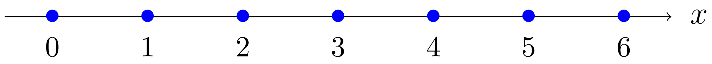
Chapter 23

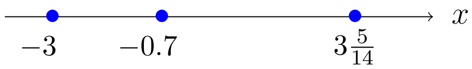
Continuous Random Variables

23.1 Introduction

So far, we have only considered **discrete** random variables, which means that the values of the random variables can be written in a list, and have gaps between them. In particular, we have usually had $X = 0,\ 1,\ 2,\ \ldots$.



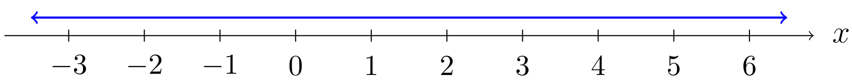
It is, however, possible for a discrete variable to have other values, such as negative values, or fractional values, or decimal values. For example, we could have a discrete random variable which takes the values $-3,\ -0.7,\ 3\frac{5}{14}$.



The important aspect of being a **discrete** variable is that there are

gaps between its possible values.

In contrast, if X is a **continuous** random variable, then X can be any real number within some interval (or within some union of intervals). For example, a continuous random variable might take any value from the interval $\mathbf{R}$ (of all real numbers), with no gaps between its possible values.



Alternatively, a continuous random variable might take any value from some smaller interval.

Example 1.

Let X be the height (in centimetres) of a randomly chosen university student. Then X is a continuous random variable.

Note that, if we assume that every university student has height between (say) 80 cm and 260 cm, then X can take any value from the interval $[80, 260]$ (with no gaps between its possible values).

Also, note that, for example,

$$\Pr(X = 170) = \Pr(X = 170.0000000\ldots) = 0.$$

This probability is zero since it is so incredibly unlikely that there is a student with height equal to *exactly* 170 cm, accurate to infinitely many decimal places.

In contrast, probabilities such as $\Pr(169 < X < 170)$ are (probably) non-zero.

$\square$

Result:

For every *continuous* random variable X and for every real number b, we have

$$\boxed{\mathbf{Pr}(\boldsymbol{X} = \boldsymbol{b}) = \mathbf{0}.}$$

- Notice that this is very different from the situation for *discrete* random variables.
- Because of this result, in examples involving a continuous random variable X, we will usually be more interested in finding probabilities such as

$$\Pr(a \le X \le b)$$

  (of X lying in an *interval*), rather than the trivial $\Pr(X = b)$.
- Another consequence of this result is that it does not matter whether we write $\le$ or $<$ when we write down probabilities for a continuous random variable. For example, if X is a continuous random variable, we can write

$$\begin{aligned}
\Pr(a \le X \le b) &= \Pr(X = a \quad \text{or} \quad a < X < b \quad \text{or} \quad X = b) \\
&= \Pr(X = a) + \Pr(a < X < b) + \Pr(X = b) \\
&= 0 + \Pr(a < X < b) + 0 \\
&= \Pr(a < X < b)\,.
\end{aligned}$$

$$\begin{aligned} \text{Similarly, we can write} \quad \Pr(X \leq b) &= \Pr(X < b \quad \text{or} \quad X = b) \\ &= \Pr(X < b) + \Pr(X = b) \\ &= \Pr(X < b) + 0 \\ &= \Pr(X < b) \end{aligned}$$

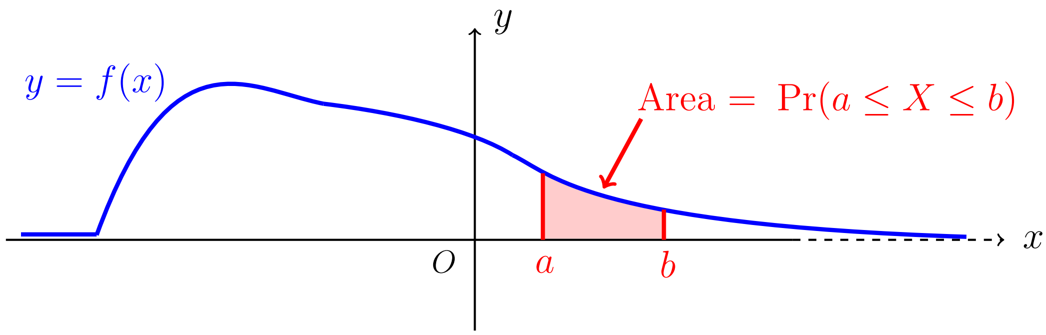
$$\begin{aligned} \text{and} \quad \Pr(X \geq a) &= \Pr(X > a \quad \text{or} \quad X = a) \\ &= \Pr(X > a) + \Pr(X = a) \\ &= \Pr(X > a) + 0 \\ &= \Pr(X > a)\,. \end{aligned}$$

If X is a continuous random variable, then probabilities of the form

$$\Pr(a \leq X \leq b)$$

are found by calculating the

area under a curve $y = f(x)$, for x between a and b.



The function f is known as the **probability density function** of X, and we will learn about such functions next.

23.2 Probability Density Functions

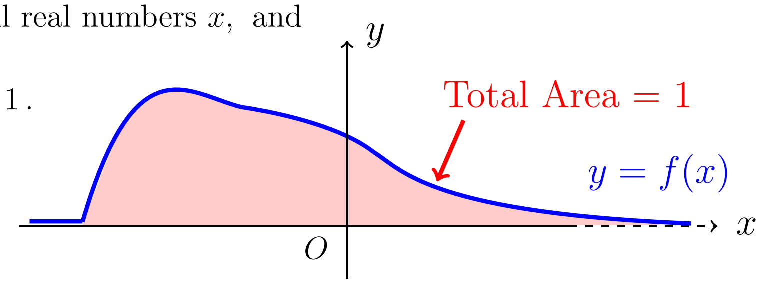
Definition:

A **probability density function** (often referred to as a **pdf**) is a function f with the following properties:

(a) The domain of f is the set of all real numbers; and

(b) $f(x) \geq 0$ for all real numbers x, and

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



Recall (from Chapter 20) that for each *discrete* probability distribution,

the overall sum of the probabilities must be 1.

