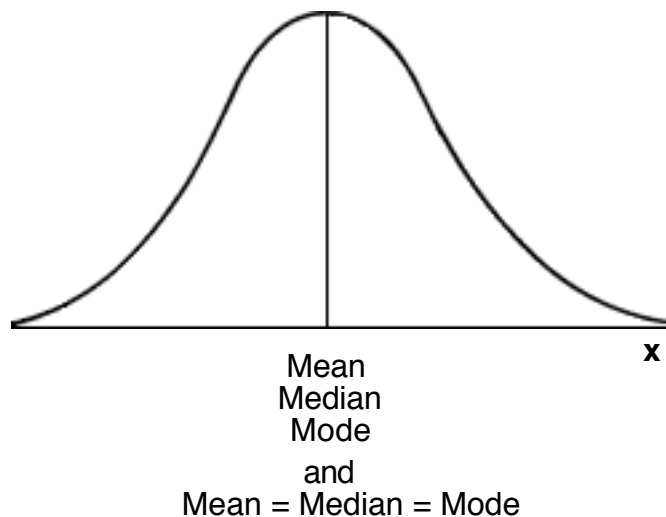


Section 6 – 2A: Continuous Normal Probability Distributions Introduction

The last section listed several facts about Continuous Probability Distributions. The graph of each distribution is a smooth unbroken curve above the x axis. **The total area under the curve is equal to 1. The value of any $P(x)$ for any single value of x is $0+$.** This section will develop the use of bell shaped curves called **Normal Curves** or **Normal Distributions** to answer questions involving many common probability questions.

A continuous probability distribution that is **normal** has the **mean, median and mode equal to each other** and they all **occur at the center** where the maximum height occurs. The center of the graph is the line of symmetry for the graph.



Important Note:

In most of the remaining sections of this course you **cannot use the procedures introduced in that section** unless the data is **normally distributed**. The requirements for many sections start with the requirement that the population data is **normally distributed**. Each section that has this requirement will have a rule that will help you determine if the data is normally distributed or not.

Normal Probability Distributions

Every data set that has a normal distribution is represented by a normal curve. Some normal curves are tall and skinny and some are short and more spread out.



Parameters

The general form of an equation of a line is $y = mx + b$ where m is the slope and b is the y-intercept. The values of m and b cause the line to cross the y axis at different points and a different steepness. We call m and b the **parameters for a line**. For any specific data set the values of m and b completely describe the line.

Any normal distribution can be completely described by two parameters: the mean μ_x and the standard deviation σ_x . **If the mean and standard deviation are known, then you essentially know as much about the data set as if you had access to every point in the data set.**

All Normal Probability Distribution Curves have a general equation

$$y \approx \frac{(2.72) \frac{-(x-\mu)^2}{2\sigma^2}}{2.51 \cdot \sigma}$$

The **variables** for the equation above are x and y .

The **parameters** for the equation above are μ_x and σ_x

The values of μ_x and σ_x determine the shape of the Normal Probability Distribution Curve

For any specific data set the values for the mean μ_x and the standard deviation σ_x are known or can be calculated. For each of these normal curves, the **exact shape and location** of the curve on the x axis **depends on the values of the mean μ_x and standard deviation σ_x** that represent the data set. The two parameters mean μ_x and standard deviation σ_x act like the parameters m and b for linear equations. They determine the **placement** and **steepness** of the graph.

The value of **the mean** will determine **where on the x axis the high point is**.

The value of **the Standard Deviation** will determine how **“wide” or “skinny”** the curve is.

