Notes on Normal Distribution

Measures of Central Tendency

- **Mean**: the balance point for a set of data. It is found by taking the sum of the data divided by the number of items in the data set (the average of the data).

"new" μ represents a population mean and $\bar{x}$ represents a sample mean.

- **Median**: the middle number of the ordered data, or the mean of the middle 2 numbers.
- **Mode**: the number(s) that occur most often.

***Example A*: Given the data set: {5, 7, 7, 7, 3, 11, 7, 7, 7, 9, 7, 7}, find each.** *(Does this problem look familiar?)*

a) Find the mean.

$\frac{84}{12} = 7$

b) Find the median. 7

c) Find the mode. 7

d) What did you notice about your findings?

they are all the same

e) What is the name of a data set with this type of results?

Normal Distribution

Standard Deviation

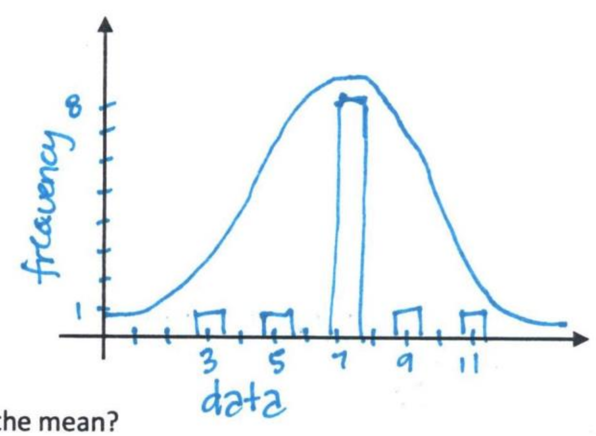
- **Standard Deviation** (σ) "sigma" describes how closely a set of data clusters about the mean.

The greater the σ, the more spread out the data is about the mean.

The less (closer to 0) the value of the σ, the more clustered the data is about the mean.

***Example B*: For the same data set as Example A: {5, 7, 7, 7, 3, 11, 7, 7, 7, 9, 7, 7}, do the following.**

a) Construct a histogram of the data.



b) Using a calculator (see notes on the back of this sheet), find the standard deviation of the data set.

$\sigma = 1.8257... \approx 1.83$

c) Which values are within 1 standard deviation of the mean?

None

d) Are any data values more than 2 standard deviations from the mean?

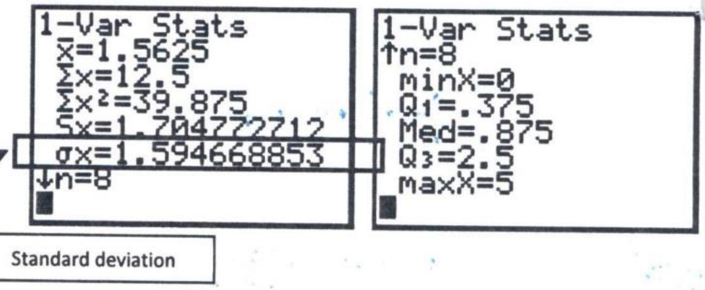
3 & 11

CALCULATOR - To find mean, standard deviation, median, etc. for a list of data:

STAT > edit > L_1 > ENTER (enter the data)

STAT > calc > 1-var Stats > ENTER > ENTER

Example: {0, 0.5, 1, 5, 0.75, 2, 0.25, 3}



ex on front

$\bar{x} = 7$ $\sigma x = 1.8257$

$\Sigma x = 84$ $n = 12$

$\Sigma x^2 = 628$ $\min x = 3$

$Sx = 1.9069$ $Q_1 = 7$

med = 7

$Q_3 = 7$ max = 11

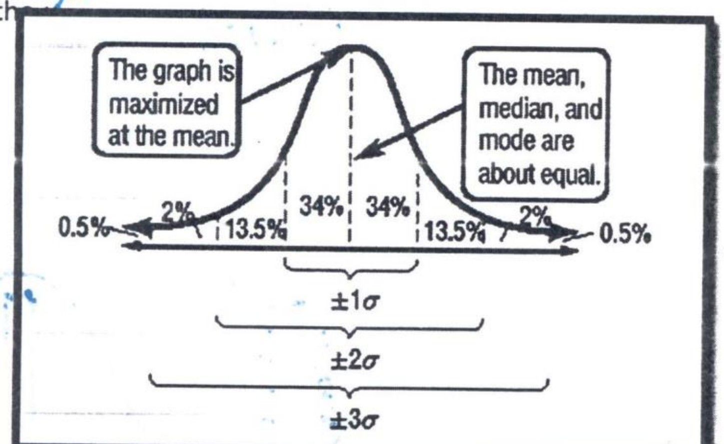
5 Characteristics of a Normal Distribution

1. The mean, median, and mode are always equal to each other.
2. The graph is called a normal curve and the maximum occurs at the mean.
3. A normal curve is bell shaped and symmetrical about the mean.
4. A normal curve never touches, but it gets closer and closer to the ____ - __________ as it gets farther from the mean.
5. Because the area under the normal curve represents probabilities, this area will always be one.

FYI: The Standard Normal Curve has a mean of 0 and a standard deviation of 1.

Empirical Rule 68-95-99

- **Empirical Rule 68-95-99** states that approximately 68 % of the data fall within 1σ of the mean(μ), approximately 95 % of the data fall within 2σ of the μ, and approximately 99 % of the data fall within 3σ of the μ.



