm{P}(x \leq c) = \mathrm{P}(0 \leq x \leq c) = \mathrm{P}(-\infty \leq x \leq c)$.

When $F(x)$ is graphed against x, the curve formed is continuous across the range of the probability function.



As $F(x) = \int_{0}^{x} \left(\frac{x}{2}\right) dx$ for $0 \leq x \leq 2$,

$$F(x) = \left[\frac{x^2}{4}\right]_0^x.$$

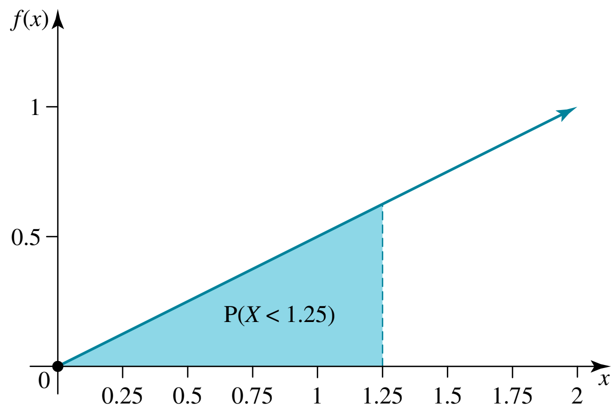
Thus, $F(x) = \frac{x^2}{4}$ for $0 \leq x \leq 2$.

The cumulative distribution function for the continuous random variable X will, in this case, be such that:

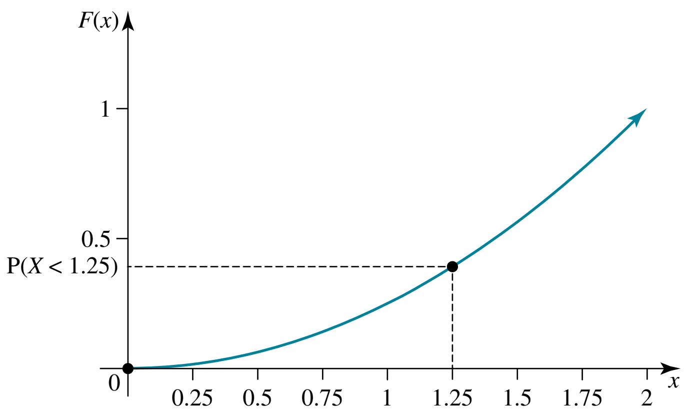
$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x^2}{4}, & 0 < x \le 2. \\ 1, & x > 2 \end{cases}$$

It is important to remember that the area under the probability density curve for an interval will be equivalent to a point on the cumulative distribution curve.

Probability density function



Cumulative distribution function



WORKED EXAMPLE 5

A continuous random variable Y has a probability density function such that

$$f(y) = \begin{cases} y, & 0 \le y \le 1 \\ 2 - y, & 1 < y \le 2. \\ 0, & \text{elsewhere} \end{cases}$$

a. **Determine the cumulative distribution function, $F(y)$, of Y.**

b. **Determine:**

i. **$P(Y \le 0.8)$**

ii. **$P(0.6 \le Y < 1.2)$**

THINK	WRITE
a. 1. Integrate $f(y)$ to determine $F(y)$ over the domains $0 \le y \le 1$ and $1 < y \le 2$.	**a.** For $0 \le y \le 1$, $F(y) = \int_0^y y\,dy = \frac{y^2}{2}$. For $1 < y \le 2$, $F(y) = P(Y \le y) = P(Y \le 1) + P(1 < Y \le y)$ $F(y) = \int_0^1 y\,dy + \int_1^y (2 - y)\,dy$ $= \frac{1}{2} + \left[2y - \frac{y^2}{2}\right]_1^y$ $= \frac{1}{2} + \left[\left(2y - \frac{y^2}{2}\right) - \left(2 \times 1 - \frac{1^2}{2}\right)\right]_1^y$ $= 2y - \frac{y^2}{2} - 1$
2. Write $F(y)$ as a hybrid function for all $y \in R$.	$F(y) = \begin{cases} 0, & y \le 0 \\ \frac{y^2}{2}, & 0 < y \le 1 \\ 2y - \frac{y^2}{2} - 1, & 1 < y \le 2 \\ 1, & y > 2 \end{cases}$
b. i. 1. Identify the appropriate equation and substitute the appropriate values.	**b. i.** $\mathrm{P}(Y \le 0.8) = F(Y \le 0.8)$ $= \frac{(0.8)^2}{2}$
2. Evaluate.	$= 0.32$
3. Answer the question.	The probability of Y being less than 0.8 is 0.32
ii. 1. Identify the appropriate equation.	**ii.** $\mathrm{P}(a \le Y < b) = F(b) - F(a)$
2. Evaluate.	$\mathrm{P}(0.6 \le Y < 1.2) = F(1.2) - F(0.6)$ $= \left[2(1.2) - \frac{(1.2)^2}{2} - 1\right] - \left[\frac{(0.6)^2}{2}\right]$ $= 0.68 - 0.18$ $= 0.5$
3. Answer the question.	The probability of Y being between 0.6 and 1.2 is 0.5.

on Resources

Interactivity: Cumulative density functions (int-6434)

studyon

Units 3 & 4 > Area 7 > Sequence 1 > Concept 1

Probability density functions Summary screen and practice questions

Exercise 11.3 Cumulative distribution functions

Technology free

1. WE4 The continuous random variable Z has a probability density function given by

$$f(z) = \begin{cases} -z+1, & 0 \le z < 1 \\ z-1, & 1 \le z \le 2 \\ 0, & \text{elsewhere} \end{cases}.$$

   a. Sketch the graph of f.
   b. Determine $P(Z < 0.75)$.
   c. Determine $P(Z > 0.5)$.

2. The continuous random variable X has a probability density function given by

$$f(x) = \begin{cases} 4x^3, & 0 \le x \le a \\ 0, & \text{elsewhere} \end{cases}$$

   where a is a constant.
   a. Determine the value of the constant a.
   b. Sketch the graph of f.
   c. Determine $P(0.5 \le X \le 1)$.

3. WE5 The continuous random variable X has a uniform rectangular probability density function defined by

$$f(x) = \begin{cases} \frac{1}{5}, & 1 \le x \le 6 \\ 0, & \text{elsewhere} \end{cases}.$$

   a. Determine the cumulative distribution function, $F(x)$, for X.
   b. Determine:
      i. $P(x \le 4)$
      ii. $P(2.2 < x \le 4.5)$

4. Let X be a continuous random variable with a probability density function defined by

$$f(x) = \begin{cases} \frac{1}{2}\sin(x), & 0 \le x \le \pi \\ 0, & \text{elsewhere} \end{cases}.$$

   a. Determine the cumulative distribution function $F(x)$ for X.
   b. Determine $P\left(X \le \frac{\pi}{2}\right)$.
   c. Determine $P\left(\frac{\pi}{4} < X < \frac{3\pi}{4}\right)$.
   d. Determine $P\left(X > \frac{\pi}{4} | X < \frac{3\pi}{4}\right)$.

5. A probability density function is defined by the rule

$$f(x) = \begin{cases} k(2+x), & -2 \le x < 0 \\ k(2-x), & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

where X is a continuous random variable and k is a constant.

a. Sketch the graph of f.

