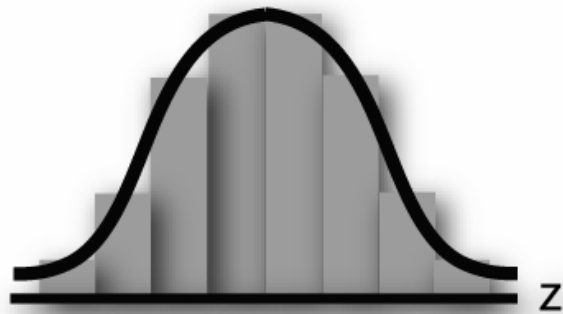


Pure Math 30:

STATISTICS



LESSON THREE

THE NORMAL DISTRIBUTION

Pure Math
30:

EXPLAINED!

By
Barry
Mabillard

STATISTICS LESSON 3

PART I: MEAN AND STANDARD DEVIATION

Finding Mean & Standard Deviation:

Using the TI-83, it is possible to find several important pieces of information from a set of data.

- Mean = The average value in the data set.
- Standard Deviation = The amount of dispersion in the data set.
(The bigger the standard deviation, the further apart the data values are.)

Example 1: Finding mean & standard deviation from raw data.

The following set of numbers represents the temperatures for a week in July.

23, 27, 30, 33, 25, 19, 29

Use your TI-83 to find the mean & standard deviation.

In your calculator, go stat → edit. Put all the values from above into list 1.
Once the values are entered, type 2nd → quit to return to the main screen.
Next type Stat → Calc → 1 Var Stats → Enter.
In the screen that appears,
 $\bar{x} = 26.57$ is the mean, and
 $\sigma x = 4.3378$ is the standard deviation.

Example 2: Finding mean & standard deviation from a frequency table.

A particular store sells computers. Over a period of ten days, the number of computers sold each day was recorded in a table.

Computers Sold	Frequency
0	4
1	2
2	1
3	0
4	2
5	1

To type a frequency table into the calculator, do the following:

First go stat → edit to bring up your lists, then clear them. Type the first column into L1, and the second column into L2.

Type 2nd → quit to return to the main screen.

Next type: Stat → Calc → 1 Var Stats → Enter → 2nd → List → L1 → Enter → Comma → Enter → 2nd → List → L2 → Enter

Based on this frequency table,
 $\bar{x} = 1.7$ is the mean, and
 $\sigma x = 1.85$ is the standard deviation.

STATISTICS LESSON 3

PART I: MEAN AND STANDARD DEVIATION

Example 3: Finding mean & standard deviation from a histogram.

A student performs 39 trials of flipping a coin 10 times. The number of heads in each trial is recorded and the results are displayed in a histogram. Use your TI-83 to find the mean & standard deviation.



Treat the histogram the same way as a frequency table, and put the following information into your lists:

Number of Heads (L1)	Frequency (L2)
0	2
1	3
2	7
3	3
4	5
5	5
6	6
7	5
8	2
9	0
10	1

Use the same calculator commands as in Example 2 to determine the mean & standard deviation:

$\bar{x} = 4.31$ is the mean, and $\sigma_x = 2.42$ is the standard deviation.

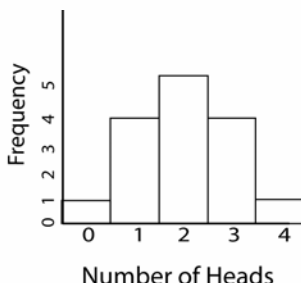
Questions:

1) Find the mean & standard deviation of the following exam marks: 98, 45, 58, 54, 76, 96, 95, 87.

2) Find the mean & standard deviation from the following frequency table:

Number of Goals	Frequency
0	4
1	6
2	5
3	6

3) Find the mean & standard deviation from the following histogram:



Answers:

1)
 $\bar{x} = 76.1$
 $\sigma_x = 19.8$

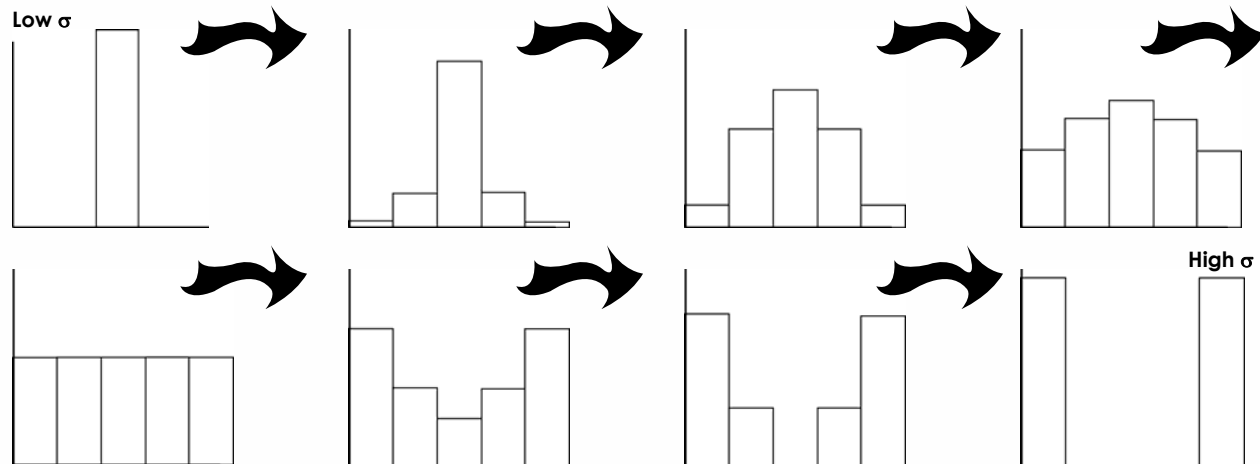
2)
 $\bar{x} = 1.62$
 $\sigma_x = 1.09$

3)
 $\bar{x} = 2$
 $\sigma_x = 1.03$

STATISTICS LESSON 3

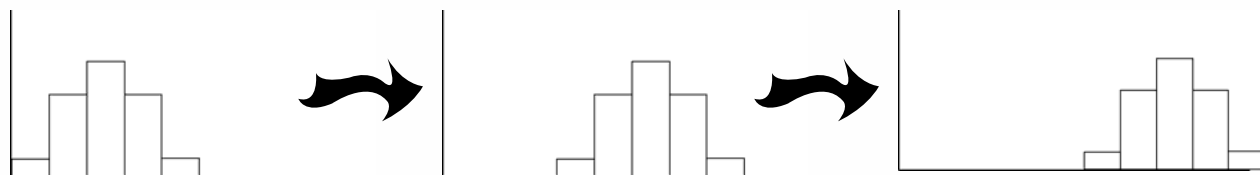
PART I: MEAN AND STANDARD DEVIATION

Example 1: What happens to a data distribution as the standard deviation is increased?



When the standard deviation is low, all of the data values are very close together. As the standard deviation increases, the data becomes more spread out.

Example 2: What happens to a data distribution as the mean is changed?



Mean Increasing \rightarrow

As the mean increases, the centre of the distribution is shifted to the right.

Questions:

- 1) A teacher adjusts the marks of an examination by raising each score by 5 percent. What happens to the mean and the standard deviation?
- 2) Would you expect the following exam marks to have a small or large standard deviation?

92, 93, 92, 94, 92, 91, 92

Answers:

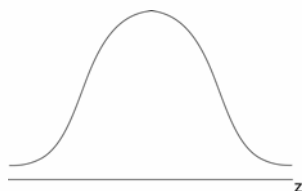
1) The mean will increase since all the marks are increased, but the standard deviation will remain the same. All marks are changed by the same amount, leaving the spread identical.

2) All the data is very close together, so the standard deviation should be small.

STATISTICS LESSON 3

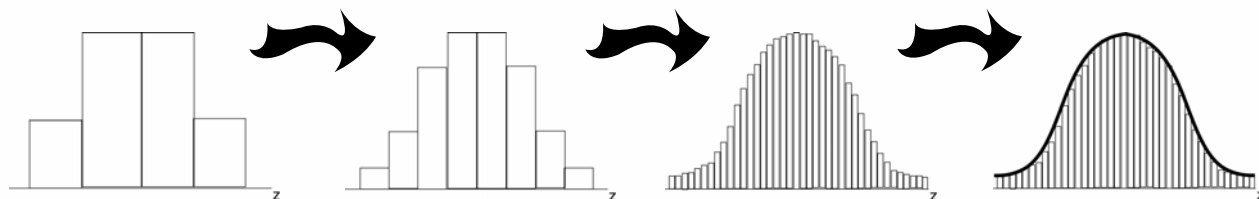
PART II: THE NORMAL DISTRIBUTION

The Normal Distribution Curve is a continuous graph that can be used to find probability.



