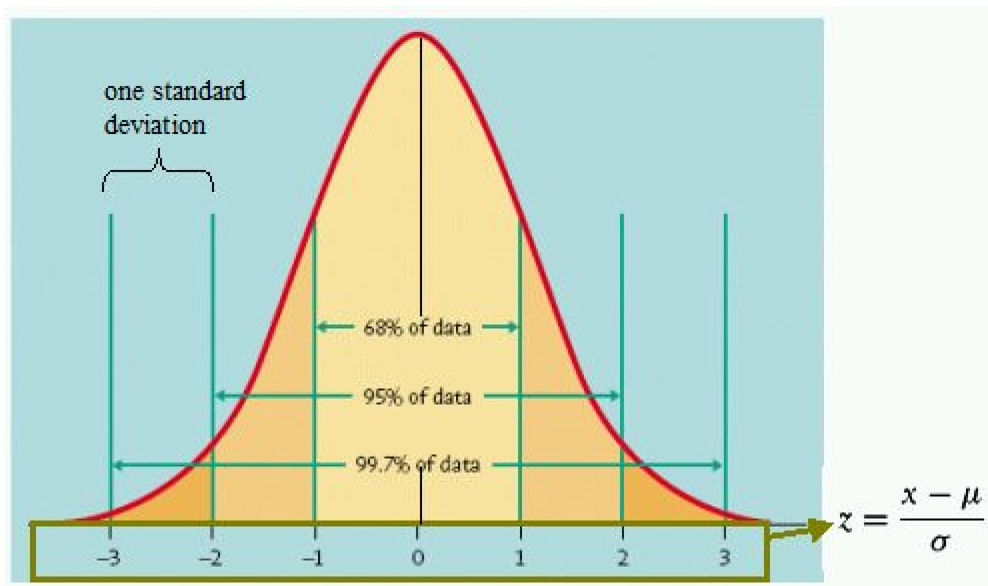
# THE NORMAL DISTRIBUTION



We’ve already seen a few attempts to define what is known as the *normal distribution.*

We've already seen a few attempts to define what is known as the *normal distribution.*

1. A bell-shaped distribution that is symmetrical.

We've already seen a few attempts to define what is known as the *normal distribution.*

1. A bell-shaped distribution that is symmetrical.

2. A symmetrical bell-shaped distribution that has 68% of its scores within 1 standard deviation of the mean, 95% within 2 standard deviations, and 99.7% within 3 standard deviations.

And now we are ready for the definitive definition.

1. A bell-shaped distribution that is symmetrical.

2. A symmetrical bell-shaped distribution that has 68% of its scores within 1 standard deviation of the mean, 95% within 2 standard deviations, and 99.7% within 3 standard deviations.

And now we are ready for the definitive definition.

1. A bell-shaped distribution that is symmetrical.

2. A symmetrical bell-shaped distribution that has 68% of its scores within 1 standard deviation of the mean, 95% within 2 standard deviations, and 99.7% within 3 standard deviations.

3. A normal distribution a function of the form:

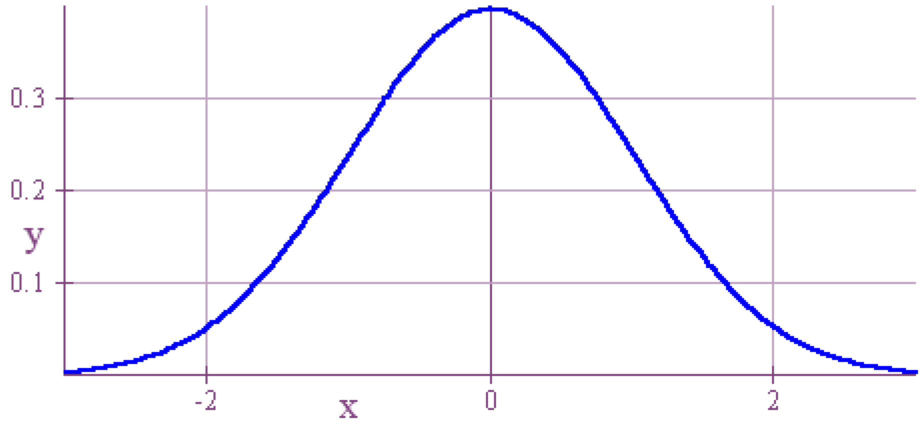
$$y = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