Correspondingly, for each *continuous* probability distribution,

the total area under the curve of the probability density function must be 1

(hence the third condition in the definition of a probability density function). That is, we must have

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

Let X be a continuous random variable and let f be a probability density function. Then f is called "the probability density function ***of X***" if probabilities for X can be calculated by finding areas under the curve $y = f(x)$. That is, we have

$$\boxed{\Pr(a \leq X \leq b) = \int_a^b f(x)\,dx.}$$

Notice that, using this idea, we have

$$\begin{aligned}\Pr(X = b) &= \Pr(b \leq X \leq b)\\ &= \int_b^b f(x)\,dx\\ &= 0 \text{ (using Property (f) on page 12 of Chapter 11),}\end{aligned}$$

which (of course) agrees with the result that we saw back on page 2 of this current chapter.

Language Alert:

The word "continuous" is used for two different ideas within the Mathematics 1 course (and so we need to ensure that we do not confuse these two ideas).

- Back in Chapter 7 we learnt about *continuous functions.*
  (Recall that a function f is continuous at $x = a$ if $\lim_{x \to a} f(x) = f(a)$.)
- In contrast, in this current chapter we are learning about *continuous random variables* (which are random variables that can take any value from some interval).

In particular, if f is the probability density function of a continuous random variable X, then f is *not* necessarily a continuous function. In fact, in the examples, we will see that often f is *not* a continuous function.

Finding probabilities for continuous random variables often involves calculating improper integrals (as studied in Chapter 13). In particular, we have

$$\boxed{\Pr(X \le b) \quad = \quad \int_{-\infty}^{b} f(x)\, dx}$$

and also

$$\boxed{\Pr(X \ge a) \quad = \quad \int_{a}^{\infty} f(x)\, dx.}$$

Recall from Chapter 13.2 that we must *use limits* to evaluate improper integrals.
For example, we will need to calculate

$$\begin{aligned}
\Pr(X \le b) &= \int_{-\infty}^{b} f(x)\, dx \\
&= \lim_{l \to -\infty} \int_{l}^{b} f(x)\, dx \\
&= \ldots
\end{aligned}$$

and

$$\begin{aligned}
\Pr(X \ge a) &= \int_{a}^{\infty} f(x)\, dx \\
&= \lim_{n \to \infty} \int_{a}^{n} f(x)\, dx \\
&= \ldots .
\end{aligned}$$

The only situation for which we do *not* need to introduce limits to calculate improper integrals is if we are integrating *zero*. In particular, it will be very useful to remember that

an integral of the form $\displaystyle\int_a^b 0\,dx$ is

always equal to $\mathbf{0}$,

regardless of whether a and b are finite or infinite.

In particular, we can always straightaway write $\displaystyle\int_{-\infty}^b 0\,dx = 0$ and $\displaystyle\int_a^{\infty} 0\,dx = 0$, without introducing limits.

Fortunately, we will often encounter this situation. The reason why so many examples involve integrals of the form $\displaystyle\int_a^b 0\,dx$ (where a and b might be finite or infinite) is explained next:

> Suppose that a continuous random variable X can take any value within some interval I (which might be smaller than the interval $\mathbf{R}$). If f is the probability density function of X, then we must have $f(x) = 0$ for all $x \in \mathbf{R} \setminus I$.

Example 2. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 1 \\ \dfrac{1}{6} & \text{if } 1 \le x \le 7 \\ 0 & \text{if } x > 7\,. \end{cases}$$

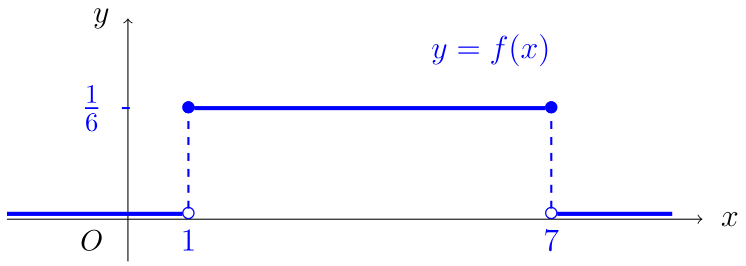
(a) Sketch a graph of $y = f(x)$.

(b) Find $\Pr(1 \le X \le 4)$.

(c) Find $\Pr(X < 5)$.

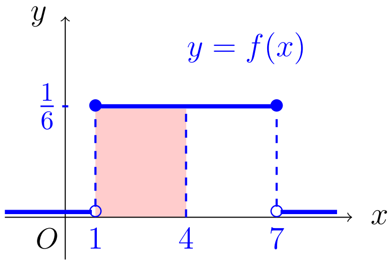
Solution:

(a)

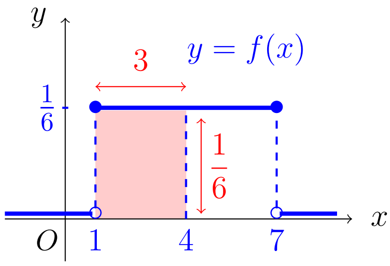


Notice that even though we are told that X is a *continuous* random variable, we can easily see that this probability density function f is *not* a continuous function (since clearly we lift our pens off the page at $x = 1$ and again at $x = 7$).

(b) We see that all x-values from the interval $[1, 4]$ satisfy the middle condition of the function f. Therefore, we will only need to use the middle part of the function f to calculate this probability.



$$\begin{aligned}
\Pr(1 \le X \le 4) &= \int_1^4 f(x)\,dx \\
&= \int_1^4 \frac{1}{6}\,dx \\
&= \left[\frac{1}{6}x\right]_1^4 \\
&= \frac{4}{6} - \frac{1}{6} \\
&= \frac{1}{2}.
\end{aligned}$$



Alternatively, since the region of interest (as shaded in the adjacent diagram) is very simple for this example, we can easily find the required area without integrating:

We see that the region under the graph for x between 1 and 4 is a rectangle, with width 3 and height $\frac{1}{6}$. Thus we have

$$\begin{aligned}
\Pr(1 \le X \le 4) &= \text{area under the graph for } x \text{ between 1 and 4} \\
&= \text{width} \times \text{height (of the rectangle)} \\
&= 3 \times \frac{1}{6} \\
&= \frac{1}{2} \quad \text{(as seen previously).}
\end{aligned}$$

(c) First, notice that $\Pr(X < 5) = \int_{-\infty}^5 f(x)\,dx$. However, since

- some x-values from the interval $(-\infty, 5)$ satisfy the *first* condition of f, whereas
- other x-values from that interval satisfy the *middle* condition of f,

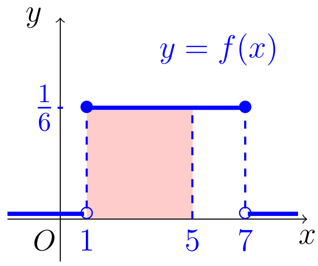
we will need to split the integral $\int_{-\infty}^5 f(x)\,dx$ into two integrals.

