# Concepts and notation for continuous random variables

# Continuous random variables

Continuous random variables are generally used to measure properties which are continuous in nature, e.g.,

- measurements of distance, like length, height, width
- measurements of time

Every continuous random variable $X$ has the following property:

$$P(X = a) = 0, \text{ for all values of } a.$$

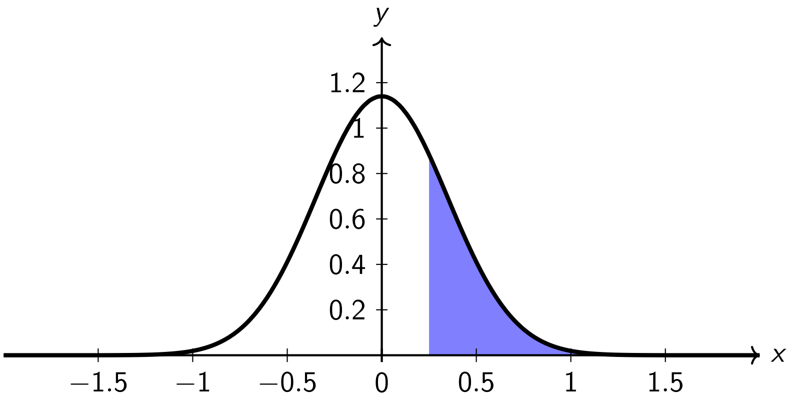
A lot of people have height *approximately* equal to 1.75m, but there is no chance that someone is exactly 1.7500000000000000...m.

So it is necessary to quantify continuous random variables by using cumulative probabilities:

$$P(X \leqslant a) \qquad P(X \geqslant a) \qquad P(a \leqslant X \leqslant b)$$

# Area under graphs

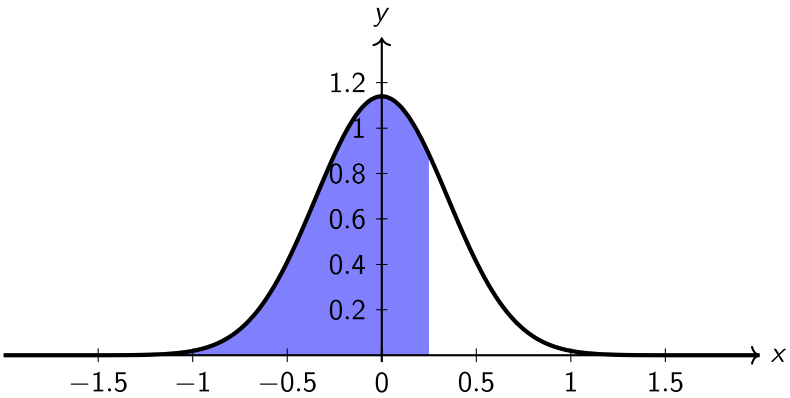
A key feature of continuous random variables is an understanding of intervals of area beneath certain graphs.




$P(X \geqslant 0.25)$

# Area calculations

Area calculations can be difficult. If straight lines are involved, we can just use knowledge of triangles/rectangles/etc. But for more detailed graphs, specialised techniques are required.




$P(X \leqslant 0.25)$

# Notation

If $f(x) \geqslant 0$ for all values of $x$, then:

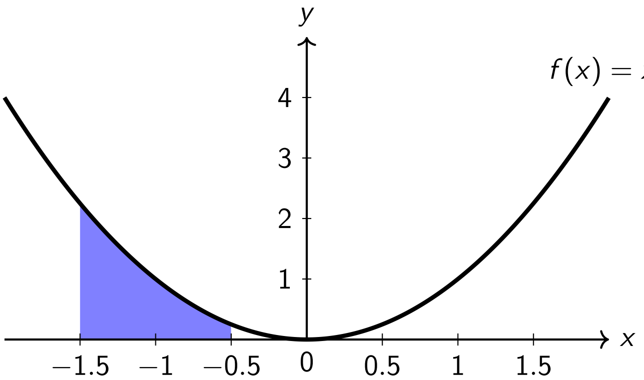
- $\int_a^b f(x)\,dx$ represents the area under the graph between $a$ and $b$ (as long as $a \leqslant b$).
- $\int_a^\infty f(x)\,dx$ represents the area under the graph from $a$ to $\infty$ (i.e. from $a$, all the way to the right).
- $\int_{-\infty}^b f(x)\,dx$ represents the area under the graph from $-\infty$ to $b$ (i.e. from all the way on the left up to $b$).