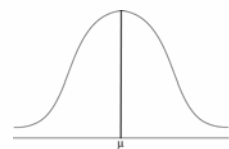
The "bell shape" arises because the outcomes of random situations will frequently be clustered around the mean, with only a few outcomes far away from the mean.

The normal distribution is a limiting case of the binomial distribution, where more and more trials are conducted until the shape of all the rectangles is a near perfect fit to the curve.

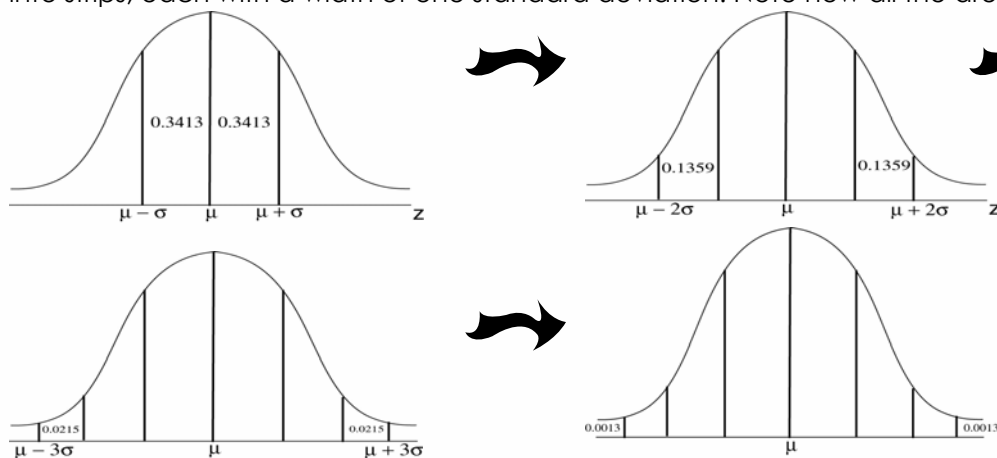


The normal distribution curve extends left and right forever, so don't connect the curve to the axis. (It's the same idea as having a horizontal asymptote). The area under the curve represents probability, and the full curve has a total area of 1.

Mean: The mean value occurs at the very centre of the distribution, and is denoted by the Greek letter Mu (μ)



Standard Deviation: is a measure of data spread, and is denoted by the Greek letter Sigma (σ). The following graphs show approximate areas of the curve when it is divided up into strips, each with a width of one standard deviation. Note how all the areas sum up to 1!



Notice that on the extreme ends, the area is very small. This is because the probability of an event occurring this far away from the mean is very low!

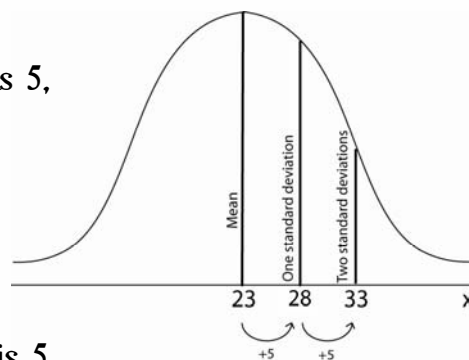
STATISTICS LESSON 3

PART II: THE NORMAL DISTRIBUTION

Z-scores represent the number of standard deviations a data point is away from the mean.

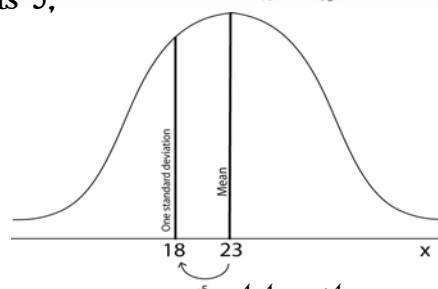
Example 1: If the mean is 23 & the standard deviation is 5, what data point will occupy a z-score of +2?

A z-score of 2 means that we'll go two full standard deviations to the right in order to find the data point. Since each standard deviation is 5, we'll move 10 units right, giving us a data point of 33.



Example 2: If the mean is 23 & the standard deviation is 5, what data point will occupy a z-score of -1?

