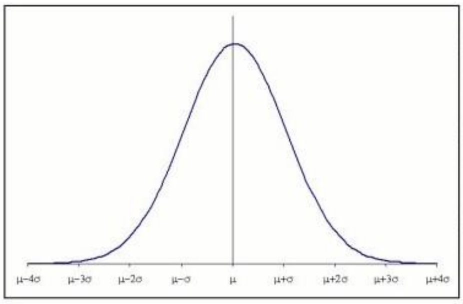
Statistics Sector 2: Normal Distribution

A normal distribution is a continuous function that has a high probability density close to the mean and decreases the further away you get from the mean, this is the most common continuous distribution used.

Examples of variables that are likely to follow a normal distribution are:
- Heights of male adults in the UK.
- The length of leaves from an elm tree.
- Time spent wait for a bus at a particular stop.

Properties of Normal Distribution



- Bell shaped
- Symmetrical about the mean (the mode, median and mean are equal)
- Total area under the curve from $-\infty$ to $+\infty$ is 1.
- Approximately two thirds (67%) of the values lie within one standard deviation of the mean ($\mu \pm \sigma$ or $\bar{x} \pm s$). There are **points of inflection** one standard deviation from the mean.
- Approximately 95% of the values lie within two standard deviations of the mean ($\mu \pm 2\sigma$ or $\bar{x} \pm 2s$).
- Approximately 99.8% of the values lie within three standard deviations of the mean ($\mu \pm 3\sigma$ or $\bar{x} \pm 3s$).

These properties can be used to determine whether or not data is likely to follow a normal distribution.

<u>Example 1</u>

A random sample of 50 garden centres was selected as part of a survey on the use pesticides. The annual expenditure in pounds, x, on pesticides was recorded and the results are summarised below.

$$\sum x = 3212.0 \qquad \sum(x - \bar{x})^2 = 54302.45$$

a) Calculate the values for $\bar{x}$ and s, where s^2 denotes the unbiased estimator of σ^2.

b) Hence show why the annual expenditure, X, on pesticides is unlikely to be normally distributed. Give numerical support for your answer.

Example 2

The time taken, T minutes, for phone calls to be answered in a call centre is recorded for a random sample of 80 calls. Salvador was asked to summarise the results, they were as follows.

Mean, $\bar{t} = 4.32$ Variance (unbiased estimate), $s^2 = 20.4$

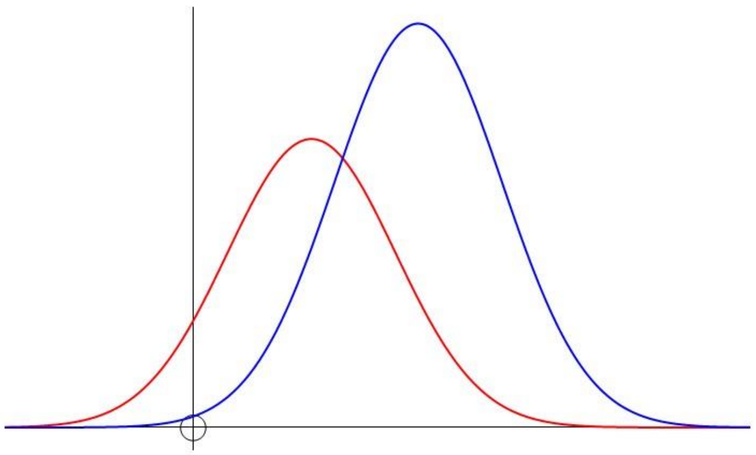
Salvador states that, from his summarised results, T was not normally distributed. Explain, using the value of $\bar{t} - 2s$, why Salvador's conclusion that the data is not normally distributed is likely to be correct.

A normal distribution has two parameters, μ and σ^2, the mean and variance respectively. The distribution is written as

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

This is read as the random variable, X, has a normal distribution with mean μ and variance σ^2.

Changes in the value of μ change the location of the curve on the x axis. Changes in the value of σ change the shape of the curve however the curve will always remain bell-shaped and symmetrical.



Which curve has a larger mean?

Which curve has a larger standard deviation?