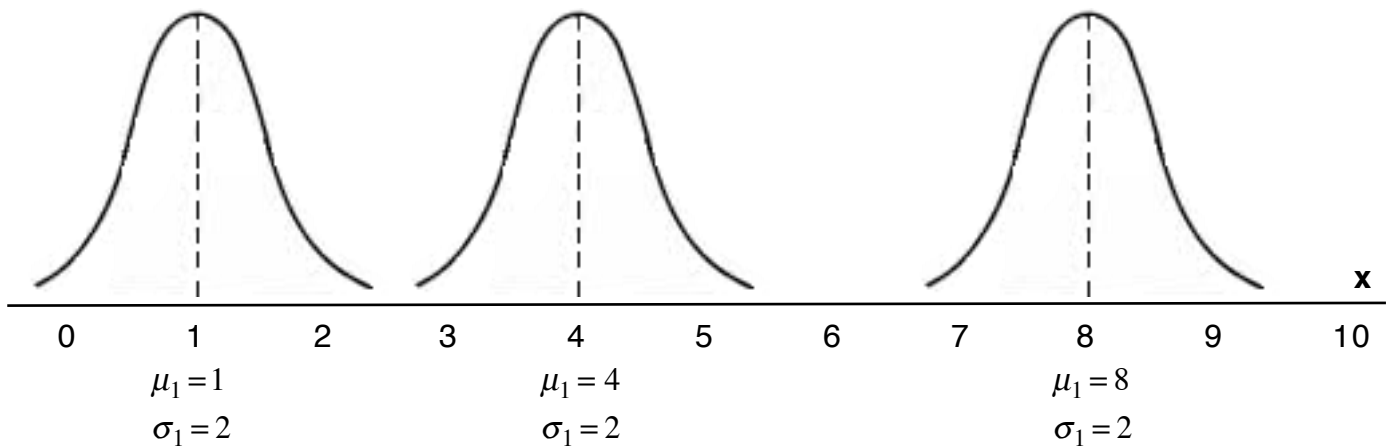
The value for **the mean**
determines the location on the x axis
of the highest point of the graph.

The parameter μ_x is the value of the mean for a specific data set that the curve represents. **The value of the mean tells us where on the x axis the highest point or peak on the graph will be located.** It does not tell us the height of the peak only that the location of the peak will be at $x = \mu_x$ on the x axis. The larger the value of the mean the farther to the right of 0 on the x axis the peak is located.

Each of the data sets below has a different mean. $\mu_1 = 1$ and $\mu_2 = 4$ and $\mu_3 = 8$.

The location (right or left) on the x axis of each of the 3 curves is different because each of the data sets has **a different mean**.

Each of the normal curves below has the same “shape” because all three of the data sets have the **same standard deviation of 2**.



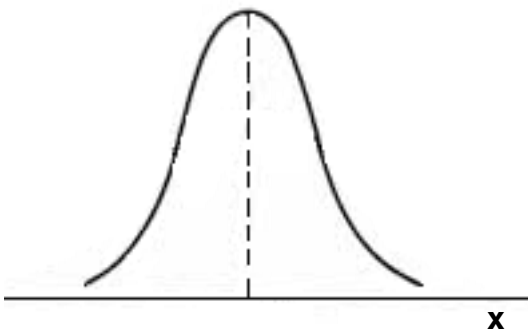
The value of the Standard Deviation determines the shape of the normal curve
“tall and skinny” or “short and spread out”

The parameter σ_x is the value of the standard deviation for a specific data set. The value of σ_x tells us how spread out the data is.

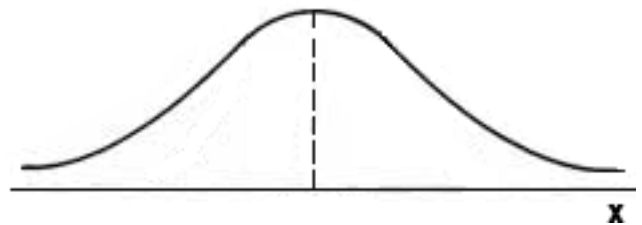
If the value of the **Standard Deviation is SMALL** then the graph is tall and skinny. **MOST of the data is very close to the mean.** The data close to the mean occurs very frequently and data farther from the mean occurs less frequently. This will create a normal curve that has a high narrow peak and a rapid decline in the height of the curve towards the tails on both ends.

If the value of the **Standard Deviation is LARGE** then the graph is short and more spread out. **LESS of the data is very close to the mean.** The data close to the mean occurs less frequently than the curve shown above and data farther from the mean occurs more frequently than the curve shown above. This will create a normal curve that has a lower narrow peak and a gradual decline in the height of the curve towards the tails.