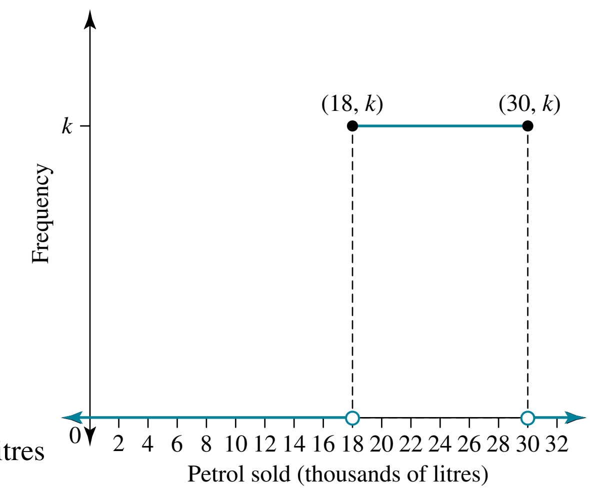
b. Determine the value of k.

c. Determine the cumulative distribution function $F(x)$.

d. Determine $P(-1 \le X \le 1)$.

e. Determine $P\left(X \ge -1 | X \le 1\right)$.

6. The amount of petrol sold daily by a busy service station is a uniformly distributed probability density function. A minimum of 18 000 litres and a maximum of 30 000 litres are sold on any given day. The graph of the function is shown.



a. Determine the value of the constant k.

b. Determine the formula (including cases) for the probability density function $f(x)$.

c. Determine the formula (including cases) for the cumulative distribution function $F(x)$.

d. Determine the probability that between 20 000 and 25 000 litres of petrol are sold on a given day.

e. Determine the probability that as much as 26 000 litres of petrol were sold on a particular day, given that it was known that at least 22 000 litres were sold.

Technology active

7. The continuous random variable X has a probability density function defined by

$$f(x) = \begin{cases} \frac{3}{8}x^2, & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}.$$

a. Determine the cumulative distribution function for $f(x)$.

b. Determine $P(X > 1.2)$.

c. Determine the value of n such that $P(X \le n) = 0.75$.

8. The continuous random variable Z has a probability density function defined by

$$f(z) = \begin{cases} \frac{1}{2z}, & 1 \le z \le e^2 \\ 0, & \text{elsewhere} \end{cases}.$$

a. Sketch the graph of f and shade the area that represents $\int_1^{e^2} f(z)\,dz$.

b. Determine $\int_1^{e^2} f(z)\,dz$. Explain your result.

The continuous random variable U has a probability function defined by

$$f(u)=\begin{cases} e^{4u}, & u \geq 0 \\ 0, & \text{elsewhere} \end{cases}.$$

c. Sketch the graph of f and shade the area that represents $\int_0^a f(u)\,du$, where a is a constant.

d. Determine the exact value of the constant a if $\int_1^{e^2} f(z)\,dz$ is equal to $\int_0^a f(u)\,du$.

9. The continuous random variable Z has a probability density function defined by

$$f(z)=\begin{cases} \frac{1}{2}\cos(z), & -\frac{\pi}{2} \leq z \leq \frac{\pi}{2} \\ 0, & \text{elsewhere} \end{cases}.$$

Determine $\text{P}\left(-\frac{\pi}{6} \leq Z \leq \frac{\pi}{4}\right)$, correct to 3 decimal places.

10. The continuous random variable U has a probability density function defined by

$$f(u)=\begin{cases} 1-\frac{1}{4}(2u-3u^2), & 0 \leq u \leq a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant. Determine:

a. the value of the constant a
b. $\text{P}(U < 0.75)$
c. $\text{P}(0.1 < U < 0.5)$
d. $\text{P}(U = 0.8)$.

11. The continuous random variable Z has a probability density function defined by

$$f(z)=\begin{cases} e^{-\frac{z}{3}}, & 0 \leq z \leq a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant. Determine:

a. the value of the constant a such that $\int_0^a f(z)\,dz = 1$
b. $\text{P}(0 < Z < 0.7)$
c. $\text{P}(Z < 0.7|Z > 0.2)$, correct to 4 decimal places
d. the value of α, correct to 2 decimal places, such that $\text{P}(Z \leq \alpha) = 0.54$.

12. The continuous random variable X has a probability density function given as

$$f(x)=\begin{cases} 3e^{-3x}, & x \geq 0 \\ 0, & \text{elsewhere} \end{cases}.$$

a. Determine $\text{P}(0 \leq X \leq 1)$, correct to 4 decimal places
b. Determine $\text{P}(X > 2)$, correct to 4 decimal places.

13. The continuous random variable X has a probability density function defined by

$$f(x) = \begin{cases} \log_e(x^2), & x \geq 1 \\ 0, & \text{elsewhere} \end{cases}.$$

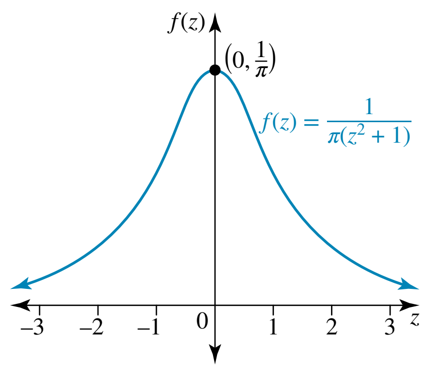
a. Differentiate $x\log_e\left(x^2\right)$ and hence determine an antiderivative of $\log_e(x^2)$

b. Determine the value of the constant a correct to 4 decimal places if $\int_1^a f(x)\ dx = 1$

c. Calculate $P(1.25 \leq X \leq 2)$ correct to 4 decimal places.

14. The graph of the probability function $f(z) = \dfrac{1}{\pi(z^2+1)}$ is shown.

Using technology, determine $P(-0.25 < Z < 0.25)$ correct to 4 decimal places.



11.4 Measures of centre and spread

The commonly used measures of central tendency and spread in statistics are the mean, median, variance, standard deviation and range. These same measures are appropriate for continuous probability functions.

11.4.1 Measures of central tendency: the mean

Remember that for a discrete random variable,

$$E(X) = \mu = \sum_{x=1}^{x=n} x_n P(X = x_n).$$

This definition can also be applied to a continuous random variable.

We define $E(X) = \mu = \displaystyle\int_{-\infty}^{\infty} xf(x)\,dx$.

The expected value of a continuous random variable

If $f(x) = 0$ everywhere except for $x \in [a, b]$, where the function is defined, then

$$E(X) = \mu = \int_a^b xf(x)\,dx.$$

Consider the continuous random variable, X, which has a probability density function defined by

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ 0, & \text{elsewhere} \end{cases}.$$

For this function,

$$\begin{aligned}
\mathrm{E}(X) = \mu &= \int_0^1 xf(x)\,dx \\
&= \int_0^1 x(x^2)\,dx \\
&= \int_0^1 x^3 dx \\
&= \left[\frac{x^4}{4}\right]_0^1 \\
&= \frac{1^4}{4} - 0 \\
&= \frac{1}{4}
\end{aligned}$$

Similarly, if the continuous random variable X has a probability density function of

$$f(x) = \begin{cases} 7e^{-7x}, & x \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

then

$$\begin{aligned}
\mathrm{E}(X) = \mu &= \int_0^\infty xf(x)dx \\
&= \lim_{k\to\infty} \int_0^k 7xe^{-7x}dx \\
&= 0.1429
\end{aligned}$$

where the method of *integration by recognition* is required to determine the integral.

The mean of a function of X is similarly found.

The mean of a continuous random variable

The function of X, $g(x)$, has a mean defined by

$$\mathrm{E}(g(x)) = \mu = \int_{-\infty}^{\infty} g(x)f(x)\,dx.$$

So if we again consider

$$f(x) = \begin{cases} x^2, & 0 \leq x \leq 1 \\ 0, & \text{elsewhere} \end{cases}$$

then

$$\begin{aligned}
E(X^2) &= \int_0^1 x^2 f(x)\,dx \\
&= \int_0^1 x^4 dx \\
&= \left[\frac{x^5}{5}\right]_0^1 \\
&= \frac{1^5}{5} - 0 \\
&= \frac{1}{5}
\end{aligned}$$

This definition will be important when we investigate the variance of a continuous random variable.

