Normal Distribution

Colby Community College

Example 1

Consider making popcorn.

You put some oil and corn kernels in a pan and start heating.

For the first few minutes nothing happens, then a few kernels start to pop.

A little while later more and more start to pop.

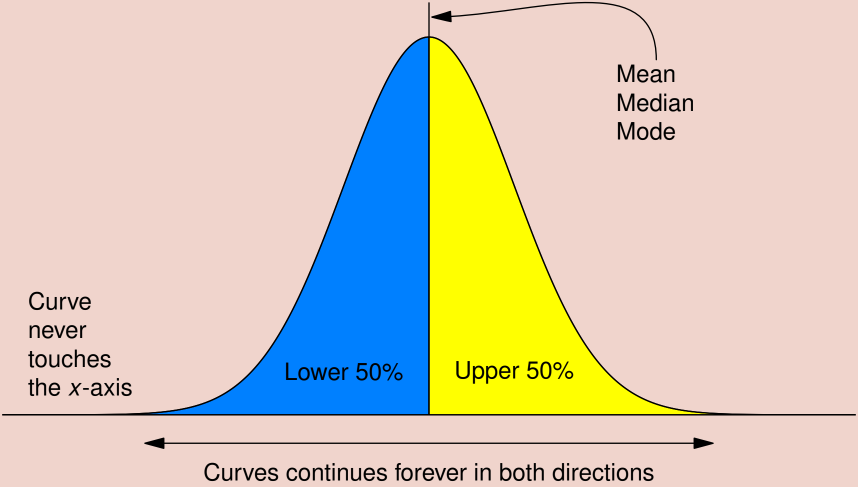
This goes one for a minute or so, and the popping gradually tappers off.

Most of the popping happens in that brief, noisy moment.

This demonstrates a typical pattern that is part of many phenomena.

Definition

A **normal distribution** is a perfectly symmetric, bell-shaped distribution. It is also referred to as a **normal curve** or a **bell curve**.

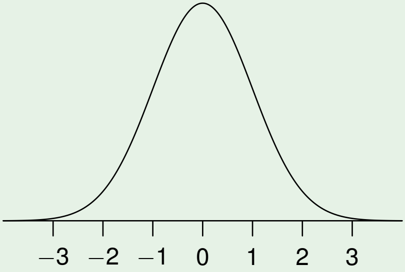


Note

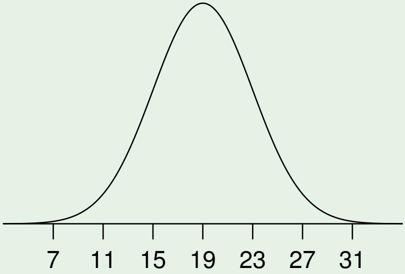
The normal distribution with mean μ and standard distribution σ is denoted $N(\mu, \sigma)$, where μ and σ are called the parameters.

Example 2

Both are normal distributions, but with different center and spread.



(a) $N(\mu = 0, \sigma = 1)$



(b) $N(\mu = 19, \sigma = 4)$

Definition

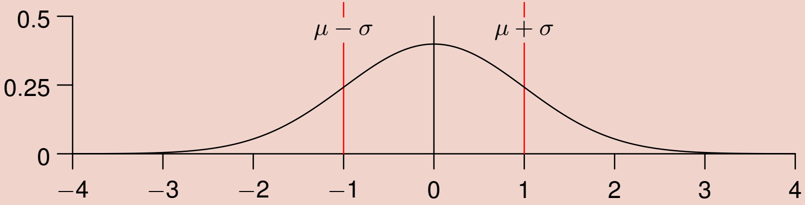
The graph of any continuous probability distribution is called a **density curve** if the total area under the curve is exactly 1.

Note

This means there is a correspondence between the area under a density curve and probabilities.

