UNIT 4: STATISTICS

Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/h9194WdG8zw

Statistics is the name given to the branch of applied mathematics concerned with the collection, analysis, and interpretation of numerical data.

The collection of data, by various methods, has been covered in previous math courses. Data can be collected from a **population** or from a **sample** of the population.

Once the data has been collected, it can be analyzed in a variety of ways. In previous courses, we have represented a collection of data by some kind of "average" value.

For example, when students receive their unit test marks, they often ask the teacher, "What was the class average?"

Measures of Central Tendency

Once data has been collected, it is useful to be able to describe the data by one single central value.

In previous courses we have studied three different types of central value or "average".

The **mean, median,** and **mode** are measures of central tendency.

The **mean, median,** and **mode** are measures of central tendency.

Measures of Central Tendency

Mean

- The mean of a **population** is denoted by the Greek symbol μ (mu).
- The mean of a **sample** is denoted by the symbol $\bar{x}$.
- The mean of a population and the mean of a sample are both calculated using the same method.
- The mean is calculated by adding all the data values and then dividing by the number of data values.

Median

- The median is determined by first sorting the data into numerical order.
- If there are an odd number of data values, the median is the middle data value.
- If there are an even number of data values, the median is the mean of the two middle data values.

Mode

- The mode is the data value which occurs most often, i.e. it has the highest frequency of occurrence.
- It is possible to have multiple modes or no mode.

Example 1:

The data below represents the time taken, to the nearest minute, for a high school student to drive to school on each of the last ten school days.

11 10 19 16 15 14 13 12 17 15

a) Calculate the mean driving time for the data.

b) State the mode of the data.

Example 2:

Hideki's unit test marks in English 20 and Math 20 are given below.

English 20	67	71	58	78	65	80	78
Math 20	74	78	61	86	84	72	

Calculate his median mark in each of the subjects.

Example 3:

If the mean of the data $10, 4, 7, 12, 9, 4, x,\ 6$ is seven, find the value of x.

Example 4:

The data below represents the number of touchdowns thrown by Jason, the Centre High School quarterback, during his high school football career.

0 1 0 2 1 2 0 1 2 2 1 3 1 2 1 2 1 2 3

a) Calculate, to one decimal place if necessary, the mean, median, and mode for the number of touchdowns per game.

b) Suzanne drew a frequency table to represent the data.

Number of Touchdowns per Game	Number of Games
0	3
1	7
2	7
3	2

Show how Suzanne can use the data in the frequency table to calculate the mean, median, and mode for the number of touchdowns per game.

Calculating Measures of Central Tendency Using the STAT Feature

A graphing calculator can be used to determine the mean and median of a set of data values. Use the following method for a TI-83/84 Plus.

We will use the following data values as an illustration.

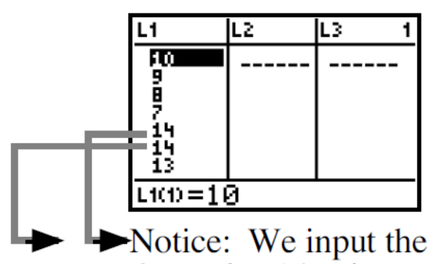
10 9 8 7 14 14 13 5 11 16

1. Clear all previous lists by pressing $\boxed{\text{2ND}}$ $\boxed{+}$ $\boxed{4}$ and then $\boxed{\text{ENTER}}$.

2. Access the list feature by pressing $\boxed{\text{STAT}}$ and then "Edit".

3. Enter the data values in list L_1 .

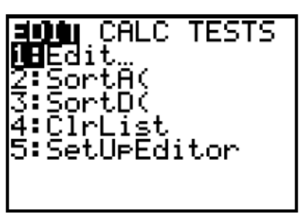
4. Exit the lists using the "QUIT" command.

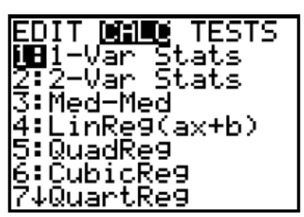


Notice: We input the data value 14 twice.

5. To calculate the descriptive measures

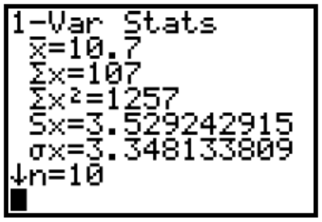
- press $\boxed{\text{STAT}}$ then scroll to "CALC".

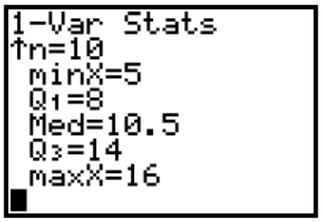




- select the command "1-Var Stats" and press $\boxed{\text{ENTER}}$.

Scroll up or down to view the descriptive measures.



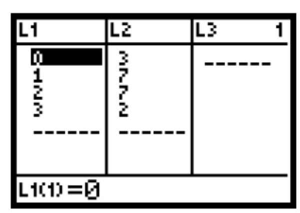


- The mean is represented by "$\bar{x}$ ".
- The median is represented by "Med".
- The calculator does not calculate the mode.

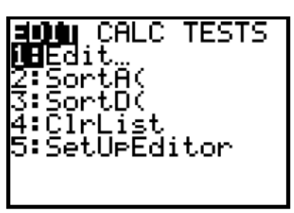
The following procedure is used when the data is represented in a frequency table.

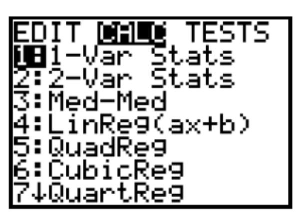
Number of Touchdowns per Game	Number of Games
0	3
1	7
2	7
3	2

1. Clear all previous lists by pressing [2ND] [+] [4] and then [ENTER].
2. Access the list feature by pressing [STAT] and then "Edit".
3. Enter the data values (Number of Touchdowns per Game) in list L_1, and the frequencies (Number of Games) in list L_2.



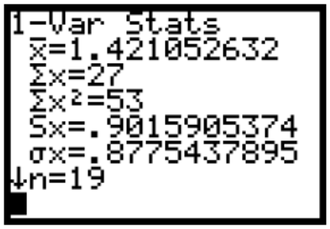
4. Exit the lists using the "QUIT" command.
5. To calculate the descriptive measures,
   - press [STAT] then scroll to "CALC".

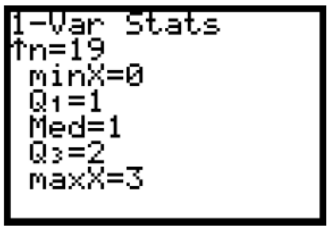




   - select the command "1-Var Stats", type in L_1, L_2 and press [ENTER].

6. Scroll up or down to view the descriptive measures.





- The mean is represented by "$\bar{x}$".
- The median is represented by "Med".
- The calculator does not calculate the mode.

Assignment:

1.

The minimum daily temperatures in Edmonton during an 11 day period are shown.

8.6, 7.3, 10.7, 15.2, 9.3, 8.6, 7.3, 8.5, 7.3, 5.9, 1.0

Calculate, to the nearest tenth where necessary, the mean, median, and mode.

2.

In the annual teachers vs. students golf challenge, the scores of the ten teachers were

74, 74, 77, 78, 79, 81, 85, 85, 86, 146.

a) Calculate the mean, median, and mode.

b) Which of these three measures of central tendency best represent the data? Explain.

3.

If the mean of the data $20, 10, 15, 14, 9, 9, x, 8$ is twelve, find the value of x.

4.

A student registered in a Grade 11 autobody class has been assessed on the following four modules and his marks are shown:

Metal Repair 80%
Surface Preparation 76%
Trim Replacement 73%
Refinishing 86%

What mark must he achieve in the last module, Touch-Up and Finishing, to complete the course with an average (mean) of 80%?

5.

The frequency table below represents the number of students absent from class during the month of January. Calculate, to the nearest hundredth where necessary, the mean, the median, and the mode of the number of students absent per day.