In this formula, *mu* is the mean of the distribution and *sigma* is the standard deviation

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

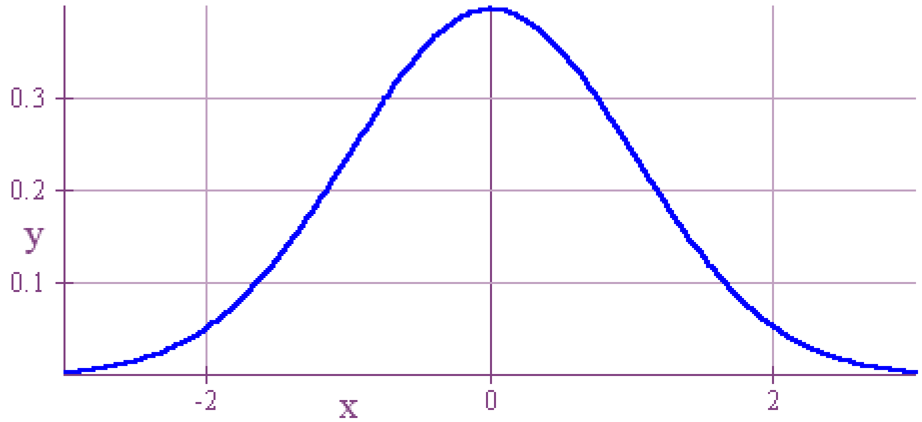
The normal distribution with *mu = 0* and *sigma = 1* is called the *standard normal distribution*.

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



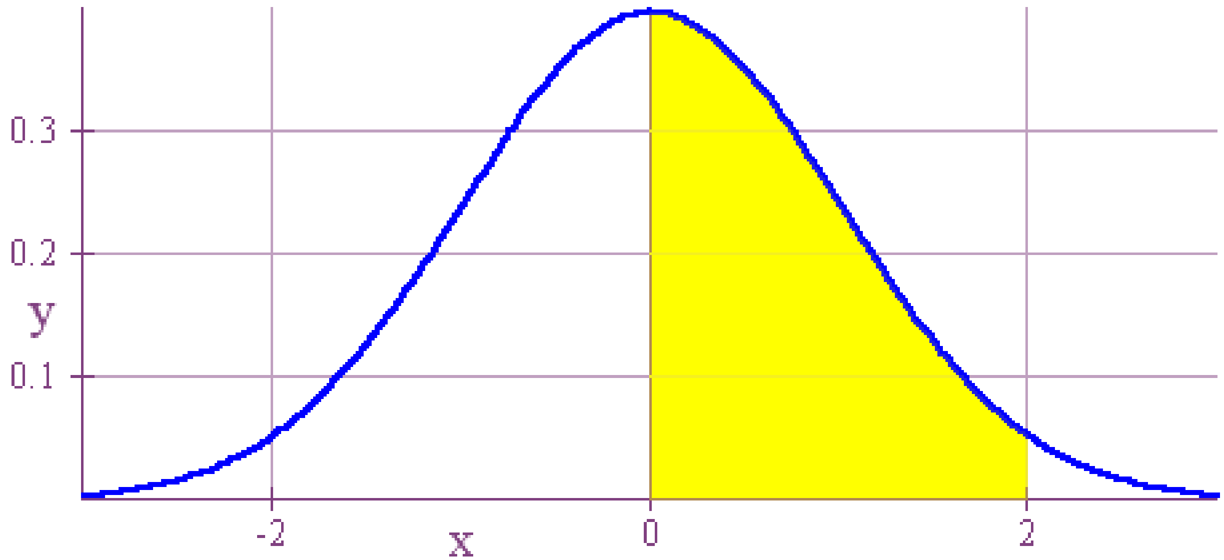
For all normal distributions, the area under the curve is equal to 1.