Definition

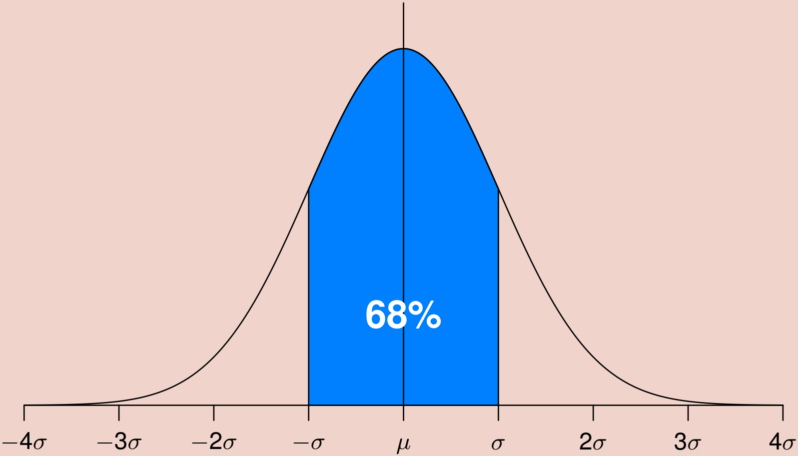
The special case $N(\mu = 0, \sigma = 1)$ is called the **standard normal distribution.** The total area under the curve is exactly equal to 1.



Definition

The **empirical rule**, or **68-95-99.7 rule**, states approximately how much of the area is contained when stepping one, two, or three standard deviations from the mean.

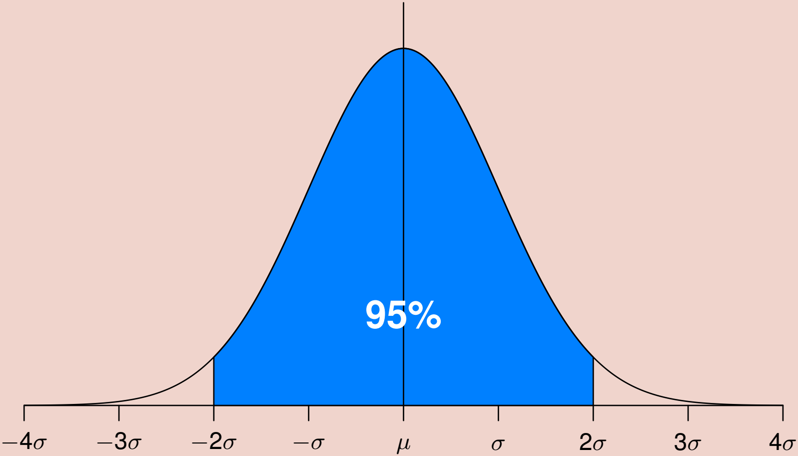
One standard deviation from the mean.



Definition

The **empirical rule**, or **68-95-99.7 rule**, states approximately how much of the area is contained when stepping one, two, or three standard deviations from the mean.

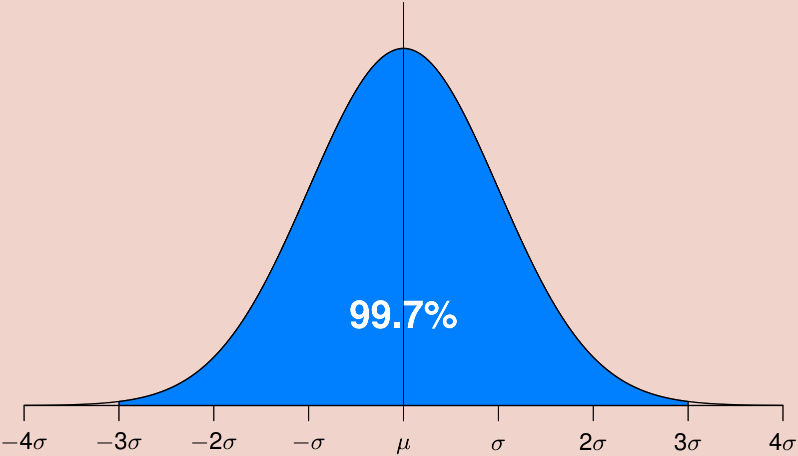
Two standard deviations from the mean.