Number of Absent Students	**Number of Days**
0	8
1	4
2	7
3	0
4	2
5	1

Standard Deviation

Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/hl6cm7DE6Us

One disadvantage of using the range as a measure of dispersion about the mean is that it only uses two of the data values and can be influenced by one extreme value. A better measure of dispersion would be a measure which uses all the data values in its calculation. Such a measure is the **standard deviation**.

Standard Deviation is a measure that describes the variation, or spread, between the data values and the mean of the data.

A low standard deviation means that most of the data values are close to the mean, and hence the data values are more consistent.

A high standard deviation means that most of the data values are scattered further from the mean, and hence the data values are less consistent.

Standard deviation is often used to compare two sets of data.

The standard deviation of a **population** is denoted by the symbol $\boldsymbol{\sigma}$ (sigma).
The standard deviation of a sample is denoted by the symbol s.

In this course, we will assume that all data comes from populations and we will use the symbols μ for mean and σ for standard deviation.

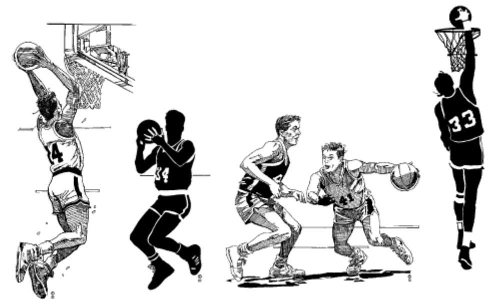
The population standard deviation (σ) can be calculated using $\sigma = \sqrt{\dfrac{\sum(x-\mu)^2}{n}}$.

Follow the procedure below to use the formula:

1. For each data value, calculate the deviation (or difference) from the mean, $x - \mu$.
2. Square each of the deviations from the mean, $(x - \mu)^2$.
3. Add up the all the values in 2, $\sum(x-\mu)^2$.
4. Divide the answer in 3 by the number of data values, $\dfrac{\sum(x-\mu)^2}{n}$.
5. Take the square root of the answer in 4, $\sqrt{\dfrac{\sum(x-\mu)^2}{n}}$.

Example 1:

The heights of the Cobras basketball team are given in the diagram.



170 182 193 193 212

Height in centimetres

a) Calculate the mean height, μ.

b) Calculate the standard deviation, σ, using the formula $\sigma = \sqrt{\dfrac{\sum(x-\mu)^2}{n}}$.

The table below is included as an aid in the calculation.

Height x	**Deviation from the mean** $(x-\mu)$	**Deviation squared** $(x-\mu)^2$
170		
182		
193		
193		
212		
Totals		

Using a Graphing Calculator to Determine Standard Deviation

The mean of the data, μ, is represented in the calculator window by $\bar{x}$.
The standard deviation of the data, σ, is represented in the calculator window by σx.

Example 2:

Verify the mean and standard deviation of the above example using a graphing calculator.

Mean and Standard Deviation from a Frequency Table

A graphing calculator can be used to determine the mean and standard deviation of data given in a frequency table.

The method is identical to the method on the previous page except for the following:

- In step 3, enter the data values in L_1 and the corresponding frequencies in L_2.
- In step 4, after selecting "1-Var Stats", enter L_1, L_2, then press ENTER twice.

Example 3:

The frequency table shows the number of hits per game by a baseball player during the course of one month.

Calculate the mean and standard deviation, to the nearest tenth, of the number of hits per game.

# of Hits	0	1	2	3	4
# of Games (Frequency	5	10	4	3	1

Assignment:

1.

The maximum daily temperatures in Calgary over a 7 day period are shown.

7.2, 4.8, 4.8, 2.0, 4.1, 12.7, 16.8

Calculate, to the nearest tenth, the mean, median, mode, range, and standard deviation.

2.

The data shows the amount of principal left on home mortgage loans handled by a loan officer at a bank.

Calculate, to the nearest dollar, the mean and standard deviation of the data.

10 000	39 500	51 140
13 000	25 900	43 200
75 400	30 900	123 800

3.

ITA, an independent testing agency, was contracted to check the lifetimes of two brands of light bulbs. The lifetimes, in hundreds of hours, of eight bulbs of each brand are given in the table.

Slim	*Oval*
8.5	9.3
7.1	7.0
8.0	7.7
9.2	6.6
8.4	8.5
7.6	9.8
7.7	6.7
7.8	10.0

a) Determine, to the nearest hour, the mean and standard deviation of each type of bulb.

b) Which bulb, on average, lasts longer?

c) Which bulb is more reliable in terms of the number of hours it will last? Explain.

4.

Cricket, a game which is popular in Australia, the Indian sub-continent, South Africa, and the West Indies, originated in England. The data shows the number of runs scored by two players, Wasim and Mushtaq, in various innings.

a) On average, who is the better player? Why?

b) Which player is more consistent? Why?

	Wasim	**Mushtaq**
	37	56
	1	20
	24	12
	33	21
	106	77
	68	34
	82	49
	31	38
	5	14
	45	45
Totals	432	366

5.

A set of data has a mean value of 15 with a standard deviation of 3. State the mean and standard deviation if

a) each data value is increased by 7 **b)** each data value is decreased by 5

6.

The “Postage Stamp”, the 8th hole at Royal Troon Golf Club in Scotland, is one of the most difficult short par threes in championship golf. During a recent tournament, the scores at that hole in the first round are shown.

Calculate the mean and standard deviation, to the nearest tenth, of the number of strokes taken.

# of strokes	Frequency
2	3
3	58
4	47
5	8
6	3
7	1

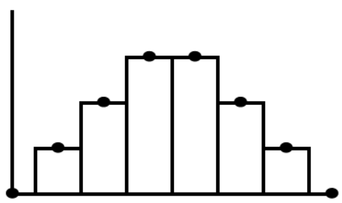
Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/8zTEg4T3Jts

The Normal Distribution

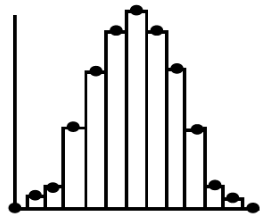
It has been discovered that many observations of physical measurements such as length, volume, mass, time, etc. all have common characteristics in how their data is distributed.

The examples below show histograms developed from data with these common characteristics.

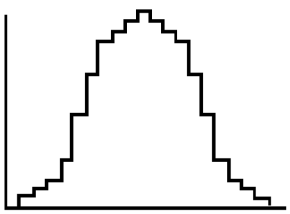
Weights of Infants



Time Taken to Complete an Assignment



Height of Students



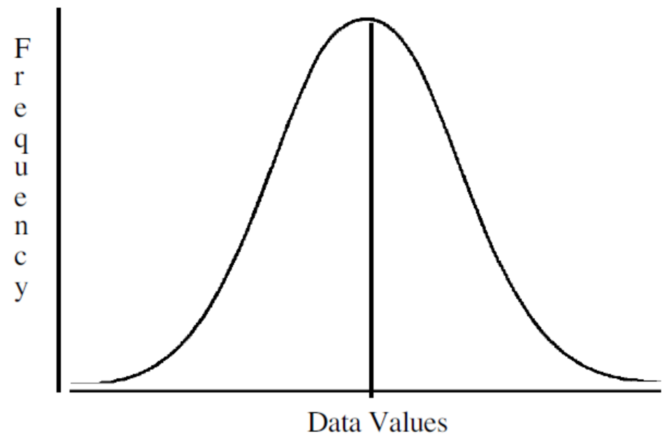
- Draw a frequency polygon by connecting all the midpoints of the top of each bar with a smooth curve.

In each case, the frequency polygon of the data results in a bell shape.

Data which results in a frequency polygon with a bell shaped curve is said to be **normally distributed**.

The curve is referred to as the **Normal Distribution Curve**, the **Normal Curve**, or the **Bell Curve** and is widely used in making predictions in statistics.

