Normal Distribution

Department of Mathematics

Salt Lake Community College

Example 1

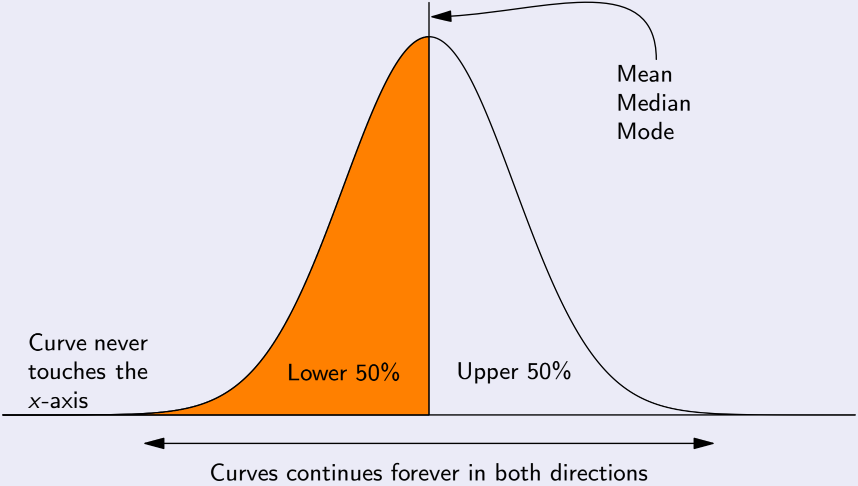
Consider making popcorn. You put some oil and corn kernels in a pan and start heating.

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



Note

It is common to use the notation:

- $\mu =$ mean
- $\sigma =$ standard deviation

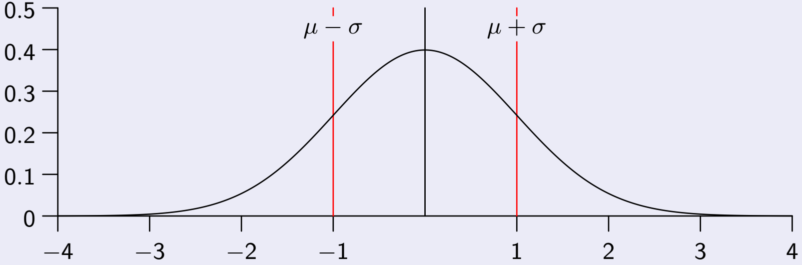
Note

It is common to use the notation:

- μ = mean
- σ = standard deviation

Definition

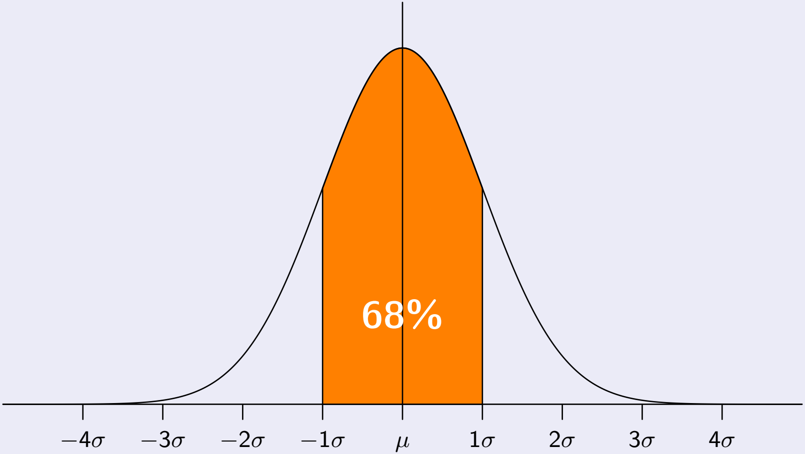
When $\mu = 0$ and $\sigma = 1$, the curve is called the **standard normal distribution**.



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.

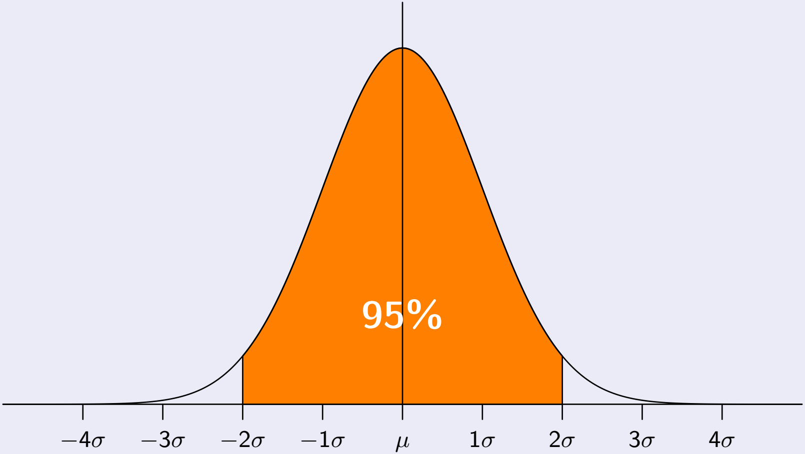
One standard deviation from the mean.



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.