Definition

The **empirical rule**, or **68-95-99.7 rule**, states approximately how much of the area is contained when stepping one, two, or three standard deviations from the mean.

Three standard deviations from the mean.



Definition

A ***z*-score** is a measure of the number of standard deviations a particular data point is away from the mean.

$$z = \frac{\text{(data point)} - \text{(mean)}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

Example 3

On a college entrance exam, the mean was 70, and the standard deviation was 8. Rose scored a 85, what is her z-score?

$$z = \frac{x - \mu}{\sigma} = \frac{85 - 70}{8} \approx 1.875$$

Example 4

On the same exam, George has a z-score of -1.3. What was his score?

$$z = \frac{x - \mu}{\sigma} \quad \Rightarrow \quad z\sigma = x - \mu \quad \Rightarrow \quad x = z\sigma + \mu = (-1.3)(8) + 70 = 59.6$$

Example 5

The mean on a exam was 82, with a standard deviation of 7 points. An "A" on the exam is a 93, what is the z-score?

$$z = \frac{x - \mu}{\sigma} = \frac{93 - 82}{7} \approx 1.57$$

Note

We know from the empirical rule that roughly 68% of the scores fall within one standard deviation of the mean.

This means that 68% of the students scored between 75 and 89.

Moreover, we know that roughly 95% of the scores fall within two standard deviations of the mean.

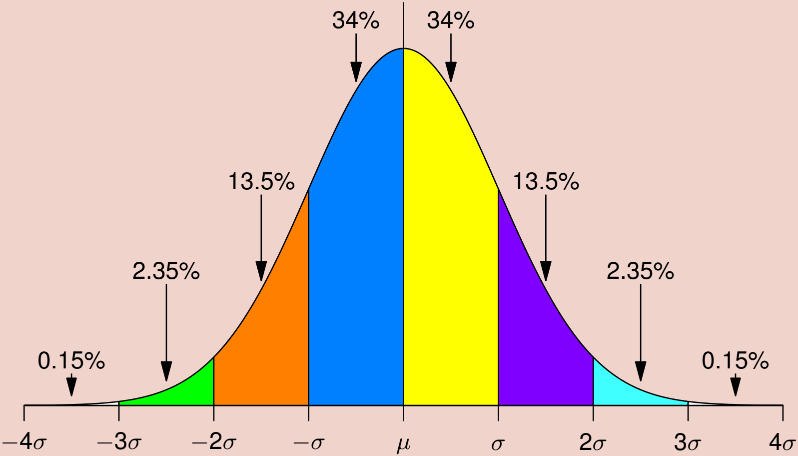
Which means that $95\% - 68\% = 27\%$ of the scores are more than one standard deviation from the mean, but less than two.

Since the curve is symmetric, we know that 13.5% of the students scored between 89 and 96, as well as 13.5% between 68 and 75

Definition

The **empirical rule**, or **68-95-99.7 rule**, states approximately how much of the area is contained when stepping one, two, or three standard deviations from the mean.

For each standard deviation.