Example 3

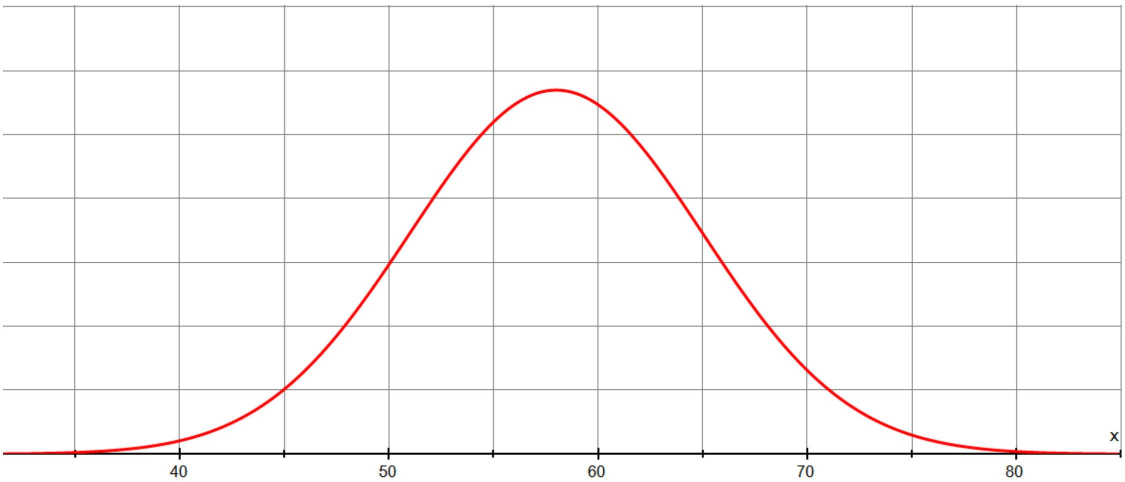
It is known that the heights of females of Chinese origin are normally distributed with a mean of 155.8 cm and a standard deviation of 7.11 cm. It is also known that the heights of males of Chinese origin are also normally distributed with a mean of 167.1 cm and a standard deviation of 7.42.

a) Compare the means and standard deviations of the two populations in context.

b) Use the properties of the normal distribution to sketch both normal distribution curves on the same axes.