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

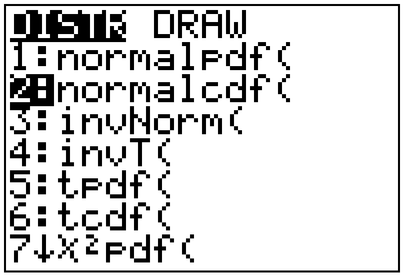


To find the probability that an *x-value* below is between zero and two, we just need to know the area under the curve from zero to two.

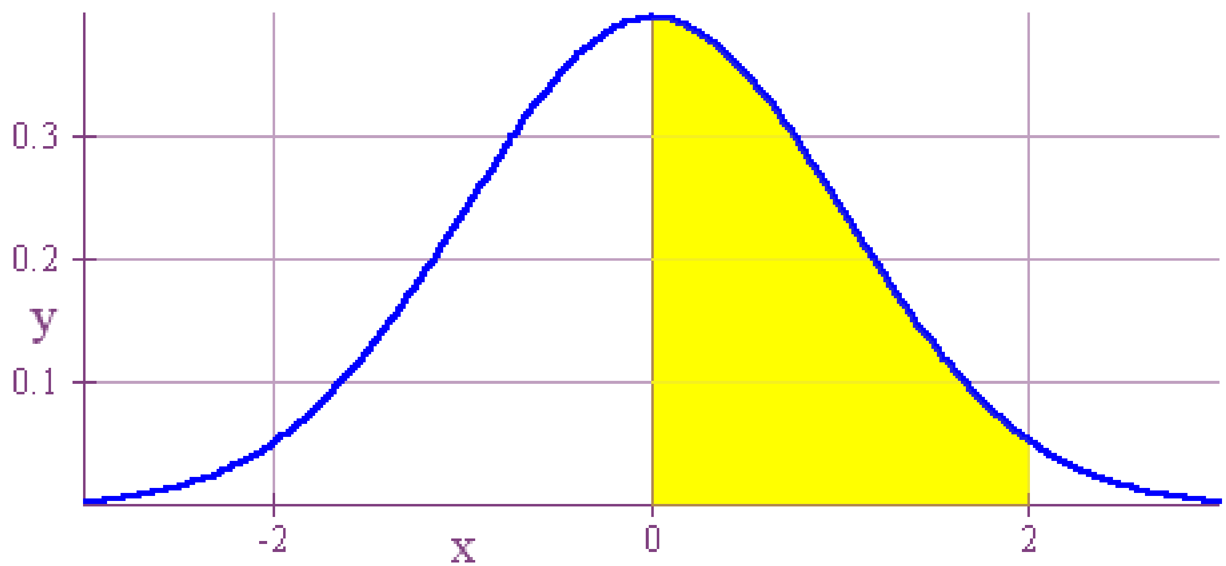
$$y = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$



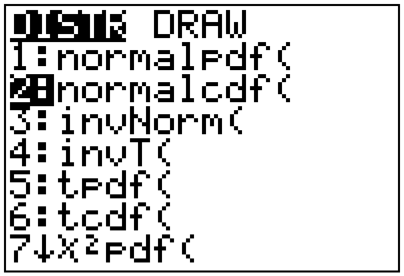
There's a nice tool for doing this under the DISTR menu.