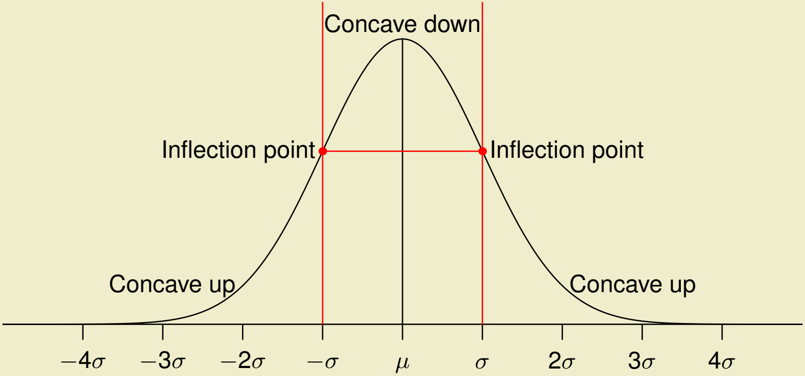


Definition

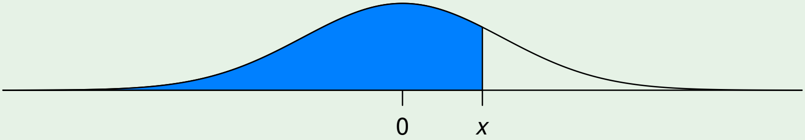
An **inflection point** is where a curve changes from being concave up to concave down, or vice versa

Note

A normal density curve always has two inflection points, each one standard deviation from the mean.



Example 6

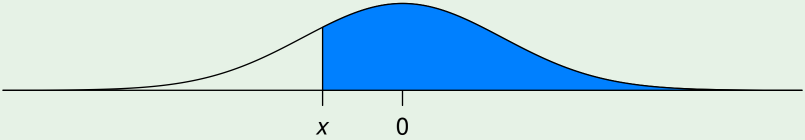


The area of the shaded region is the probability that a z score is less than or equal to x, $P(z \leq x)$.

Note

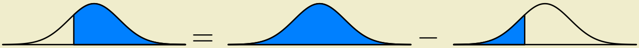
Most statistical software, programming languages, spreadsheets programs, and calculators are able to calculate the area for you.

Example 7



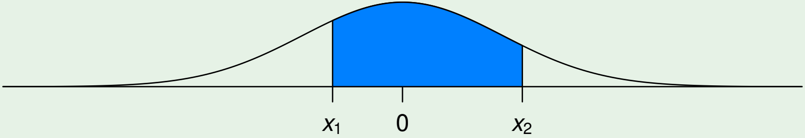
The area of the shaded region is the probability that a z score is greater than or equal to x, $P(z \geq x)$.

Note



$$P(z \geq x) = 1 - P(z \leq x)$$

Example 8

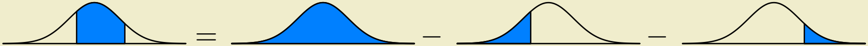


The area of the shaded region is the probability that a z score lies between x_1 and x_2, $P(x_1 \leq z \leq x_2)$.

Note

$$P(x_1 \leq z \leq x_2) = P(z \leq x_2) - P(z \leq x_1)$$

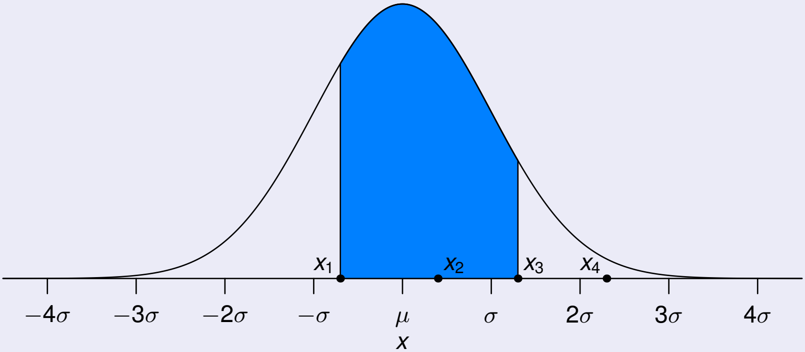
Note



$$P(x_1 \leq z \leq x_2) = 1 - P(z \leq x_1) - P(z \geq x_2)$$

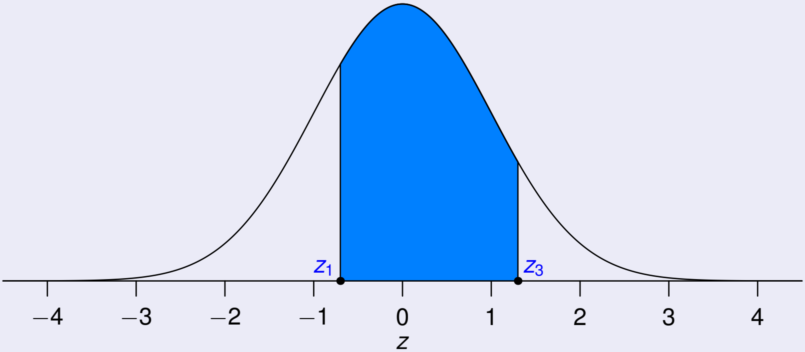
Procedure for Finding Areas with a Nonstandard Normal Distribution