The split occurs at $x = 1$ (which is the x-value where f changes from its first piece to its second piece).

Thus we have $\Pr(X < 5)$

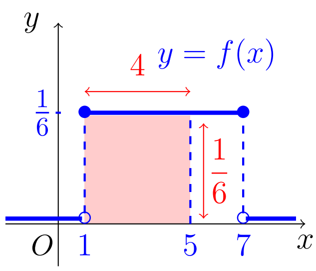
These steps can be omitted if you feel comfortable going directly to the next step.

$$
\begin{aligned}
&= \int_{-\infty}^{5} f(x)\,dx \\
&= \int_{-\infty}^{1} f(x)\,dx + \int_{1}^{5} f(x)\,dx \\
&= \int_{-\infty}^{1} 0\,dx + \int_{1}^{5} \frac{1}{6}\,dx \\
&= 0 + \int_{1}^{5} \frac{1}{6}\,dx \\
&= \int_{1}^{5} \frac{1}{6}\,dx \\
&= \left[\frac{1}{6}x\right]_{1}^{5} \\
&= \frac{5}{6} - \frac{1}{6} \\
&= \frac{2}{3}\,.
\end{aligned}
$$



Notice that we immediately replaced the improper integral $\int_{-\infty}^{1} 0\,dx$ with its answer 0 (without bothering to introduce a limit). We will use this type of idea repeatedly, and without further comment, throughout the rest of the chapter.

Also, again notice that, since the region is very simple, we can easily find the required area without integrating. We see that the region under the graph, for x on the left of 5, is a rectangle with width 4 and height $\frac{1}{6}$. Thus we have



$$
\begin{aligned}
&\Pr(X < 5) \\
&\quad = \text{area under the graph for } x \text{ on the left of } 5 \\
&\quad = \text{width} \times \text{height (of the rectangle)} \\
&\quad = 4 \times \frac{1}{6} \\
&\quad = \frac{2}{3} \quad \text{(as seen previously).}
\end{aligned}
$$

$\square$

Example 3. Consider the function $f(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{3}{4}\left(1 - x^2\right) & \text{if } -1 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$

Show that f is a probability density function.

Solution:

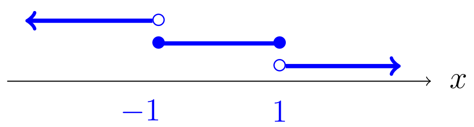
By the definition given on page 4 of this chapter, we need to show that the following three properties are satisfied:

(a) The domain of f is the set of all real numbers.

(b) $f(x) \ge 0$ for all real numbers x.

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

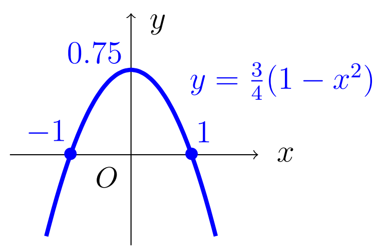
- For property (a), we simply observe that the three intervals for x that were listed as part of the given piecewise function f, combine to give the entire number-line.



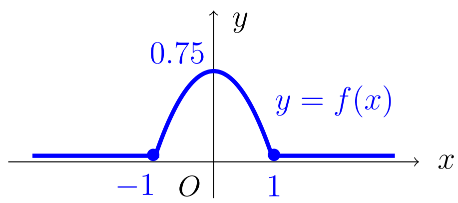
That is,
$$(-\infty, -1) \cup [-1, 1] \cup (1, \infty) = \mathbf{R}.$$
Thus f is defined for all real numbers, as required.

- To see that property (b) is satisfied, it is easiest to consider the graph of $y = f(x)$.

We know that the parabola $y = \frac{3}{4}(1 - x^2)$ looks like



and hence the graph of $y = f(x)$ is



from which we see that $f(x) \ge 0$ for all $x \in \mathbf{R}$, as required.

- Finally, for property (c), we need to show that $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx = 1.$

  Notice that

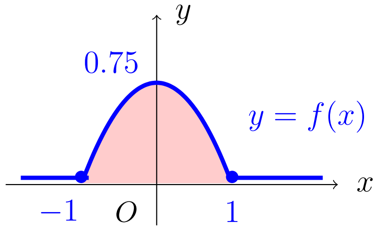
  - some x-values from the interval $(-\infty, \infty)$ satisfy the *first* condition of f, whereas
  - other x-values from that interval satisfy the *second* or *third* conditions of f.

  Therefore, we will need to split the integral $\displaystyle\int_{-\infty}^{\infty} f(x)\,dx$ into *three* integrals, with the splits occurring at $x = -1$ and at $x = 1$ (matching the values at which f changes between its various pieces). Thus we have

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

These steps can be omitted if you feel comfortable going directly to the next step.

$$\begin{aligned}
&= \int_{-\infty}^{-1} f(x)\,dx + \int_{-1}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx \\
&= \int_{-\infty}^{-1} 0\,dx + \int_{-1}^{1} \frac{3}{4}\left(1 - x^2\right)\,dx + \int_{1}^{\infty} 0\,dx \\
&= 0 + \int_{-1}^{1} \frac{3}{4}\left(1 - x^2\right)\,dx + 0
\end{aligned}$$



$$\begin{aligned}
&= \frac{3}{4}\int_{-1}^{1} \left(1 - x^2\right)\,dx \\
&= \frac{3}{4}\left[x - \frac{x^3}{3}\right]_{-1}^{1} \\
&= \frac{3}{4}\left(\left\{1 - \frac{1^3}{3}\right\} - \left\{(-1) - \frac{(-1)^3}{3}\right\}\right) \\
&= \frac{3}{4}\left(1 - \frac{1}{3} + 1 - \frac{1}{3}\right) \\
&= \frac{3}{4} \times \frac{4}{3} \\
&= 1, \quad \text{as required.}
\end{aligned}$$

$\square$

Example 4. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < -1 \\ x+1 & \text{if } -1 \le x < 0 \\ \dfrac{1}{2}\,e^{-x} & \text{if } x \ge 0\,. \end{cases}$$

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

(b) Find $\Pr(X = 0)$.