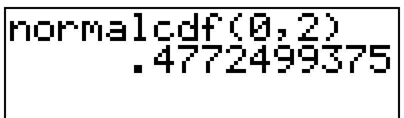

$$y = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$



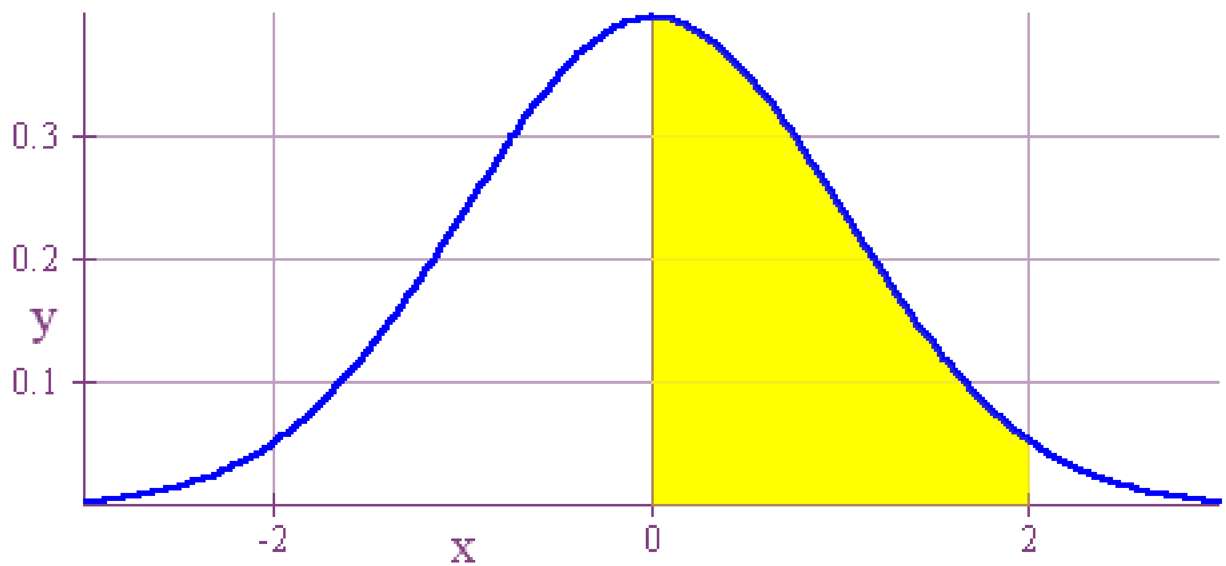
The syntax to follows is *normalcdf(start,stop)*.




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




$$P(0 \le z \le 2)$$



Here are some examples of finding probabilities using a standard normal distribution (*mean=0, standard deviation=1*).

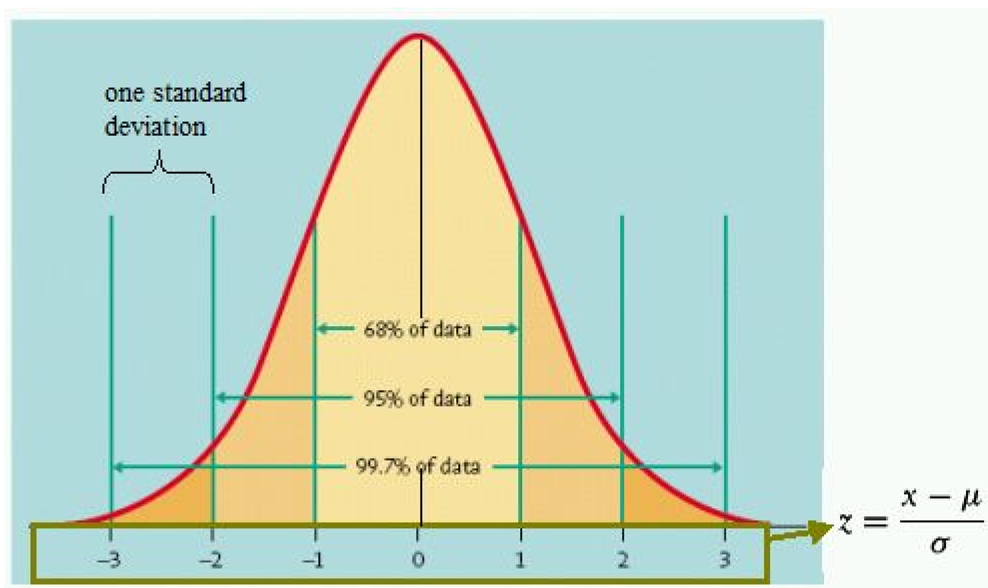
The probability that a score is within 1 standard deviation of the mean.