1. Sketch a normal curve, label the mean and any specific x values, and then shade the region representing the desired probability.
2.
3.



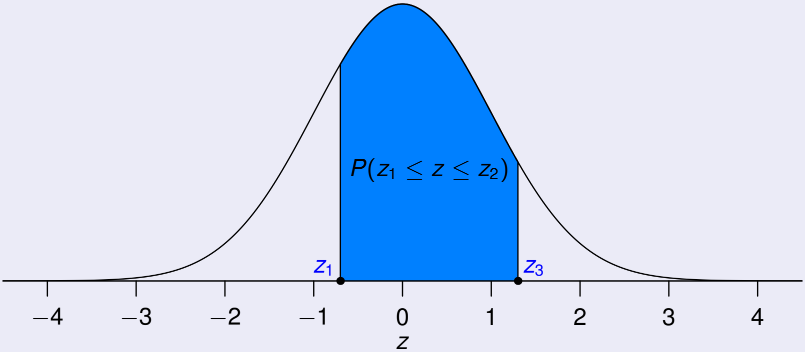
Procedure for Finding Areas with a Nonstandard Normal Distribution

1. Sketch a normal curve, label the mean and any specific x values, and then shade the region representing the desired probability.
2. For each relevant x value that is a boundary for the shaded region, convert that value to the equivalent z score.
3.



Procedure for Finding Areas with a Nonstandard Normal Distribution

1. Sketch a normal curve, label the mean and any specific x values, and then shade the region representing the desired probability.
2. For each relevant x value that is a boundary for the shaded region, convert that value to the equivalent z score.
3. Use technology to find the area of the shaded region.



Example 9

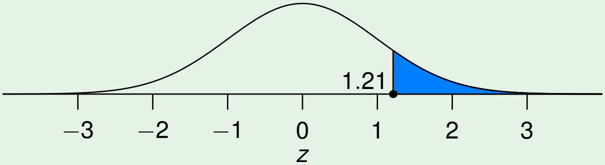
The heights of men are normally distributed with a mean of 68.6 in. and a standard deviation of 2.8 in.

Let's find the percentage of men who are taller than a shower head installed at a height of 72 in.

We start by finding the *z*-value of the shower head.

$$z = \frac{x - \mu}{\sigma} = \frac{72 - 68.6}{2.8} = 1.21$$

Next, we sketch a picture and shade the area we wish to find:



We can then use technology to compute:

$$P\,(z \geq 1.21) \approx 0.1123 \quad \text{(rounded)}$$

So, about 11.23% of men are taller than the shower head.

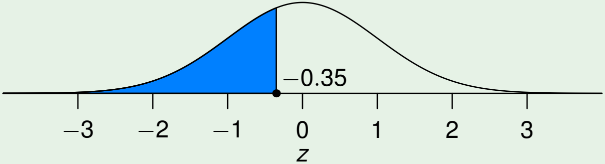
Example 10

Cumulative SAT scores are approximately normal with a mean of 1100 and standard distribution of 200. Edward earned a 1030 on his SAT.

The z-value of his score is:

$$z = \frac{x - \mu}{\sigma} = \frac{1030 - 1100}{200} = -0.35$$

Recall that the percentile of a data value is the percentage of data less than the data value.



This means $P(z \leq -0.35)$ is the percentile of Edwards SAT score.

We can then use technology to compute:

$$P(z \leq -0.35) \approx 0.3632 \quad \text{(rounded)}$$

So, Edward is in the 36$^{\text{th}}$ percentile.

