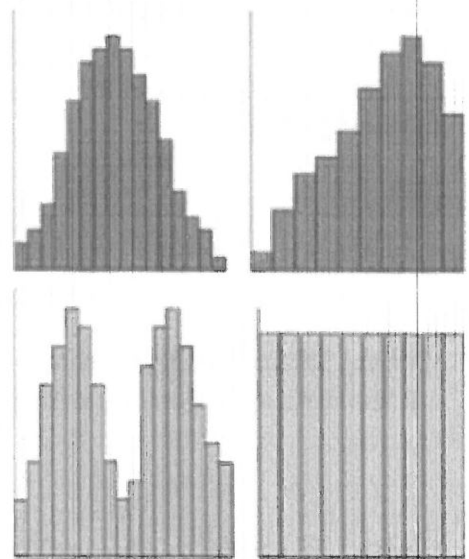
5.4 The Normal Distribution (Concept #27 and 28)

Once we organize data into a histogram, we can see how the data has been distributed.



Looking at the left the first (purple) histogram represents **normal distribution.** Can you list any characteristics of this histogram? like a hill

- What causes the difference between the graph with one peak vs the one with two peaks?

  One mode vs 2 modes in the data

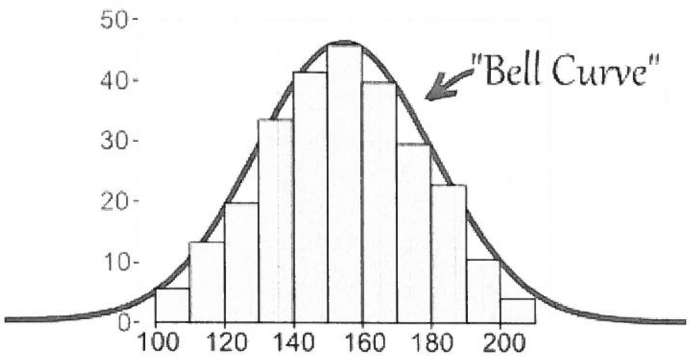
- What causes the second graph to have its one peak not in the center but on the right side?

  The data is skewed to the right

- What causes the last graph to be flat?

  There is no fluctuation in the data

But there are many cases where the data tends to be around a central value with no bias left or right, and it gets close to a "Normal Distribution" like this:

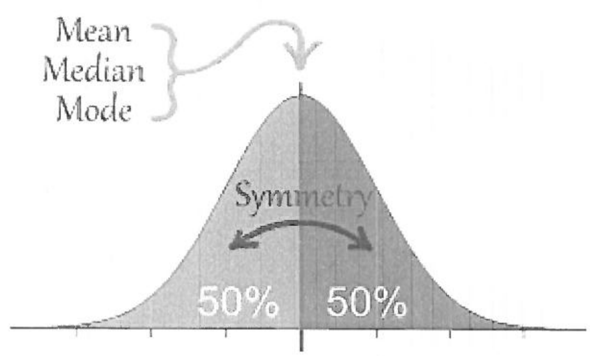


It is often called a "Bell Curve" because it looks like a bell.

The "Bell Curve" is a Normal Distribution. Many things closely follow a Normal Distribution: example : Heights of people, blood pressure and marks on a test

The Normal Distribution has:

- mean = median = mode (or close to equal)
- symmetry about the center
- the mean is the axis of symmetry
- 50% of values less than the mean and 50% greater than the mean

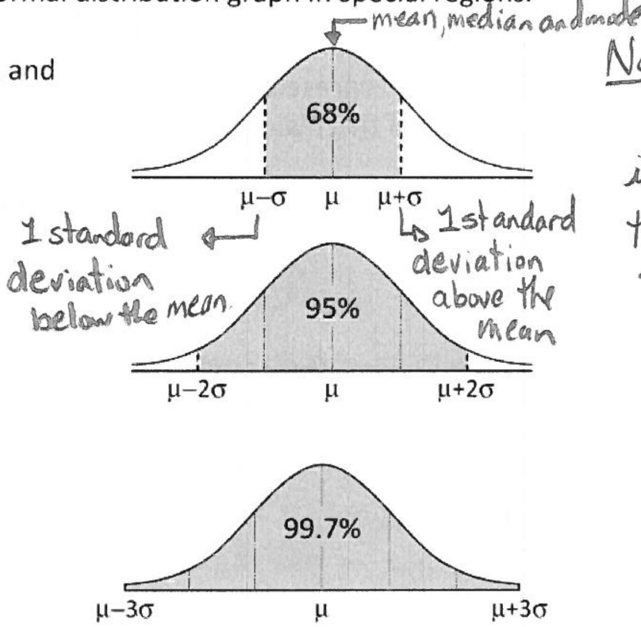


68-95-99 RULE

This rule estimates areas under the normal distribution graph in special regions.