11.4.2 Median and percentiles

The median is also known as the 50th **percentile**, Q_2, the halfway mark or the middle value of the distribution.

The median of a continuous random variable

For a continuous random variable, X, defined by the probability function f, the median, m, can be found by solving $\int_{-\infty}^{m} f(x)\,dx = 0.5$.

Other percentiles that are frequently calculated are the 25th percentile or lower **quartile**, Q_1, and the 75th percentile or upper quartile, Q_3.

The interquartile range

The **interquartile range** is calculated as:

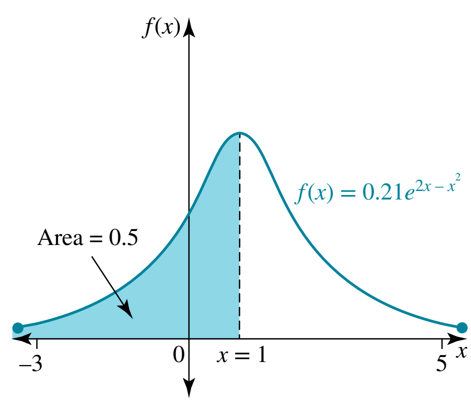
$$\text{IQR} = Q_3 - Q_1$$

Consider a continuous random variable, X, that has a probability density function of

$$f(x) = \begin{cases} 0.21e^{2x-x^2}, & -3 \le x \le 5 \\ 0, & \text{elsewhere} \end{cases}.$$

To determine the median, m, we solve for m as follows:

$$\int_{-3}^{m} 0.21e^{2x-x^2} dx = 0.5$$



The area under the curve is equated to 0.5, giving half of the total area and hence the 50th percentile. Solving using technology, the result is that $m = 0.9897 \simeq 1$. This can be seen on a graph as shown.

Consider the continuous random variable X, which has a probability density function of

$$f(x) = \begin{cases} \dfrac{x^3}{4}, & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}.$$

The median is given by $P(0 \le x \le m) = 0.5$:

$$\begin{aligned} \int_0^m \frac{x^3}{4} dx &= 0.5 \\ \left[\frac{x^4}{16}\right]_0^m &= \frac{1}{2} \\ \frac{m^4}{16} - 0 &= \frac{1}{2} \\ m^4 &= 8 \\ m &= \pm\sqrt[4]{8} \\ m &= 1.6818\,(0 \le m \le 2) \end{aligned}$$

To determine the lower quartile, we make the area under the curve equal to 0.25. Thus the lower quartile is given by $P(0 \le x \le a) = 0.25$:

$$\begin{aligned} \int_0^a \frac{x^3}{4} dx &= 0.25 \\ \left[\frac{x^4}{16}\right]_0^a &= \frac{1}{4} \\ \frac{a^4}{16} - 0 &= \frac{1}{4} \\ a^4 &= 4 \\ a &= \pm\sqrt[4]{4} \\ a = Q_1 &= 1.4142 \ (0 \le a \le m) \end{aligned}$$

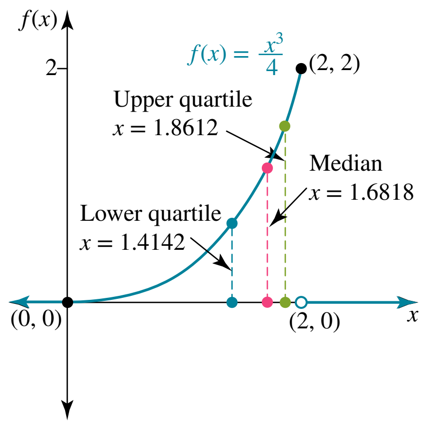
Similarly, to determine the upper quartile, we make the area under the curve equal to 0.75. Thus, the upper quartile is given by $P(0 \le x \le n) = 0.75$:

$$\begin{aligned} \int_0^n \frac{x^3}{4} dx &= 0.75 \\ \left[\frac{x^4}{16}\right]_0^n &= \frac{3}{4} \\ \frac{n^4}{16} - 0 &= \frac{3}{4} \\ n^4 &= 12 \\ n &= \pm\sqrt[4]{12} \\ n = Q_3 &= 1.8612 \ (m \le x \le 2) \end{aligned}$$

So, the interquartile range is given by

$$\begin{aligned} Q_3 - Q_1 &= 1.8612 - 1.4142 \\ &= 0.4470 \end{aligned}$$

These values are shown on the following graph.



WORKED EXAMPLE 6

A continuous random variable, Y, has a probability density function, f, defined by

$$f(y) = \begin{cases} ky, & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

where k is a constant.

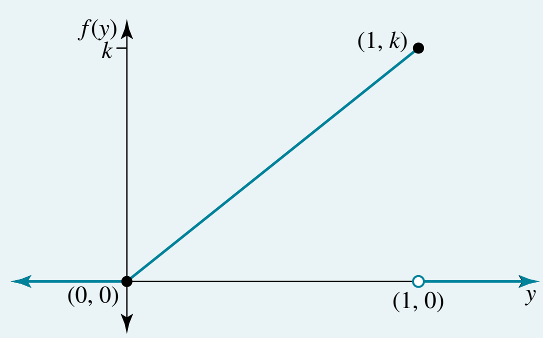
a. **Sketch the graph of f.**
b. **Determine the value of the constant k.**
c. **Determine:**
   i. **the mean of Y**
   ii. **the median of Y.**
d. **Determine the interquartile range of Y.**

THINK

a. The graph $f(y) = ky$ is a straight line with end points at $(0, 0)$ and $(1, k)$. Remember to include the lines $f(y) = 0$ for $y > 1$ and $y < 0$.

WRITE

a.



b. Solve $\int_0^1 ky\,dy = 1$ to determine the value of k.

b.
$$\int_0^1 ky\,dy = 1$$
$$k\int_0^1 y\,dy = 1$$
$$k\left[\frac{y^2}{2}\right]_0^1 = 1$$
$$\frac{k(1)^2}{2} - 0 = 1$$
$$\frac{k}{2} = 1$$
$$k = 2$$

Using the area of a triangle also enables you to find the value of k.

$$\frac{1}{2} \times 1 \times k = 1$$
$$\frac{k}{2} = 1$$
$$k = 2$$

c. i. 1. State the rule for the mean.

c. i. $\mu = \int_0^1 y(2y)\,dy$

$$= \int_0^1 2y^2\,dy$$

2. Antidifferentiate and simplify.

$$= \left[\frac{2}{3}y^3\right]_0^1$$
$$= \frac{2(1)^3}{3} - 0$$
$$= \frac{2}{3}$$

ii 1. State the rule for the median.

ii $\int_0^m f(y)\,dy = 0.5$

$$\int_0^m 2y\,dy = 0.5$$

2. Antidifferentiate and solve for m. Note that m must be a value within the domain of the function, so within $0 \le y \le 1$.

$$\left[y^2\right]_0^m = 0.5$$
$$m^2 - 0 = 0.5$$
$$m = \pm\sqrt{0.5}$$

3. Write the answer.

Median $= 0.7071$ (as $0 < m < 1$)

d. 1. State the rule for the lower quartile, Q_1.

d. $\int_0^a f(y)\,dy = 0.25$

$$\int_0^a 2y\,dy = 0.25$$

2. Antidifferentiate and solve for Q_1.

$$\left[y^2\right]_0^a = 0.25$$
$$a^2 - 0 = 0.25$$
$$a = \pm\sqrt{0.25}$$
$$a = Q_1 = 0.5 \quad (\text{as } 0 < Q_1 < 0.7071)$$

3. State the rule for the upper quartile, Q_3.

$$\int_0^n f(y)\,dy = 0.75$$
$$\int_0^n 2y\,dy = 0.75$$

4. Antidifferentiate and solve for Q_3.

$$\left[y^2\right]_0^n = 0.75$$
$$n^2 - 0 = 0.75$$
$$n = \pm\sqrt{0.75}$$
$$n = Q_3 = 0.8660 \quad (\text{as } 0.7071 < Q_3 < 1)$$

5. State the rule for the interquartile range.

$$\text{IQR} = Q_3 - Q_1$$

6. Substitute the appropriate values and simplify.

$$= 0.8660 - 0.5$$
$$= 0.3660$$

11.4.3 Measures of spread: variance, standard deviation and range

The variance and standard deviation are important measures of spread in statistics. From previous calculations for discrete probability functions, we know the following.