Here are some examples of finding probabilities using a standard normal distribution (*mean=0, standard deviation=1*).

The probability that a score is within 1 standard deviation of the mean.

```
normalcdf(-1,1)
       .6826894809
```

$$P(-1 \le z \le 1)$$

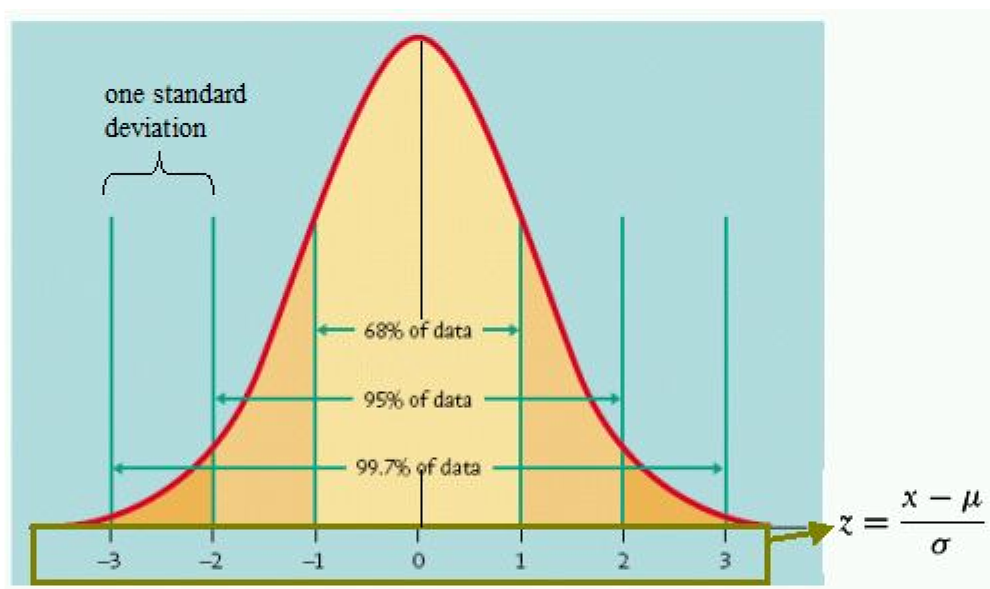


Here are some examples of finding probabilities using a standard normal distribution (*mean=0, standard deviation=1*).

The probability that a score is within 2 standard deviations of the mean.

```
normalcdf(-2,2)
        .954499876
```