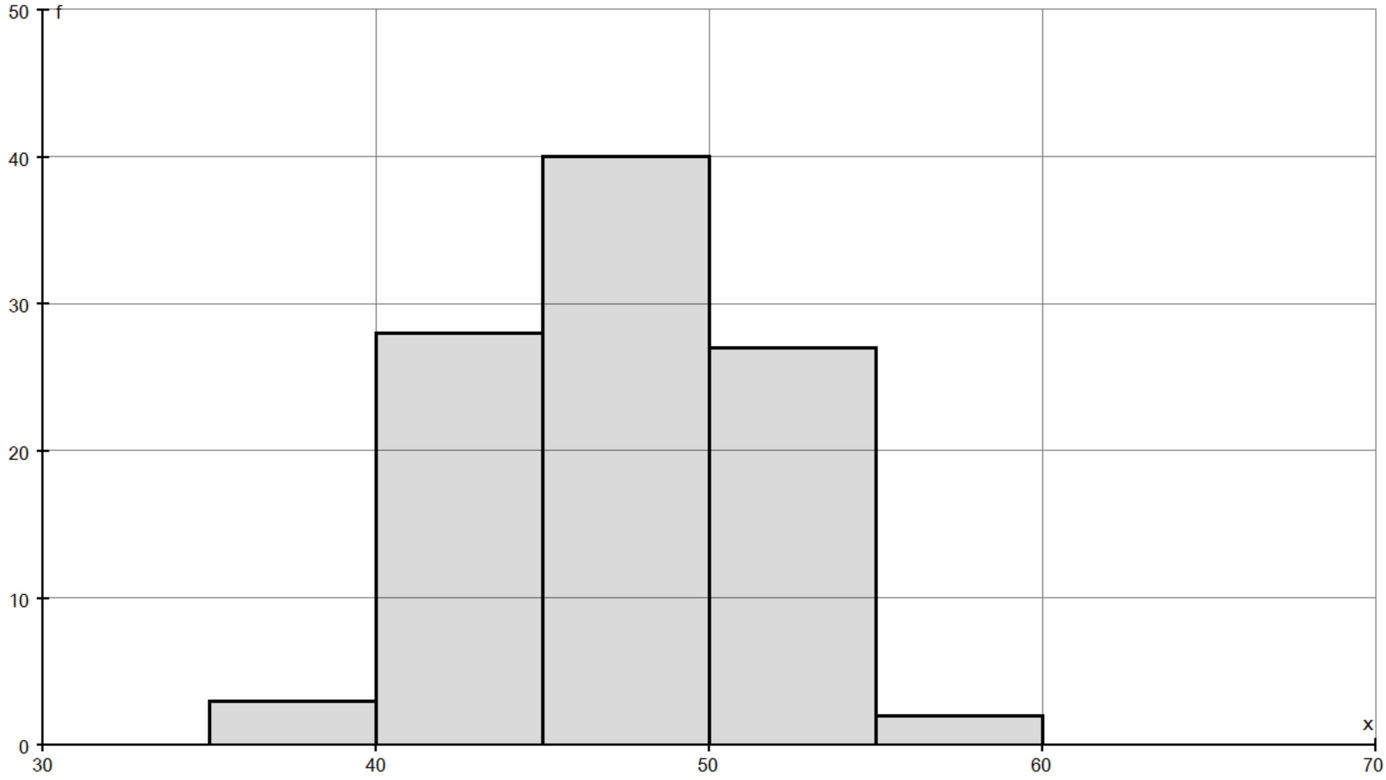
Example 4

The time taken by students to complete a test, x minutes, are normally distributed. The data is shown on the diagram below, estimate the mean and standard deviation.



Example 5

The histogram shows the speeds of 100 randomly selected cars on a stretch of road.

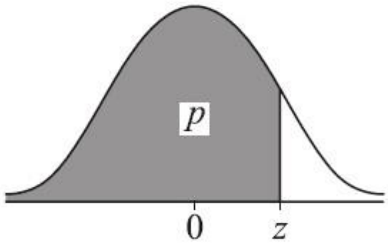


a) Explain why it would be reasonable to use a normal distribution to model the cars' speeds.

b) Estimate the mean and standard deviation of the speeds.

c) Hence find the points of inflection.

Calculating Probabilities

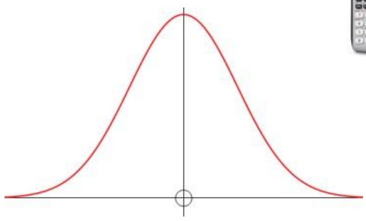


Probabilities for the normal distribution are found by calculating the required area under the curve, we use graphical calculators to find these probabilities. Since Z is a continuous random variable $(Z = z) = 0$. Also $(Z < z) = (Z \leq z)$ and $P(Z > z) = P(Z \geq z)$.

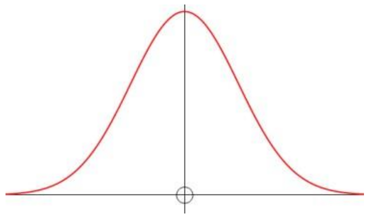
<u>Example 6</u>

Given that $Z \sim (0,1^2)$, determined:

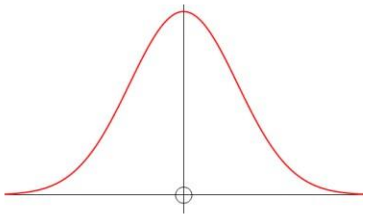
a) $P(Z < 1.25)$



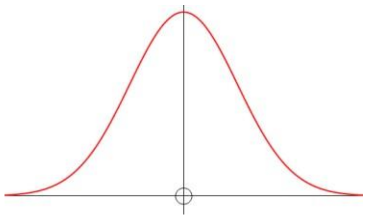
b) $P(Z > 0.76)$



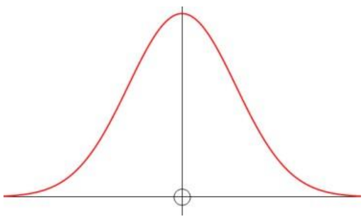
c) $(Z < -0.65)$



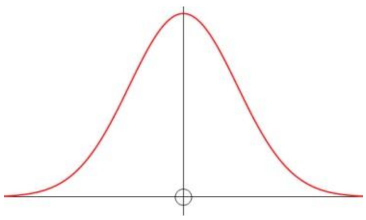
d) $(Z > -1.49)$



e) $P(0.42 < Z < 1.32)$



f) $P(-1.84 < Z < 0.95)$



Standardising

It is unlikely that a distribution will follow a standard normal distribution with $\mu = 0$ and $\sigma^2 = 1$. We can standardise the given values so that we are able to use standard normal tables.

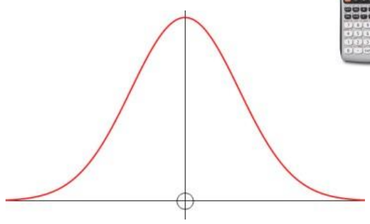
If $X\sim(\mu, \sigma^2)$ then we can standardise the values of X to the values of $Z\sim N(0, 1^2)$ by:

$$\boxed{\mathrm{Z} = \frac{X - \mu}{\sigma}}$$

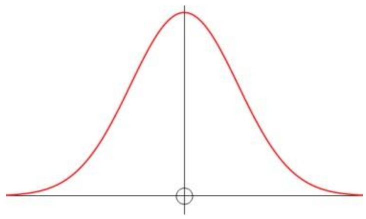
<u>Example 7</u>

The weight of salt, delivered by a machine into cardboard tubs may be assumed to be a normal random variable with mean 355 grams and standard deviation 5.2 grams.

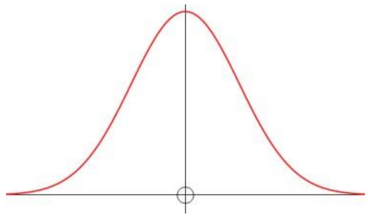
a) Determine the probability that the weight of salt in a randomly selected tub is:



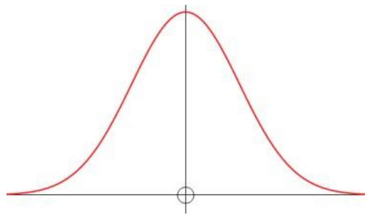
i) More than 352 grams



ii) Less than 345 grams



iii) Between 354 grams and 358 grams

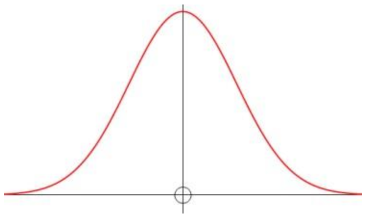


iv) Within one standard deviation from the mean

b) A customer randomly selects 8 tubs of salt. Calculate the probability that all 8 tubs selected contain more than 352g of salt.

Example 8

The plums from a particular variety of plum tree have masses which can be modelled by a normal distribution with mean 24 g and variance 25 g. Plums weighing more than 28 g are graded as large. What proportion of the plums is graded as large?



Exam Questions

3 Shower-cleaner liquid is sold in spray bottles. The volume of liquid in a bottle may be modelled by a normal distribution with mean 955 millilitres and standard deviation 5 millilitres.

(a) Determine the probability that the volume of liquid in a particular bottle is:

(i) at most 960 millilitres;

(ii) more than 946 millilitres;

(iii) exactly 950 millilitres;

(iv) between 946 millilitres and 960 millilitres.

[7 marks]

7. The volume of liquid in bottles of sparkling water from one producer is believed to be normally distributed with a mean of 704 ml and a variance of 3.2 ml^2.

Calculate the probability that a randomly chosen bottle from this producer contains

(a) more than 706 ml, **(3 marks)**

(b) between 703 and 708 ml. **(4 marks)**

The bottles are labelled as containing 700 ml.

(c) In a delivery of 1 200 bottles, how many could be expected to contain less than the stated 700 ml?