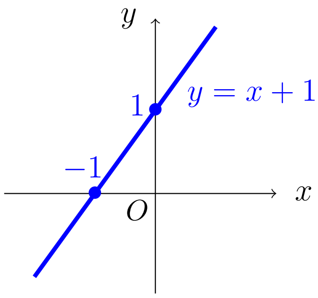
(c) Find $\Pr(X \le 0)$.

(d) Find $\Pr(-1 < X < 1)$.

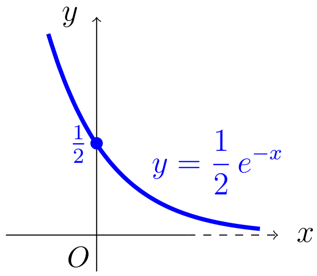
(e) Find $\Pr(X > 2)$.

Solution:

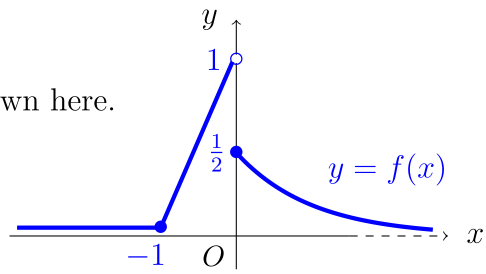
(a) We know that the graphs of $y = x + 1$ and $y = \dfrac{1}{2}\,e^{-x}$ look like



and



and so the graph of $y = f(x)$ is as shown here.



(b) Since X is a continuous random variable, we immediately conclude that $\Pr(X = 0) = 0$.

(c) First, notice that $\Pr(X \leq 0) \;=\; \Pr(X < 0) \;=\; \displaystyle\int_{-\infty}^{0} f(x)\,dx$.

Furthermore, notice that,

- some x-values from the interval $(-\infty, 0)$ satisfy the *first* condition of f,

and

- the other x-values from that interval satisfy the *second* condition of f.

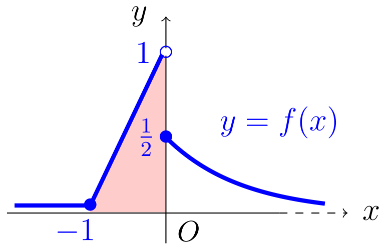
Thus we will split the integral $\displaystyle\int_{-\infty}^{0} f(x)\,dx$ into *two* integrals.

- Note that we do *not* need to include a third integral to deal with the single value $x = 0$, as $\displaystyle\int_{0}^{0} f(x)\,dx = 0$.

The split is chosen to occur at $x = -1$ (to match with where f changes from its first piece to its second piece).

These steps can be omitted if you feel comfortable going directly to the next step.

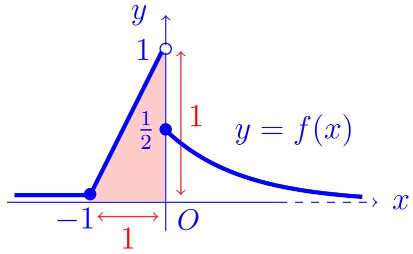
$$
\begin{aligned}
&\Pr(X \leq 0) \\
&= \int_{-\infty}^{0} f(x)\,dx \\
&= \int_{-\infty}^{-1} f(x)\,dx \;+\; \int_{-1}^{0} f(x)\,dx \\
&= \int_{-\infty}^{-1} 0\,dx \;+\; \int_{-1}^{0} (x+1)\;dx \\
&= 0 \;+\; \int_{-1}^{0} (x+1)\;dx \\
&= \int_{-1}^{0} (x+1)\;dx \\
&= \left[\frac{x^2}{2} \;+\; x\right]_{-1}^{0} \\
&= \left(\frac{0^2}{2} \;+\; 0\right) \;-\; \left(\frac{(-1)^2}{2} \;+\; -1\right) \\
&= 0 \;-\; \frac{1}{2} \;+\; 1 \\
&= \frac{1}{2}\,.
\end{aligned}
$$



An alternative method for calculating this probability is shown next.

Notice that, since this region is very simple, we can (again) easily find the required area without integrating.

In particular, since the region under the graph for x on the left of 0 is a triangle, with width 1 and height 1, we have



$$\begin{aligned}
\Pr(X \le 0) &= \text{area under the graph for } x \text{ on the left of } 0 \\
&= \frac{1}{2} \times \text{width} \times \text{height (of the triangle)} \\
&= \frac{1}{2} \times 1 \times 1 \;=\; \frac{1}{2}\,.
\end{aligned}$$

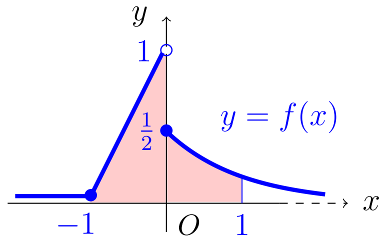
(d) We have $\Pr(-1 < X < 1) = \int_{-1}^{1} f(x)\,dx$. Since

- some x-values from the interval $(-1, 1)$ satisfy the *second* condition of f, whereas
- other x-values from that interval satisfy the *third* condition of f,

we will need to split the integral $\int_{-1}^{1} f(x)\,dx$ into two integrals.

The split is chosen to occur at $x = 0$ (which is the x-value where f changes from its second piece to its third piece). We have

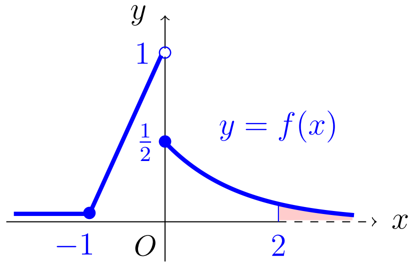
These steps can be omitted if you feel comfortable going directly to the next step.