Normal Distribution Curve
(The Bell Curve)



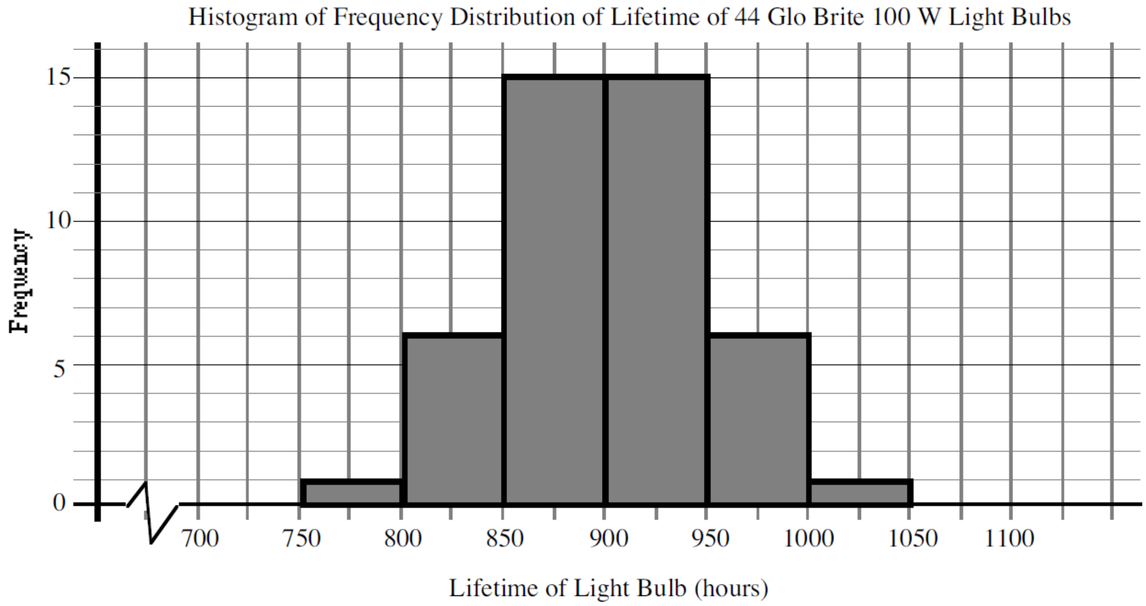
The normal distribution curve can best be understood by a concrete example.

A manufacturer who makes 100W light bulbs for Glow Brite Inc. is interested in determining the distribution of the lifetimes of the bulbs. After testing 44 bulbs, he calculates the mean life of the bulbs to be approximately 900 hours and the standard deviation to be approximately 50 hours.
The lifetimes, in hours, of the 44 bulbs tested are shown.

767	849	845	830.5	835.1	840	849.9	851	851.4	854.8	860
899	898	894	872	874	875	880	881	882.3	885	899
901	903	903	905	908	910	915	919.8	920	922	925
949	932	903	922	950.4	950.9	962.7	975	980.3	997.4	1049

The mean and standard deviation of the data is approximately $\mu = 900$ and $\sigma = 50$.

a) Verify the histogram to the data.



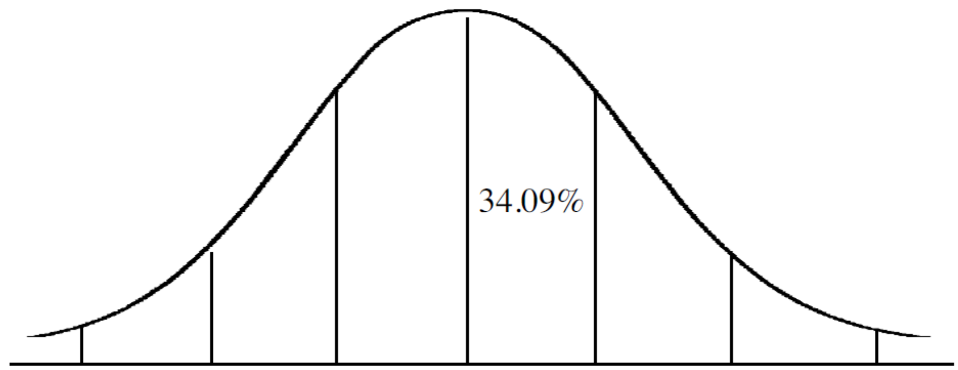
b) Use the histogram to complete the following table.

Interval	<800	800 - 850	850 - 900	900 - 950	950 - 1000	>1000
# of Bulbs in the interval						
% of Bulbs in the Interval				$\frac{15}{44}$ $= 34.09\%$		

c) Using $\mu = 900$ and $\sigma = 50$, write the following numbers in terms of μ and σ.

i) 950 **ii)** 850 **iii)** 800 **iv)** 1050

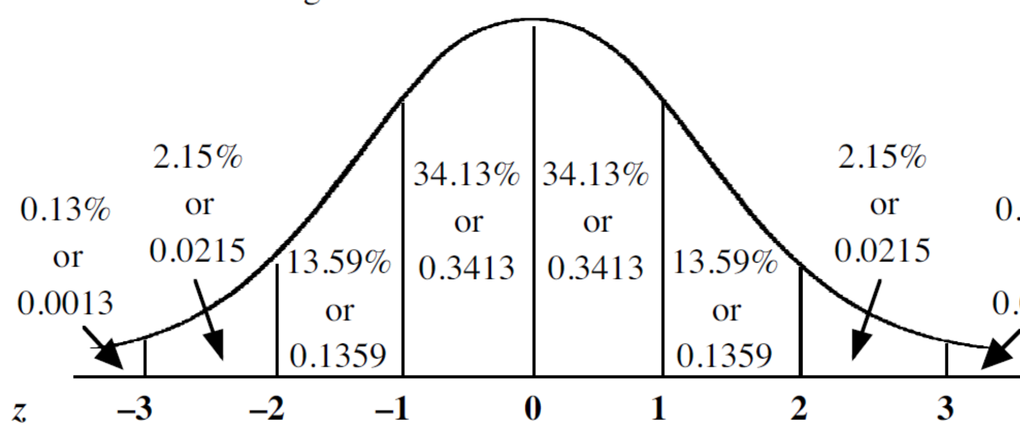
d) Complete the normal distribution curve based on the results for this example.



Data Value	900	950
Number of Standard Deviations from the Mean		1
Data Value in Terms of μ and σ		$\mu + \sigma$

The Standard Normal Distribution Curve

The diagram shows the approximate area under the standard normal distribution curve sub-divided into regions each of width equal to one standard deviation. The percentage of the area under the curve in each region is indicated.

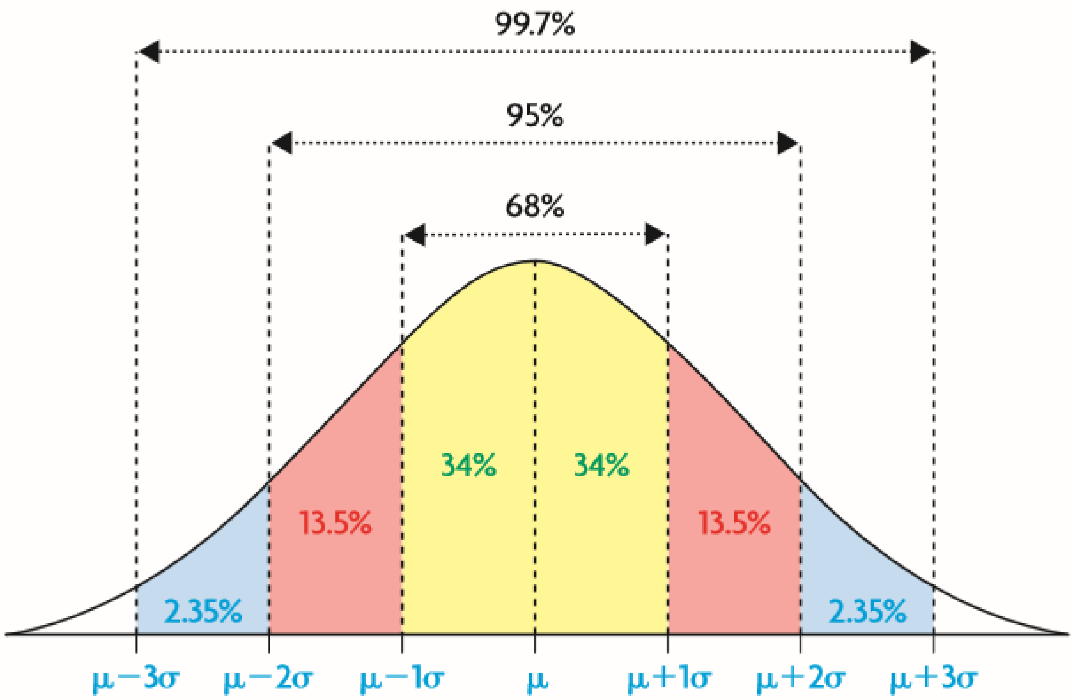


Complete the following:

1. The mean of the data is _____ .
2. The standard deviation of the data is _____ .
3. ______% of the data is above the mean.
4. ______% of the data is within one standard deviation of the mean.
5. ______% of the data is within two standard deviations of the mean.
6. ______% of the data is within three standard deviations of the mean.
7. ______% of the data is between 1 and 2 standard deviations above the mean.
8. ______% of the data is between 1 and 2 standard deviations below the mean.

Need to Know

- The properties of a normal distribution can be summarized as follows:
  - The graph is symmetrical. The mean, median, and mode are equal (or close) and fall at the line of symmetry.
  - The normal curve is shaped like a bell, peaking in the middle, sloping down toward the sides, and approaching zero at the extremes.
  - About 68% of the data is within one standard deviation of the mean.
  - About 95% of the data is within two standard deviations of the mean.
  - About 99.7% of the data is within three standard deviations of the mean.
  - The area under the curve can be considered as 1 unit, since it represents 100% of the data.

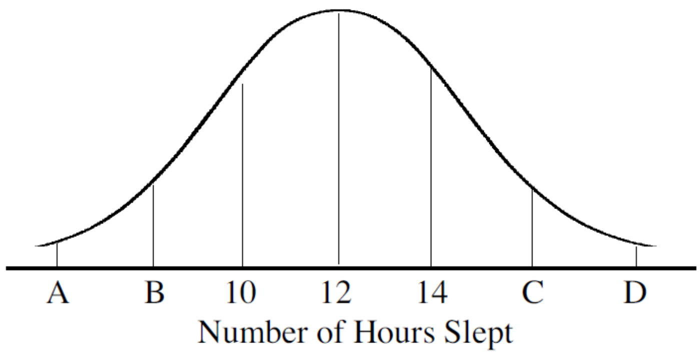


- Generally, measurements of living things (such as mass, height, and length) have a normal distribution.

Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/zRJIvJ6CsAg

Example 1:

A nurse records the number of hours an infant sleeps during a day. He then records the data on a normal distribution curve shown below. The values shown on the horizontal axis differ by one standard deviation.



a) What is the mean of the data?

b) What is the standard deviation?

c) What are the values for A, B, C, and D?

d) What percentage of days, to the nearest hundredth, does the infant sleep:

i) between 12 and 14 h? **ii)** between 8 and 16 h? **iii)** less than 6 h?

e) Why is it not possible at this time to determine the percentage of days that the infant sleeps for less than 13 hours?

To answer part e) we need the concept of ***z*-scores**.

z-scores

It is not possible with the given curve in Class Ex. #1 to determine the percentage of days the infant sleeps for less than 13 hours because the model of the curve used is limited to integer values of standard deviations from the mean.

The concept of **z-scores** can be used to determine this percentage.

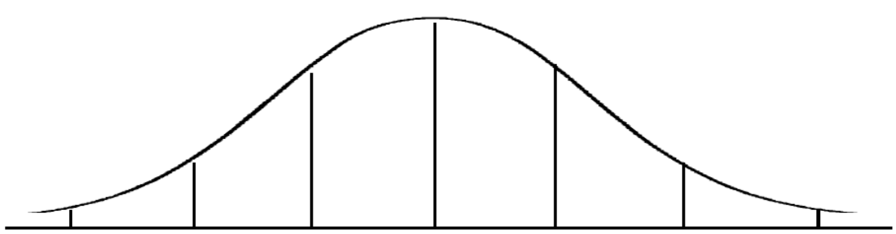
A **z-score** for a particular data value is the number of standard deviations the data value is above or below the mean.

Calculating and Displaying z*-scores *on the Normal Curve

a) Consider Class Example #1 with a mean value of 12 and a standard deviation of 2.

- The data value 18 is ___ standard deviations above the mean and has a z-score of ___ .
- The data value 8 is ___ standard deviations below the mean and has a z-score of ___ .
- The data value 12 is ___ standard deviations away from the mean and has a z-score of ___ .
- The data value 13 is ____ standard deviations _______ the mean and has a z-score of ___ .

b) Complete the rows below the normal curve.



Data Value (x) $\Rightarrow$	6	8	10	12	14	16	18
In Terms of μ and σ $\Rightarrow$							
z-score $\Rightarrow$							

c) How can the z-score for a data value of 13 be determined?
Calculate the z-score for the data value 13, and list both the data value and the z-score on the normal curve above.

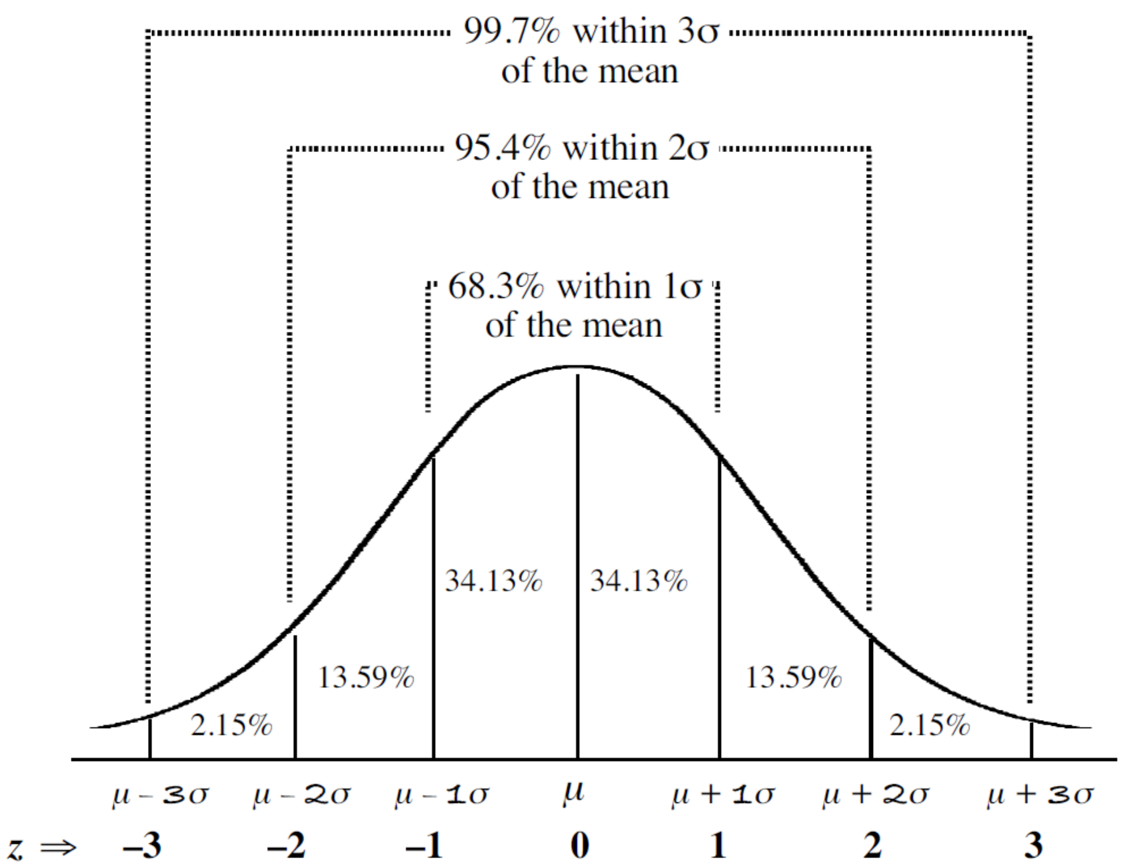
z-score *Formula*

- A z-score for a data value can be calculated using the formula

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

where, z is the z-score of the data value
x is the particular data value
μ is the mean
σ is the standard deviation

- A z-score indicates the number of standard deviations that a data value lies from the mean.
- A positive z-score indicates that the data value lies above the mean.
- A negative z-score indicates that the data value lies below the mean.