Positively Skewed Distribution

Negatively Skewed Distribution

- **Z-score:** the measure of how many standard deviations an element falls above or below the mean of a set of data.

$$z - score = \frac{X-\mu}{\sigma} \quad \ldots X \text{ is an element of the data set}$$

show ex. on front w/ a bell curve

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

***Example C*: A data set has a mean, μ = 10, and a standard deviation, σ = 2. Answer the following questions.**

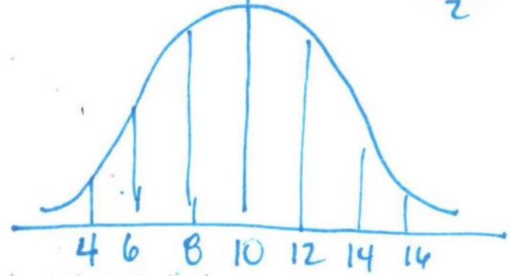
a) Why would a data value of 12 have a z-score of 1 and a data value of 8 have a z-score of -1?

1 is to the right of the mean & -1 is to the left

$\frac{12-10}{2} = \frac{2}{2} = 1$

$\frac{8-10}{2} = \frac{-2}{2} = -1$

b) What z-score would be assigned to 6? -2 14? 2



c) What data value would have a z-score of -3? 4 4? 18

$4 = \frac{x-10}{2}$ $\quad 8 = x - 10$ $\quad 18 = x$

d) Write a formula to determine the z-score for any value, X, in this data set.

$z = \frac{x-10}{2}$ or $2z = x - 10$ or $2z + 10 = x$

***Example D*: Amy scored a 31 on the mathematics portion of her 2009 ACT® (μ=21 σ=5.3). Stephanie scored a 720 on the mathematics portion of her 2009 SAT® (μ=515 σ=116.0). Whose achievement was higher on the mathematics portion of their national achievement test?**

Amy

$x = \frac{31-21}{5.3} = 1.89$

Stephanie

$x = \frac{720-515}{116} = 1.77$

*Amy - she was further to the right of μ!

Area Under a Normal Curve

- **Population Density Function**: used to graph the bell curve.
- **Normal Cumulative Density Function**: used to calculate probabilities.
- **Finding area under a normal curve**: Area can be found under a normal curve using the 68-95-99 rule if the areas are bounded at places where an exact standard deviation occurs. Areas that aren't bounded at specific standard deviation units can be found using a calculator or a z-table.
- ***CALCULATOR*** - **To calculate the area under a normal curve**:

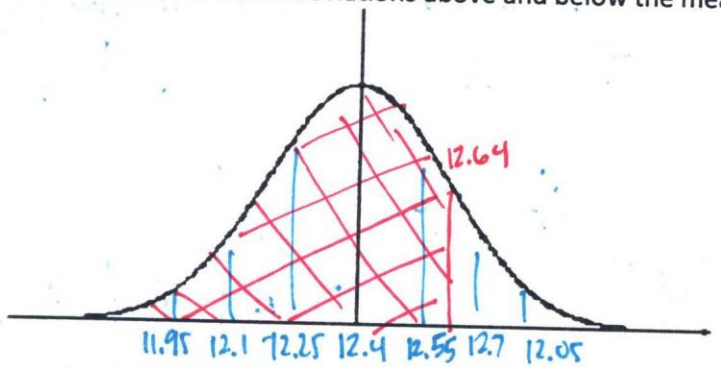
2^{nd} VARS > normalcdf (L, U, M, S) > ENTER

L, U, M, S
Lower limit of the random variable,
Upper limit of the random variable,
Mean of the distribution,
Standard deviation of the distribution

Hint: use -99999999 and 99999999 as needed for really small and really large limits!

Example E: A corn chip factory packs chips in bags with normally distributed weights with a mean of 12.4 oz and a standard deviation of 0.15 oz.

a) Label the mean and 3 standard deviations above and below the mean on the graph provided.



b) On the curve in a), shade the region that indicates the percentage of bags that contain less than 12.64 oz.

Red shaded region

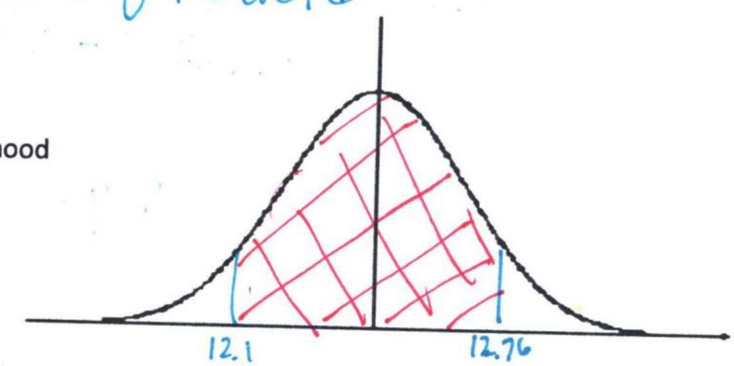
c) Determine the z- score corresponding to 12.64 oz.

$$z - score = \frac{X-\mu}{\sigma} \qquad z = \frac{12.64-12.4}{0.15} = 1.60$$

d) Use the Standard Normal Probabilities Table to find the area associated with the z-score obtained in c and interpret your result.

- 1.60 = .9452 = 94.52 % of the data

e) Label and shade the region that represents the likelihood a bag will contain between 12.1 and 12.76 oz.



f) Calculate the z-scores corresponding to both 12.1 and 12.76 and find the Standard Normal Probabilities for each using a calculator and the Standard Normal Probabilities Table.

$z = \frac{12.76-12.4}{.15} = 2.4$ table .9918

$z = \frac{12.1-12.4}{.15} = -2$ table .0228

g) Determine how you would use those values to determine the probability a bag chosen at random would contain between 12.1 and 12.76 chips. $P(12.1 < X < 12.76)$