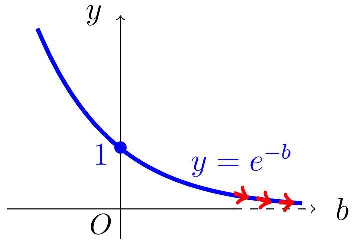


$$\begin{aligned}
&\Pr(-1 < X < 1) \\
&= \int_{-1}^{1} f(x)\,dx \\
&= \int_{-1}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx \\
&= \int_{-1}^{0} (x+1)\,dx + \int_{0}^{1} \left(\frac{1}{2}\,e^{-x}\right) dx \\
&= \left[\frac{x^2}{2} + x\right]_{-1}^{0} + \left[-\frac{1}{2}\,e^{-x}\right]_{0}^{1} \\
&= \left(\frac{0^2}{2} + 0\right) - \left(\frac{(-1)^2}{2} + -1\right) + \left(-\frac{1}{2}e^{-1}\right) - \left(-\frac{1}{2}\,e^{0}\right) \\
&= 0 - \left(\frac{1}{2} - 1\right) - \frac{1}{2e} + \frac{1}{2} \times 1 \\
&= \frac{1}{2} - \frac{1}{2e} + \frac{1}{2} \times 1 \\
&= 1 - \frac{1}{2e}\,.
\end{aligned}$$

(e) We see that *all* the x-values from the interval $(2, \infty)$ satisfy the third condition of the function f. Therefore, we will *only* need to use the third part of the function f to calculate this probability. Thus we have

$$
\begin{aligned}
\Pr(X > 2) &= \int_2^\infty f(x)\,dx \\
&= \int_2^\infty \frac{1}{2}\,e^{-x}\,dx \\
&= \lim_{b\to\infty} \int_2^b \frac{1}{2}\,e^{-x}\,dx \\
&= \lim_{b\to\infty} \left[-\frac{1}{2}\,e^{-x}\right]_2^b \\
&= \lim_{b\to\infty} \left(\left(-\frac{1}{2}\,e^{-b}\right) - \left(-\frac{1}{2}e^{-2}\right) \right) \\
&= \lim_{b\to\infty} \left(-\frac{1}{2}\,e^{-b} + \frac{1}{2e^2}\right) .
\end{aligned}
$$





We see, from the graph of $y = e^{-b}$ (shown here) that $\lim\limits_{b\to\infty} e^{-b} = 0$, and so we can use Limit Laws to conclude that

$$
\begin{aligned}
\Pr(X > 2) &= -\frac{1}{2} \times \lim_{b\to\infty} \left(e^{-b}\right) + \lim_{b\to\infty} \left(\frac{1}{2e^2}\right) \\
&= -\frac{1}{2} \times 0 + \frac{1}{2e^2} = \frac{1}{2e^2} .
\end{aligned}
$$

An alternative method, which avoids the need to introduce a limit, is as follows.

$$
\begin{aligned}
\Pr(X > 2) &= 1 - \Pr(X \le 2) \\
&= 1 - \int_{-\infty}^2 f(x)\,dx \\
&= 1 - \left(\int_{-\infty}^0 f(x)\,dx + \int_0^2 f(x)\,dx \right) \\
&= 1 - \left(\frac{1}{2} \ \textit{(from (c))} + \int_0^2 \frac{1}{2}\,e^{-x}\,dx \right) \\
&= 1 - \frac{1}{2} - \frac{1}{2} \int_0^2 e^{-x}\,dx \\
&= \frac{1}{2} - \frac{1}{2} \left[-\,e^{-x}\right]_0^2 \\
&= \frac{1}{2} - \frac{1}{2} \left((-e^{-2}) - (-e^0)\right) \\
&= \frac{1}{2} + \frac{1}{2}e^{-2} - \frac{1}{2} \times 1 = \frac{1}{2}e^{-2}
\end{aligned}
$$

$\square$

Exercises for Section 23.2

1. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < -4 \\ k & \text{if } -4 \le x \le 3 \\ 0 & \text{if } x > 3\,, \end{cases}$$

   where k is a constant.

   (a) Find the value of k .

   (b) Find $\Pr(X > -1)$.

2. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < -1 \\ \dfrac{3}{4}\left(1 - x^2\right) & \text{if } -1 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$

   (a) Find $\Pr\left(X = \dfrac{1}{2}\right)$.

   (b) Find $\Pr\left(-\dfrac{1}{2} < X < \dfrac{1}{2}\right)$.

3. Consider the continuous random variable X with probability density function

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

   (a) Sketch a graph of $y = f(x)$.

   (b) Find $\Pr(X > 1)$.

   (c) Find $\Pr\left(X > 1 \mid X > \dfrac{1}{3}\right)$.

4. Consider the probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < k \\ \dfrac{1}{4} & \text{if } k \le x < 1 \\ \dfrac{1}{4x^2} & \text{if } x \ge 1\,, \end{cases}$$

   where k is a negative constant.

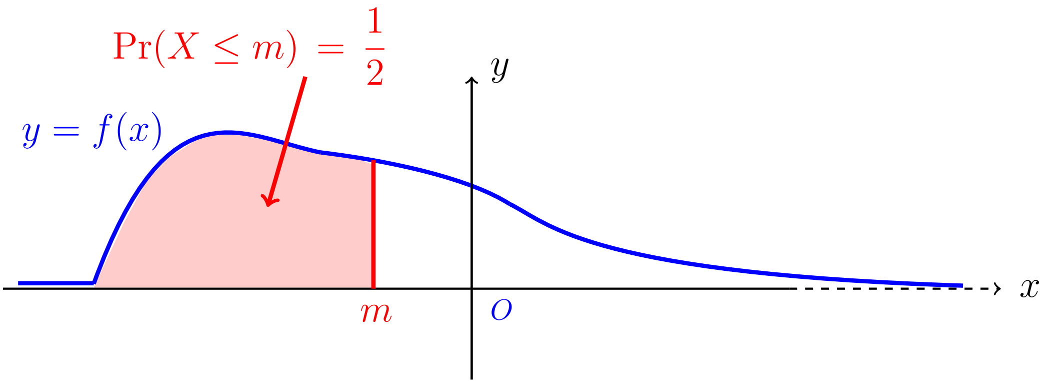
   (a) Sketch a graph of $y = f(x)$.

   (b) Find the value of k .

23.3 Median

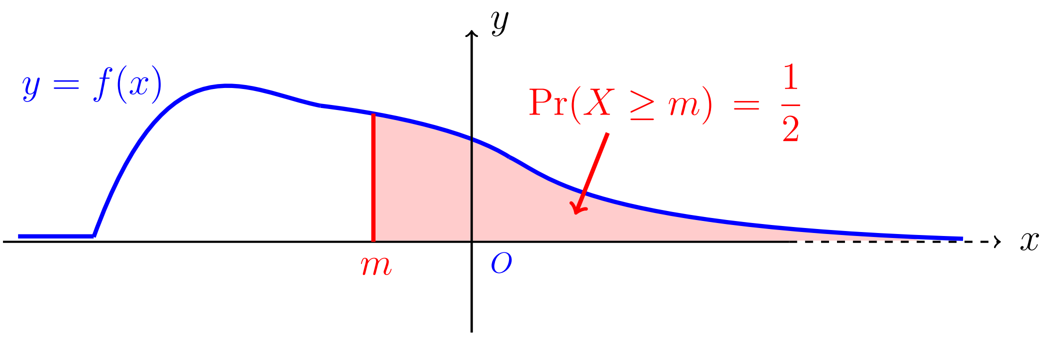
If X is a continuous random variable with probability density function f, then the **median** is

$$\boxed{\text{the number } m \text{ with the property that } \Pr(X \le m) = \frac{1}{2}\,.}$$



Of course, this is equivalent to saying that the **median** is

$$\boxed{\text{the number } m \text{ with the property that } \Pr(X \ge m) = \frac{1}{2}\,.}$$



Thus, the median is the *middle value* of the distribution.