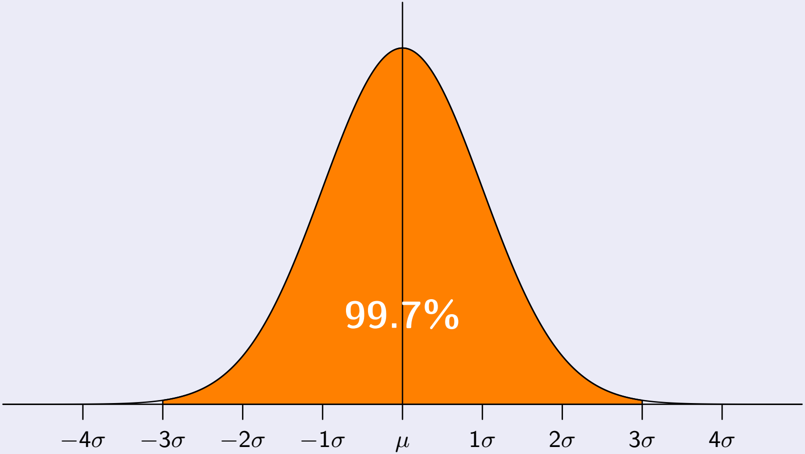
Two standard deviations from the mean.



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.

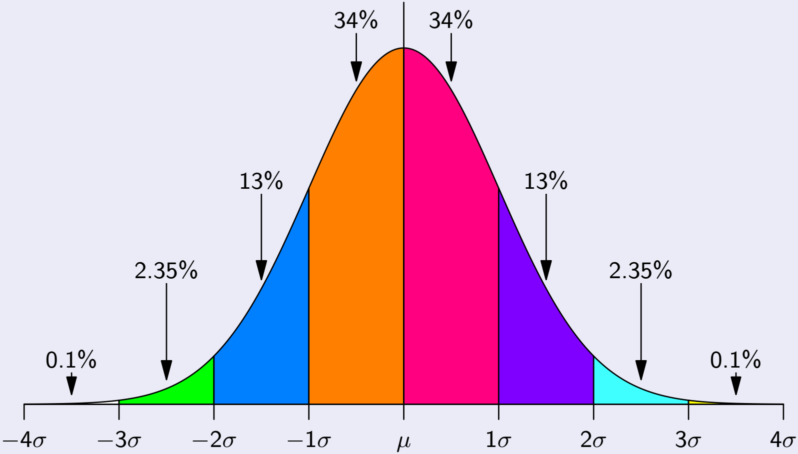
Three standard deviations from the mean.



Definition

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 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. It is calculated with:

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

Definition

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

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

Example 2

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

Definition

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

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

Example 2

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

$$z = \frac{x - \mu}{\sigma}$$

Definition

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

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

Example 2

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

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

Definition

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

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

Example 2

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

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

Definition

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

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

Example 2

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 3

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

Definition

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

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

Example 2

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 3

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

$$z = \frac{x - \mu}{\sigma}$$

Definition

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

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

Example 2

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 3

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

$$z = \frac{x - \mu}{\sigma} \quad \Rightarrow \quad z\sigma = x - \mu$$

Definition

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

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

Example 2

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 3

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

$$z = \frac{x - \mu}{\sigma} \quad \Rightarrow \quad z\sigma = x - \mu \quad \Rightarrow \quad x = z\sigma + \mu$$

Definition

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

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

Example 2

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 3

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 4

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?

Example 4

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?

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

Example 4

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.

Example 4

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 $95\% - 68\% = 32\%$ of the scores are more than one standard deviation from the mean, but less than two.

Example 4

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 $95\% - 68\% = 32\%$ of the scores are more than one standard deviation from the mean, but less than two.

Since the curve is symmetric, we know that 16% of the students scored between 89 and 96, as well as 16% between 68 and 75

Note

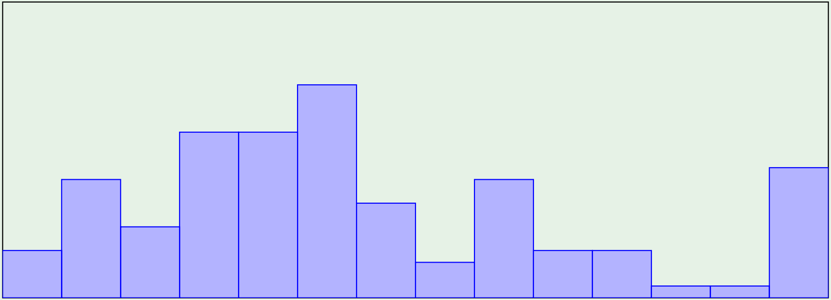
Visual representations of data, such as histograms can tell you if data is roughly normal or not.

Note

Visual representations of data, such as histograms can tell you if data is roughly normal or not.

Example 5

This data is not a normal distribution.



Definition

A **density curve** is an idealized representation of a distribution in which the area under the curve is defined to be one.

Definition

A **density curve** is an idealized representation of a distribution in which the area under the curve is defined to be one.

Note

A density curve does not need to be normal, but the normal density curve will be of the most use.

Definition

A **density curve** is an idealized representation of a distribution in which the area under the curve is defined to be one.

Note

A density curve does not need to be normal, but the normal density curve will be of the most use.