For every normal curve with mean, σ, and standard deviation , σ.:

- About 68% of the data is within 1 standard deviation of the mean.
- About 95% of the data is within 2 standard deviation of the mean.
- About 99.7% of the data is within 3 standard deviation of the mean.



Note: μ (pronounced mu) is a symbol used for the mean of the entire population

$\bar{X}$ = sample mean

The graph of a normal distribution depends on two factors:

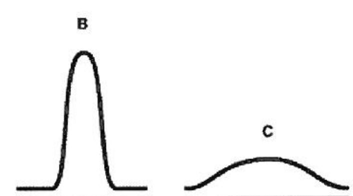
- the mean, μ, which is a measure of central tendency and
- the standard deviation, σ, which is a measure of dispersion.

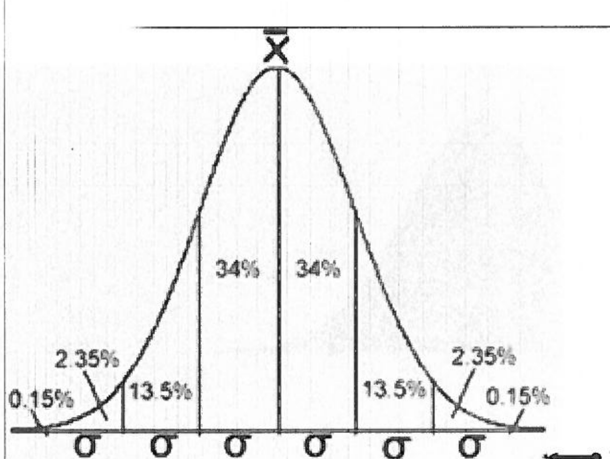
Remember, the total area under the graph must be 1. or 100%

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})}{n}}$$

Note: The standard deviation determines the width of the graph

- The standard deviation determines the width and height of the curve.
- Larger S.D. means larger intervals and a shorter wider graph (graph C)
- Smaller S.D. means smaller intervals and a taller, more narrow graph (graph B)





Standard deviation away from the mean

~~The normal distribution has certain properties:~~

1) ~~68% of the data are within one standard deviation of the mean.~~
2) ~~95% of the data are within two standard deviations of the mean.~~
3) ~~99.7% of the data are within three standard deviations of the mean.~~

Example #1:

Shirley wants to buy a new cell phone. She researches the cell phone she is considering buying and finds the following data on its longevity in years.

2.0	2.4	3.3	1.7	2.5	3.7	2.0	2.3	2.9	2.2	2.3	2.7	2.5	2.7	1.9	2.4	2.6	2.7
2.8	2.5	1.7	1.1	3.1	3.2	3.1	2.9	2.9	3.0	2.1	2.6	2.6	2.2	2.7	1.8	2.4	2.5
2.4	2.3	2.5	2.6	3.2	2.1	3.4	2.2	2.7	1.9	2.9	2.6	2.7	2.8				

$n=50$

a) Does the data approximate a normal distribution?

Step 1- Find the mean, median and mode. (Use Calculator)

$\bar{x}=2.53$ Median $=2.55$ Mode $=2.7$

Step 2- Find the standard deviation(Use Calc)

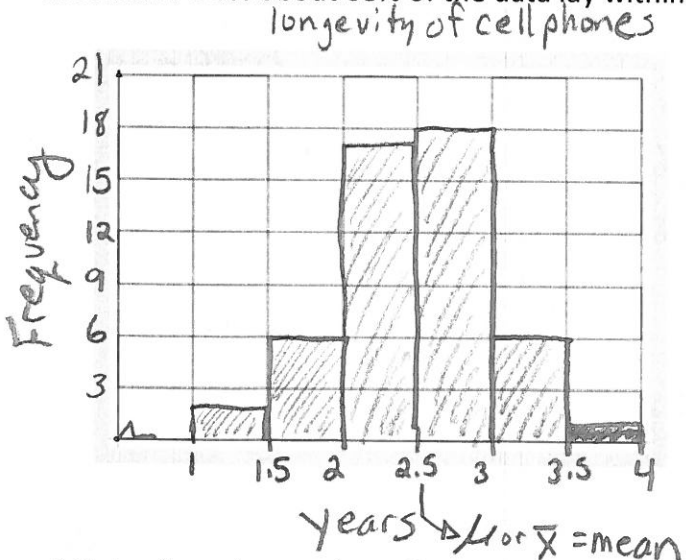
$\sigma=0.5$

Step 3- Create a frequency table, using the standard deviation as the interval width

Longevity (years)	Tally	Frequency
1 - 1.5	II	2
1.5 - 2	~~IIII~~ I	6
2 - 2.5	~~IIII~~ ~~IIII~~ ~~IIII~~ II	17
2.5 - 3	~~IIII~~ ~~IIII~~ ~~IIII~~ III	18
3 - 3.5	~~IIII~~ I	6
3.5 - 4	I	1