To find the median, we solve for m in either of the following integral equations:

$$\boxed{\int_{-\infty}^{m} f(x)\,dx = \frac{1}{2} \qquad \text{or} \qquad \int_{m}^{\infty} f(x)\,dx = \frac{1}{2}\,.}$$

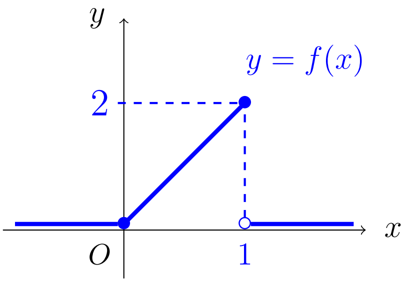
Example 5. Consider the continuous random variable X with probability density function

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

Find the median of X .

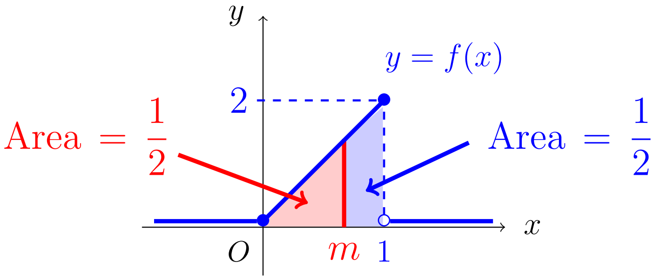
Solution:

It is useful to consider the graph of $y = f(x)$, given below.



We want to find the number m such that $\displaystyle\int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}$,

$\left(\text{or we could find the number } m \text{ such that } \displaystyle\int_{m}^{\infty} f(x)\,dx = \frac{1}{2}\right)$.



We can see from the above graph that we must have

$$\boldsymbol{0 < m < 1},$$

since

- if m satisfied $m \le 0$, then the area on the left of m would be 0 (instead of $\frac{1}{2}$), and, similarly,
- if m satisfied $m \ge 1$, then the area on the right of m would be 0 (instead of $\frac{1}{2}$).

Suppose we choose to work with the requirement $\displaystyle\int_{-\infty}^{m} f(x)\,dx = \frac{1}{2}$.

Since

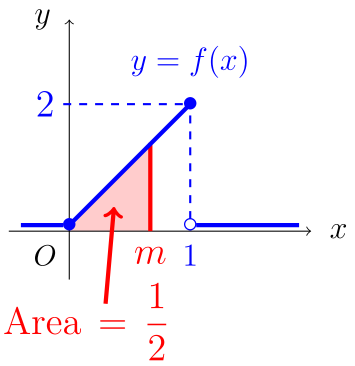
- some of the x-values from the interval $(-\infty, m]$ satisfy the *first* condition of f, whereas
- other x-values from that interval satisfy the *second* condition of f,

we need to split the integral $\displaystyle\int_{-\infty}^{m} f(x)\,dx$ into two integrals. The split is chosen to occur at $x = 0$ (which is the x-value where f changes from its first piece to its second piece).

We have

$$\begin{aligned} \int_{-\infty}^{m} f(x)\,dx &= \frac{1}{2} \\ \Rightarrow \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{m} f(x)\,dx &= \frac{1}{2} \\ \Rightarrow \int_{-\infty}^{0} 0\,dx + \int_{0}^{m} 2x\,dx &= \frac{1}{2} \\ \Rightarrow 0 + \int_{0}^{m} 2x\,dx &= \frac{1}{2} \end{aligned}$$

These steps can be omitted if you feel comfortable going directly to the next step.



$$\begin{aligned} \Rightarrow \int_{0}^{m} 2x\,dx &= \frac{1}{2} \\ \Rightarrow \Big[x^2\Big]_0^m &= \frac{1}{2} \\ \Rightarrow m^2 - 0^2 &= \frac{1}{2} \\ \Rightarrow m^2 &= \frac{1}{2} \\ \Rightarrow m &= \pm\frac{1}{\sqrt{2}}. \end{aligned}$$

However, recall that we previously noticed that, for this example, we need

$$0 < m < 1.$$

Thus, we *cannot* have $m = -\dfrac{1}{\sqrt{2}}$, and so we conclude that

$$\text{the median of } X \text{ is } \frac{1}{\sqrt{2}}.$$

□

23.4 Expected Value

The expected value $E(X)$, or mean μ, of a random variable X is the *long-run average* of X, in the sense that, if we repeat the experiment many times, the average of the outcomes will approach $E(X)$.
Recall (from Chapter 20.2) that for a *discrete* random variable, X, we have

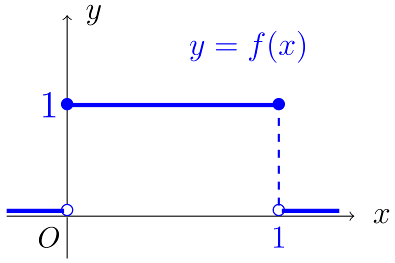
$$E(X) \;=\; \sum x \, \Pr(X = x)\,.$$

In contrast, if X is a *continuous* random variable with probability density function f, then we define the **expected value** of X as follows.

$$\boxed{E(X) \;=\; \int_{-\infty}^{\infty} x\, f(x)\, dx.}$$

Example 6. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$



Find $E(X)$.

