THE NORMAL DISTRIBUTION

What is the normal distribution and why does it pop-up everywhere?

Let's take a quick look at the definition.

Probability Density Function is given by

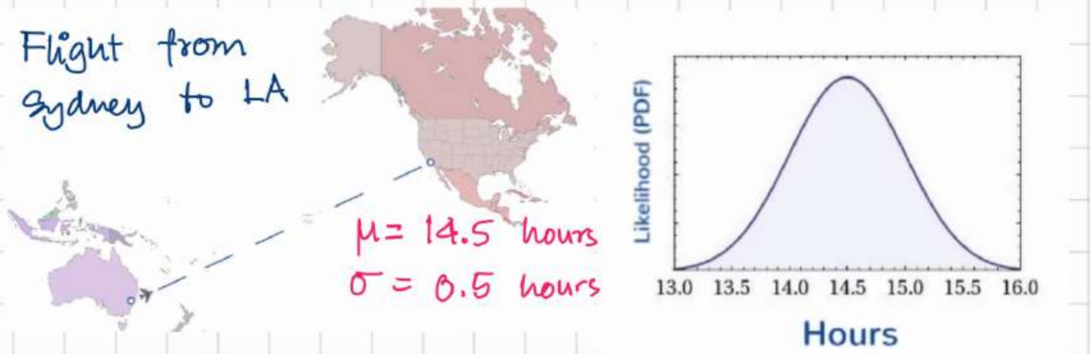
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

It needs two parameters μ and σ.

It defines the height and width of all bell curves that we see.

Question : Where does this bell curve pop up in real life?

Example 1:



It takes 14.5 hours generally but may be shorter or longer depending on other factors.

Example 2 : Heights in a classroom .

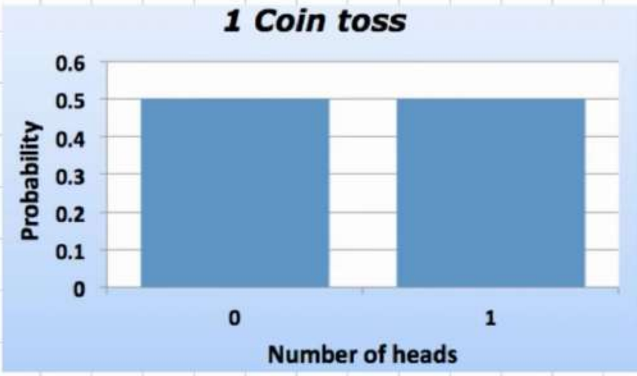
$\mu = 168$ cm
$\sigma = 10$ cm

Now it is obvious that flight times or heights are not always normally distributed . But in most cases it is safe to assume a normal distribution .

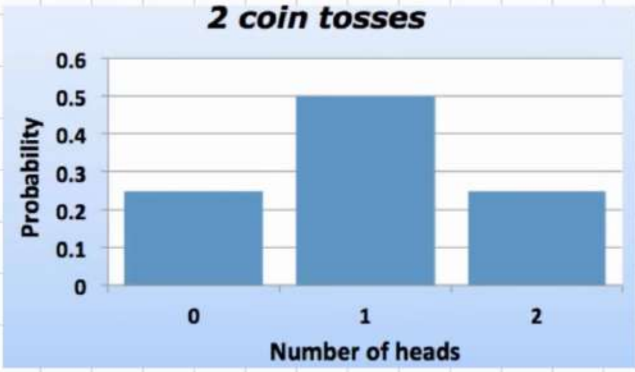
There is a statistical reason for that :

* Consider tossing a coin once .

there is 0.5 probability that there is no head and 0.5 probability that there is one head .