(4 marks)

Inverse Normal Distribution

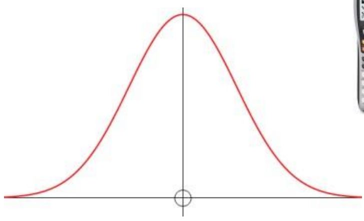
Sometimes given a probability we need to find the corresponding z value such that $(Z < z) = p$ where $Z \sim (0, 1^2)$. We use inverse normal on the calculator to work backwards to find z, μ or σ.

Example 9

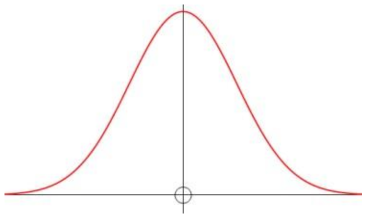
The random variable Z has a normal distribution with mean zero and standard deviation 1. Find the value of z that satisfies the following probabilities.

Given that $Z \sim (0,1^2)$, determined:

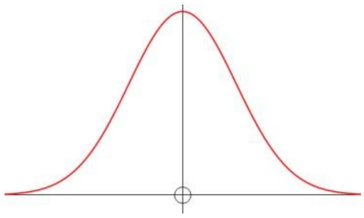
i) $(Z < z) = 0.76$



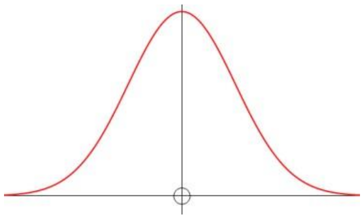
ii) $(Z > z) = 0.39$



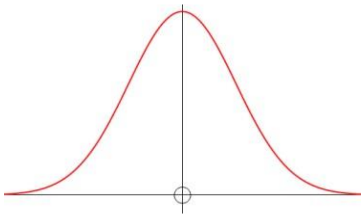
iii) $(Z > z) = 0.962$



iv) $(Z < z) = 0.035$



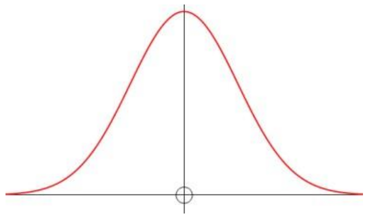
v) $(-z < Z < z) = 0.98$



Example 10

Bags of sugar packed by a machine have masses which can be modelled by a normal distribution with mean 1000 g and standard deviation 4 g, if the machine is working correctly.

If the machine is working correctly, 1% of the bags are rejected because they are underweight. Calculate the minimum acceptable mass of a bag of sugar.

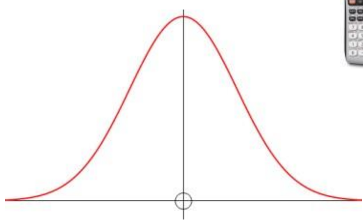


Example 11

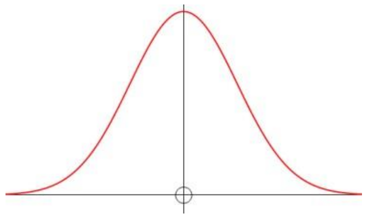
Louisa is employed to fit aerials in new houses being built by a developer. Her time, X minutes, to install an aerial can be assumed to be normally distributed with a mean of 32 and a standard deviation 14.

Determine:

a) The time, k minutes, such that $(X < k) = 0.77$.



b) Find the time exceeded on 90% of her jobs.



Exam Question

5. Police measure the speed of cars passing a particular point on a motorway.
The random variable X is the speed of a car.

X is modelled by a normal distribution with mean 55 mph (miles per hour).

(a) Draw a sketch to illustrate the distribution of X. Label the mean on your sketch.

(2)

The speed limit on the motorway is 70 mph. Car drivers can choose to travel faster than the speed limit but risk being caught by the police.

The distribution of X has a standard deviation of 20 mph.

(b) Find the percentage of cars that are travelling faster than the speed limit.

(3)

The fastest 1% of car drivers will be banned from driving.

(c) Show that the lowest speed, correct to 3 significant figures, for a car driver to be banned is 102 mph. Show your working clearly.