Variance and standard deviation for discrete probability functions

$\text{Var}(X) = \text{E}(X^2) - [\text{E}(X)]^2$ and $\text{SD}(X) = \sqrt{\text{Var}(X)}$.

For continuous probability functions,

$$\begin{aligned}
\text{Var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \\
&= \int_{-\infty}^{\infty} (x^2 - 2x\mu + \mu^2) f(x)\,dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \int_{-\infty}^{\infty} 2xf(x)\mu\,dx + \int_{-\infty}^{\infty} \mu^2 f(x)\,dx \\
&= \text{E}(X^2) - 2\mu \int_{-\infty}^{\infty} xf(x)\,dx + \mu^2 \int_{-\infty}^{\infty} 1f(x)\,dx \\
&= \text{E}(X^2) - 2\mu \times \text{E}(X) + \mu^2 \\
&= \text{E}(X^2) - 2\mu^2 + \mu^2 \\
&= \text{E}(X^2) - \mu^2 \\
&= \text{E}(X^2) - [\text{E}(X)]^2
\end{aligned}$$

Two important facts were used in this proof: $\int_{-\infty}^{\infty} f(x)\,dx = 1$ and $\int_{-\infty}^{\infty} xf(x)\,dx = \mu = \mathrm{E}(X)$.

Substituting this result into $\mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$ gives us

$$\mathrm{SD}(X) = \sqrt{\mathrm{E}(X^2) - [\mathrm{E}(X)]^2}.$$

The range is calculated as the highest value minus the lowest value. So, for the probability density function given by $f(x) = \begin{cases} \frac{1}{5}, & 1 \le x \le 6 \\ 0, & \text{elsewhere} \end{cases}$, the highest possible x-value is 6 and the lowest is 1. Therefore, the range for this function $= 6 - 1 = 5$.

WORKED EXAMPLE 7

For a continuous random variable, X, with a probability density function, f, defined by

$$f(x) = \begin{cases} \frac{1}{2}x + 2, & -4 \le x \le -2 \\ 0, & \text{elsewhere} \end{cases}$$

determine:

a. **the mean**

b. **the median**

c. **the variance**

d. **the standard deviation, correct to 4 decimal places.**

THINK

a. 1. State the rule for the mean and simplify.

2. Antidifferentiate and evaluate.

b. 1. State the rule for the median.

WRITE

a. $\mu = \int_{-4}^{-2} xf(x)\,dx$

$= \int_{-4}^{-2} x\left(\frac{1}{2}x + 2\right) dx$

$= \int_{-4}^{-2} \left(\frac{1}{2}x^2 + 2x\right) dx$

$= \left[\frac{1}{6}x^3 + x^2\right]_{-4}^{-2}$

$= \left(\frac{1}{6}(-2)^3 + (-2)^2\right) - \left(\frac{1}{6}(-4)^3 + (-4)^2\right)$

$= -\frac{4}{3} + 4 + \frac{32}{3} - 16$

$= -\frac{8}{3} = -2\frac{2}{3}$

b. $\int_{-4}^{m} f(x)\,dx = 0.5$

$\int_{-4}^{m} \left(\frac{1}{2}x + 2\right) dx = 0.5$

2. Antidifferentiate and solve for m. The quadratic formula is needed as the quadratic equation formed cannot be factorised. Alternatively, use technology to solve for m.

$$\begin{aligned}
\left[\frac{1}{4}x^2 + 2x\right]_{-4}^{m} &= 0.5\\
\left(\frac{1}{4}m^2 + 2m\right) - \left(\frac{(-4)^2}{4} + 2(-4)\right) &= 0.5\\
\frac{1}{4}m^2 + 2m + 4 &= 0.5\\
m^2 + 8m + 16 &= 2\\
m^2 + 8m + 14 &= 0
\end{aligned}$$

$$\begin{aligned}
\text{So,}\quad m &= \frac{-8 \pm \sqrt{(8)^2 - 4(1)(14)}}{2(1)}\\
m &= \frac{-8 \pm \sqrt{8}}{2}\\
&= -4 \pm \sqrt{2}\\
\therefore m &= -4 + \sqrt{2} \text{ as } m \in [-4, 2].
\end{aligned}$$

3. Write the answer.

The median is $-4 + \sqrt{2}$.

c. 1. Write the rule for variance.

c. $\text{Var}(X) = \text{E}(X^2) - [\text{E}(X)]^2$

2. Determine $\text{E}\left(X^2\right)$ first.

$$\begin{aligned}
\text{E}\left(X^2\right) &= \int_a^b x^2 f(x)\,dx\\
&= \int_{-4}^{-2} x^2\left(\frac{1}{2}x + 2\right)dx\\
&= \int_{-4}^{-2}\left(\frac{1}{2}x^3 + 2x^2\right)dx\\
&= [\frac{1}{18}x^4 + \frac{2}{3}x^3]_{-4}^{-2}\\
&= \left(\frac{1}{8}(-2)^4 + \frac{2}{3}(-2)^3\right) - \left(\frac{1}{8}(-4)^4 + \frac{2}{3}(-4)^3\right)\\
&= 2 - \frac{16}{3} - 32 + \frac{128}{3}\\
&= -30 + \frac{112}{3}\\
&= \frac{22}{3}
\end{aligned}$$

3. Substitute $\text{E}(X)$ and $\text{E}\left(X^2\right)$ into the rule for variance.

$$\begin{aligned}
\text{Var}(X) &= \text{E}(X^2) - [\text{E}(X)]^2\\
&= \frac{22}{3} - \left(-\frac{8}{3}\right)^2\\
&= \frac{22}{3} - \frac{64}{9}\\
&= \frac{66}{9} - \frac{64}{9}\\
&= \frac{2}{9}
\end{aligned}$$

d. 1. Write the rule for standard deviation.

2. Substitute the variance into the rule and evaluate.

d. $\text{SD}(X) = \sqrt{\text{Var}(X)}$

$= \sqrt{\frac{2}{9}}$

$= 0.4714$

Resources

Interactivities: Mean (int-6435)

Median and percentiles (int-6436)

Variance, standard deviation and range (int-6437)

studyon

Units 3 & 4 > Area 7 > Sequence 1 > Concepts 4 & 5

Mean and median Summary screen and practice questions

Variance and standard deviation Summary screen and practice questions

Exercise 11.4 Measures of centre and spread

Technology free

1. **WE6** The continuous random variable Z has a probability density function of

$$f(z) = \begin{cases} 4, & 1 \le z \le a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

a. Determine the value of the constant a.

b. Determine:

i. the mean of Z

ii. the median of Z.

2. The continuous random variable, Y, has a probability density function of

$$f(y) = \begin{cases} 2y, & 0 \le y \le a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

Determine:

a. the value of the constant a

b. $\text{E}(Y)$

c. the median value of Y.

3. **WE7** For the continuous random variable Z, the probability density function is

$$f(z) = \begin{cases} 2z-4, & 2 \le z \le 3 \\ 0, & \text{elsewhere} \end{cases}.$$

Determine the mean, median, variance and standard deviation.