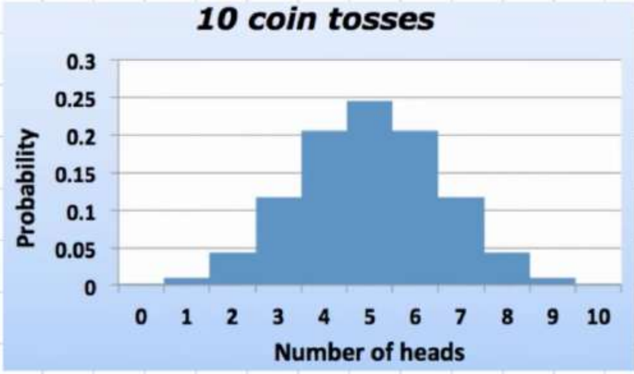
What about two coin tosses ?

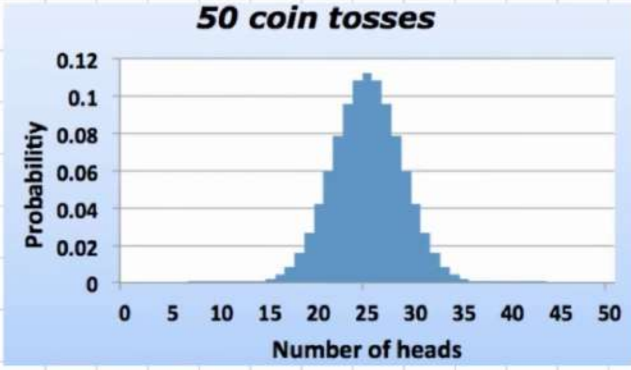


I can get upto 10 heads .

What about 10 coin tosses ?

I can get upto 10 heads.





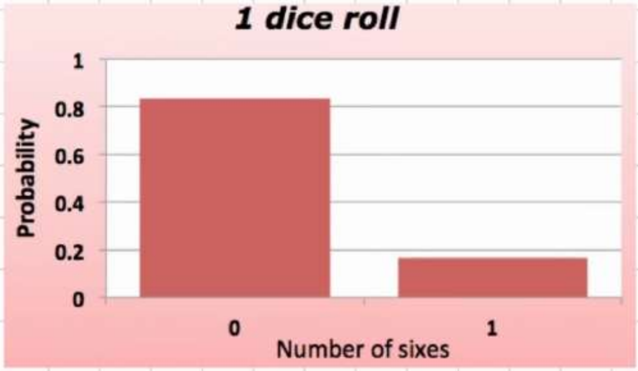
What about 50 tosses?

Does it look familiar ?

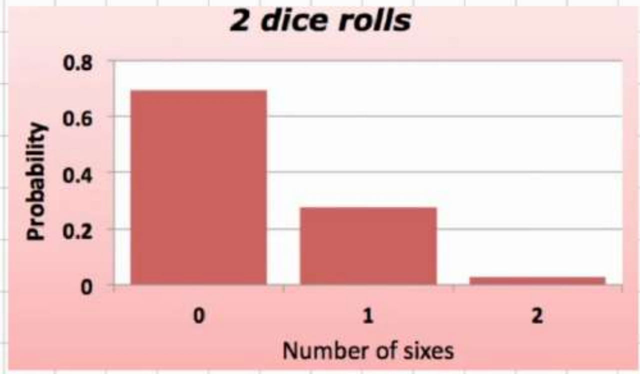
Let's take another example.

Consider the probability of rolling a 6 on one roll of a dice.

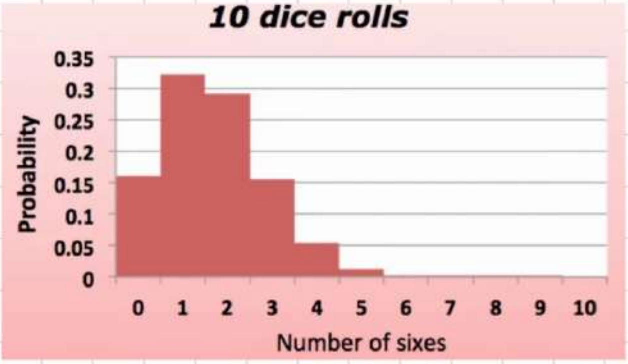
The distribution will look like this.