(3)

Car drivers will just be given a caution if they are travelling at a speed m such that

$$P(70 < X < m) = 0.1315$$

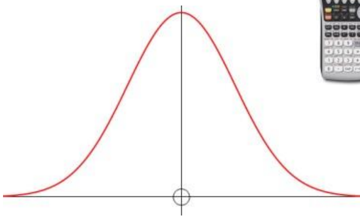
(d) Find the value of m. Show your working clearly.

(4)

Finding Missing Parameters

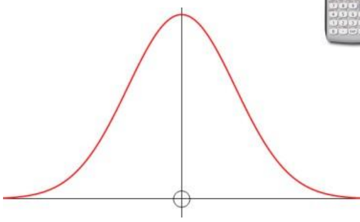
Example 12

The weight, X grams, of tins of floor cleaner may be assumed to be normally distributed with mean 245 and standard deviation σ. Assuming the mean remains unchanged; determine the value of σ necessary to ensure that 96% of the tins contain more than 238 grams of floor cleaner.



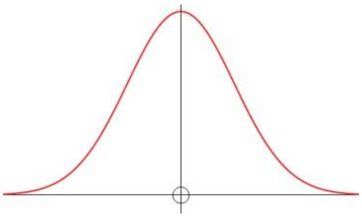
Example 13

The lengths of rods produced in a workshop have a mean μ and standard deviation 2. 10% of rods are less than 17.4cm. Find the value of μ.



Example 14

Two gauges are used to test the thickness of panes of glass. Over a long period it is found that 1.2% of the panes pass through the 1.4 mm gauge but 95.4% pass through the 1.6 mm gauge. Assuming that the thickness of the glass can be modelled by a normal distribution, find, to the nearest 0.01 mm, the mean and the standard deviation of the thickness.



Exam Questions

2 (a) Tim rings the church bell in his village every Sunday morning. The time that he spends ringing the bell may be modelled by a normal distribution with mean 7.5 minutes and standard deviation 1.6 minutes.

Determine the probability that, on a particular Sunday morning, the time that Tim spends ringing the bell is:

(i) at most 10 minutes;

(ii) more than 6 minutes;

(iii) between 5 minutes and 10 minutes.

[6 marks]

(b) June rings the same church bell for weekday weddings. The time that she spends, in minutes, ringing the bell may be modelled by the distribution $\mathrm{N}(\mu, 2.4^2)$.

Given that 80 per cent of the times that she spends ringing the bell are less than 15 minutes, find the value of μ.

[4 marks]

6. The weight, in grams, of beans in a tin is normally distributed with mean μ and standard deviation 7.8

Given that 10% of tins contain less than 200 g, find

(a) the value of μ

(3)

(b) the percentage of tins that contain more than 225 g of beans.

(3)

The machine settings are adjusted so that the weight, in grams, of beans in a tin is normally distributed with mean 205 and standard deviation σ.

(c) Given that 98% of tins contain between 200 g and 210 g find the value of σ.

(4)

Extension

6. A geologist is analysing the size of quartz crystals in a sample of granite. She estimates that the longest diameter of 75% of the crystals is greater than 2 mm, but only 10% of the crystals have a longest diameter of more than 6 mm.

The geologist believes that the distribution of the longest diameters of the quartz crystals can be modelled by a normal distribution.

(a) Find the mean and variance of this normal distribution. **(9 marks)**

The geologist also estimated that only 2% of the longest diameters were smaller than 1 mm.

(b) Calculate the corresponding percentage that would be predicted by a normal distribution with the parameters you calculated in part *(a)*.

(3 marks)

(c) Hence, comment on the suitability of the normal distribution as a model in this situation.

(2 marks)

7. The heights of adult females are normally distributed with mean 160 cm and standard deviation 8 cm.

(a) Find the probability that a randomly selected adult female has a height greater than 170 cm.

(3)

Any adult female whose height is greater than 170 cm is defined as tall.

An adult female is chosen at random. Given that she is tall,

(b) find the probability that she has a height greater than 180 cm.

(4)

Half of tall adult females have a height greater than h cm.

(c) Find the value of h.

(5)