4. The function

$$f(x) = \begin{cases} 3e^{-3x}, & x \ge 0 \\ 0 & \text{elsewhere} \end{cases}$$

Defines the probability density function for the continuous random variable X.

a. Determine the median
b. Differentiate xe^{-3x} and hence determine an antiderivative of $3xe^{-3x}$
c. Determine the mean using the answer from part **b**.
d. Differentiate x^2e^{-3x} and hence determine an antiderivative of $3x^2e^{-3x}$
e. Determine the variance using the answers from parts **c**. and **d**.
f. Determine the standard deviation.

5. Let X be a continuous random variable with a probability density function of

$$f(x) = \begin{cases} \dfrac{1}{2\sqrt{x}}, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}.$$

a. Prove that f is a probability density function.
b. Determine $\mathrm{E}(X)$.
c. Determine the median value of f.

6. The time in minutes that an individual must wait in line to be served at the local bank branch is defined by

$$f(t) = 2e^{-2t}, t \ge 0$$

where T is a continuous random variable.

a. Determine the median waiting time in the queue, correct to 2 decimal places
b. Differentiate xe^{-2x} and hence determine an antiderivative of $2xe^{-2x}$
c. What is the mean waiting time for a customer in the queue?
d. Differentiate x^2e^{-2x} and hence determine an antiderivative of $2x^2e^{-2x}$
e. Calculate the standard deviation for the waiting time in the queue, correct to 1 decimal place.

7. a. Determine the derivative of $\sqrt{4-x^2}$.
b. Hence, determine the mean value of the probability density function defined by

$$f(x) = \begin{cases} \dfrac{3}{\pi\sqrt{4-x^2}}, & 0 \le x \le \sqrt{3} \\ 0, & \text{elsewhere} \end{cases}.$$

8. Consider the continuous random variable X with a probability density function of

$$f(x) = \begin{cases} h(2-x), & 0 \le x \le 2 \\ h(x-2), & 2 < x \le 4 \\ 0, & \text{elsewhere} \end{cases}$$

where h is a positive constant.

a. Determine the value of the constant h.

b. Determine $\mathrm{E}(X)$.

c. Determine $\mathrm{Var}(X)$.

9. Consider the continuous random variable X with a probability density function of

$$f(x) = \begin{cases} k, & a \le x \le b \\ 0, & \text{elsewhere} \end{cases}$$

where a, b and k are positive constants.

a. Sketch the graph of the function f.

b. Show that $k = \dfrac{1}{b-a}$.

c. Determine $\mathrm{E}(X)$ in terms of a and b.

d. Determine $\mathrm{Var}(X)$ in terms of a and b.

Technology active

10. The continuous random variable Y has a probability density function defined by

$$f(y) = \begin{cases} \dfrac{y^2}{3}, & 0 \le y \le \sqrt[3]{9} \\ 0, & \text{elsewhere} \end{cases}.$$

Determine, correct to 4 decimal places:

a. the expected value of Y

b. the median value of Y

c. the lower and upper quartiles of Y

d. the interquartile range of Y.

11. The continuous random variable Z has a probability density function defined by

$$f(y) = \begin{cases} \dfrac{a}{z}, & 1 \le z \le 8 \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

a. Determine the value, correct to 4 decimal places, of the constant a.

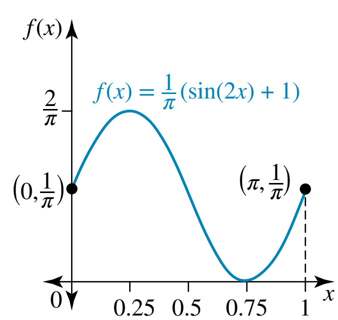
b. Determine $\mathrm{E}(Z)$ correct to 4 decimal places.

c. Determine $\mathrm{Var}(Z)$ and $\mathrm{SD}(Z)$.

d. Determine the interquartile range for Z.

e. Determine the range for Z.

12. X is a continuous random variable. The graph of the probability density function $f(x) = \dfrac{1}{\pi}(\sin(2x) + 1)$ for $0 \le x \le \pi$ is shown.



a. Show that $f(x)$ is a probability density function.

b. Calculate $\mathrm{E}(X)$ correct to 4 decimal places.

c. Calculate, correct to 4 decimal places:

i. $\mathrm{Var}(X)$ **ii.** $\mathrm{SD}(X)$

d. Determine the median value of f correct to 4 decimal places.

13. The continuous random variable X has a probability density function defined by

$$f(x) = \begin{cases} ax - bx^2, & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}.$$

Determine the values of the constants a and b if $E(X) = 1$.

14. The continuous random variable Z has a probability density function of

$$f(z) = \begin{cases} \dfrac{3}{z^2}, & 1 \le z \le a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

a. Show that the value of a is $\dfrac{3}{2}$.
b. Determine the mean value and variance of f correct to 4 decimal places.
c. Determine the median and interquartile range of f.

15. The continuous random variable Y has a probability density function

$$f(y) = \begin{cases} 0.2\log_e\left(\dfrac{y}{2}\right), & 2 \le y \le 7.9344 \\ 0, & \text{elsewhere} \end{cases}$$

a. Using technology, verify that f is a probability density function.
b. Using technology, determine $E(Y)$ correct to 4 decimal places.
c. Using technology, determine $\text{Var}(Y)$ and $\text{SD}(Y)$ correct to 4 decimal places.
d. Using technology, determine the median value of Y correct to 4 decimal places.
e. State the range.

16. The continuous random variable Z has a probability density function

$$f(z) = \begin{cases} \sqrt{z-1}, & 1 \le z \le a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

a. Determine the value of the constant a correct to 4 decimal places.
b. Determine, correct to 4 decimal places:
 i. $E(Z)$ ii. $E(Z^2)$ iii. $\text{Var}(Z)$ iv. $\text{SD}(Z)$

11.5 Review: exam practice

A summary of this chapter is available in the Resources section of your eBookPLUS at www.jacplus.com.au.

Simple familiar

1. Which of the following represent continuous random variables?
 a. The number of goals scored at a football match
 b. The heights of students in a Maths B class
 c. Shoe sizes
 d. The number of girls in a five-child family
 e. The time taken to run a distance of 10 kilometres in minutes

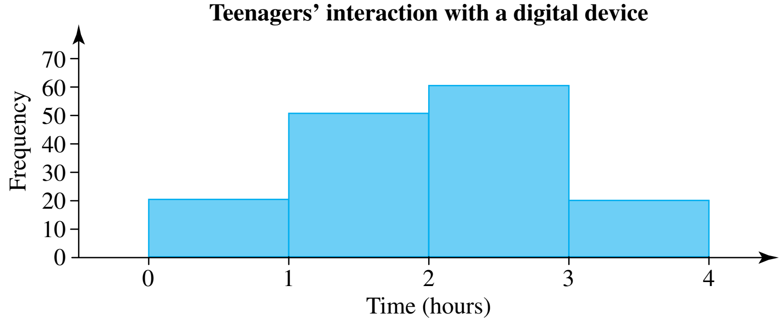
2. X is a continuous random variable with a probability function defined by

$$f(x) = \begin{cases} 2\sin(2x), & a \le x \le b \\ 0, & \text{elsewhere} \end{cases}.$$

Given that $a = 0$ and $0 \le b \le \pi$, what is the value of b?