Each of these is called an *integral*.

Note: the stipulation $f(x) \geqslant 0$ is important.

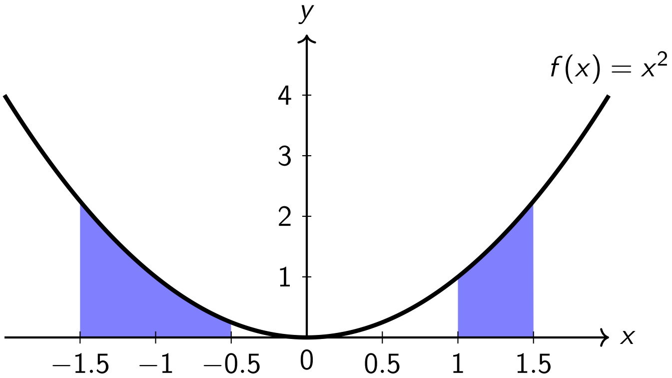
# Example

Express the shaded area using integral notation.



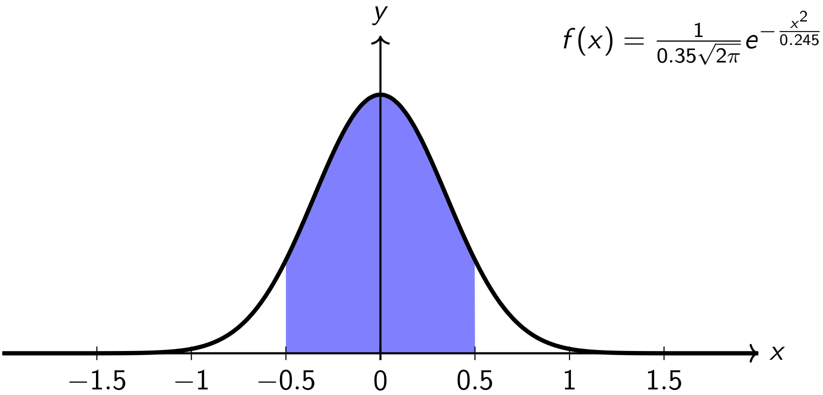
# Example

Express the shaded area using integral notation.



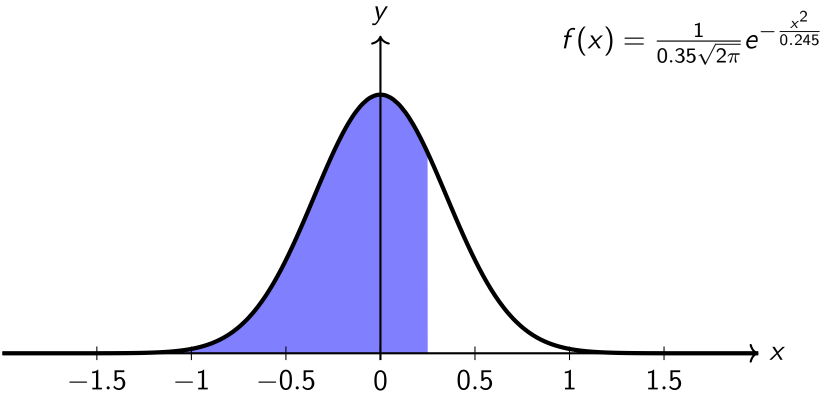
# Example

Express the shaded area using integral notation.



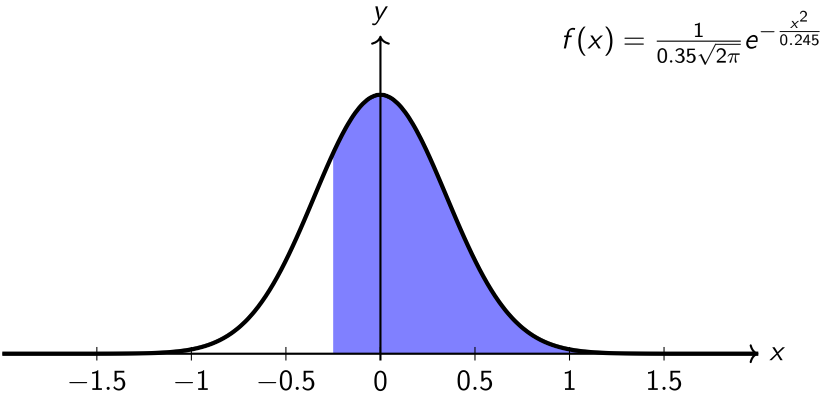
# Example

Express the shaded area using integral notation.