$$P(-2 \le z \le 2)$$

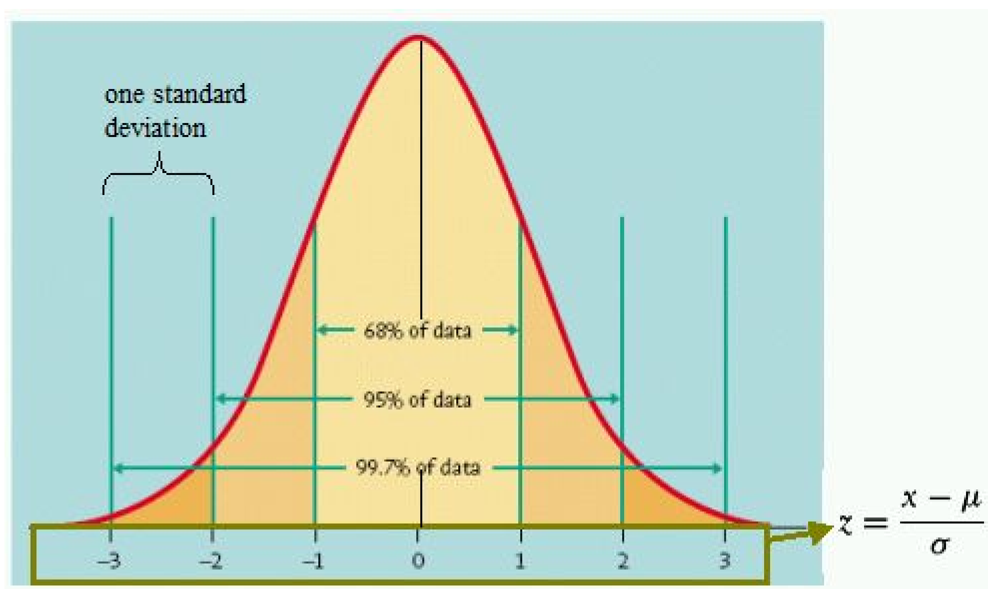


Here are some examples of finding probabilities using a standard normal distribution (*mean=0, standard deviation=1*).

The probability that a score is within 3 standard deviations of the mean.

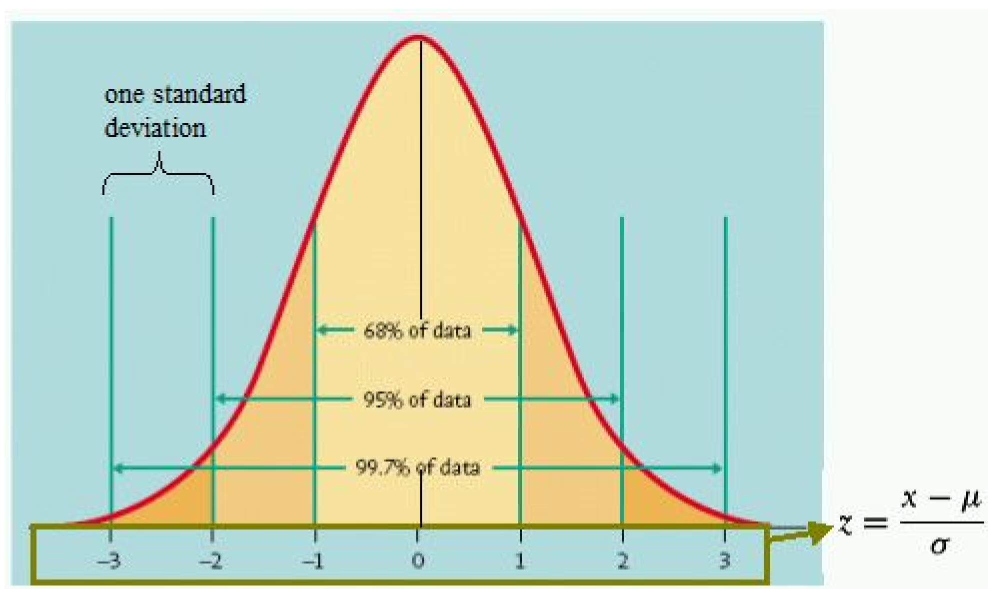
```
normalcdf(-3,3)
          .9973000656
```

$$P(-3 \leq z \leq 3)$$



What if we want to find the probability that a score's deviation from the mean is greater than 1 (*mean=0, standard deviation=1*)?