Solution:

$$E(X) \;=\; \int_{-\infty}^{\infty} x\, f(x)\, dx$$

These steps can be omitted if you feel comfortable going directly to the next step.

$$\begin{aligned}
&= \int_{-\infty}^{0} x\, f(x)\, dx \;+\; \int_{0}^{1} x\, f(x)\, dx \;+\; \int_{1}^{\infty} x\, f(x)\, dx \\
&= \int_{-\infty}^{0} x \times 0\, dx \;+\; \int_{0}^{1} x \times 1\, dx \;+\; \int_{1}^{\infty} x \times 0\, dx \\
&= \int_{-\infty}^{0} 0\, dx \;+\; \int_{0}^{1} x \times 1\, dx \;+\; \int_{1}^{\infty} 0\, dx \\
&= 0 \;+\; \int_{0}^{1} x \times 1\, dx \;+\; 0
\end{aligned}$$

$$\begin{aligned}
&= \int_{0}^{1} x \times 1\, dx \\
&= \int_{0}^{1} x\, dx \\
&= \left[\frac{x^2}{2}\right]_0^1 \\
&= \frac{1^2}{2} - \frac{0^2}{2} \\
&= \frac{1}{2}\,.
\end{aligned}$$

$\square$

Properties of Expected Value

If X is a continuous random variable, and if a and b are constants, then

$$\boxed{E\,(aX+b) \;=\; a\,E(X)\,+\,b.}$$

- This is the same result as was stated back in Chapter 20.2 (when we were learning about the expected value of *discrete* random variables).

Example 7. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$

Find $E(2X+1)$.

Solution:

By Example 6, we already know that this random variable X has $E(X) = \dfrac{1}{2}$. We can use this to help us find $E(2X+1)$ very quickly.

$$\begin{aligned} E(2X+1) &= 2\,E(X)\,+\,1 \\ &= 2 \times \frac{1}{2} + 1 \\ &= 2. \end{aligned}$$

$\square$

If X is a continuous random variable, with probability density function f, then, as in the case of a discrete random variable, we usually have

$$E(X^2) \;\neq\; \Big[E(X)\Big]^2\,.$$

Instead, to calculate $E(X^2)$ for a continuous random variable X, we must use the formula

$$\boxed{E\left(X^2\right) \;=\; \int_{-\infty}^{\infty} x^2\,f(x)\,dx.}$$

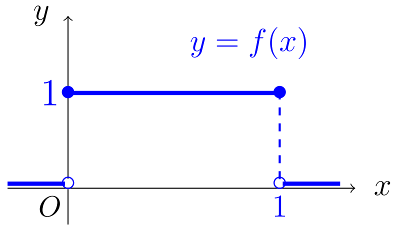
In general, for any function g, and for any continuous random variable X, we have

$$\boxed{E\Big(g(X)\Big) \;=\; \int_{-\infty}^{\infty} g(x)\,f(x)\,dx.}$$

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

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$

Find $E\,(X^2)$.



Solution:

$$E\left(X^2\right) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx$$

These steps can be omitted if you feel comfortable going directly to the next step.

$$\begin{aligned}
&= \int_{-\infty}^{0} x^2\, f(x)\; dx \;+\; \int_{0}^{1} x^2\, f(x)\; dx \;+\; \int_{1}^{\infty} x^2\, f(x)\; dx \\
&= \int_{-\infty}^{0} x^2 \times 0\; dx \;+\; \int_{0}^{1} x^2 \times 1\; dx \;+\; \int_{1}^{\infty} x^2 \times 0\; dx \\
&= \int_{-\infty}^{0} 0\; dx \;+\; \int_{0}^{1} x^2 \times 1\; dx \;+\; \int_{1}^{\infty} 0\; dx \\
&= 0 \;+\; \int_{0}^{1} x^2 \times 1\; dx \;+\; 0
\end{aligned}$$

$$\begin{aligned}
&= \int_{0}^{1} x^2 \times 1\; dx \\
&= \int_{0}^{1} x^2\, dx \\
&= \left[\frac{x^3}{3}\right]_0^1 \\
&= \frac{1^3}{3} - \frac{0^3}{3} \\
&= \frac{1}{3}\,.
\end{aligned}$$

□

We have seen, in Examples 6 and 8, that this continuous random variable X has

$$E(X) = \frac{1}{2} \quad \text{and} \quad E(X^2) = \frac{1}{3}\,.$$

Thus, it is evident that, for this particular continuous random variable X, we have

$$E(X^2) \neq \Big[E(X)\Big]^2\,.$$

23.5 Variance and Standard Deviation

Let X be a continuous random variable with mean μ. The **variance** of X is defined in exactly the same way as seen previously (in Chapter 20.3) for *discrete* random variables.

$$\boxed{\operatorname{Var}(X) \;=\; E\Big((X-\mu)^2\Big)\,.}$$

Alternatively (and again as seen previously for *discrete* random variables), the **variance** of a *continuous* random variable can be found by using the following formula (which is simpler to calculate).

$$\boxed{\operatorname{Var}(X) \;=\; E\left(X^2\right) - \mu^2.}$$

Let X be a continuous random variable with mean μ and probability density function f.

If we use the formulae from above, together with the property of expected value seen just prior to Example 8 of this chapter, we see that the variance of X can be found by calculating

$$\begin{aligned}\operatorname{Var}(X) &= E\Big((X-\mu)^2\Big)\\ &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx\,,\end{aligned}$$

or by calculating

$$\begin{aligned}\operatorname{Var}(X) &= E\left(X^2\right) - \mu^2\\ &= \left(\int_{-\infty}^{\infty} x^2 f(x)\,dx\right) \;-\; \mu^2.\end{aligned}$$

As was the case for discrete random variables, the **standard deviation** of a continuous random variable X is defined by

$$\boxed{\sigma \;=\; \sqrt{\operatorname{Var}(X)}\,.}$$

Of course, since σ is the *positive* square root of the variance, it follows immediately that $\sigma \geq 0$ and $\sigma^2 \;=\; \operatorname{Var}(X)$.

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

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$

Find the variance and standard deviation of X .

Solution:

This is the same random variable that we considered in Examples 6 and 8.

In those examples, we found that

$$E(X) = \frac{1}{2} \quad \text{and} \quad E\left(X^2\right) = \frac{1}{3}\,.$$

Using those results, we have

$$\begin{aligned} \mathrm{Var}(X) &= E\left(X^2\right) - \left[E(X)\right]^2 \\ &= \frac{1}{3} - \left(\frac{1}{2}\right)^2 \\ &= \frac{1}{3} - \frac{1}{4} \\ &= \frac{1}{12}\,, \end{aligned}$$

and therefore we have

$$\begin{aligned} \sigma &= \sqrt{\mathrm{Var}(X)} \\ &= \sqrt{\frac{1}{12}} \\ &= \frac{1}{2\sqrt{3}}\,. \end{aligned}$$

□

Properties of Variance and Standard Deviation

If X is a continuous random variable, and if a and b are constants, then

$$\boxed{\mathrm{Var}(aX+b) \;=\; a^2\,\mathrm{Var}(X).}$$

- This is the same result as was stated back in Chapter 20.3 (when we were learning about the variance of *discrete* random variables).