# Example

Express the shaded area using integral notation.



# Area calculations

Some areas can be determined exactly using calculus techniques. In STM4PSD, we will not use these techniques directly, although you may already be familiar with the ideas.

For example, it can be shown that

$$\int_a^b x^2 \, dx = \frac{1}{3}(b^3 - a^3)$$

On the other hand, for most values of $a$ and $b$, calculating the following integral exactly (using elementary functions) is mathematically impossible:

$$\int_a^b e^{-x^2} \, dx = ???$$

We will instead do calculations like this numerically using R.

# Probability density functions

# Probability density functions

Recall that a continuous random variable $X$ has the following property:

$$P(X = a) = 0, \text{ for all values of } a.$$

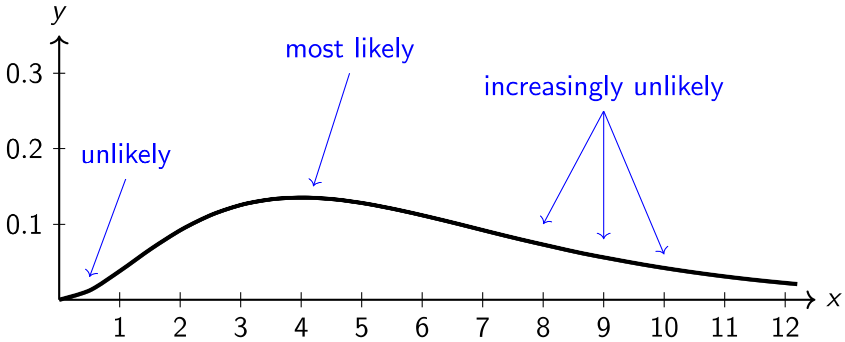
This means we cannot utilise the techniques we did for discrete random variables.

Probabilities for continuous random variables are defined by using a **probability density function**. Any function can be a probability density function so long as it meets both of the following conditions:

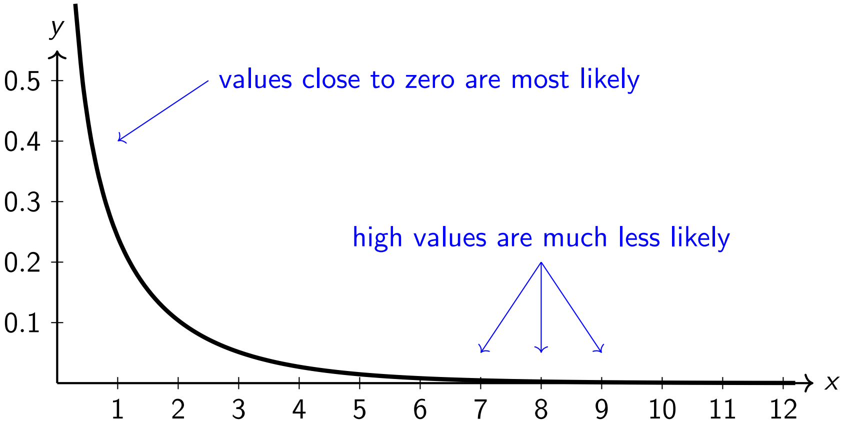
1. $f(x) \geqslant 0$ for all $x$
2. $\displaystyle\int_{-\infty}^{\infty} f(x)\, dx = 1$

# Density functions

Probability density functions measure relative likelihood:



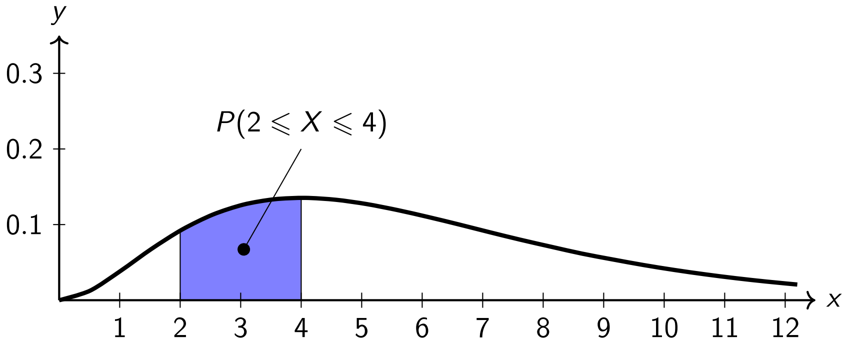
# Density functions




But probabilities can't be directly read from the graph.