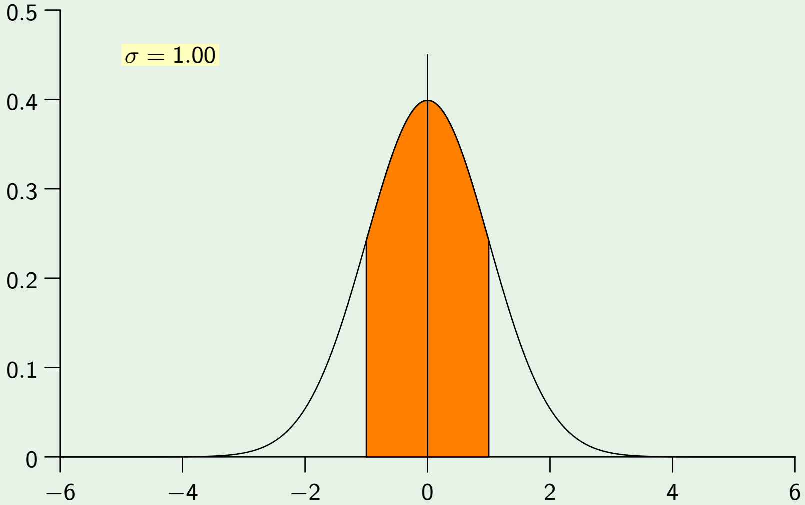
Note

We know from the Empirical rule that roughly two-thirds of the data in a normal distribution falls within one standard deviation of the mean.

With a normal density curve, this means that about 68% of the total area under the curve is within z-scores of ± 1.

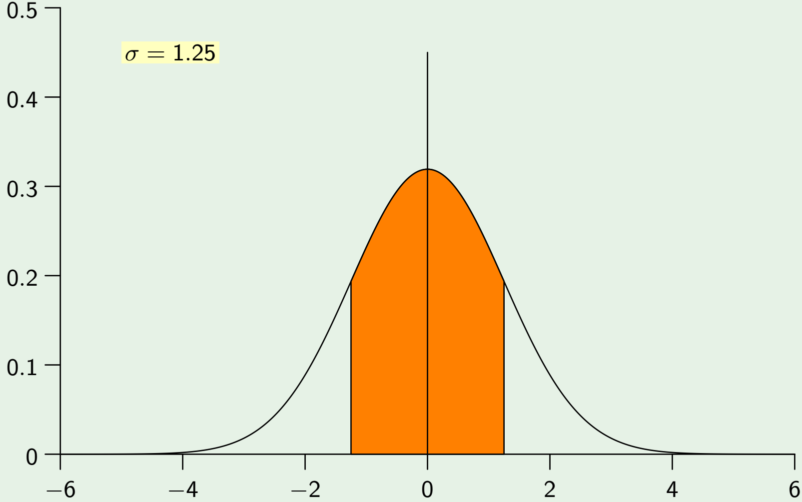
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



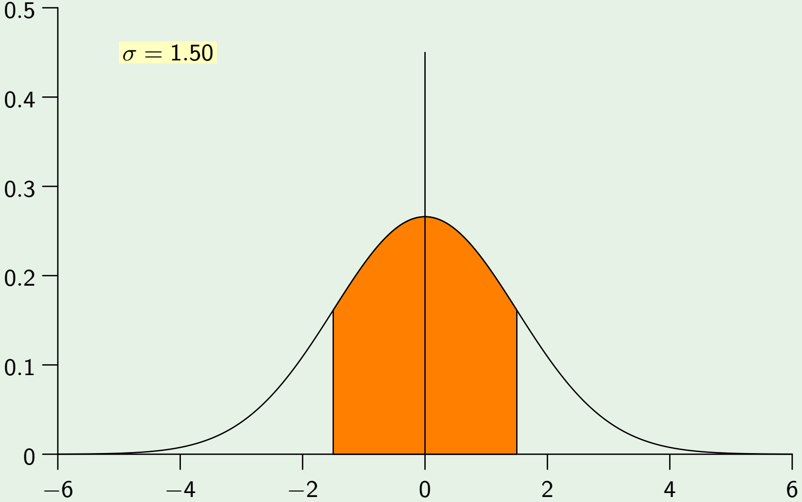
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



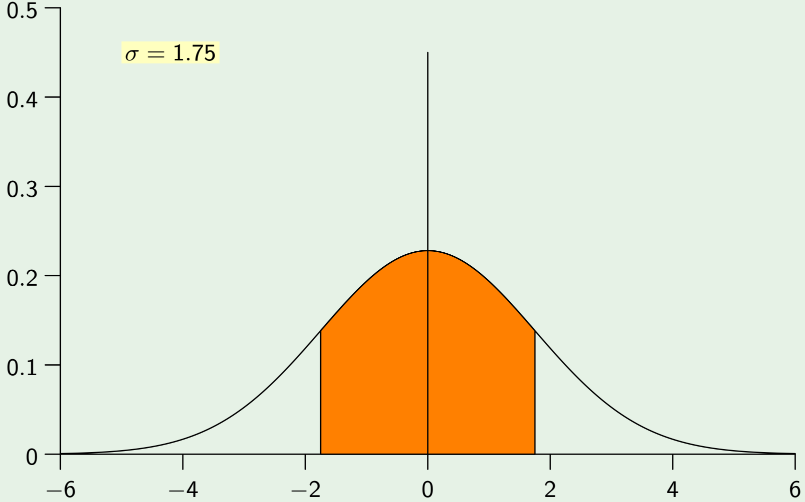
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