Example 11

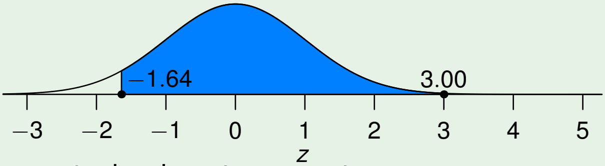
The U.S. Air Force requires that pilots have heights between 64 and 77 in. The heights of men are normally distributed with a mean of 68.6 in. and a standard deviation of 2.8 in.

Let's find the percentage of men meet that requirement.

We start by finding the z-values of the height requirements.

$$z_1 = \frac{x-\mu}{\sigma} = \frac{64-68.6}{2.8} = -1.64 \text{ and } z_2 = \frac{x-\mu}{\sigma} = \frac{77-68.6}{2.8} = 3.00$$

Next, we sketch a picture and shade the area we wish to find:



We can then use technology to compute:

$$P(-1.64 \leq z \leq 3.00) \approx 0.9484 \quad \text{(rounded)}$$

So, we see that about 94.84% of men meet the requirements.

Example 12

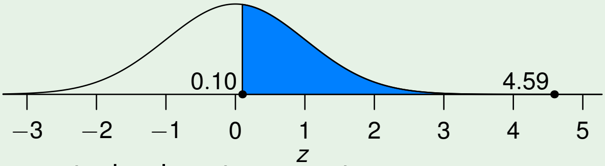
The U.S. Air Force requires that pilots have heights between 64 and 77 in. The heights of women are normally distributed with a mean of 63.7 in. and a standard deviation of 2.9 in.

Let's find the percentage of women meet that requirement.

We start by finding the z-values of the height requirements.

$$z_1 = \frac{x-\mu}{\sigma} = \frac{64-63.7}{2.9} = 0.10 \quad \text{and} \quad z_2 = \frac{x-\mu}{\sigma} = \frac{77-63.7}{2.9} = 4.59$$

Next, we sketch a picture and shade the area we wish to find:



We can then use technology to compute:

$$P\,(0.10 \leq z \leq 4.59) \approx 0.4601 \quad \text{(rounded)}$$

So, we see that only about 46% of women meet the requirements.

When Finding Values from Known Areas

- Draw a sketch of the graph.
- Don't confuse z scores and areas.
- Choose the correct side of the graph.
- A z score must be negative whenever it is located in the left half of the normal distribution.
- Areas are always between 0 and 1, and are never negative.

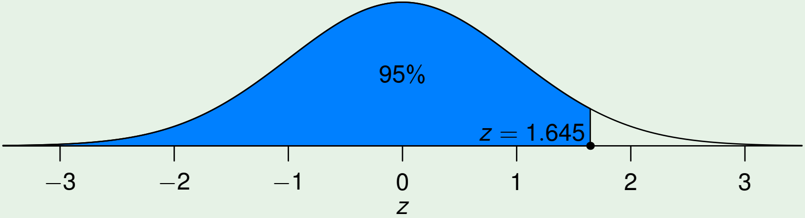
Procedure

1. Sketch the normal distribution curve, write the given probability or percentage in the appropriate region of the graph, and identify the x values being sought.
2. Either use technology or a table to identify the z scores corresponding to that area.
3. Convert to x values: $x = \mu + z \cdot \sigma$
4. Use your sketch to verify that the solution makes sense.

Example 13

When designing equipment, one common criterion is to use a design that accommodates 95% of the population. In Example 12 we saw that only 46% of women satisfy the U.S. Air Force pilot height requirement.

What would be the maximum acceptable height of a woman if the requirements were changed to allow the shortest 95% of women?



Using either technology or a table, we find that $z = 1.645$.

We then need to convert to the x value.

$$x = \mu + z \cdot \sigma = 63.7 + 1.645 \cdot 2.9 = 68.4705$$

A requirement of a height less than or equal to 68.5 in. would allow 95% of women to be eligible.