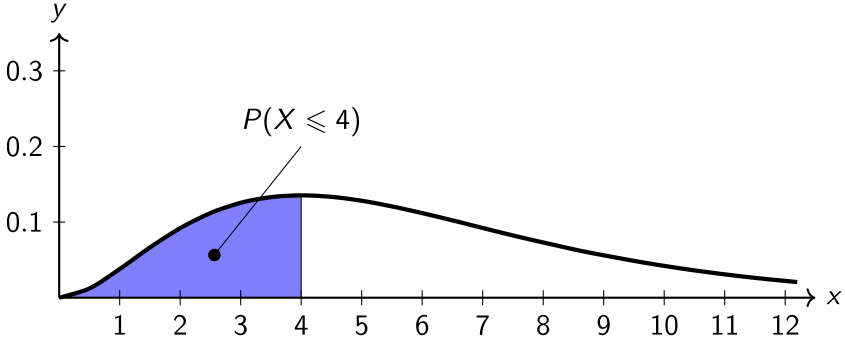
# Probability is area

The area under a graph is used to determine probabilities.



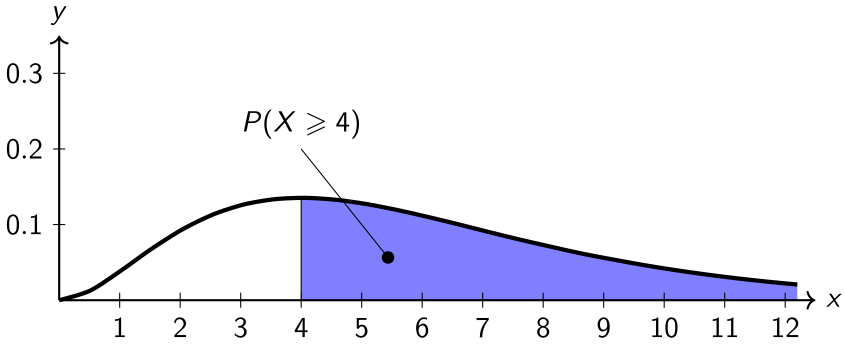
# Probability is area

The area under a graph is used to determine probabilities.



# Probability is area

The area under a graph is used to determine probabilities.



# Probability for continuous random variables?

If $X$ is a continuous random variable with density function $f_X(x)$, then

$$P(a \leqslant X \leqslant b) = \int_a^b f_X(x)\,dx$$

$$P(X \leqslant a) = \int_{-\infty}^a f_X(x)\,dx$$

And since $P(X = a) = 0$, a consequence of this is

$$P(X \leqslant a) = P(X < a).$$

Thus, for continuous random variables, strict and non-strict inequalities can be interchanged.

# Example

A random variable $X$ has probability density function $f_X(x)$ given by

$$f_X(x) = \begin{cases} \frac{2}{3}(x+1) & \text{if } -1 \leqslant x \leqslant 0 \\ \frac{1}{3}(2-x) & \text{if } 0 < x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

1. Sketch the graph of $f_X(x)$ and verify that $f_X(x)$ is a valid probability density function.
2. Calculate $P(X \leqslant 0)$, $P(X \leqslant 1)$ and $P(0 \leqslant X \leqslant 1)$.
3. Determine $P(X = 0)$, $P(X < 1)$ and $P(X > 1)$.
4. Determine a generic formula for $P(X \leqslant x)$.

# Probability distribution functions

# Probability distribution functions

A probability **distribution** function is different to a probability **density** function.

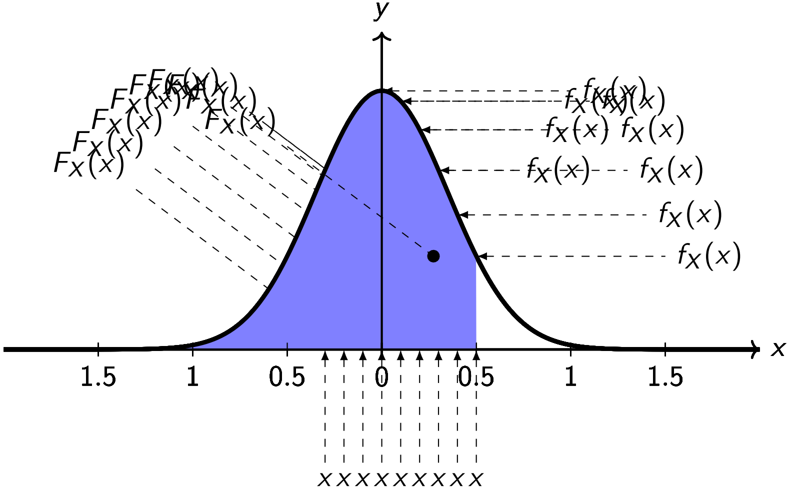
If $X$ is a random variable, then the probability distribution function is the function $F_X(x)$ defined by:

$$F_X(x) = P(X \leqslant x).$$

Therefore, the probability distribution function answers the question, "what is the probability that $X$ is less than or equal to a given value"?