$$P(z > 1)$$

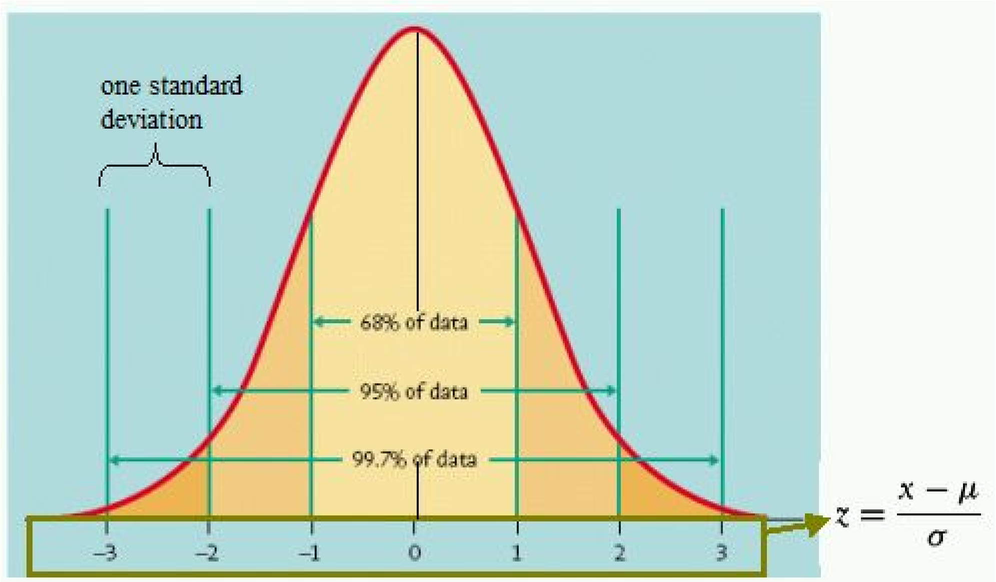


What if we want to find the probability that a score's deviation from the mean is greater than 1 (*mean=0, standard deviation=1*)?

We could do it two ways.  The less accurate but good enough way is illustrated below.

```
normalcdf(1,9999
99)
        .1586552596
```

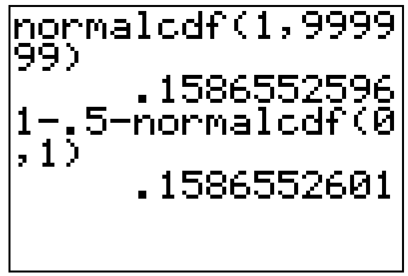
$$P(z>1)$$

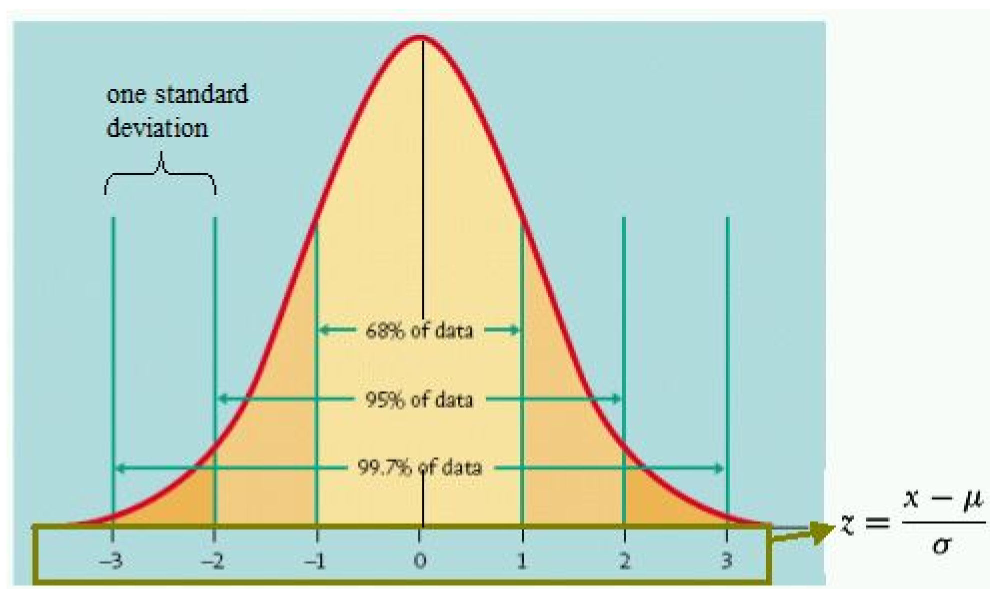


What if we want to find the probability that a score's deviation from the mean is greater than 1 (*mean=0, standard deviation=1*)?

And this is followed by the more accurate method.

$$P(z > 1)$$



What if we want to find areas or probabilities with respect to some other normal distribution where the *mean* is not zero and the *standard deviation* is not one?

What if we want to find areas or probabilities with respect to some other normal distribution where the *mean* is not zero and the *standard deviation* is not one?