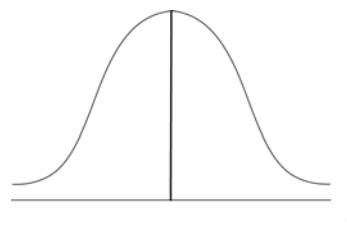
A z-score of -1 means that we'll go one standard deviation to the left in order to find the data point. Since each standard deviation is 5, we'll move 5 units left, giving us a data point of 18.



Drawing the normal curve: Raw data points are expressed by the variable x , and z-scores by the variable z .

Reserve the space immediately below the line for your raw data values, and the space below that for your z-scores.

This will help you organize the information so you know where everything goes!

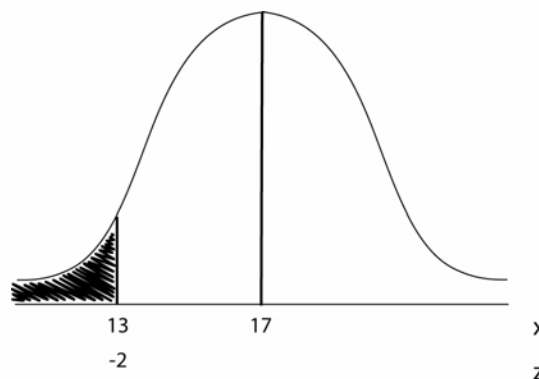


Example 3: The mean height of an action figure is 17 cm, with a standard deviation of 2 cm. Shade in the region of the normal curve corresponding to action figures with a height less than 13 cm, and indicate the z-score value.

Place the mean value of 17 at the centre.

The data point 13 is two standard deviations left, so indicate these values on the graph.

Since we want the values less than 13, shade in the left of this line.



STATISTICS LESSON 3

PART II: THE NORMAL DISTRIBUTION

The formula $z = \frac{x - \mu}{\sigma}$ is used to solve z-score questions.

Example 4: A particular CD has a running length of 72 minutes. This corresponds to a z-score of 2 and a standard deviation of 3 minutes.

a) What is the mean running length of a CD?

This is the raw data point, so it's the x-value.

First manipulate the formula to get μ by itself:

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

$$z\sigma = x - \mu$$

$$z\sigma + \mu = x$$

$$\mu = x - z\sigma$$

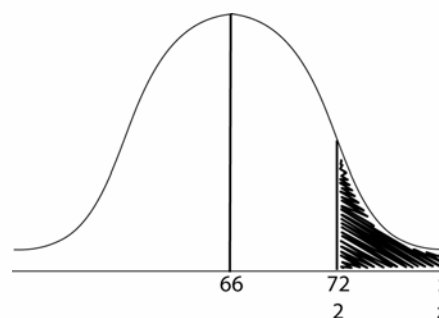
Next plug in the numbers and solve for the mean.

$$\mu = x - z\sigma$$

$$\mu = 72 - (2)(3)$$

$$\mu = 72 - 6$$

$$\mu = 66 \text{ minutes}$$



b) Shade in the area representing the probability a CD lasts longer than 72 minutes.

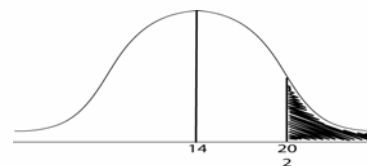
***Note that x, μ , and σ all have the same units, but z has no units.**

Questions:

- 1) An average household watches 14 hours of television in a week, with a standard deviation of 3 hours.
 - a) What z-score would correspond to 20 hours of television being watched?
 - b) Shade in the area representing the probability a household watches more than 20 hours of television per week.

Answers:

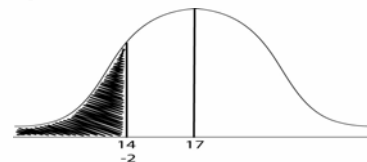
1) $z = 2$



2) The mean height of an action figure is 17 cm, and a particular action figure has a height of 14 cm & a z-score of -2.

- a) What is the standard deviation of the action figures?
- b) Shade in the region of the normal curve representing heights less than 14 cm.

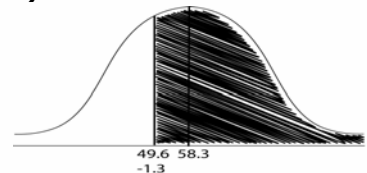
2) $\sigma = 1.5 \text{ cm}$



3) The average mark on a test was 58.3% with a standard deviation of 6.7%. The z-score of a particular test was -1.3.

