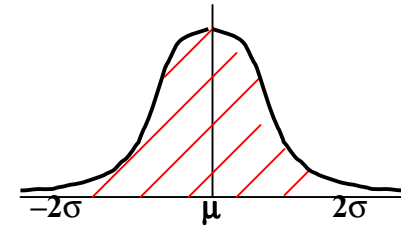


The Normal Distribution

The normal distribution is used as a model for continuous data such as height, weight, time etc. The probability cannot be set out in a table as there are an infinite number of values the variable can take. Instead, it is given by a formula and an associated sketch graph.

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



All normal distributions are symmetric and have bell-shaped probability density curves with a single peak. The distribution is defined by its mean μ {mew} and its standard deviation σ {sigma}.

Almost all the graph lies within 2 standard deviations of the mean. (About 95%)

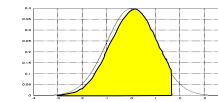
The area under the graph is 1.

So if a particular type of fruit has mean weight 150g and standard deviation 10g. The variance would be $\sigma^2 = 10^2$.

X is the random variable **weight of fruit**. We write: $X \sim N(150, 10^2)$ putting the **mean** and **standard deviation squared** in the bracket.

We now wish to evaluate a probability:

The probability that a fruit picked at random has a weight under 165g.



$\Pr(x \leq 165)$: One way to find this probability would be to find the area to the left of 165 by integration.

A better way would be to standardize the weight of 165 by subtracting the mean and dividing by the standard deviation, thereby converting to a z – value and looking up this probability in Standardized Normal Tables.

The graph connected to the tables would be the Standardized Normal graph and it would have a mean of zero with a standard deviation of 1.

$$z = \frac{165 - 150}{10} = 1.5$$

We have converted the value 165 into standard deviation units –

165 is 1.5 standard deviations to the right of the mean. The tables would give us $\Pr(z \leq 1.5) = 0.9332$

If we want the probability of a fruit weighing **over** 165g. we look at the other side of the graph. $\Pr(z \geq 1.5) = 1 - 0.9332 = 0.0668$

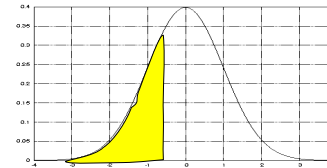
Mean μ unknown.



So, if a particular type of berry has unknown mean weight and known standard deviation 10g. We need some further information before we can find the unknown mean.

If the question tells us that 60 out of 200 of these berries have been found to have a weight less than 100 grams, we translate the information to say that 30% have weight less than 100 grams or $\Pr(x < 100) = 0.3$

Standardising: $\Pr(z < \frac{100 - \mu}{10}) = 0.3$



But from tables (looking up tables backwards) $\Pr(z < -0.53) = 0.3$ { since $\Pr(z < 0.53) = 0.7$ }

The mean can be found by equating $\frac{100 - \mu}{10}$ to -0.53 . $\frac{100 - \mu}{10} = -0.53$

$$\underline{\underline{\mu = 100 + 10 \times (-0.53) = 105.3}}$$

Standard deviation σ unknown.

Suppose the **mean** weight of a particular type of berry is known to be **100g** with unknown standard deviation σ . It is also known that 40% of a sample of berries have a weight of over 120g.

$\Pr(x > 120) = 0.4$ i.e. $\Pr(z > \frac{120 - 100}{\sigma}) = 0.4$ But from tables $\Pr(z > 0.26) = 0.4$

$$\frac{120 - 100}{\sigma} = 0.26$$

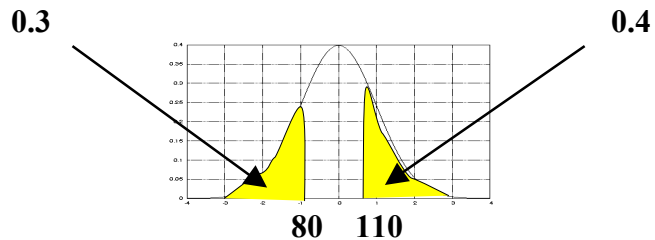
$$\sigma = 77 \text{ (to the nearest whole number).}$$

Mean μ and standard deviation σ unknown.

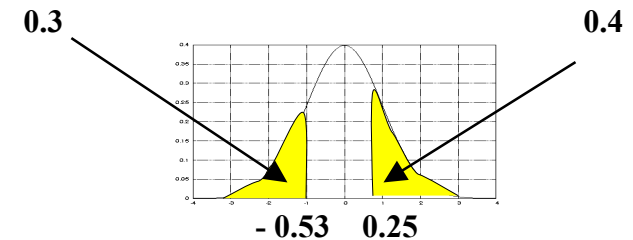
We need some further information before we can find the unknown mean and standard deviation. In fact, we need **two** pieces of information with **two** unknowns just like in simultaneous equations.

Suppose that it is known that 30% of the berries have weight under 80g and 40% have weight over 110g.

Sketch the graph of $N(\mu, \sigma^2)$



and the corresponding graph of $N(0, 1)$



Since the graph on the right gives the values in standard deviation units, we can make two statements:

$$80 = \mu - 0.53\sigma$$

$$110 = \mu + 0.25\sigma$$

Solving these simultaneously by subtracting the top equation from the bottom gives: $30 = 0.78\sigma$

$\sigma = 38.5$ and after substituting: $\mu = 100.4$ The mean is 100.4 and the standard deviation is 38.5.

CHECK: Find the value of the weight such that 30% of berries have a weight less than this value.

Let the value be k , then:

$$\Pr(x < k) = 0.3$$

Standardize:

$$\Pr\left(z < \frac{k - 100.4}{38.5}\right) = 0.3$$

But from tables:

$$\Pr(z < -0.53) = 0.3$$

Equating:

$$\frac{k - 100.4}{38.5} = -0.53$$

$$K = 100.4 - 0.53 \times 38.5 = 80 \text{ (as stated originally).}$$