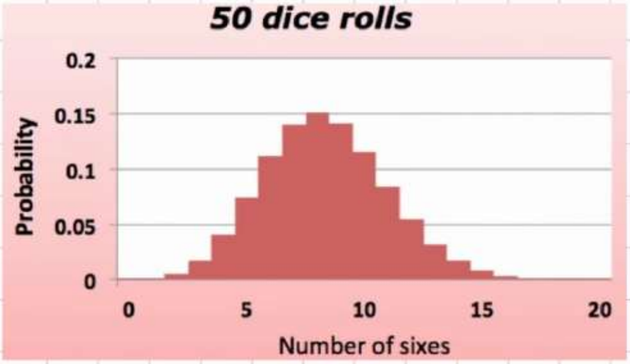


Two dice rolls..



Ten dice rolls.





50 dice rolls.

Does it look familiar now?

Although it didn't start out symmetrically like the previous example, it is a lot more symmetric now after 50 rolls.

Distributions approaching a normal distribution on increasing the no. of trials is an extremely useful property. We give it a name -

CENTRAL LIMIT THEOREM

This is the reason why the normal distribution is very prevalant in natural phenomena.

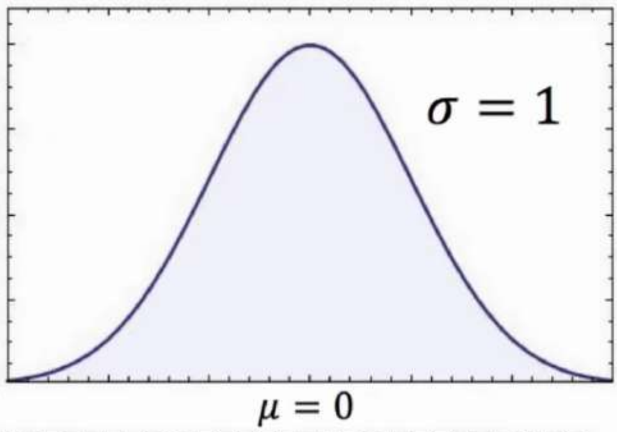
Simply stated the central limit theorem reads:

as n increases, the distribution of the sample mean or sum approaches a normal distribution

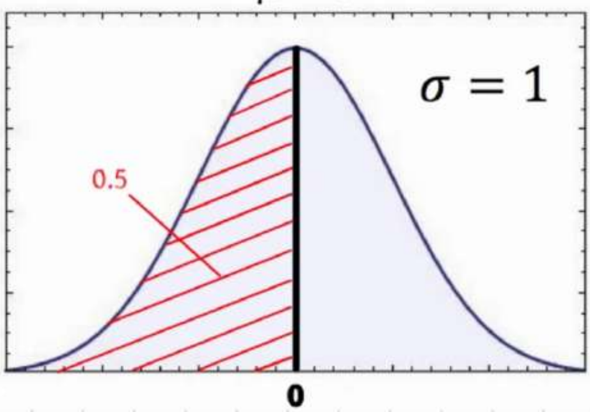
In these examples we have taken the sum, but if we take the mean, their distribution will also tend to a normal distribution with increase in 'n'.

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

STANDARD NORMAL DISTRIBUTION.

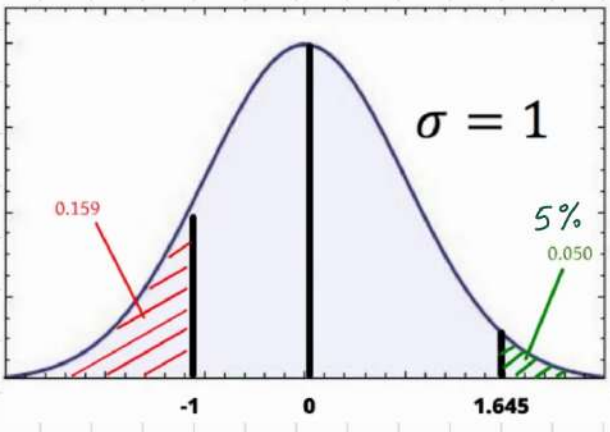


Standard normal distribution is a distribution whose mean is 0 and the std. deviation is 1.