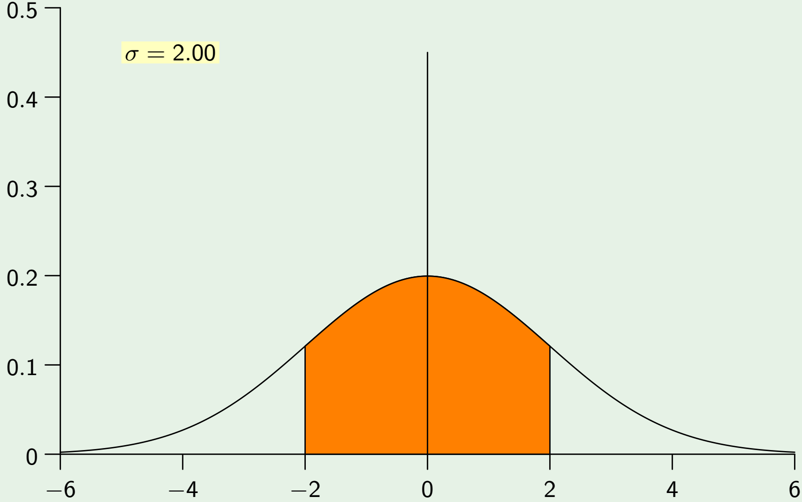
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



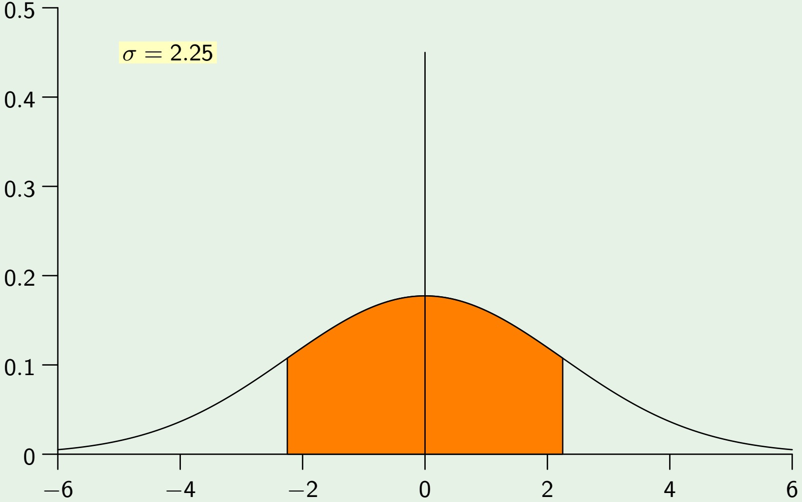
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



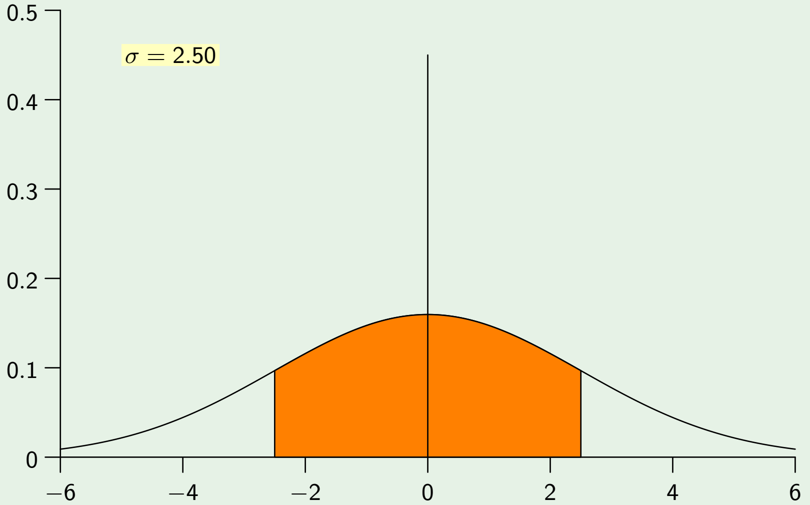
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



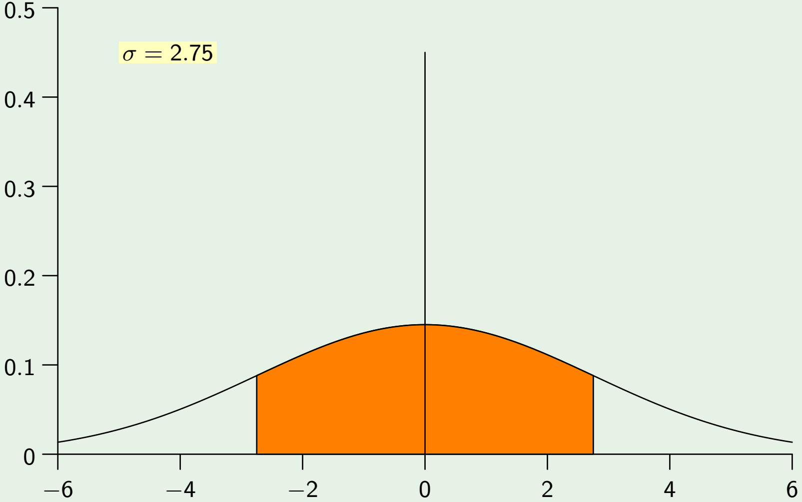
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