Note: Normal Distribution
Mean = Median = Mode
These are all very close to equal

Step 4: Decide if the data in the table would represent a Bell Curve on a histogram. Things to check before you can conclude: a) Does it increase up until the mean and then decrease on the other side? b) Is it roughly symmetrical abouth the mean? c) Does about 68% of the data lay within one standard deviation of the mean($\pm\sigma$)?



a) Approximately yes

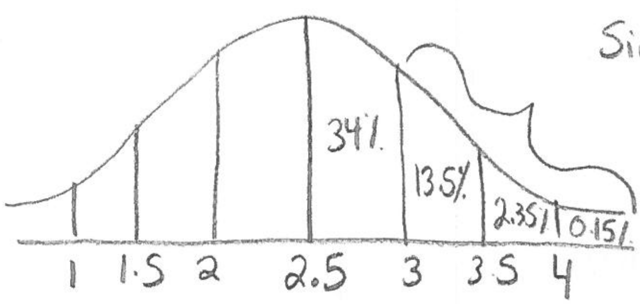
b) Yes

c)
$$\underset{-1\sigma}{\overset{2}{|}}\text{—17—}\underset{\bar{x}}{\overset{2.5}{|}}\text{—18—}\underset{+1\sigma}{\overset{3}{|}}$$

$$\frac{18+17}{50}=\frac{35}{50}=70\%$$

$\therefore$ approx. 68% of the data is $\pm 1\sigma$ of the mean. Therefore the data will approx reflect a normal distribution.

d) If she does choose this cell phone what are the chances that it will last more than 3 years?



Since it follows a normal distributions

$13.5+2.35+0.15$

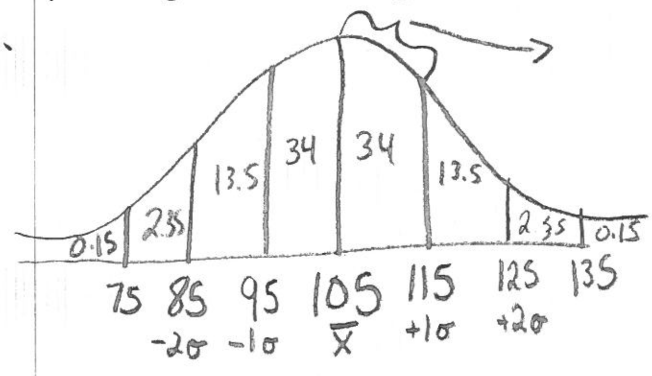
$=16\%$ chance the cell phone will last more than 3 years

Example #2:

The speeds of 1000 cars were recorded by photo radar. If the data collected was normally distributed with mean of 105 km/h and a standard deviation of 10 km/h, determine:

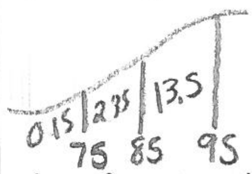
$\bar{x} = 105$ Km/h $\sigma = 10$ Km/h

a) The percentage of cars traveling between 105 and 115 km/h



34% of cars were recorded travelling between 105-115 Km/h

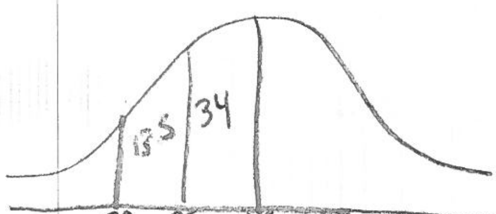
b) The percentage of cars traveling less than 95 km/h



$13.5 + 2.35 + 0.15$

= 16% of cars were recorded travelling less than 95 Km/h

c) <u>Number</u> of cars traveling between 85 and 105 km/h



① % of cars = $13.5 + 34$ = 47.5%

② # of cars

47.5% of 1000

0.475×1000

= 475 cars

∴ 475 cars were travelling between 85 - 105 Km/h.

5.4 Assignment Pg 251 # 1-4 ,6, 10, 11, 13

1, 3, 4 $\boxed{\text{Tell student no \#2}}$

In Summary

- Graphing a set of grouped data can help you determine whether the shape of the frequency polygon can be approximated by a normal curve.
- You can make reasonable estimates about data that approximates a normal distribution, because data that is normally distributed has special characteristics.
- Normal curves can vary in two main ways: the mean determines the location of the centre of the curve on the horizontal axis, and the standard deviation determines the width and height of the curve.

- The curve is called a bell curve
- The curve is symmetrical
- The mean, median and mode are equal (or close) and fall at the center of the curve
- The total area under the curve is 1
- Generally, measurements of living things (mass, height and length) have normal distribution
- Normal distribution can be helpful in answering probability questions