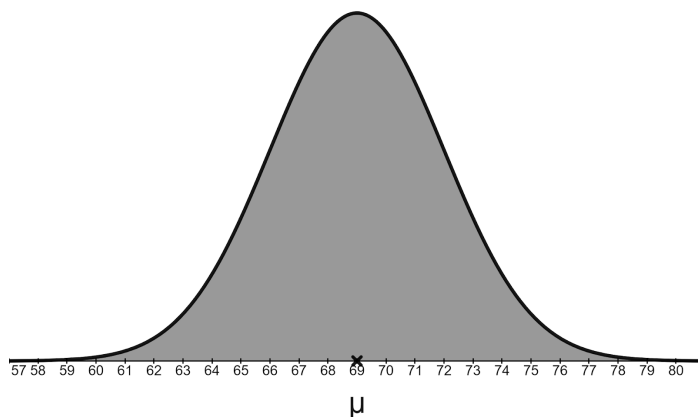


5.2: Characteristics of the Normal Distribution and The Empirical Rule

In the last lesson, we learned about continuous random variables and their probability distributions. A continuous random variable is special because it can take on any of an infinite number of possible values which cannot be listed in any order (or counted). Continuous probability distributions can take on a variety of shapes including uniform, bi- or multi-modal, or bell-shaped, etc., as long as the total area under the curve is 1.

A bell-shaped curve is one that is symmetric, and has a single mode in the center, and has two skinny tails. It is used to represent situations where the random variable is more likely to take on values closer to its average and less likely to take on extreme values. One distribution that we will be using often is called the **normal distribution** for which there is a formula allowing one to find the precise areas (probabilities) for any range of values of the continuous random variable. There are many examples of situations that warrant the use of the normal distribution.

For example, the heights of randomly selected men have a distribution that is approximately normal. The mean of this population is $\mu = 69$ inches, and the population standard deviation is $\sigma = 3$ inches.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

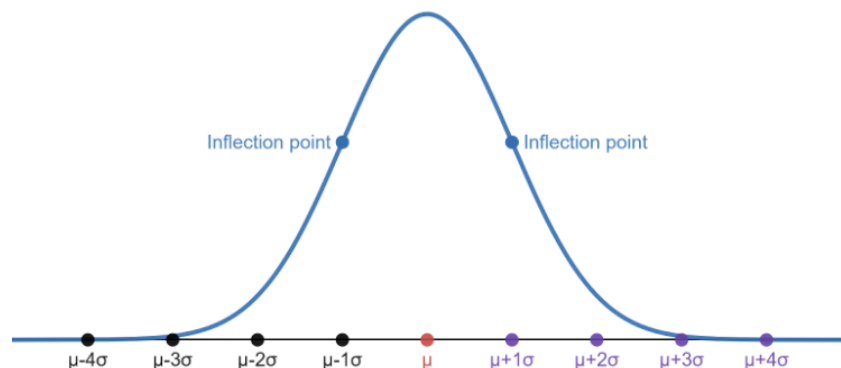
Some symbols to note:

- μ ("mu") represents a population mean.
- \bar{x} ("x-bar") represents a sample mean.
- σ ("sigma") represents a population standard deviation.
- s ("s") represents a sample standard deviation.

Characteristics of the Normal Distribution

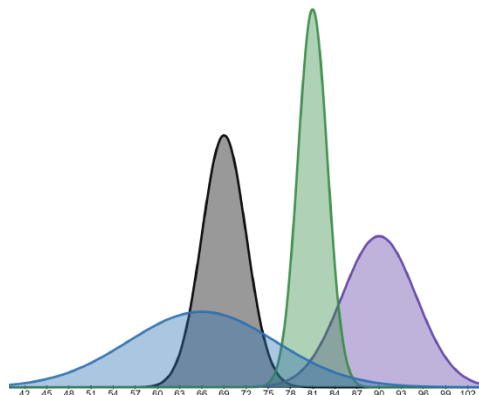
A **normal distribution** with mean μ and standard deviation has the following characteristics:

- The mean, median, and mode are equal. 50% of all values are below the mean and 50% are above it.
- The normal curve is bell-shaped and symmetric about its mean μ .
- The total area under the normal curve is 1.
- Normal curves extend endlessly in both directions, but the curve becomes so close to zero for values of the random variable that are more than 4 standard deviations above or below the mean that the area is negligible.
- The normal curve changes its curvature from a “cup” shape to a “cap” shape or vice-versa at one standard deviation below and above the mean. These points are called inflection points.