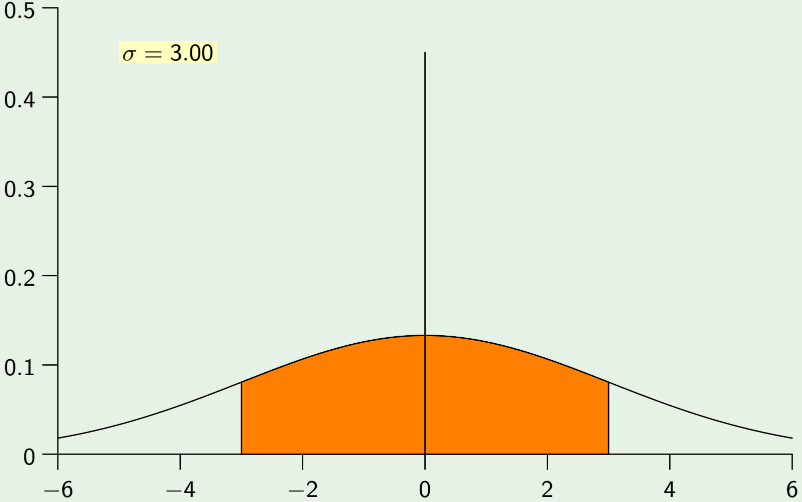
Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



Example 6

Notice that as the standard deviation increases, the curve gets flatter. This is because 68% of the area always falls within a single standard deviation of the mean.



Definition

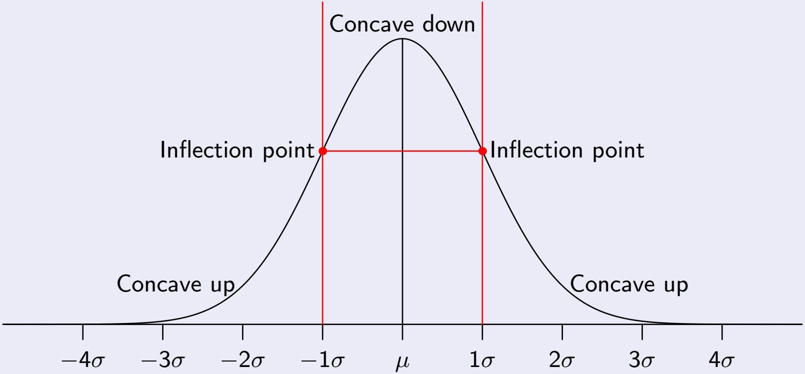
An **inflection point** is where a curve changes from being concave up to concave down, or vice versa

Definition

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

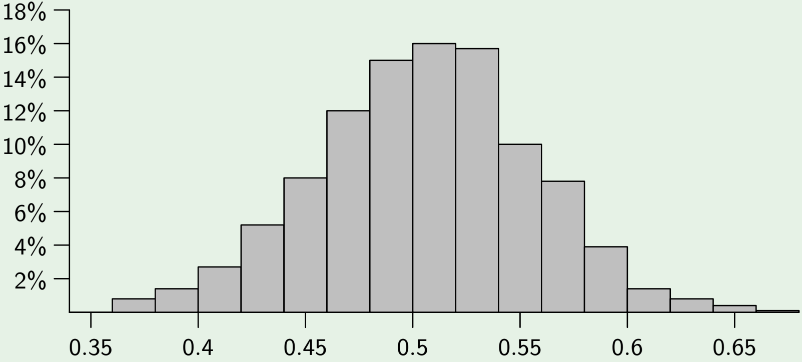
Note

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



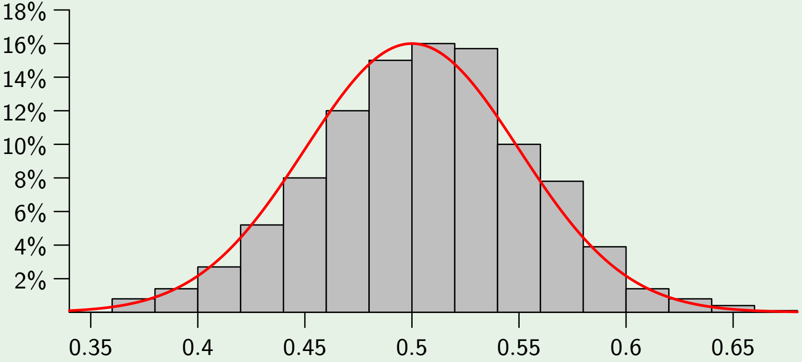
Example 7

Estimate the standard deviation of the distribution represented by the histogram.



Example 7

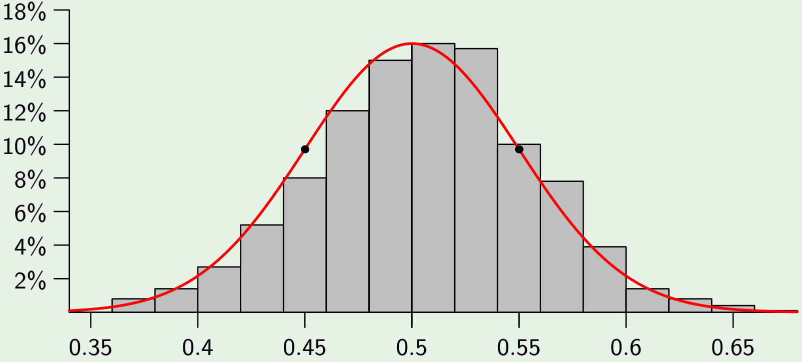
Estimate the standard deviation of the distribution represented by the histogram.