- a) What was the mark on the test?
- b) Shade in the region of the normal curve representing marks greater than the mark in part a).

3) $x = 49.59\%$



STATISTICS LESSON 3

PART III: AREA AND PROBABILITY

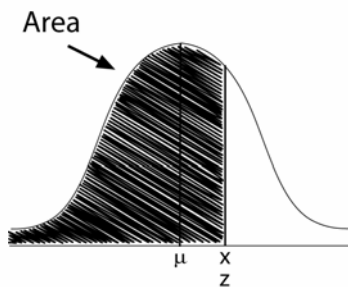
Area & Probability: The area underneath the graph represents probability.

Area can be obtained by using a z-score chart. (See last two pages of this lesson)

The z-score can be found by scanning the left column for the whole number / first decimal place, then lining it up with the second decimal place from the top.

The area (probability) is found in all the boxes in the middle of the table.

The diagram below serves as a reminder that the area you read off the chart is from the absolute left of the graph to the z-score.



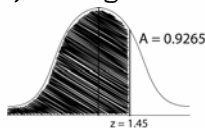
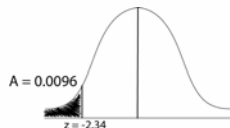
z-score (Second Decimal Place)											
z-score (Whole Number / First Decimal Place)	0	1	2	3	4	5	6	7	8	9	

AREA
(Probability)

Example 1: Find the area corresponding to a z-score of:

a) -2.34

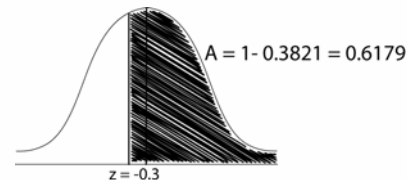
b) 1.45



Example 2: Find the area greater than a z-score of -0.3

Find the area to the left of $z = -0.30$ first. (0.3821)

Then subtract this area from 1 to get the area on the right.

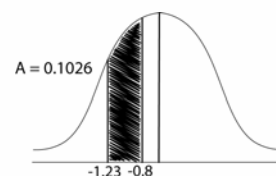


Example 3: Find the area between z-scores of -1.23 & -0.8

Find the area to the left of -1.23 Area = 0.1093

Find the area to the left of -0.80 Area = 0.2119

The difference will give the area in between. Area = 0.1026

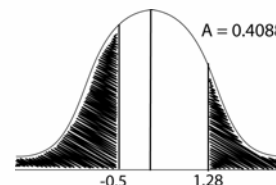


Example 4: Find the total area left of the z-score -0.5 and to the right of the z-score 1.28

Find the area to the left of -0.50 Area = 0.3085

Now find the area to the right of 1.28 by subtracting the table area from 1. ($1 - 0.8997 = 0.1003$) Area = 0.1003

Add the two results. Area = 0.4088



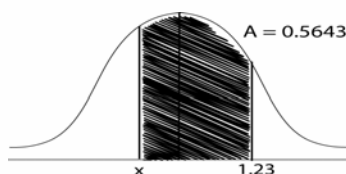
STATISTICS LESSON 3

PART III: AREA AND PROBABILITY

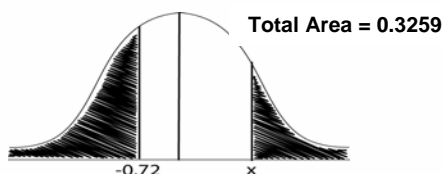
Questions:

- 1) Find the area to the left of the z-score -2.32
- 2) Find the area to the left of the z-score 1.3
- 3) Find the area to the right of the z-score -0.92
- 4) Find the area to the right of the z-score 2.01
- 5) Find the area between the z-scores -2.6 & -0.9
- 6) Find the area between the z-scores -1.2 & 0.53
- 7) Find the area to the left of the z-score -1.93 and to the right of the z-score -0.72
- 8) Find the area to the left of the z-score 0.22 and to the right of the z-score 0.7
- 9) If the area to the left of a z-score is 0.9920, what is the z-score?
- 10) If the area to the right of a z-score is 0.5040, what is the z-score?

- 11) Given the following diagram, solve for x



- 12) Given the following diagram, solve for x



9) Find area on the chart, then line up with z-score.

$z = 2.41$

10) First go $1 - \text{area}$. Now find the area on the chart, then line up with z-score.