Taking the positive square root (to obtain the standard deviation of $aX+b$) gives

$$\begin{aligned}
\sigma_{aX+b} &= \sqrt{\mathrm{Var}(aX+b)} \\
&= \sqrt{a^2\,\mathrm{Var}(X)} \quad \text{(using the result from above)} \\
&= \sqrt{a^2}\,\sqrt{\mathrm{Var}(X)} \\
&= \big|a\big|\;\sigma_X\,.
\end{aligned}$$

Example 10. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{if } x > 1\,. \end{cases}$$

Find the variance and standard deviation of $-3X+1$.

Solution:

We have seen (in Example 9) that this random variable X has

$$\mathrm{Var}(X) \;=\; \frac{1}{12} \quad \text{and} \quad \sigma_X \;=\; \frac{1}{2\sqrt{3}}\,.$$

Then, using the property from above, we can quickly find $\mathrm{Var}(-3X+1)$.

We have

$$\begin{aligned}
\mathrm{Var}(-3X+1) &= (-3)^2\,\mathrm{Var}(X) \\
&= 9 \times \frac{1}{12} \\
&= \frac{3}{4}\,.
\end{aligned}$$

It immediately follows that
$$\begin{aligned}\sigma_{-3X+1} &= \sqrt{\operatorname{Var}(-3X+1)} \\ &= \sqrt{\frac{3}{4}} \\ &= \frac{\sqrt{3}}{2}\,.\end{aligned}$$

Alternatively, we can find σ_{-3X+1} as follows.
$$\begin{aligned}\sigma_{-3X+1} &= |-3|\,\sigma_X \\ &= 3 \times \frac{1}{2\sqrt{3}} \\ &= \frac{\sqrt{3}}{2} \quad \text{(the same answer as obtained previously).}\end{aligned}$$

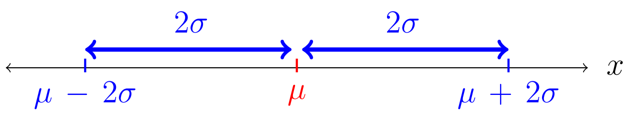
□

Recall (from Chapter 20.3) that the standard deviation and variance of a random variable measure

the *spread* of its distribution.

Just like in Chapter 20.3, this is illustrated by the following result:

Many continuous random variables have the property that
approximately 95% of the distribution lies within
two standard deviations of the mean.



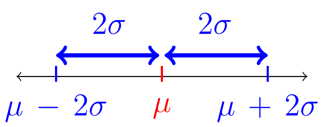
We can re-state this result in symbols as follows.

Many continuous random variables X satisfy

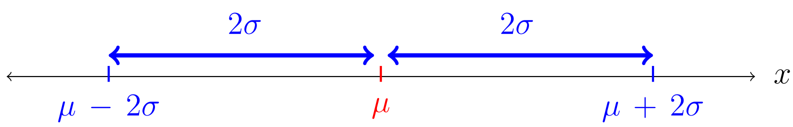
$$\boxed{\Pr\Big(\mu - 2\sigma \le X \le \mu + 2\sigma\Big) \approx 0.95\,.}$$

Using this result, we see that a larger standard deviation corresponds to a more spread-out distribution.

Smaller σ :



Larger σ :



Exercises for Sections 23.3 – 23.5

1. Consider the continuous random variable X with probability density function

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

   Find $E(X)$.

2. Let X be the lifetime (in years) of a light-globe manufactured by a company. Suppose that X has the probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ \dfrac{1}{7} e^{-\frac{x}{7}} & \text{if } x \ge 0\,. \end{cases}$$

   (a) Find the median lifetime of a light-globe. Answer in a sentence, and give your answer to 3 decimal places.

   (b) Find the value of p such that 90% of the light-globes last less than p years. Answer in a sentence, and give your answer to 3 decimal places.

3. Consider the continuous random variable X with probability density function

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ \dfrac{1}{4} x & \text{if } 0 \le x \le 2\sqrt{2} \\ 0 & \text{if } x > 2\sqrt{2}\,. \end{cases}$$

   (a) Find $E(X)$. (b) Find $E(X^2)$. (c) Find $\mathrm{Var}(X)$.

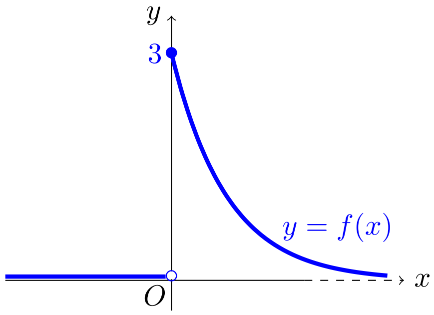
   (d) Find σ_X . (e) Find $E(3X + 1)$. (f) Find $\mathrm{Var}(3X + 1)$.

23.6 Answers for the Chapter 23 Exercises

23.2 1. (a) $\frac{1}{7}$ (b) $\frac{4}{7}$

2. (a) 0 (b) $\frac{11}{16}$

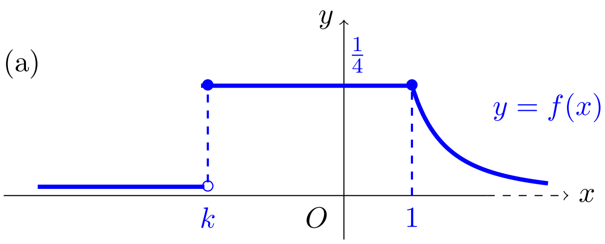
3. (a)



(b) e^{-3}

(c) e^{-2}

4. (a)



(b) $k = -2$

23.3–23.5 1. 1

2. (a) The median lifetime of a light-globe is 4.852 years (3 d.p.).

(b) Ninety percent of the light-globes last less than 16.118 years (3 d.p.).

3. (a) $\frac{4\sqrt{2}}{3}$ (b) 4 (c) $\frac{4}{9}$ (d) $\frac{2}{3}$ (e) $4\sqrt{2}+1$ (f) 4