First, draw a density curve that fits the histogram fairly well.

Example 7

Estimate the standard deviation of the distribution represented by the histogram.



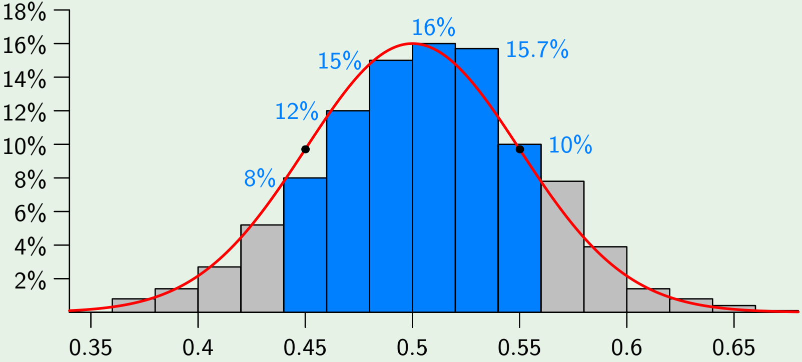
Next, estimate the inflections points, as well as the mean.

In this case, we have $\mu \approx 0.5$ and $\sigma \approx 0.05$.

(The actual statistics are: mean $= 0.04988$ and std. dev. $= 0.4997$.)

Example 7

Estimate the standard deviation of the distribution represented by the histogram.



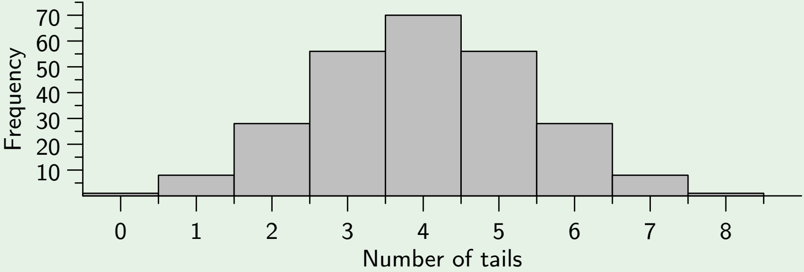
We can estimate how much of the graph is within one standard deviation by estimating the heights of the colored bins.

$$\frac{8}{2} + 12 + 15 + 16 + 15.75 + \frac{10}{2} = 67.7$$

This is remarkably close to the 68% the Empirical rule gives.

Example 8

Suppose you flip 8 coins and record the results. The histogram shows the results of performing this experiment 256 times.



Example 8

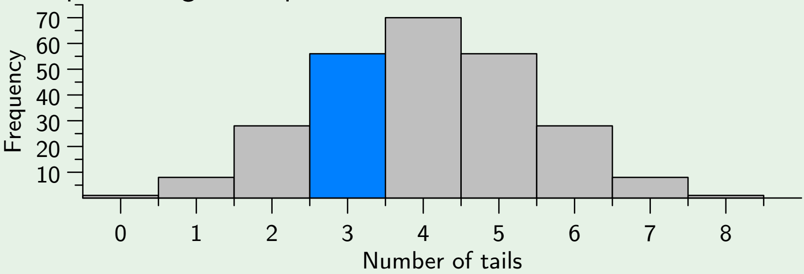
Suppose you flip 8 coins and record the results. The histogram shows the results of performing this experiment 256 times.

One way to calculate the probability that exactly three tails are flipped is:

$$P\left(3 \text{ tails}\right) = \frac{{}_8C_3}{2^8} = \frac{56}{256} \approx 0.2188$$

Example 8

Suppose you flip 8 coins and record the results. The histogram shows the results of performing this experiment 256 times.

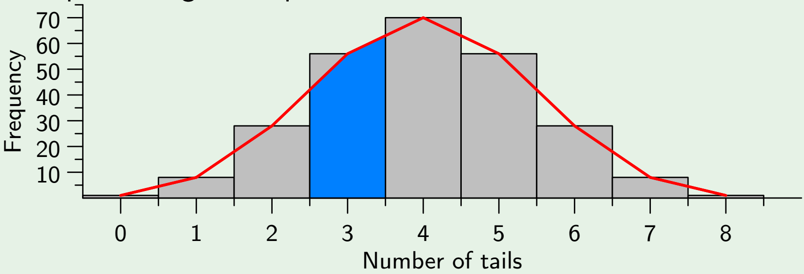


A second way is to approximate the heights of each box and divide the area of the colored box by the areas of all the boxes:

$$P\,(3 \text{ tails}) = \frac{55}{1+8+28+55+70+55+28+8+1} = \frac{55}{254} \approx 0.2165$$

Example 8

Suppose you flip 8 coins and record the results. The histogram shows the results of performing this experiment 256 times.

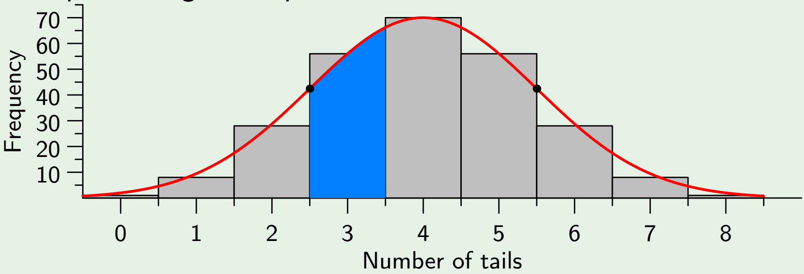


A third way is to draw an approximate normal curve. Then, the area from 2.5 to 3.5 relates to the probability of the event.

$$P\,(3 \text{ tails}) = \frac{\text{Area from 2.5 to 3.5}}{\text{Total area under curve}} = \frac{61.25}{324} \approx 0.1890$$

Example 8

Suppose you flip 8 coins and record the results. The histogram shows the results of performing this experiment 256 times.



