MATH 3307
Lesson 14

The Normal Distribution

A density curve that is symmetric, single peaked and bell shaped is called a **normal distribution**.
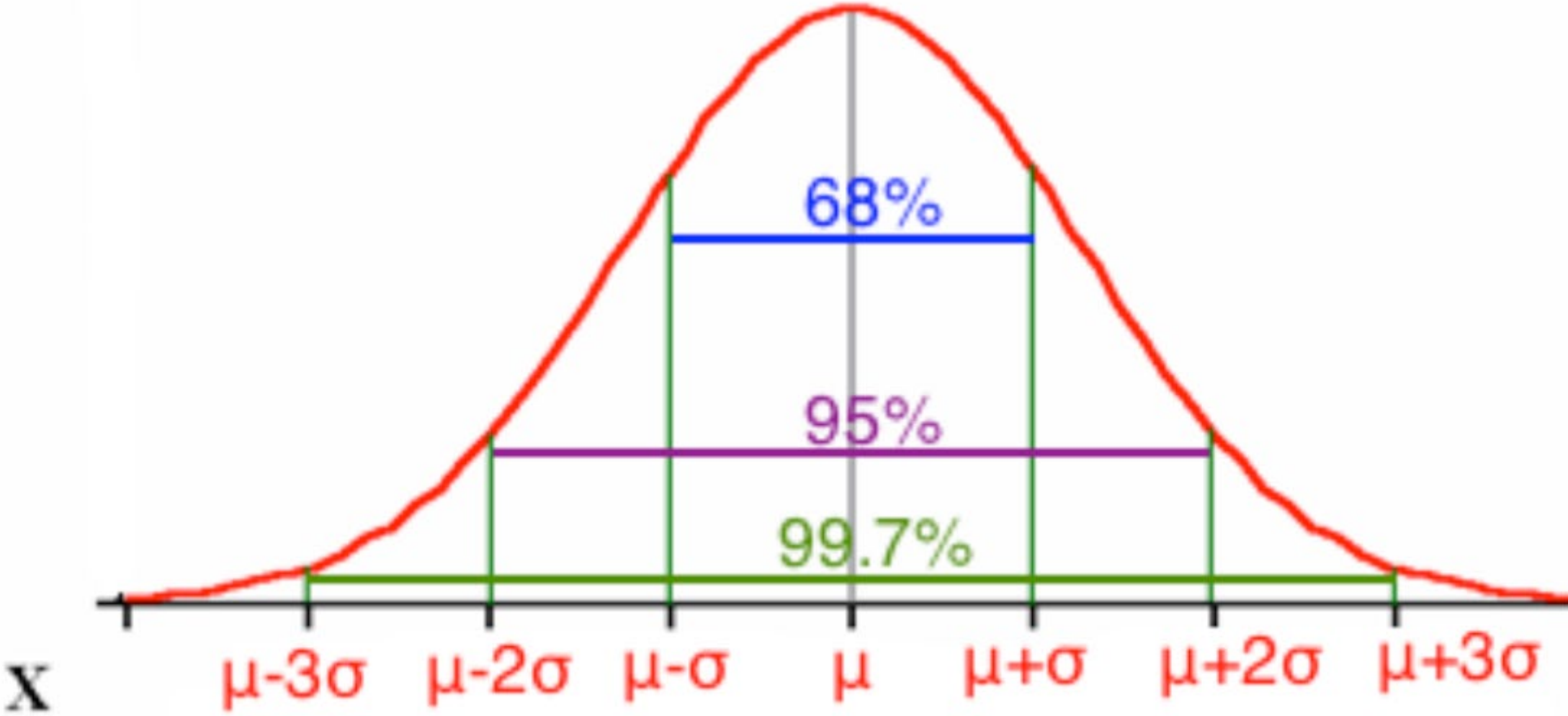
The normal distribution with mean μ and standard deviation σ is represented by **N(μ, σ)**.

The Empirical Rule:

The Empirical Rule states if a distribution has a normal distribution,

1. Approximately 68% of all observations fall within one standard deviation of the mean.

2. Approximately 95% of all observations fall within two standard deviations of the mean.

3. Approximately 99.7% of all observations fall within three standard deviations of the mean.
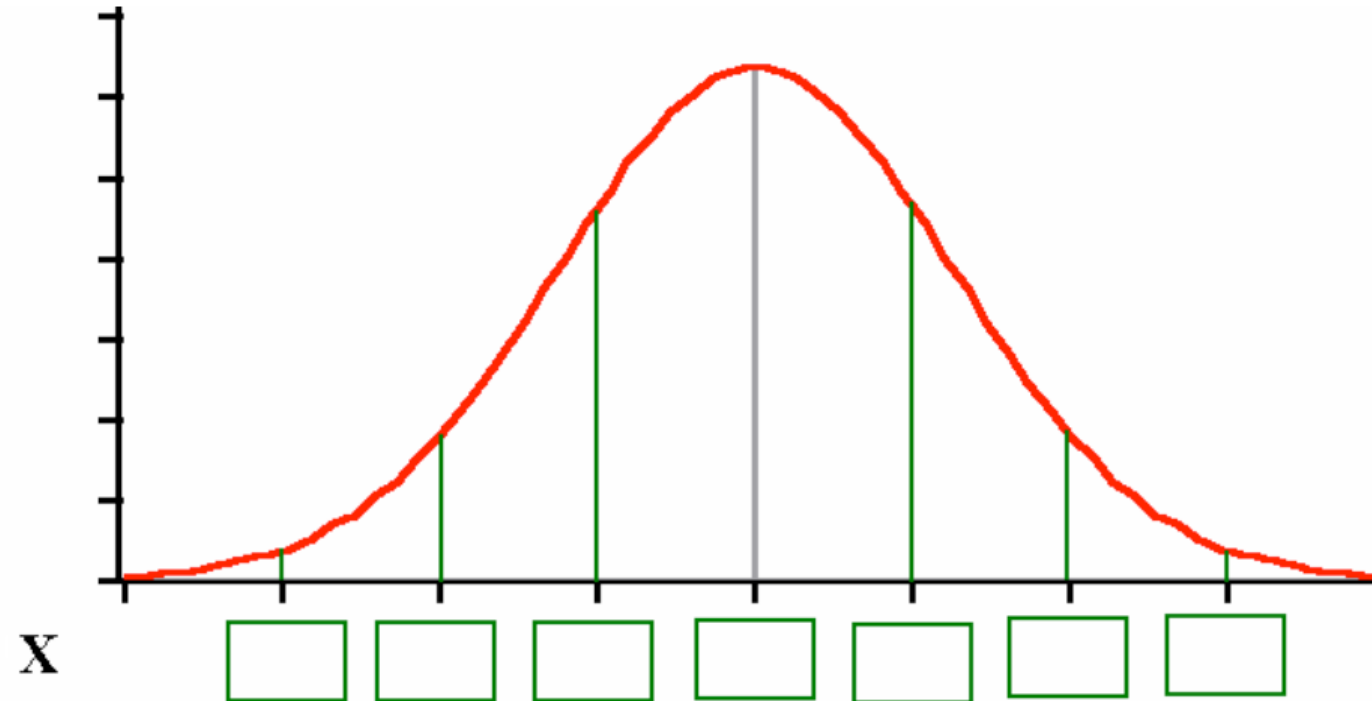
The Empirical Rule:

Example:

The length of time needed to complete a certain test is normally distributed with mean 60 minutes and standard deviation 10 minutes.

What is the probability that someone will take between 40 and 80 minutes to complete the test? Sketch the distribution and shade in the area in question.

Example:

The length of time needed to complete a certain test is normally distributed with mean 60 minutes and standard deviation 10 minutes.

What is the probability that someone will take between 40 and 80 minutes to complete the test? Sketch the distribution and shade in the area in question.

Example:

The length of time needed to complete a certain test is normally distributed with mean 60 minutes and standard deviation 10 minutes.

What is the probability that someone will take between 40 and 80 minutes to complete the test? Sketch the distribution and shade in the area in question.

Find the interval that contains the middle 68% of completion times for all people taking the test.

What percent of people take more than 80 minutes to complete the test?

X
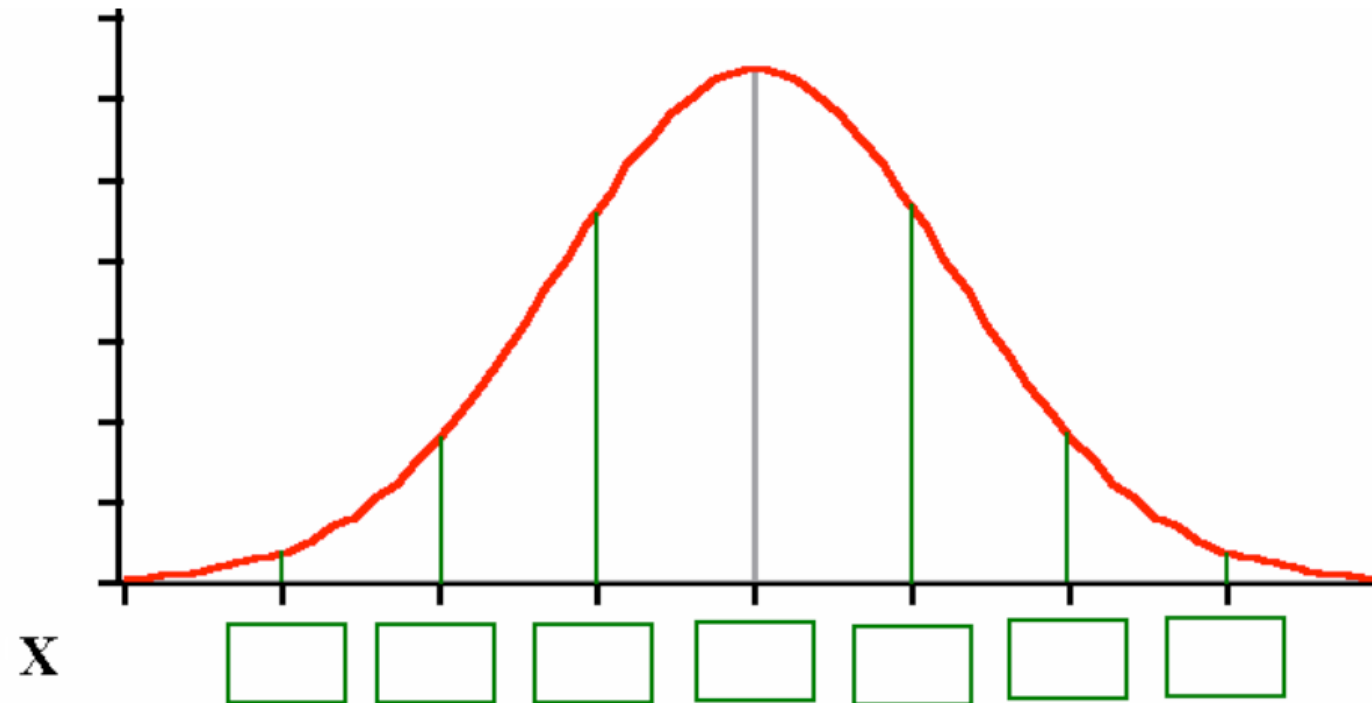
Other Probabilities:

What if our values are not exactly within one, two or three standard deviations from the mean? Probabilities for these can still be found a number of ways, one of which we will explore in the next section. Using R, the probability can be found with the command

$$\text{pnorm}(X, \mu, \sigma).$$

Note that the command in R only gives the probability that X is less than a given value. If we need to find the probability that X is greater than the given value, we will need to subtract the answer from 1.

With the TI-83 and TI-84 calculator, the command is normalcdf(lower_limit, upper_limit, μ, σ).

What is the probability that someone will take less than 45 minutes to complete the test?

What is the probability that someone will take more than 30 minutes to complete the test?

How long would it take someone to finish the test if they are in the top 10% of the times? (Approximate Answer)

Another Example:

The Rstudio dataset, faithful, contains data on the duration of eruptions (eruptions) and waiting time between eruptions (waiting) of Old Faithful Geyser in Yellowstone National Park.

Assuming that the wait time falls into a normal distribution, what is the probability that you will have to wait (a) less that 60 minutes, (b) between 45 and 75 minutes, and (c) longer than 120 minutes, to see an eruption?

Is the following a Normal Distribution?
Why or Why Not? *(Both graphs have a total enclosed area of 1)*

f(x)
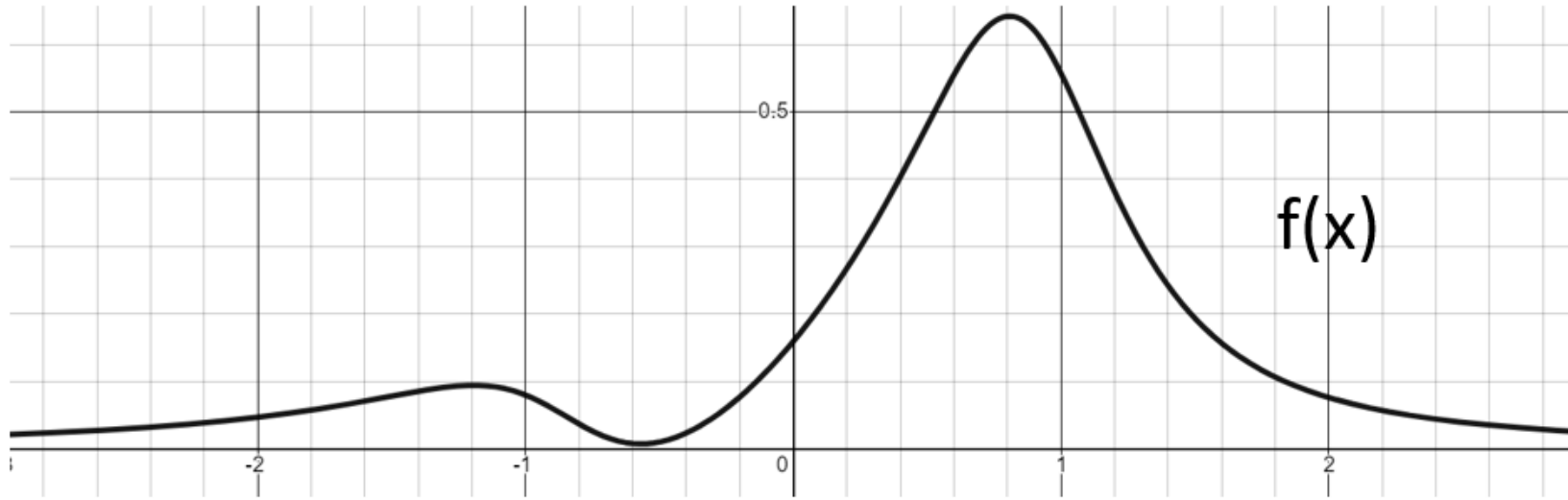
g(x)

Mean: 0; Standard Deviation: 1

Area from x = 0 to x = a:

$$\frac{1}{\pi}\tan^{-1} a$$