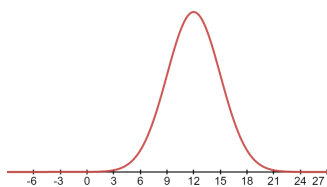
Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

A normal distribution is determined by its **mean** and its **standard deviation**. All normal distributions have the same shape but one distribution might have a mean of 69 whereas another has a mean of 90. This manifests on the graph as a horizontal translation. One normal distribution might have a standard deviation of 10 whereas another might have a standard deviation of 2. Standard deviation is a measure of spread so the higher the standard deviation, the more spread out the data is, and we see a flatter curve. The lower the standard deviation, the less spread there is, resulting in a curve that is narrow with a tall peak in the middle. Standard deviation manifests in the graph as horizontal stretching or shrinking. Explore these characteristics using [this desmos graph](#). Click the play button left of m to see how changing the mean changes the graph of the normal distribution. Click pause to stop the animation. Click the play button next to s to see how changing the standard deviation changes the graph of the normal distribution.

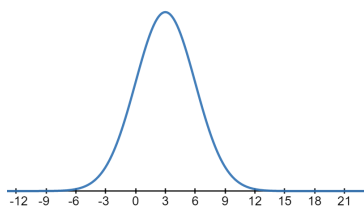


Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

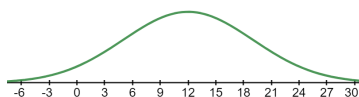
1. Match the following normal distributions with the appropriate means and standard deviations. Explain your reasoning for your choices.



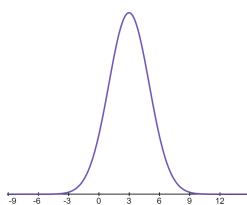
a.



b.



c.



d.

Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

i. $\mu = 12, \sigma = 7$

ii. $\mu = 3, \sigma = 3$

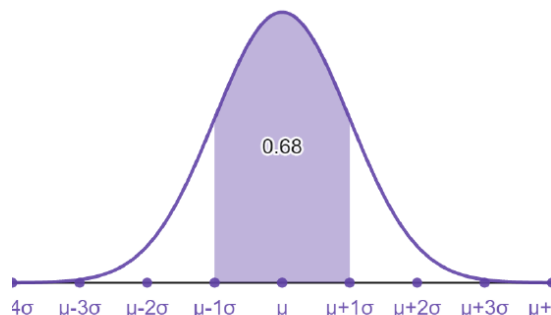
iii. $\mu = 3, \sigma = 2$

iv. $\mu = 12, \sigma = 3$

The Empirical Rule

All normal distributions are the same with respect to their mean and standard deviation. Using the normal distribution as the model, about 68% of values lie within one standard deviation of the mean. About 95% of values lie within two standard deviations from the mean, and about 99.7% of values lie within three standard deviations of the mean. These approximations are collectively known as the **Empirical Rule**.

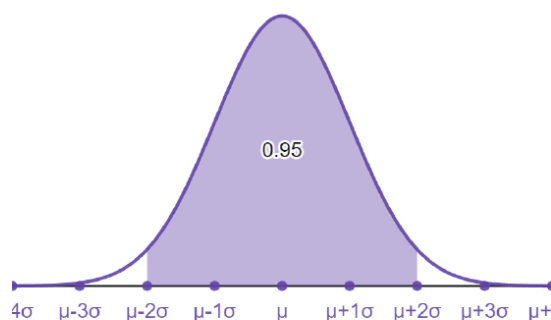
The graph below is of a normal distribution with the mean μ , located in the center of the graph. We see that about 68% of the area under the curve is between $\mu - 1\sigma$ (one standard deviation below the mean) and $\mu + 1\sigma$ (one standard deviation above the mean).



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

In other words, this tells us that about 68% of all values in a normal population lie within one standard deviation of the mean.

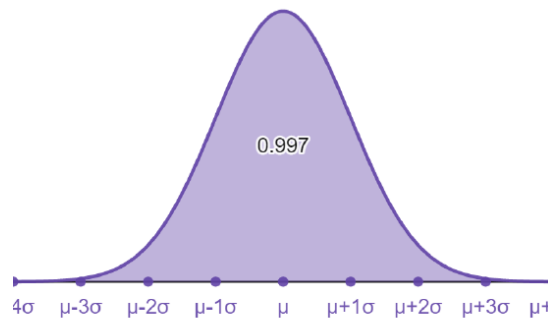
For the region between $\mu - 2\sigma$ (two standard deviations below the mean) and $\mu + 2\sigma$ (two standard deviations above the mean), the area under the curve is about 95%.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

In other words, this tells us that about 95% of all values in a normal population lie within two standard deviations of the mean.

For the region between $\mu - 3\sigma$ (three standard deviations below the mean) and $\mu + 3\sigma$ (three standard deviations above the mean), the area under the curve is about 99.7%.