The better we draw the density curve, the better our approximation.

$$P\,(3\text{ tails}) = \frac{\text{Area from 2.5 to 3.5}}{\text{Total area under curve}} = \frac{55.478}{262.075} \approx 0.2117$$

Definition

The area of a region under the standard normal density curve is the probability of a data point falling in that region.

Definition

The area of a region under the standard normal density curve is the probability of a data point falling in that region.

Note

To actually calculate the area of density curve is not easy, this is why either a table or a computer is usually used.

Definition

The area of a region under the standard normal density curve is the probability of a data point falling in that region.

Note

To actually calculate the area of density curve is not easy, this is why either a table or a computer is usually used.

Example 9

Here is a section of a Standard Normal Probability table.

z	0.00	0.01	0.02	0.03	0.04	0.05
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678

What is the probability that $z \leq 1.63$?

Definition

The area of a region under the standard normal density curve is the probability of a data point falling in that region.

Note

To actually calculate the area of density curve is not easy, this is why either a table or a computer is usually used.

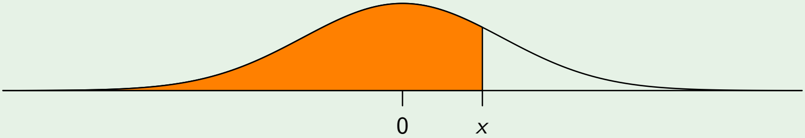
Example 9

Here is a section of a Standard Normal Probability table.

z	0.00	0.01	0.02	0.03	0.04	0.05
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678

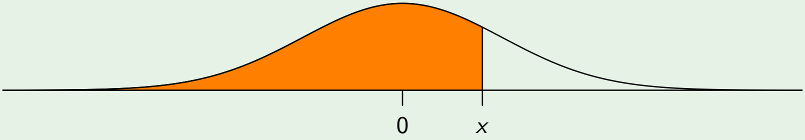
What is the probability that $z \leq 1.63$? $P(z \leq 1.63) = 0.9484$

Example 10



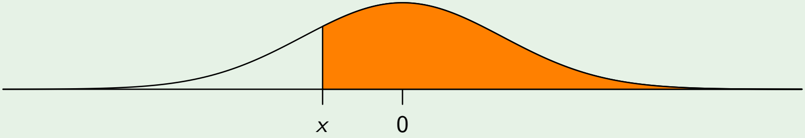
The area of the shaded region is the probability that a z score is less than or equal to x, $P(z \leq x)$.

Example 10



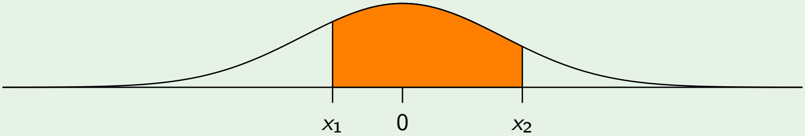
The area of the shaded region is the probability that a z score is less than or equal to x, $P(z \leq x)$.

Example 11



The area of the shaded region is the probability that a z score is greater than or equal to x, $P(z \geq x)$. Can also be calculated as $1 - P(z \leq x)$.

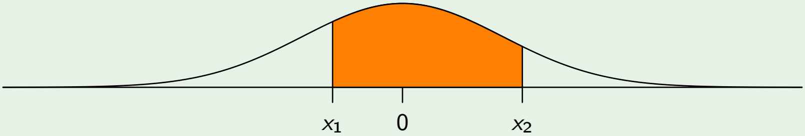
Example 12



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

Can also be calculated as $P\left(x \leq x_2\right) - P\left(z \leq x_1\right)$.

Example 12



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

Can also be calculated as $P(x \leq x_2) - P(z \leq x_1)$.

Note

It is always a good idea to sketch a picture of the normal density curve you are dealing with, so that you make sure you're finding the proper area.

Note

Always remember that Standard Normal Probability tables and programs are dealing with the standard normal distribution, that is $\mu = 0$ and $\sigma = 1$.

Note

Always remember that Standard Normal Probability tables and programs are dealing with the standard normal distribution, that is $\mu = 0$ and $\sigma = 1$.

To use such a table on an arbitary density function, you need to **standardize** the distribution by calculating z-scores.

Note

Always remember that Standard Normal Probability tables and programs are dealing with the standard normal distribution, that is $\mu = 0$ and $\sigma = 1$.

To use such a table on an arbitary density function, you need to **standardize** the distribution by calculating *z*-scores.

This works because all normal distribution curves have the same basic shape and the *z*-score gives a relative distance from the mean.

Note

Always remember that Standard Normal Probability tables and programs are dealing with the standard normal distribution, that is $\mu = 0$ and $\sigma = 1$.

To use such a table on an arbitary density function, you need to **standardize** the distribution by calculating *z*-scores.

This works because all normal distribution curves have the same basic shape and the *z*-score gives a relative distance from the mean.