Example 2:

Complete the following:

1. Approximately 68.3% of the data lies between $z =$ _____ and $z =$ _____ .
2. Approximately 95.4% of the data lies between $z =$ _____ and $z =$ _____ .
3. Approximately 99.7% of the data lies between $z =$ _____ and $z =$ _____ .

Example 3:

The heights, in centimetres, of five starting members of the Wolfhounds basketball team are 170, 182, 193, 195, and 212. If the mean height is 190 cm and the standard deviation is 15 cm, calculate the z-score, to the nearest hundredth, for the tallest player. The calculation of the z-score for the shortest player is shown.

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

$$z_{170} = \frac{170 - 190}{15}$$

$$z_{170} = -1.33$$

Example 4:

z-scores can be used to compare data from different normal distributions by converting the distributions to the standard normal distribution.

Tony's midterm marks are shown below, together with the class mean and standard deviation for each subject. By calculating z-scores, determine in which subject Tony performed best relative to the rest of the class.

Subject	Tony's Mark	Mean Mark	Standard Deviation
Mathematics	74	68	12
Chemistry	79	73	14
Physics	68	66	11

Example 5:

The average mark on an English exam was 63 and the standard deviation was 12. If the marks were normally distributed and a student's z-score was 1.5, then what was the student's actual mark?

Example 6:

The weights of a large shipment of cantaloupes are normally distributed with a mean of 2.3 kg. The weight of a particular cantaloupe is 1.7 kg, which is 1.01 standard deviations below the mean weight. Determine the standard deviation to the nearest hundredth.

Example 7:

The marks on a math exam at a university were found to have a mean of 52 with a standard deviation of 12. A professor who thought the exam was too difficult decided to adjust the original marks by raising the mean to 65, while reducing the standard deviation to 10 and leaving the z-scores unchanged. What would the new mark be for a student who received an original mark of 34?

Practice:

1.

The goals scored by a major league hockey player over 12 seasons are shown.

11, 18, 23, 27, 21, 30, 28, 24, 17, 21, 19, 24

a) Use the statistical features of a graphing calculator to determine, to the nearest hundredth, the mean and standard deviation of the data.

b) Assuming the data is normally distributed, calculate the z-scores, to the nearest hundredth, for the highest and lowest number of goals.

2.

Pat's unit test marks are shown below, together with the class mean and standard deviation for each unit test.

Subject	Pat's Marks	Mean Mark	Standard Deviation
Measurement	79	70	12
Reasoning	71	61	14
Trigonometry	78	68	13
Statistics	76	65	13

By calculating z-scores, determine, relative to the rest of the class, in which unit test Pat performed

a) best

b) worst

3.

At the Growers Apple Festival, a panel of ten judges award points, on a scale of 1 to 10, in order to determine the most appealing apples. One particular apple had an overall score of 89 with a z-score of 2.35. If the data was normally distributed with a standard deviation of 8.5, determine the overall mean of the data to the nearest whole number.

4.

On a nursing proficiency exam at a Canadian college the mean score was 63 and the standard deviation was 10. If Nicole's z-score was 1.7, then what was her actual exam mark?

5.

Mark owns Prime Fruits Ltd., a company which sells fruit by setting up fruit stands at various road junctions. BC cherries are amongst the many types of fruit that he sells. During a particular week, his fruit stands sold an average of 575 kg of cherries. One of his fruit stands sold 478 kg of cherries, which was 1.73 standard deviations below the mean. If the data is normally distributed, then determine the standard deviation to the nearest tenth of a kg.

6.

The weights of a large shipment of watermelons are normally distributed with a standard deviation of 1.2 kg. The weight of one watermelon picked at random from the shipment is 3.1 kg with a z-score of -1.19. What is the mean to the nearest tenth of a kg?

7.

A university test was given where the scores are normally distributed. A student has a score of 63%, which is 2.13 standard deviations above the mean. If the mean of the exam is 57%, then what is the standard deviation to the nearest hundredth?

8.

Toni wrote four final exams.
Her exam results are shown in the table.

Subject	Toni's Marks	Mean Mark	Standard Deviation
Mathematics	67	71	12
Chemistry	59	65	14
Physics	64	64	10
English	61	67	10

By calculating z-scores, we can determine that, relative to other students in her class, Toni performed **worst** in

9. Pam's z-score of her mark on an exam is 2.31. The mean score on the exam was 58 and the standard deviation was 7. Pam's actual exam mark, to the nearest tenth, is _____ .

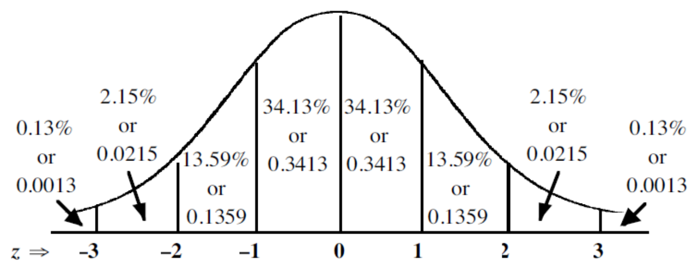
Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/cXRQ6G6nfu0

Using z-Score Tables

z-Score Tables

Recall the standard normal distribution curve shown.

The numbers on the horizontal axis represent z-scores.

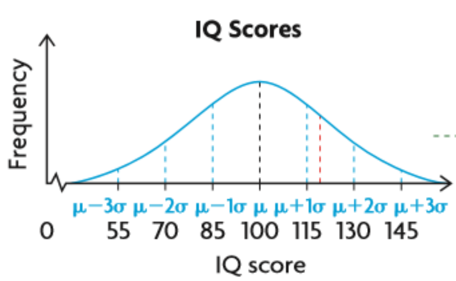


The z-score table gives you the area to the left of a particular z value.

Example 1:

IQ tests are sometimes used to measure a person's intellectual capacity at a particular time. IQ scores are normally distributed, with a mean of 100 and a standard deviation of 15. If a person scores 119 on an IQ test, how does this score compare with the scores of the general population?

Malia's Solution: Using a *z*-score table



First, I sketched a normal curve and determined the IQ scores for one, two, and three standard deviations from the mean.

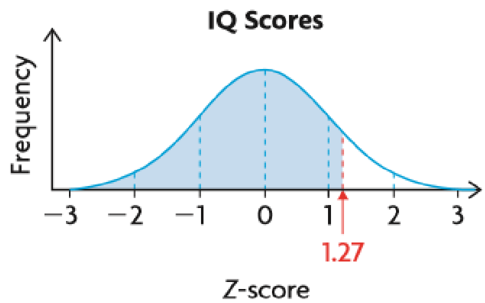
Then I drew a line that represented an IQ score of 119.

I noticed that my line was between one and two standard deviations above the mean.

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

$$z = \frac{119 - 100}{15}$$

$$z = 1.2666...$$



I determined the *z*-score for an IQ of 119.

An IQ score of 119 is about 1.27 standard deviations above the mean. I sketched this on a standard normal curve.

I knew that I needed to determine the percent of people with IQ scores less than 119. This is equivalent to the area under the curve to the left of 1.27 on the standard normal curve.

z	0.0	0.01		0.06	0.07
0.0	0.5000	0.5040		0.5239	0.5279
0.1	0.5398	0.5438		0.5636	0.5675
1.1	0.8643	0.8665		0.8770	0.8790
1.2	0.8849	0.8869		0.8962	**0.8980**
1.3	0.9032	0.9049		0.9131	0.9147

I used a ***z*-score table**.