It is also often known as a cumulative distribution function.

# Probability distribution functions



# Calculations

Remember the following:

- $P(X \geqslant a) = 1 - P(X < a)$
- $P(a \leqslant X \leqslant b) = P(X \leqslant b) - P(X < a)$

And if $X$ is continuous,

- $P(X < a) = P(X \leqslant a)$
- $P(X > a) = P(X \geqslant a)$

An important consequence of this is that, for a continuous variable, any probability can be determined using the cumulative distribution function.

# Example

Recall from the previous video:

A random variable $X$ has probability density function $f_X(x)$ given by

$$f_X(x) = \begin{cases} \frac{2}{3}(x+1) & \text{if } -1 \leqslant x \leqslant 0 \\ \frac{2}{3}(2-x) & \text{if } 0 \leqslant x \leqslant 2 \\ 0 & \text{otherwise} \end{cases}$$

4. Determine a generic formula for $P(X \leqslant x)$.

This is another way of saying, "determine the probability distribution function for $X$".

# Example

A random variable $X$ has probability density function $f_X(x)$ given by

$$f_X(x) = \begin{cases} \frac{1}{3}e^{\frac{x}{3}}, & \text{if } x \leqslant 0, \\ 0, & \text{otherwise} \end{cases}$$

You are given that the probability distribution function for $X$ is

$$F_X(x) = \begin{cases} e^{\frac{x}{3}}, & x \leqslant 0, \\ 1, & \text{otherwise.} \end{cases}$$

Use this to calculate the following:

1. $P(X < -1)$
2. $P(X > -1)$
3. $P(X > 1)$
4. $P(-2 < X < -1 \text{ or } X < -5)$

# Some continuous distributions

# Continuous uniform distribution

The simplest continuous probability distribution is the continuous uniform distribution. It takes on the same value over a given range, and is zero everywhere else. If $X$ is distributed according to a continuous uniform distribution, we write

$$X \sim U(a, b).$$

It corresponds to a situation that all outcomes between $a$ and $b$ are equally likely, and all other outcomes will not occur.

The probability density function is

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leqslant x \leqslant b, \\ 0 & \text{otherwise.} \end{cases}$$

The probability distribution function is

$$F_X(x) = \begin{cases} 0 & \text{if } x < a, \\ \frac{x-a}{b-a} & \text{if } a \leqslant x \leqslant b, \\ 1 & \text{otherwise} \end{cases}$$

# Example

Let $X \sim U(-3, 7)$.

1. Sketch the graph of the probability density function for $X$.
2. Sketch the graph of the probability distribution function for $X$.
3. Determine $P(X < 0)$, $P(X \geqslant 1)$ and $P(3 \leqslant X \leqslant 5)$.

# Normal distribution

An important probability distribution is the normal distribution. If $X$ is distributed according to a continuous uniform distribution with mean $\mu$ and standard deviation $\sigma$, we write

$$X \sim N(\mu, \sigma^2).$$

The probability density function is

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

There is no closed form expression for the probability distribution function, so numerical methods *must* be used.

The probability distribution function for a $N(\mu, \sigma^2)$ distribution is calculated in R using the `pnorm` function.

# Normal distribution

## The 68-95-99.7 rule

If $X \sim N(\mu, \sigma^2)$, then approximately 68%, 95% and 99.7% of values lie within one, two and three standard deviations of the mean, respectively. In symbols,

$$P(\mu - 1\sigma \leqslant X \leqslant \mu + 1\sigma) \approx 0.68$$
$$P(\mu - 2\sigma \leqslant X \leqslant \mu + 2\sigma) \approx 0.95$$
$$P(\mu - 3\sigma \leqslant X \leqslant \mu + 3\sigma) \approx 0.997$$

# Example

Give approximate sketches of the density function for a $N(\mu, \sigma^2)$ distribution with:

- $\mu = 2$, $\sigma = 1$
- $\mu = -2$, $\sigma = 1$
- $\mu = -2$, $\sigma = 3$