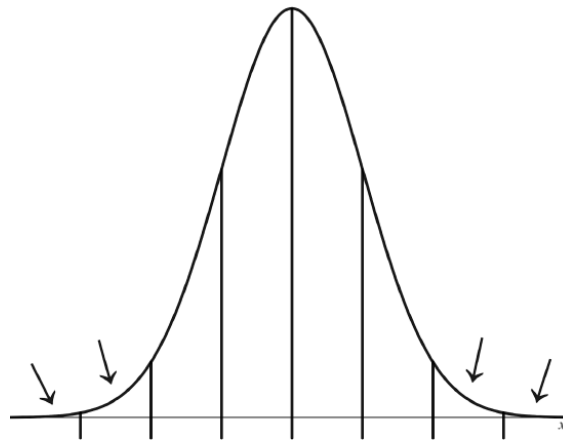
Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

In other words, this tells us that about 99.7% of all values in a normal population lie within three standard deviations of the mean. The remaining 0.3% are split among the more extreme values that lie beyond three standard deviations above or below the mean.

The Empirical Rule tells us that only about 5% of values in a normal distribution are more than two standard deviations from the mean. We will say that any value that is more than two standard deviations from the mean are considered **unusual**.

Computing Probabilities Using the Empirical Rule

2. Recall that adult male heights are approximately normal. The average adult male height is 69 inches and the standard deviation is 3 inches.
 - a. Use this information and what you know about the normal distribution to label the tick marks on the horizontal axis, noting that the distance between each tick mark is 1 standard deviation.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

- b. Remembering that the total area under the probability curve is always 1, find the area of each of the eight regions above.

c. Suppose that x represents a random adult man's height. We want to know the probability that the man is between 72 and 78 inches tall. Using probability notation, this is written as $P(72 < x < 78)$. Use the completed graph above to compute this probability.

d. Find $P(x \geq 75)$

e. Find $P(x < 66)$

f. Find $P(60 \leq x \leq 69)$

g. Find $P(66 < x < 75)$

h. Do we have enough information to find the proportion of adult men who are shorter than 64 inches tall? Explain.

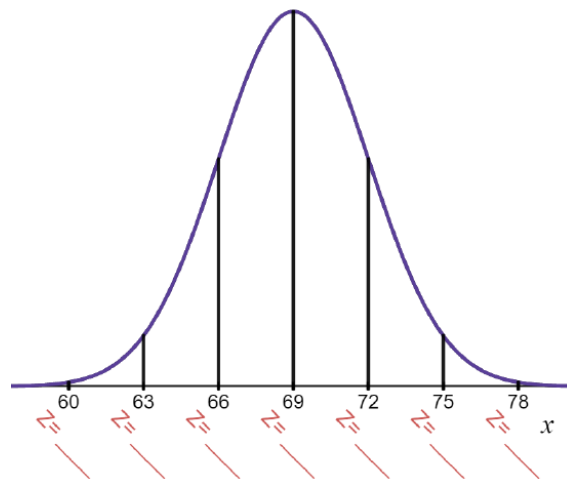
Z-scores

Determining if a value is unusual or not depends on the number of standard deviations it is from the mean. We measure the distance between values using the standard deviation to define the ruler. We define a **Z-score** to be the number of standard deviations a value is from the mean. We compute a Z-score for a given value of x using the following formula:

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

- When Z is negative, x is below the mean.
- When Z is positive, x is above the mean.
- When Z is zero, x is equal to the mean.

3. Label the horizontal axis with Z-scores below each x on the distribution of adult male heights.



Images are created with the graphing calculator, used with permission from Desmos Studio PBC.

4. We have said that a value in a normal distribution is unusual if it is more than two standard deviations from the mean.

- What range of heights belongs to unusually tall men?
- What Z-scores correspond to these heights?
- What range of heights belongs to unusually short men?
- What Z-scores correspond to these heights?

Z-scores provide a standardized way to measure values. For example, when a value has a Z-score of 0.5, we know it is half of a standard deviation above the mean, even if we do not know the values of the mean and standard deviation. Z-scores can be used to compare values from two different populations by measuring their distance relative to each population's mean.

5. A company is hiring a candidate for a high-level research position in a Chemistry lab. Only one position is available. The hiring committee narrows the choice down to two outstanding candidates. Both individuals are highly qualified for the job and have earned a doctorate in Chemistry from two different institutions. Each candidate was required to take a qualifying exam to earn their degrees. Their universities graded the qualifying exams on different scales. Candidate A earned 95 points on their qualifying exam where the average score was 73 and the standard deviation was 9. Candidate B earned 23 points on their qualifying exam where the average score was 18 and the standard deviation was 2.5 points. Using this information and Z-scores, who should be recommended for the position? Justify your answer.

This page titled [5.2: Characteristics of the Normal Distribution and The Empirical Rule](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Hannah Seidler-Wright](#).

- [Current page](#) by Hannah Seidler-Wright is licensed [CC BY-NC-SA 4.0](#).
- [1.2: The Statistical Analysis Process](#) by Hannah Seidler-Wright is licensed [CC BY-NC-SA 4.0](#).