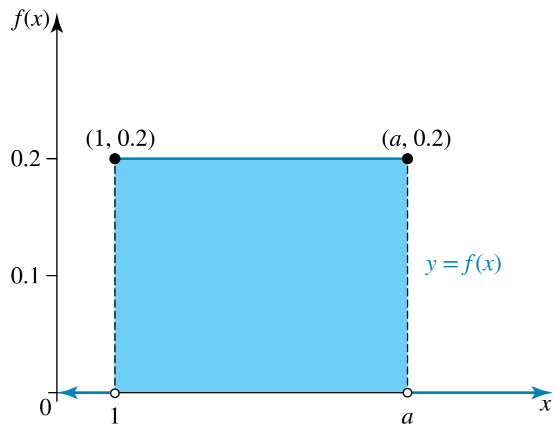
3. A survey was taken to determine the amount of time, X hours, that teenagers spend interacting with digital devices during a 24-hour period. The table of findings and histogram are shown.

Time in hours	Frequency
$0 \le x \le 1$	20
$1 < x \le 2$	50
$2 < x \le 3$	60
$3 < x \le 4$	20



a. How many teenagers were surveyed?
b. Determine $P(X \le 3)$.

4. The graph of a rectangular or uniform probability density function, $f(x)$, is shown. What is the value of the constant a?



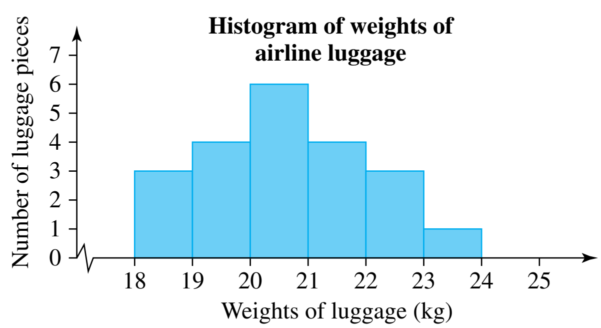
5. Y is a continuous random variable with a probability density function of

$$f(y) = \begin{cases} 3y^2, & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}.$$

a. Determine the cumulative distribution function for $f(y)$.
b. Determine $P(0.2 \le Y < 7)$.

6. A histogram was compiled as shown based upon the weights of luggage taken onto a flight from Rockhampton.

a. Explain why it is not possible to determine how many items of luggage had a weight of 19.5 kg.
b. What is the probability that a piece of luggage chosen at random from the flight weighed less than 19 kg?



7. A continuous probability density function is defined by

$$f(x) = \begin{cases} 2x, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}.$$

Calculate:

a. the mean of the distribution

b. the variance of the distribution.

8. a. Sketch the graph of

$$f(x) = \begin{cases} \frac{1}{4}(1-x), & -1 \le x < 1 \\ \frac{1}{4}(x-1), & 1 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}.$$

b. Show that $f(x)$ is a probability density function.

c. Determine $\mathrm{E}(X)$.

9. Explain the difference between the probability density function and the cumulative distribution function.

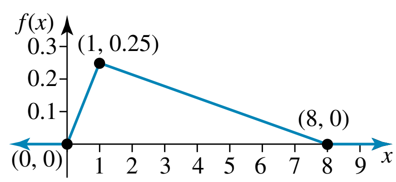
10. X is a continuous random variable with a probability density function defined by

$$f(x) = \begin{cases} \frac{1}{2}x & 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}.$$

a. Determine the cumulative distribution function for $f(x)$.

b. Sketch the cumulative distribution function, indicating on the graph where the value of $\mathrm{P}(X < 1.5)$ can be found.

11. Give a formula for the probability density function that is drawn here, given that the function is equal to 0 for $x < 0$ and $x > 8$.



12. Let X be a continuous random variable with a probability density function defined by

$f(x) = ax^2, 0 \le x \le 3$

where a is a constant.

a. Determine the value of a.

b. Determine $\mathrm{P}(1 \le X \le 2)$.

Complex familiar

13. X is a continuous random variable with a probability function defined by

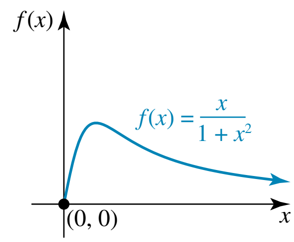
$$f(x) = \begin{cases} \dfrac{2\log_e(x)}{\sqrt{x}}, & 1 \le x \le a \\ 0, & \text{elsewhere} \end{cases}$$

where a is a constant.

a. Using technology, determine the value of the constant a. Give your answer correct to 2 decimal places.

b. Using technology, determine the mean and the median values of X. Give your answers correct to 3 decimal places.

14. The graph of the probability function $f(x) = \dfrac{x}{x^2+1}$, $x \geq 0$ for the continuous random variable X is shown.



a. Differentiate $\log_e\left(x^2+1\right)$ and hence determine an antiderivative of $\dfrac{x}{x^2+1}$

b. Determine $\text{P}(X \leq 2)$ correct to 3 decimal places.

c. If Y is a continuous random variable with a probability density function of $f(y) = \dfrac{y}{1+y^2}$ for $0 \leq y \leq a$, determine the value of the constant a correct to 1 decimal place.

d. Determine the median value of Y correct to 2 decimal places.

15. Given that the following function is a probability density function, determine the value of a.

$$f(x) = \begin{cases} ax + 0.5, & -2 \leq x < 0 \\ -ax + 1, & 0 \leq x \leq 1 \\ 0, & \text{elsewhere} \end{cases}$$

16. The continuous random variable X has a probability density function defined by

$$f(x) = \begin{cases} k\cos(2x), & 0 \leq x \leq \dfrac{\pi}{4} \\ 0, & \text{elsewhere} \end{cases}$$

where k is a constant.

a. Determine the value of k.

b. Determine the median of f.

Complex unfamiliar

17. The continuous random variable T has a probability density function defined by

$$f(t) = \begin{cases} 5e^{-5t}, & t \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

where T is a continuous random variable.

a. Determine the median value of T

b. Determine the interquartile range of T, correct to 4 decimal places.

c. Differentiate xe^{-5x} and hence determine an antiderivative of $5xe^{-5x}$

d. Determine $\text{E}(T)$

e. Differentiate x^2e^{-5x} and hence determine an antiderivative of $5x^2e^{-5x}$

f. Determine $\text{SD}(T)$

18. X is a continuous random variable such that

$$f(x) = \begin{cases} n\sin(3x)\cos(3x), & 0 \leq x \leq \dfrac{\pi}{12} \\ 0, & \text{elsewhere} \end{cases}$$

If f is known to be a probability density function, determine the value of the constant n.
(Hint: use the trigonometric identity $\sin(2kx) = 2\sin(kx)\cos(kx)$)

19. Patrick has just spread lawn seed on his nature strip. With constant watering and plenty of sunshine, the time it takes for the lawn seed to germinate, T days after the seeding, can be determined by the probability density function

$$f(t) = \begin{cases} ke^{-0.15t}, & t \geq 0 \\ 0, & \text{elsewhere} \end{cases}$$

where k is a constant.

a. Determine the value of the constant k.
b. Differentiate $te^{-0.15t}$ and hence determine an antiderivative of $0.15te^{-0.15t}$
c. What is the expected period of time for the germination of the lawn seed? Give your answer correct to the nearest day.

20. A function f is defined by the rule

$$f(x) = \begin{cases} \log_e(x), & x \geq 1 \\ 0, & \text{elsewhere} \end{cases}$$

a. Differentiate $x\log_e(x)$ and hence determine an antiderivative of $\log_e(x)$.
b. If $\int_1^a f(x)\ dx = 1$, determine the value of the real constant a.
c. Does this function define a probability density function?