Recall

The *z*-score is:

$$z = \frac{x - \mu}{\sigma}$$

Note

Always remember that Standard Normal Probability tables and programs are dealing with the standard normal distribution, that is $\mu = 0$ and $\sigma = 1$.

To use such a table on an arbitary density function, you need to **standardize** the distribution by calculating z-scores.

This works because all normal distribution curves have the same basic shape and the z-score gives a relative distance from the mean.

Recall

The z-score is:

$$z = \frac{x - \mu}{\sigma}$$

Note

In statistics, the term "unusual" often refers to an event with probability less than 5% ($p < 0.05$).

Example 13

Let us return to example 8, where we were looking for the probability that we would get exactly three tails. After fitting an approximate normal curve, we needed the area between 2.5 and 3.5.

Example 13

Let us return to example 8, where we were looking for the probability that we would get exactly three tails. After fitting an approximate normal curve, we needed the area between 2.5 and 3.5.

These values have the z-scores:

$$z_{2.5} = \frac{x - \mu}{\sigma} = \frac{2.5 - 4}{1.5} = -1 \quad \text{and} \quad z_{3.5} = \frac{3.5 - 4}{1.5} = -0.3333$$

Example 13

Let us return to example 8, where we were looking for the probability that we would get exactly three tails. After fitting an approximate normal curve, we needed the area between 2.5 and 3.5.

These values have the z-scores:

$$z_{2.5} = \frac{x-\mu}{\sigma} = \frac{2.5-4}{1.5} = -1 \quad \text{and} \quad z_{3.5} = \frac{3.5-4}{1.5} = -0.3333$$

So, then

$$\begin{aligned} P\left(-1 \leq z \leq -0.3333\right) &= P\left(z \leq -0.3333\right) - P\left(z \leq -1\right) \\ &= 0.369441 - 0.158655 \\ &= 0.210786 \end{aligned}$$

Example 13

Let us return to example 8, where we were looking for the probability that we would get exactly three tails. After fitting an approximate normal curve, we needed the area between 2.5 and 3.5.

These values have the z-scores:

$$z_{2.5} = \frac{x-\mu}{\sigma} = \frac{2.5-4}{1.5} = -1 \quad \text{and} \quad z_{3.5} = \frac{3.5-4}{1.5} = -0.3333$$

So, then

$$\begin{aligned} P\left(-1 \leq z \leq -0.3333\right) &= P\left(z \leq -0.3333\right) - P\left(z \leq -1\right) \\ &= 0.369441 - 0.158655 \\ &= 0.210786 \end{aligned}$$

Note

It's not at all unusual to flip eight coins and get three tails.

Example 14

A candy company sells small bags of candy. While they attempt to keep the number of pieces in each bag the same, small differences occur because of variation in the packaging process. A quality control expert in the factory determined that the mean number of pieces in each bag is normally distributed, with a mean of 57.3 and a standard deviation of 1.6.

Let us find the probability that we will purchase a bag with 55 or fewer pieces of candy.

Example 14

A candy company sells small bags of candy. While they attempt to keep the number of pieces in each bag the same, small differences occur because of variation in the packaging process. A quality control expert in the factory determined that the mean number of pieces in each bag is normally distributed, with a mean of 57.3 and a standard deviation of 1.6.

Let us find the probability that we will purchase a bag with 55 or fewer pieces of candy.

The z-score is

$$z = \frac{x - \mu}{\sigma} = \frac{55 - 57.3}{1.6} = -1.4375$$

Example 14

A candy company sells small bags of candy. While they attempt to keep the number of pieces in each bag the same, small differences occur because of variation in the packaging process. A quality control expert in the factory determined that the mean number of pieces in each bag is normally distributed, with a mean of 57.3 and a standard deviation of 1.6.

Let us find the probability that we will purchase a bag with 55 or fewer pieces of candy.

The z-score is

$$z = \frac{x - \mu}{\sigma} = \frac{55 - 57.3}{1.6} = -1.4375$$

Thus,

$$P\left(z \leq -1.4375\right) = 0.075288$$

A roughly 8% chance, it would be rare, but not unusual.

Example 15

Suppose that a fictional scientist measured the resting body temperature for a large number of people. She found that the mean is 98.25°F and the standard deviation is 0.73°F.

Johnny "Hothead" Rico has a resting body temperature of 100°F. Would you consider this unusual?

Example 15

Suppose that a fictional scientist measured the resting body temperature for a large number of people. She found that the mean is 98.25°F and the standard deviation is 0.73°F.

Johnny "Hothead" Rico has a resting body temperature of 100°F. Would you consider this unusual?

Johnny's z-score is

$$z = \frac{x - \mu}{\sigma} = \frac{100 - 98.25}{0.73} = 2.3972$$

Example 15

Suppose that a fictional scientist measured the resting body temperature for a large number of people. She found that the mean is 98.25°F and the standard deviation is 0.73°F.

Johnny "Hothead" Rico has a resting body temperature of 100°F. Would you consider this unusual?

Johnny's z-score is

$$z = \frac{x - \mu}{\sigma} = \frac{100 - 98.25}{0.73} = 2.3972$$

So, the probability that someone has a resting body temperature ≥ 100 is

$$P\left(z \geq 2.3972\right) = 1 - P\left(z \leq 2.3972\right) = 1 - 0.99174 = 0.00826$$

Johnny is pretty unusual, but don't let him hear you say that.