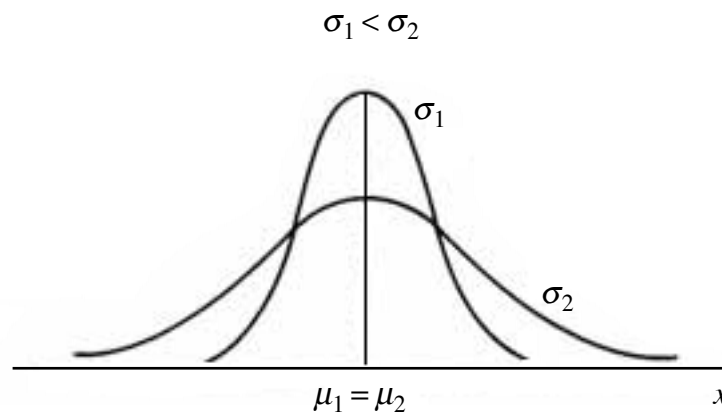
Standard Deviation σ_x is Small
tall and skinny



Standard Deviation σ_x is LARGE
short and more spread out

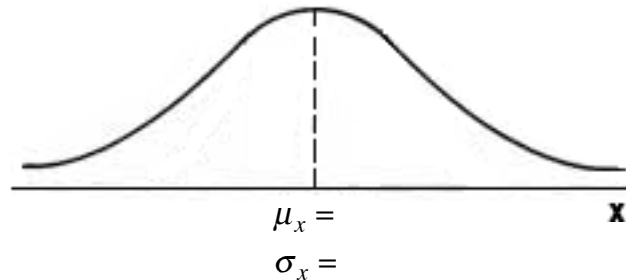


Consider the two normal curves below that are graphed on the same x axis. Their **means occur at the same location** on the x axis so the data sets have the **same mean** $\mu_1 = \mu_2$. The tall thin normal curve has a standard deviation of σ_1 and the short spread out normal curve has a standard deviation of σ_2 . The curve that is tall and skinny will have a **smaller** standard deviation than the curve that is short and spread out.



Properties of a Normal Curve

A data set with data that is normally distributed can be represented by a bell shaped curve called a **Normal Curve**. The normal curve has an x axis that represents all of the possible values in the data set. The data set has a mean and standard deviation. The values for these parameters are found by the use of a calculator. We show the mean and standard deviation on the graph at the point on the x axis that is directly under the high point (or peak) of the graph.



1. The mean, median and mode for the data are equal to each other. mean = median = mode
2. The mean occurs at a point on the x axis directly under the high point of the graph.
3. The value of the Standard Deviation determines the shape(tall and skinny) or (short and spread out) of the normal curve. As the standard deviation gets smaller the curve gets taller and skinnier. As the standard deviation gets larger the curve gets shorter and more spread out.
4. The x axis is a horizontal asymptote on both the right and left ends of the graph. As the x values increase on the right side of the graph the curve gets closer and closer to the x axis **but never touches the x axis**. As the x values decrease on the left side of the graph the curve gets closer and closer to the x axis **but never touches the x axis**.
5. The curve is symmetric about the line $x = \mu_x$
6. **The area under the curve is 1 for all normal curves.**
7. 68% of the area under the curve is within 1 standard deviation of the mean. (between $\mu \pm 1\sigma$)
95% of the area under the curve is within 2 standard deviations of the mean. (between $\mu \pm 2\sigma$)
99.7% of the area under the curve is within 3 standard deviation of the mean. (between $\mu \pm 3\sigma$)
8. 99.96% of the data falls within 3.5 standard deviation of the mean. The tables we use for this course cut the data off after 3.5 standard deviations.
9. 99.9997% of the data falls within 6 standard deviation of the mean. This means that there is almost no data in either tail after you get beyond 6 standard deviations from the mean. Cutting the curve off at 6 standard deviations from the mean still represents almost all the data.