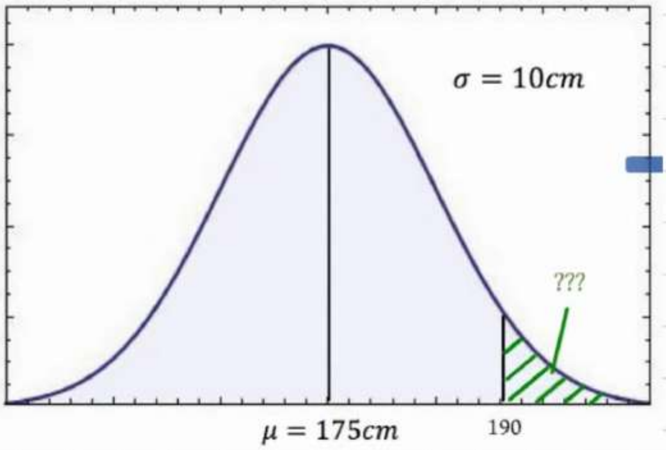
Looking at this distribution, we can tell that there is 50% probability of the given selection being less than 0.

We can make an analogy of the proportion of probability to the area of the curve.



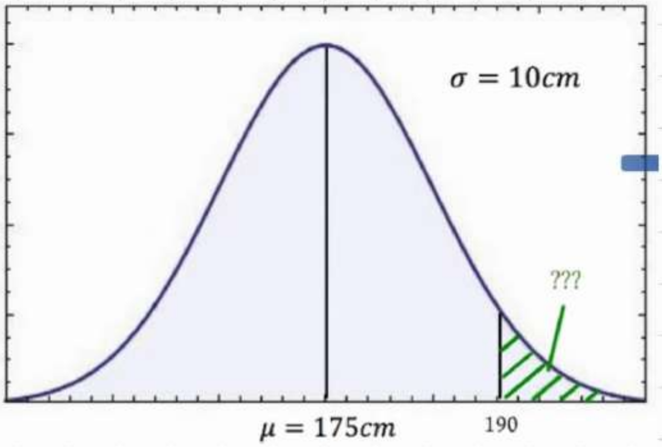
With this standard distribution we can also find more interesting probabilities such as:

For any other distibution we need to convert the standard normal distribution.



Let's take the class example. where avg. height was 175 cm and std. deviation was 10 cm.

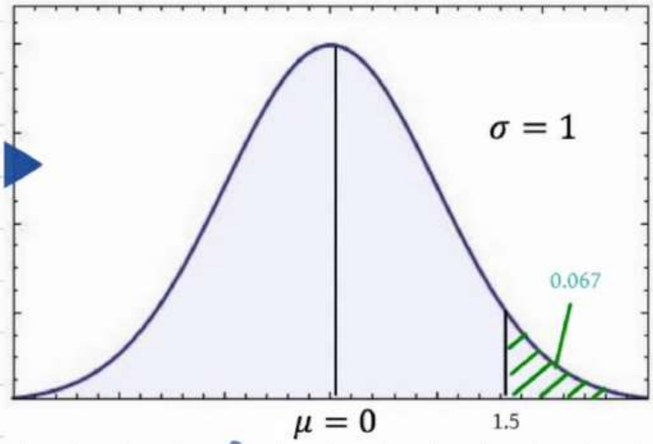
What is the probability of someone being more than 190 cm tall?



First, we are going to convert it to a point on the std. normal distribution.

$$Z = \frac{X - \mu}{\sigma} \qquad \text{Here, } Z = \frac{190-175}{10} = 1.5$$

In the standard normal distribution the Z score of 1.5 corresponds to a prob. of 0.067.



$\therefore \; P(Z > 1.5) = 0.067$

i.e. $P(\text{Height} > 190) = 0.067$