1.27 = 1.2 + 0.07

I used the 1.2 row and the 0.07 column.

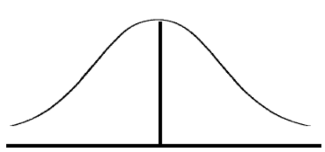
The value in the table, 0.8980, is the fraction of the area under the curve to the left of the *z*-score.

The value in the *z*-score table is 0.8980. This means that an IQ score of 119 is greater than 89.80% of IQ scores in the general population.

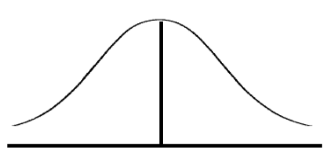
Example 2:

Determine the area under the curve for a standard normal distribution for each of the following z-score intervals. Then convert each of the areas to a percentage to the nearest hundredth.

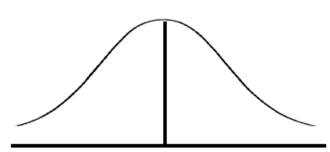
a) $z < -2.15$



b) $z > -2.15$



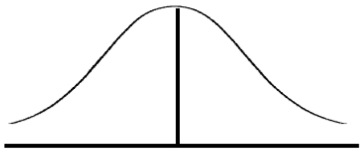
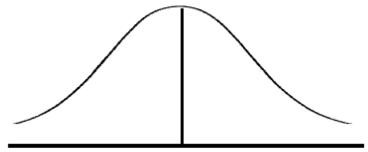
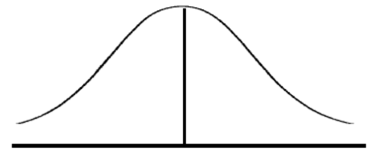
c) $-2.15 < z < 2.15$



Example 3:

Determine, to four decimal places, the area under the curve for a standard normal distribution for each of the following z-score intervals.

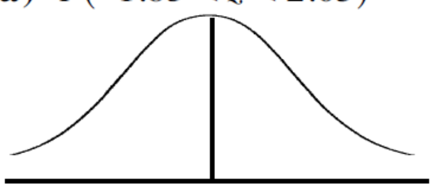
a) $z < 2.08$ **b)** $z > 0.92$ **c)** $-1.75 < z < -1.02$

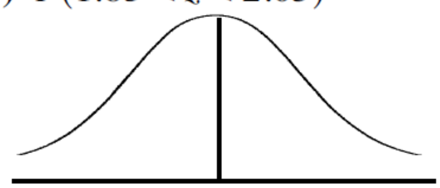
Example 4:

Determine the following probabilities to four decimal places.

a) $P(-1.83 < z < 2.65)$



b) $P(1.83 < z < 2.65)$



Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/eUhlDX7P1sw

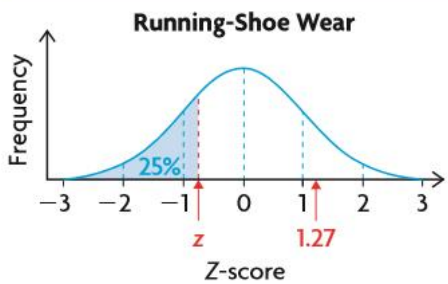
Using the Area Under the Curve to Determine z-Score Intervals

Example 1:

Athletes should replace their running shoes before the shoes lose their ability to absorb shock.

Running shoes lose their shock-absorption after a mean distance of 640 km, with a standard deviation of 160 km. Zack is an elite runner and wants to replace his shoes at a distance when only 25% of people would replace their shoes. At what distance should he replace his shoes?

Rachelle's Solution: Using a *z*-score table



I sketched the standard normal curve. I needed the *z*-score for 25% of the area under the curve, or 0.25.

z	0.09	0.08	0.07	0.06	0.05
−0.7	0.2148	0.2177	0.2206	0.2236	0.2266
−0.6	0.2451	0.2483	0.2514	0.2546	0.2578
−0.5	0.2776	0.2810	0.2843	0.2877	0.2912

I searched the *z*-score table for a value that is close to 0.25.

The *z*-score that represents an area of 0.25 is about halfway between −0.67 and −0.68, or about −0.675.

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

$$(-0.675) = \frac{x - (640)}{(160)}$$

$$-108 = x - 640$$

$$532 = x$$

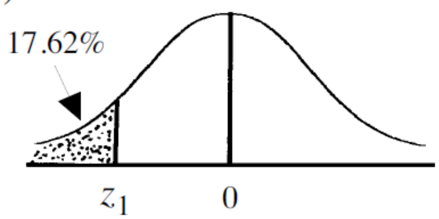
I substituted the values I knew into the *z*-score formula and solved for *x*.

Zack should replace his running shoes after 532 km.

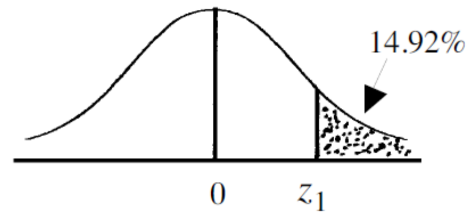
Example 2:

Determine z_1 and z_2 in the examples below.

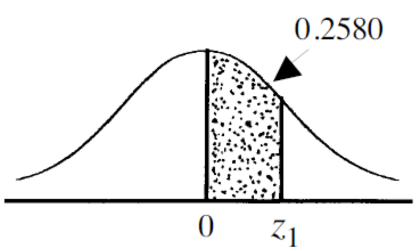
a)



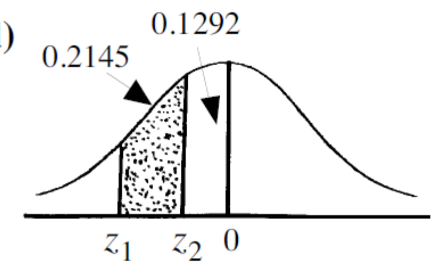
b)



c)



d)

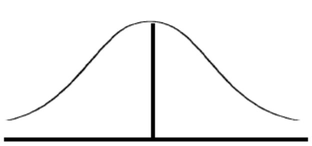


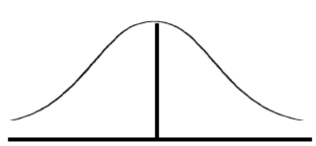
Assignment:

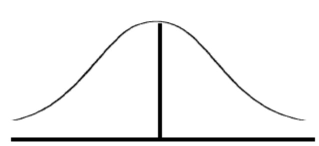
1.

Find the area under the standard normal curve for each z-score interval. Give the area as a decimal to the fourth decimal place and as a percent to the nearest hundredth.
Label the diagram.

a) $z < -1.6$ **b)** $z > -1.6$ **c)** $z > 1.6$



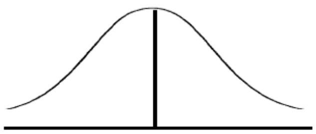


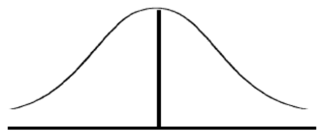


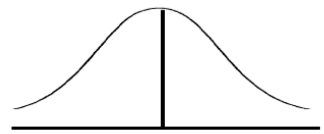
2.

Find the area under the standard normal curve for each z-score interval. Give the area as a decimal to the fourth decimal place and as a percent to the nearest hundredth. Label the diagram.

a) $-0.31 < z < 1.56$ **b)** $0.31 < z < 3.32$ **c)** $-2.00 < z < -0.41$







3.

Determine the following probabilities as a decimal to the fourth decimal place and as a percent to the nearest hundredth. Sketch a diagram in each case.