$z = -0.01$

11) The total area for $z = 1.23$ is 0.8907. The known area is 0.5643. Subtract to get 0.3264. Now find the area on the chart, then line up with z-score.

$z = -0.45$

12) The area to the left of $z = -0.72$ is 0.2358, which means the area on the right side has to be $0.3259 - 0.2358 = 0.0901$. However, to determine the z-score, we need the area on the left, so $1 - 0.0901 = 0.9099$. This gives a z-score of 1.34.

$z = 1.34$

Answers:

- 1) Area = 0.0102
- 2) Area = 0.9032
- 3) Area = 0.8212
- 4) Area = 0.0222
- 5) Area = 0.1794
- 6) Area = 0.5868
- 7) Area = 0.7910
- 8) Area = 0.8291

STATISTICS LESSON 3

PART IV: APPLICATION QUESTIONS

Application Questions:

Example 1: The mean life of a battery is 12 hours of use, with a standard deviation of 30 minutes. If the batteries are guaranteed to last for 11 hours or a replacement is given, how many batteries will likely be replaced out of a shipment of 9000?

You **must** state all values with the same units! $\mu = 12$ hours, $\sigma = 0.5$ hours, and $x = 11$ hours. The idea here is we want to find the area to the left of 11 hours, and this will represent the probability a battery fails before this amount of time.

Find the z-score first $z = \frac{x - \mu}{\sigma} = \frac{11 - 12}{0.5} = -2$

The corresponding area/probability is 0.0228

Now multiply 0.0228 by 9000 to get the expected number of batteries replaced. **Answer = 205**

Example 2: A teacher marks some exams and finds the mean is 54% and the standard deviation is 8%. The teacher then adjusts the marks by raising the mean to 60% and raising the standard deviation to 9%. The z-scores are kept constant. If a student scored 76% initially, what would their new mark be?

If the z-scores are held constant, that means we can set up the following: $\frac{x_1 - \mu_1}{\sigma_1} = \frac{x_2 - \mu_2}{\sigma_2}$

Now plug in the values given: $\frac{76 - 54}{8} = \frac{x_2 - 60}{9}$, then cross multiply & solve for x_2 .

Answer = 84.75%

Example 3: A company producing stereos find they have a mean life of 7 years and a standard deviation of 1.2 years. What guarantee should the company make so they will only have to exchange fewer than 10% of all stereos sold?

If the company plans on replacing only 10%, that means we have a probability of 0.1

Use the z-score table to find out what z-score corresponds to the area 0.1000 $\rightarrow z = -1.28$

Use this z-score and the values from the question to solve for x .

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

$$-1.28 = \frac{x - 7}{1.2}$$

$$x = 5.464$$

Answer = 5.5 years.

Example 4: The mean is $4.3x$ and the standard deviation is $3x + 2$. If the z-score is 1.4 and the raw data point is $2x$, determine the actual value of the raw data point.

This can be solved by plugging all the expressions into the z-score formula:

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

$$1.4 = \frac{2x - 4.3x}{3x + 2}$$

$$1.4 = \frac{-2.3x}{3x + 2}$$

$$4.2x + 2.8 = -2.3x$$

$$2.8 = -6.5x$$

$$x = -0.4307$$

Now plug the x-value into the raw data point, which is $2x$.

$$2(-0.4207) = -0.8615$$

Answer = -0.8615

STATISTICS LESSON 3

PART IV: APPLICATION QUESTIONS

Questions:

- 1) A computer company determines the mean life of a hard drive is 72 months with a standard deviation of 6 months. What guarantee should the company give so fewer than 9% of all hard drives sold will be returned?
- 2) A house painter determines that a coat of paint will have a mean life of 11 years with a standard deviation of 3.4 years. What is the probability the paint peels off after only 3 years?
- 3) Toothpicks are produced, and any toothpick shorter than 3.57 cm or longer than 3.68 cm will be rejected. If the mean length is 3.62 cm, and the standard deviation is 0.04 cm, how many toothpicks are expected to be rejected if 20000 are produced?
- 4) The mean life of a battery is 13 hours of use, with a standard deviation of 45 minutes. If the batteries are guaranteed to last for 11 hours or a replacement is given, how many batteries will likely be replaced out of a shipment of 10000?
- 5) A teacher marks some exams and finds the mean is 56% and the standard deviation is 9%. The teacher then adjusts the marks by raising the mean to 62% and raising the standard deviation to 11%. The z-scores are kept constant. If a student scored 73% initially, what would their new mark be?
- 6) The mean is $4.1x$ and the standard deviation is $4x - 1$. If the z-score is -2.9 and the raw data point is $2x+2$, determine the actual value of the mean.
- 7) A company producing stereos find they have a mean life of 8 years and a standard deviation of 1.7 years. What guarantee should the company make so they will only have to exchange fewer than 9% of all stereos sold?
- 8) The mean life of a car without part failure is 10 years with a standard deviation of 20 months. If the manufacturer makes 76000 cars, how many cars will have a defect within 5 years?