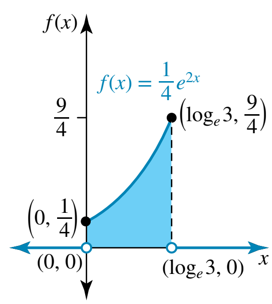
studyon
Units 3 & 4 Sit exam

Answers

11 General continuous random variables

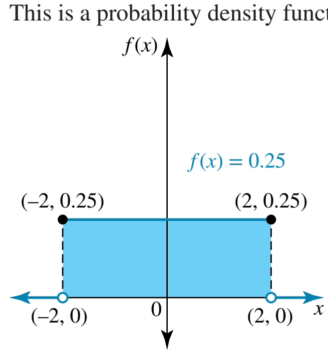
Exercise 11.2 Continuous random variables and the probability density function

1. **c**, **d**, **f**, **h**, **k** and **p** are continuous random variables.
2. a. 0.87 b. 0.17 c. 0.28
3. a. i. $\frac{9}{25}$ ii. $\frac{4}{25}$ b. i. $\frac{37}{50}$ ii. $\frac{37}{42}$
4. a. 100 batteries b. $\frac{29}{100}$ c. $\frac{41}{50}$ d. $\frac{3}{100}$
5. a.

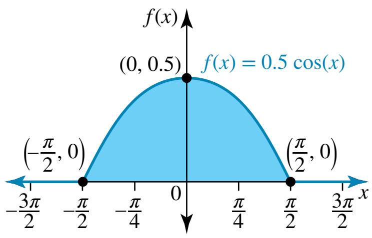


This is a probability density function.

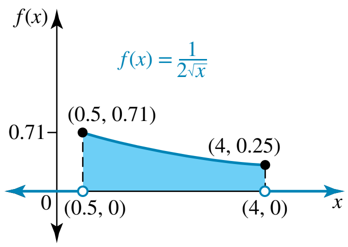
b.



This is a probability density function.

6. a. 

This is a probability density function

b. 

This is not a probability density function.

7. $n = \frac{1}{18}$

8. $a = \frac{1}{11}$

9. a. 200 b. i. $\frac{5}{8}$ ii. $\frac{31}{100}$ c. $\frac{21}{46}$

10. $\frac{5}{7}$

11. $\frac{1}{3}$

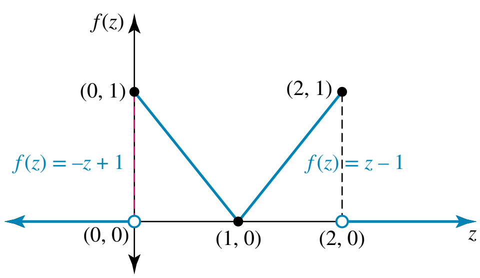
12. a. $\frac{1}{8}$ b. 2 c. $\frac{1}{4}$

13. Sample responses can be found in the worked solutions in the online resources.

14. a. $\int \log_e\left(\frac{x}{2}\right) \; dx = x\log_e\left(\frac{x}{2}\right) - x$

b. $a = 2e$

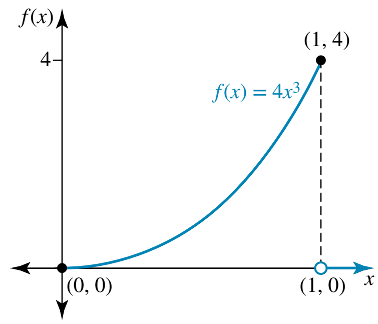
Exercise 11.3 Cumulative distribution functions

1. a. 

b. $\frac{15}{32}$

c. $\frac{5}{8}$

2. a. 1

b. 

c. $\frac{15}{16}$

3. a. $F(x) = \begin{cases} 0, & x \le 1 \\ \frac{x}{5} - \frac{1}{5}, & 1 < x \le 6 \\ 1, & x > 6 \end{cases}$

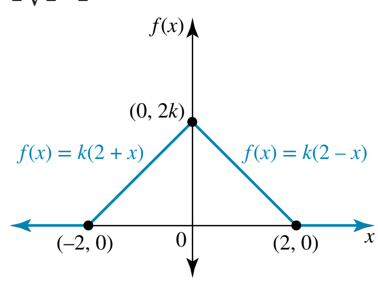
b. i. 0.6 ii. 0.46

4. a. $F(x) = \begin{cases} 0, & x \le 0 \\ \frac{1}{2}(1 - \cos(x)), & 0 < x \le \pi \\ 1, & x > \pi \end{cases}$

b. $\frac{1}{2}$

c. $\frac{\sqrt{2}}{2}$

d. $2\sqrt{2} - 2$

5. a. 

b. $\frac{1}{4}$

c. $F(x) = \begin{cases} 0, & x \le -2 \\ \frac{x}{2} + \frac{x^2}{8} + \frac{1}{2}, & -2 < x \le 0 \\ \frac{1}{2} + \frac{x}{2} - \frac{x^2}{8}, & 0 < x \le 2 \\ 1 & x > 2 \end{cases}$

d. $\frac{3}{4}$

e. $\frac{6}{7}$

6. a. $\frac{1}{12}$

b. $f(x) = \begin{cases} \frac{1}{12}, & 18 \le X < 30 \\ 0, & \text{elsewhere} \end{cases}$

c. $F(x) = \begin{cases} 0, & x < 18 \\ \dfrac{x}{12} - \dfrac{3}{2}, & 18 \le x \le 30 \\ 1, & x > 30 \end{cases}$

d. $\dfrac{5}{12}$

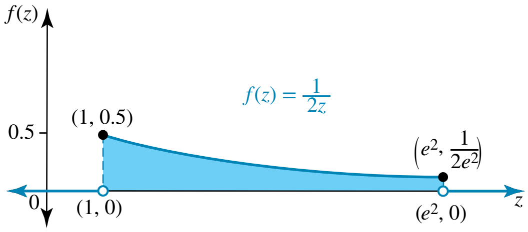
e. $\dfrac{1}{2}$

7. a. $F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x^3}{8}, & 0 \le x \le 2 \\ 1, & x > 2 \end{cases}$

b. 0.784

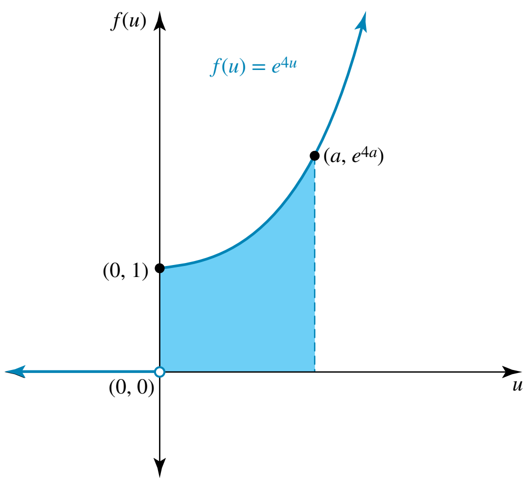
c. 1.817

8. a.



b. Sample responses can be found in the worked solutions in the online resources.

c.



d. $\dfrac{1}{4}\log_e(5)$

9. 0.604

10. a. 1 b. 0.715 c. 0.371 d. 0

11. a. $3\log_e\left(\dfrac{3}{2}\right)$ b. 0.6243 c. 0.5342 d. 0.60

12. a. 0.9502 b. 0.0025

13. a. $\int \log_e\left(x^2\right)\, dx = x\log_e\left(x^2\right) - 2x$

b. a = 2.1555 c. 0.7147

14. 0.1560

Exercise 11.4 Measures of centre and spread

1. a. $a = \dfrac{5}{4}$

b. i. $\mathrm{E}(Z) = \dfrac{9}{8}$ ii. $m = \dfrac{9}{8}$

2. a. $a = 1$ b. $\mathrm{E}(Y) = \dfrac{2}{3}$

c. $m = \dfrac{1}{\sqrt{2}}$

3. $\mathrm{E}(Z) = \dfrac{8}{3} = 2\dfrac{2}{3}$; $m = 2 + \dfrac{\sqrt{2}}{2}$; $\mathrm{Var}(Z) = \dfrac{1}{18}$; $\mathrm{SD}(Z) = \dfrac{1}{3\sqrt{2}}$