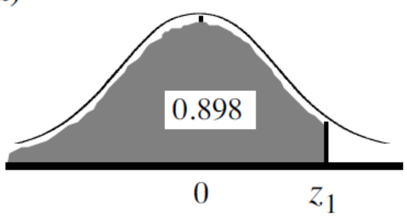
a) $P(-2.34 < z < -1.3)$ **b)** $P(z > -1.32)$ **c)** $P(z < -2.42)$

d) $P(2.34 < z < 3.00)$ **e)** $P(z > -0.09)$ **f)** $P(-1.31 < z < 0.05)$

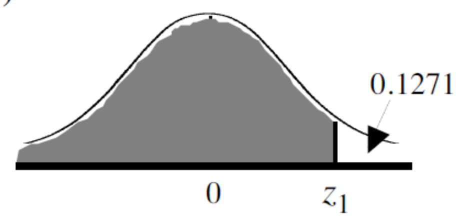
4.

Determine z_1 and z_2 in the examples below.

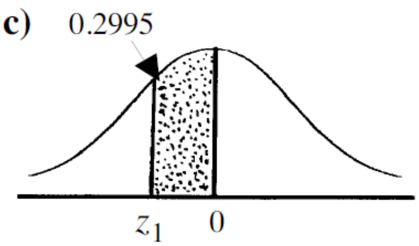
a)



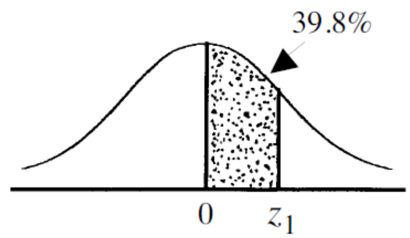
b)



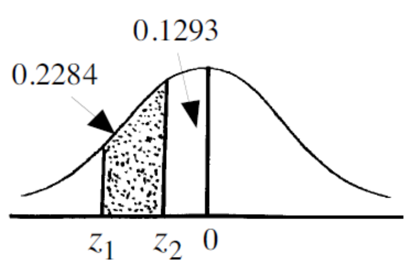
c)



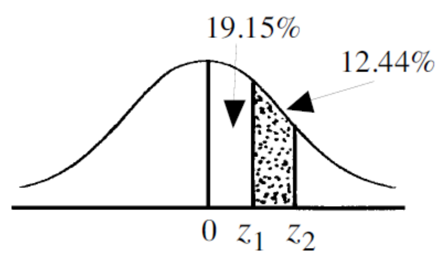
d)



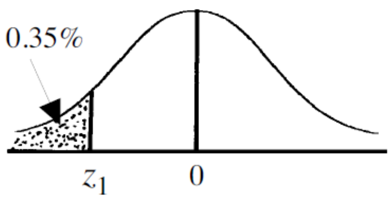
e)



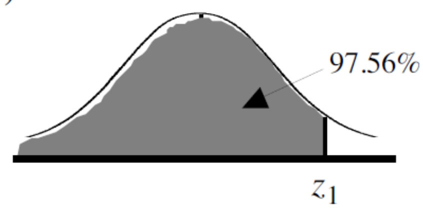
f)



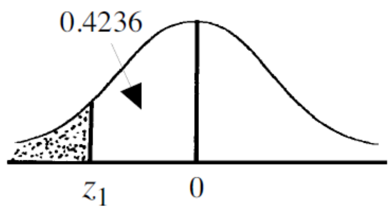
g)



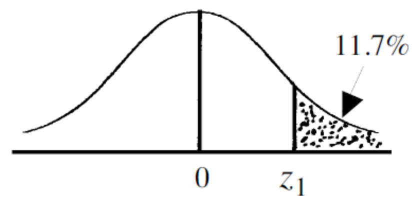
h)



i)



j)



5.

Find the value of a.

a) $P(z < a) = 0.1379$

b) $P(z > a) = 0.8508$

c) $P(0 < z < a) = 0.3907$

d) $P(a < z < 0) = 0.4306$

Watch the following instructional video. In your handout:
i) Copy down the given notes and examples
ii) Complete the assigned questions https://youtu.be/odwc-cCvGH8

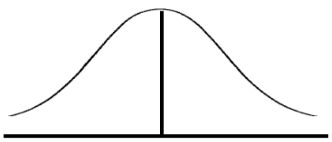
(Skip example 3. That example is explained in the previous video)

Problems Involving the Normal Distribution

Example 1:

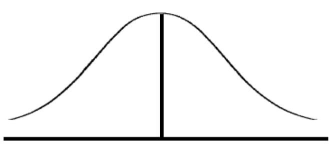
A light bulb manufacturer produces 35 000 light bulbs. From past data, the lifetimes of the bulbs are normally distributed with a mean life of 900 hours and a standard deviation of 50 hours.

a) Predict the percentage of light bulbs that will last between 825 and 875 hours.



b) How many of the 35 000 light bulbs would you expect to last between 825 and 875 hours?

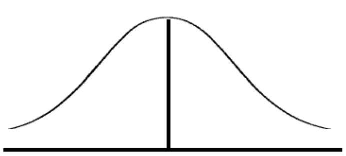
c) Determine the probability that a bulb selected at random will last less than 920 hours.



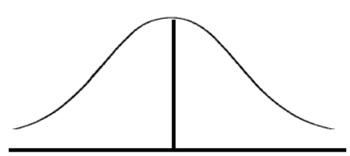
Example 2:

A study showed that the mean duration of a certain strain of flu virus was 12 days with a standard deviation of 3 days. If the data is normally distributed, and you caught this strain of flu virus, determine the probability, to the nearest hundredth, that it would last:

a) longer than 17 days



b) between 13 and 15 days

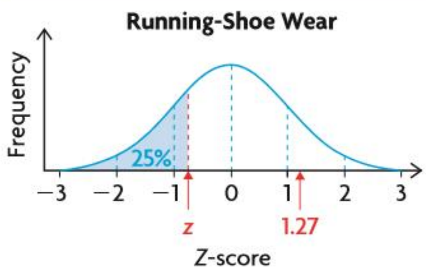


Example 3:

Athletes should replace their running shoes before the shoes lose their ability to absorb shock.

Running shoes lose their shock-absorption after a mean distance of 640 km, with a standard deviation of 160 km. Zack is an elite runner and wants to replace his shoes at a distance when only 25% of people would replace their shoes. At what distance should he replace his shoes?

Rachelle's Solution: Using a *z*-score table



I sketched the standard normal curve. I needed the *z*-score for 25% of the area under the curve, or 0.25.

z	0.09	0.08	0.07	0.06	0.05
−0.7	0.2148	0.2177	0.2206	0.2236	0.2266
−0.6	0.2451	0.2483	0.2514	0.2546	0.2578
−0.5	0.2776	0.2810	0.2843	0.2877	0.2912

I searched the *z*-score table for a value that is close to 0.25.

The *z*-score that represents an area of 0.25 is about halfway between −0.67 and −0.68, or about −0.675.

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

$$(-0.675) = \frac{x - (640)}{(160)}$$

$$-108 = x - 640$$

$$532 = x$$

I substituted the values I knew into the *z*-score formula and solved for *x*.

Zack should replace his running shoes after 532 km.

Example 4:

The ABC Company produces bungee cords. When the manufacturing process is running well, the lengths of the bungee cords produced are normally distributed, with a mean of 45.2 cm and a standard deviation of 1.3 cm. Bungee cords that are shorter than 42.0 cm or longer than 48.0 cm are rejected by the quality control workers.

a) If 20 000 bungee cords are manufactured each day, how many bungee cords would you expect the quality control workers to reject?

b) What action might the company take as a result of these findings?

Logan's Solution: Using a *z*-score table

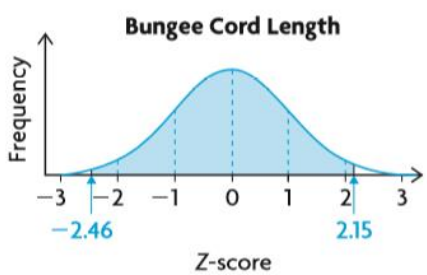
a) Minimum length = 42 cm   Maximum length = 48 cm

$$z_{min} = \frac{x - \mu}{\sigma} \qquad z_{max} = \frac{x - \mu}{\sigma}$$

$$z_{min} = \frac{42.0 - 45.2}{1.3} \qquad z_{max} = \frac{48.0 - 45.2}{1.3}$$

$$z_{min} = -2.461... \qquad z_{max} = 2.153...$$

I determined the *z*-scores for the minimum and maximum acceptable lengths.