STATISTICS LESSON 3

PART IV: APPLICATION QUESTIONS

Answers:

1) If only 9% of hard drives are returned, that means the probability (area) is 0.0900. From the z-score table, this gives a z-score of -1.34. Now plug all the known values ($\mu = 72, \sigma = 6, z = -1.34$) into the formula $z = \frac{x - \mu}{\sigma}$ and solve for x.

x = 64 months

2) Plug all the known values ($\mu = 11, \sigma = 3.4, x = 3$) into the formula $z = \frac{x - \mu}{\sigma}$, then solve for the z-score, which is -2.35. This corresponds to a probability of **0.0094**.

3) Find the z-score and area for each length of toothpick.

For 3.57 cm, the z-score is -1.25, and the probability (on left) is 0.1056

For 3.68 cm, the z-score is 1.5, and the probability (on right) is 0.0668

Together, the total probability is 0.1724, so multiplying by 20000 gives **3448 picks**

4) Plug all the known values ($\mu = 13, \sigma = 0.75, x = 11$) into the formula $z = \frac{x - \mu}{\sigma}$, then solve for the z-score, which is -2.67. The probability (on left) is 0.0038, and multiplying by 10000 gives **38 batteries**

5) Plug everything into the

formula $\frac{x_1 - \mu_1}{\sigma_1} = \frac{x_2 - \mu_2}{\sigma_2}$ since

the z-scores are held constant.

$$\frac{x_1 - \mu_1}{\sigma_1} = \frac{x_2 - \mu_2}{\sigma_2}$$

$$\frac{73 - 56}{9} = \frac{x_2 - 62}{11}$$

$$x_2 = 82.78 = \mathbf{83\%}$$

6) Plug everything into the

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

$$-2.9 = \frac{2x + 2 - 4.1x}{4x - 1}$$

$$-2.9 = \frac{-2.1x + 2}{4x - 1}$$

$$-11.6x + 2.9 = -2.1x + 2$$

$$-9.5x = -0.9$$

$$x = 0.0947$$

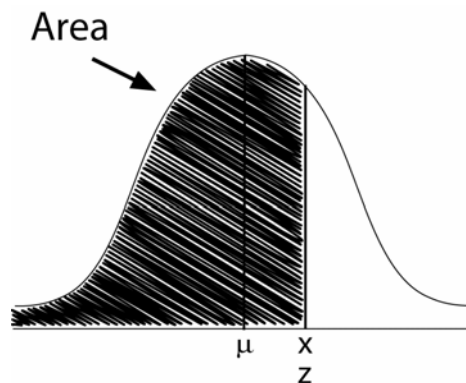
$$\text{Mean} = 4.1x = \mathbf{0.3884}$$

7) State what you know ($\mu = 8, \sigma = 1.7, \text{Area} = 0.0900$) then find the z-score, which is -1.34. Plug into the formula $z = \frac{x - \mu}{\sigma}$ and solve for x, which is **5.7 years**

8) Plug all the known values ($\mu = 120 \text{ months}, \sigma = 20 \text{ months}, x = 60 \text{ months}$) into the formula $z = \frac{x - \mu}{\sigma}$, then solve for the z-score, which is -3.00. This corresponds to a probability of 0.0013, and when multiplied by 76000 gives **99 cars**

z-score Table

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



z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
-3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
-3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
-3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
-3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
-2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
-2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
-2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
-2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
-2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
-2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
-2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
-2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
-2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
-2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
-1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
-1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
-1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
-1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
-1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
-1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
-1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
-1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
-1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
-1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
-0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
-0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
-0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
-0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
-0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
-0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
-0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
-0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
-0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
-0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

[illegible]