4. a. $m = -\dfrac{1}{3}\log_e(0.5)$

b. $\int 3xe^{-3x}dx = -\dfrac{1}{3}e^{-3x} - xe^{-3x}$

c. $\frac{1}{3}$

d. $\int 3x^2e^{-3x}dx = -\dfrac{2}{9}e^{-3x} - \dfrac{2}{3}xe^{-3x} - x^2e^{-3x}$

e. $\dfrac{1}{9}$

f. $\dfrac{1}{3}$

5. a. Sample responses can be found in the worked solutions in the online resources.

b. $\mathrm{E}(X) = \dfrac{1}{3}$

c. $m = 0.25$

6. a. 0.35 minutes

b. $\int 2xe^{-2x}dx = -\dfrac{1}{2}e^{-2x} - xe^{-2x}$

c. 0.5 minutes

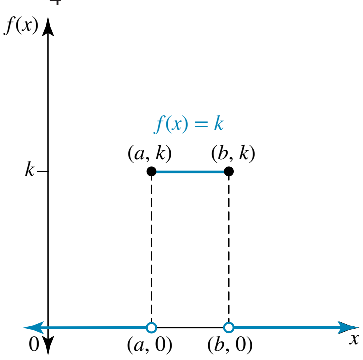
d. $\int 2x^2e^{-2x}dx = -\dfrac{1}{2}e^{-2x} - xe^{-2x} - x^2e^{-2x}$

e. 0.5 minutes

7. a. $\dfrac{dy}{dx} = -\dfrac{x}{\sqrt{4 - x^2}}$ b. $\mathrm{E}(X) = \dfrac{3}{\pi}$

8. a. $h = \dfrac{1}{4}$ b. $\mathrm{E}(X) = 2$ c. $\mathrm{Var}(X) = 2$

9. a.



b. Sample responses can be found in the worked solutions in the online resources.

c. $\mathrm{E}(X) = \dfrac{b + a}{2}$

d. $\mathrm{Var}(X) = \dfrac{(a - b)^2}{12}$

10. a. $\mathrm{E}(Y) = 1.5601$ b. $m = 1.6510$

c. $Q_1 = 1.3104$; $Q_3 = 1.8899$ d. $\mathrm{IQR} = 0.5795$

11. a. $a = 0.4809$

b. $\mathrm{E}(Z) = 3.3663$

c. $\text{Var}(Z) = 3.8195$; $\text{SD}(Z) = 1.9571$
d. $Q_1 = 1.6817$; $Q_3 = 4.7568$; IQR $= 3.0751$
e. Range $= 7$

12. a. Sample responses can be found in the worked solutions in the online resources.
b. $\text{E}(X) = 1.0708$
c. i. $\text{Var}(X) = 0.5725$ ii. $\text{SD}(X) = 0.7566$
d. $m = 0.9291$

13. $a = \dfrac{3}{2}$; $b = \dfrac{3}{4}$

14. a. Sample responses can be found in the worked solutions in the online resources.
b. $\text{E}(Z) = 1.2164$; $\text{Var}(Z) = 0.0204$
c. $m = \dfrac{6}{5}$; $Q_1 = \dfrac{12}{11}$; $Q_3 = \dfrac{4}{3}$; IQR $= \dfrac{8}{33}$

15. a. $\displaystyle\int_2^{7.9344} f(y)\,dy = 1$
b. 5.7278
c. $\text{Var}(Y) = 2.1600$, $\text{SD}(Y) = 1.4697$
d. 5.9160
e. 5.9344

16. a. $a = 2.3104$
b. i. $\text{E}(Z) = 1.7863$ ii. $\text{E}(Z^2) = 3.3085$
iii. $\text{Var}(Z) = 0.1176$ iv. $\text{SD}(Z) = 0.3430$

11.5 Review: exam practice

1. **b** and **e** are continuous random variables.

2. $\dfrac{\pi}{4}$

3. a. 150 b. $\dfrac{13}{15} = 0.8\dot{6}$

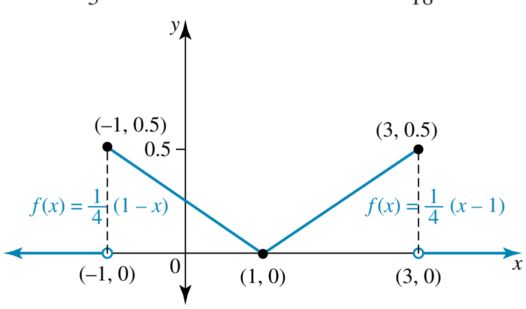
4. 6

5. a. $F(y) = \begin{cases} 0, & y < 0 \\ y^3, & 0 \le y \le 1 \\ 1, & y > 1 \end{cases}$
b. 0.992

6. a. The value of 19.5 kg does not lie at the end of an interval. As the weight is a continuous variable, the bags in the interval $19 \le W \le 20$ can have an infinite number of values.
b. $\dfrac{1}{7} \approx 0.1428$

7. a. $\text{E}(X) = \dfrac{2}{3}$ b. $\text{Var}(X) = \dfrac{1}{18} = 0.0\dot{5}$

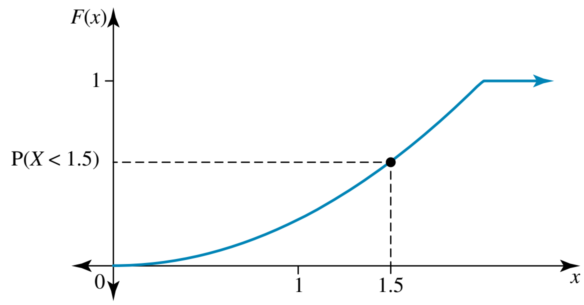
8. a.



b. Sample responses can be found in the worked solutions in the online resources.
c. 1

9. Sample responses can be found in the worked solutions in the online resources.

10. a. $F(x) = \begin{cases} 0, & x < 0 \\ \dfrac{x^2}{4}, & 0 \le x \le 2 \\ 1, & x > 2 \end{cases}$
b.



11. $f(x) = \begin{cases} \dfrac{1}{4}x, & 0 \le x \le 1 \\ \dfrac{2}{7} - \dfrac{x}{28}, & 1 < x \le 8 \\ 0, & \text{elsewhere} \end{cases}$

12. a. $\dfrac{1}{9}$ b. $\dfrac{7}{27}$

13. a. 2.37
b. $\text{E}(X) = 1.843$; Median $= 1.887$

14. a. $\displaystyle\int \frac{x}{x^2+1}\,dx = \frac{1}{2}\log_e\left(x^2+1\right)$
b. 0.805
c. 2.5
d. 1.31

15. $\dfrac{2}{5}$

16. a. 2 b. $\dfrac{\pi}{12}$

17. a. $-\dfrac{1}{5}\log_e(0.5)$ or $\dfrac{1}{5}\log_e(2)$
b. 0.2197
c. $\displaystyle\int 5xe^{-5x}\,dx = -\tfrac{1}{5}e^{-5x} - xe^{-5x}$
d. $\dfrac{1}{5}$
e. $\displaystyle\int 5x^2e^{-5x}dx = -\frac{2}{25}e^{-5x} - \frac{2}{5}xe^{-5x} - x^2e^{-5x}$
f. $\dfrac{1}{5}$

18. 12

19. a. 0.15
b. $\displaystyle\int 0.15te^{-0.15t}\,dt = -\tfrac{20}{3}e^{-0.15t} - te^{-0.15t}$
c. 7 days

20. a. $\displaystyle\int \log_e(x)\,dx = x\log_e(x) - x$
b. e
c. The function f is a probability density function because $f(x) \ge 0$ and $\displaystyle\int_1^e f(x)\,dx = 1$.