I sketched the standard normal curve. The area under the curve to the left of −2.46 represents the percent of rejected bungee cords less than 42 cm. The area under the curve to the right of 2.15 represents the percent of rejected bungee cords greater than 48 cm.

Area to left of $-2.46 = 0.0069$

Area to right of $2.15 = 1 - 0.9842$
Area to right of $2.15 = 0.0158$

I looked up each *z*-score in the *z*-score table. The *z*-score table gives the area to the left of the *z*-score, which I want for 42 cm.

Since I wanted the area to the right of the *z*-score for 48 cm, I had to subtract the corresponding area from 1.

Percent rejected = Area to the left of −2.46 + Area to the right of 2.15
Percent rejected = 0.0069 + 0.0158
Percent rejected = 0.0227 or 2.27%

I added the two areas to determine the percent of bungee cords that are rejected.

Total rejected = (0.0227)(20 000) or 454

I determined the number of bungee cords that are rejected.

b) ABC needs a more consistent process, because 454 seems like a large number of bungee cords to reject. The company should adjust its equipment so that the standard deviation is lowered.

Lowering the standard deviation will reduce the percent of rejected bungee cords.

Example 5:

A manufacturer of personal music players has determined that the mean life of the players is 32.4 months, with a standard deviation of 6.3 months. What length of warranty should be offered if the manufacturer wants to restrict repairs to less than 1.5% of all the players sold?

Sacha's Solution

$1.5\% = 0.015$

I used my graphing calculator to determine the *z*-score that corresponds to an area under the normal curve of 0.015.

$z = -2.17$

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

$$(-2.17) = \frac{x - (32.4)}{(6.3)}$$

$$-13.671 = x - 32.4$$

$$18.729 = x$$

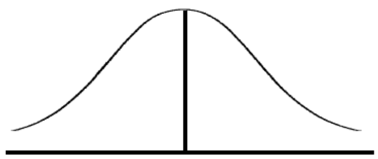
I substituted the known values into the *z*-score formula and solved for *x*.

The manufacturer should offer an 18-month warranty.

Since the manufacturer wants to repair less than 1.5% of the music players, I rounded down to 18 months.

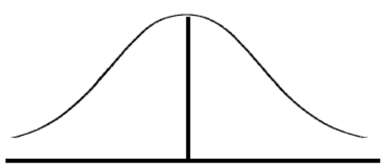
Example 6:

From extensive testing, an appliance distribution company knows that the average life of "Toasty" toasters is 4.2 years, the standard deviation is 0.65 years, and the data is normally distributed. The company does not want to replace under warranty more than 8% of the toasters that are sold.
What warranty, to the nearest necessary year, should the company offer?



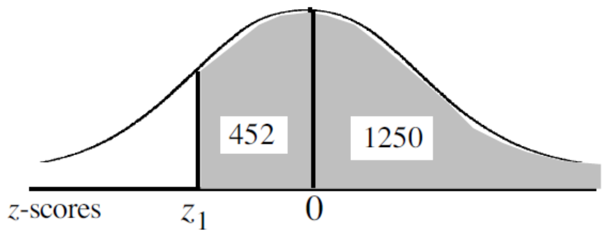
Example 7:

It was found that 62.3% of the shrimp harvested at Shrimp Harvest Farms had a mass of more than 135 grams. If the data is normally distributed, and if the mean mass of the shrimp harvested was 146 grams, determine the standard deviation to the nearest tenth.



Example 8:

The marks of a large number of students have been represented on a standard normal distribution curve. The values given represent the number of students in each area.



a) How many students are represented by the area under the standard normal curve?

b) Determine the value of z_1 to the nearest hundredth.

Assignment:

1.

The results of a provincial Grade Nine achievement test were normally distributed with a mean of 68 and a standard deviation of 12. If 8500 students wrote the test, determine

a) the percentage of students, to the nearest tenth of a percent, who scored a mark of 50 or above

b) the number of students who scored a mark of 50 or above

c) the probability that a student selected at random had a mark

i) less than 30

ii) between 50 and 60

2. The “Long Life” battery company is planning to add another 10 500 batteries to their yearly production of batteries. The mean life of “Long Life” batteries is 50 hours with a standard deviation of 10 hours. If the data is normally distributed, then how many of the new batteries produced can be expected to last less than 31 hours.

3.

The heights of 800 officers from a police force are normally distributed with a mean of 175 cm and a standard deviation of 8 cm.

a) How many of the officers are within one standard deviation of the mean height?

b) How many of the officers are between 167 and 173 cm?

c) What percentage of the officers is between 165 cm and 180 cm?

4.

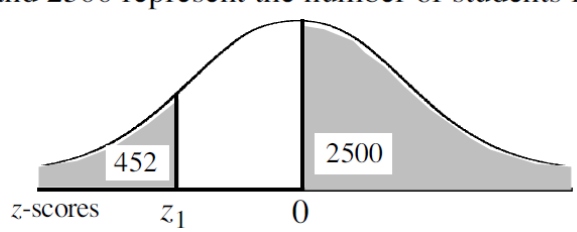
Data collected of cars passing on a road revealed that the average speed was 90 km/h with a standard deviation of 5 km/h and data which is normally distributed. A policeman is assigned to set photo radar on a road in which the posted speed limit is 80 km/h. The policeman sets the camera so that only those exceeding the speed limit by 10% are photographed and ticketed.

a) What is the lowest speed, in km/hr, for which you could be ticketed?

b) If 600 cars pass the photo radar, how many drivers can the police expect to ticket?

5.

The results of a provincial achievement test are normally distributed and are represented in the diagram below. The data under the curve represents all of the students who wrote the test. The values 452 and 2500 represent the number of students in the shaded regions.



a) How many students wrote the test? **b)** What is the value of z_1?

6.

After reviewing previous loan records, the credit manager of a bank determines that the data follows a normal distribution. The debts have a mean of \$20 000 and the probability that the loss could be greater than \$25 000 or less than \$15 000 is 0.418. Determine the standard deviation of the data to the nearest hundred dollars.

7.

A company packages rice into 10 kg bags. The machine that fills the bags can be calibrated to fill to any specified mean with a standard deviation of 0.09 kg. Any bags that weigh less than 10 kg cannot be sold and must be refilled. To what mean value, to the nearest hundredth of a kilogram, should the machine be set if the company does not want to refill more than 1.5% of the bags?

8.

The weights of a large shipment of coconuts are normally distributed with a standard deviation of 0.71 kg. The probability that a coconut weighs less than 2.1 kg is 10.2%. The mean of the shipment, to the nearest hundredth of a kg, is _____ .

9.

A medical diagnostic test counts the number of blood cells in a sample. The red blood cell count (in millions per cubic microlitre) is normally distributed, with a mean of 4.8 and a standard deviation of 0.3.

a) What percent of people have a red blood cell count that is less than 4?

b) What percent of people have a count between 4.7 and 5.0?

c) What red blood cell count would someone have if 95% of people have a lower count?

10.

An MP3 player has a one-year warranty. The mean lifespan of the player is 2.6 years, with a standard deviation of 0.9 years.

a) A store sells 4000 players. How many of these players will fail before the warranty expires?

b) Tyler is offered an extended warranty, for one extra year, when he buys a player. What is the likelihood that he will make a claim on this warranty if he takes it?

11.

A manufacturer of plasma televisions has determined that the televisions require servicing after a mean of 67 months, with a standard deviation of 7.2 months. What length of warranty should be offered, if the manufacturer wants to repair less than 1% of the televisions under the warranty?

12.

A tutor guarantees that 10% of her students will obtain an A on every test they write. For the last test, the mean mark is 68 and the standard deviation is 6. What mark is required to receive an A on the test?