In that case follow the syntax *normalcdf(start,stop,mean,deviation)*

What if we want to find areas or probabilities with respect to some other normal distribution where the *mean* is not zero and the *standard deviation* is not one?

In that case follow the syntax *normalcdf(start,stop,mean,deviation)*

```
normalcdf(85,115
,100,15)
      .6826894809
```

Another useful tool is *inverse normal* tool where we input an area and we get back the *z-score* that would have that area to the left in a standard normal curve.

The syntax is *invNorm(area)*

```
invNorm(.95)
      1.644853626
```

Another useful tool is *inverse normal* tool where we input an area and we get back the *z-score* that would have that area to the left in a standard normal curve.

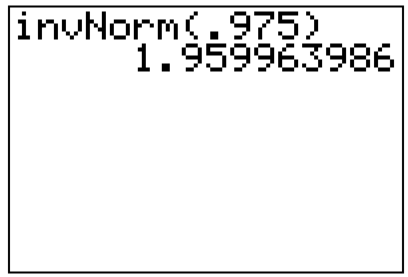
The syntax is *invNorm(area)*

```
invNorm(.95)
         1.644853626
```

However, we often use the area to the right to denote this number.

$$z_{.05} = 1.644853626$$

Notice that $-z_{.025}$ and $z_{.025}$ give you more accurate bounds for where the middle 95% of the data lies in a normal distribution.



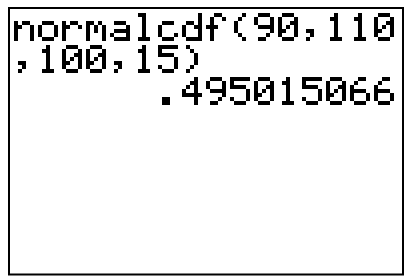

$$z_{.025} = 1.959963986$$

$$-z_{.025} = -1.959963986$$

The middle 95% of the data lies between *z=-1.96* and *z=1.96*.

For most IQ tests, the scores are distributed normally with a *mean* of 100 and a *standard deviation* of 15.  Find the following probabilities.

A person's IQ is between 90 and 110.




$$P(90 \le x \le 110) = .4950$$

For most IQ tests, the scores are distributed normally with a *mean* of 100 and a *standard deviation* of 15. Find the following probabilities.

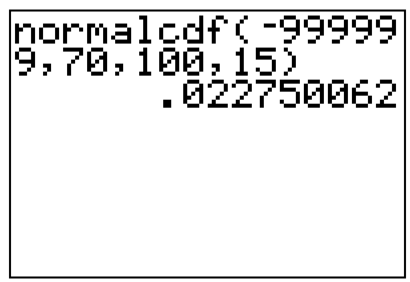
A person's IQ is between 120 and 130.

```
normalcdf(120,13
0,100,15)
      .0684612199
```

$$P(120 \le x \le 130) = .0685$$

For most IQ tests, the scores are distributed normally with a *mean* of 100 and a *standard deviation* of 15. Find the following probabilities.

A person's IQ is less than 70.




$$P(x < 70) = .0228$$

For most IQ tests, the scores are distributed normally with a *mean* of 100 and a *standard deviation* of 15.  Find the following probabilities.

A person's IQ is greater than 140.

```
normalcdf(140,99
9999,100,15)
      .0038304251
```

$$P(x > 140) = .0038$$

For most IQ tests, the scores are distributed normally with a *mean* of 100 and a *standard deviation* of 15. Find the following probabilities.

A person's IQ is less than 152.

```
normalcdf(-99999
9,152,100,15)
      .9997364758
```

$$P(x < 152) = .9997$$

We can also find the cutoff score corresponding to the percentage or proportion that we want an IQ to be greater than. We just use the *invNorm* tool with the following syntax: *invNorm(proportion, mean, standard deviation)*.

What IQ is greater than that of 95% of the population?.

```
invNorm(.95,100,
15)
      124.6728044
```

$$IQ \approx 125$$