Warm-Up

- You and your friends have measured the heights of your dogs: 600mm, 470mm, 170mm, 430mm, and 300mm. Find the variance and standard deviation for the data set.

Mean: 394 Variance: 21, 704 Standard Deviation: 147.32

Heads Up

TEST is Tuesday!

PROJECT is due Wednesday!

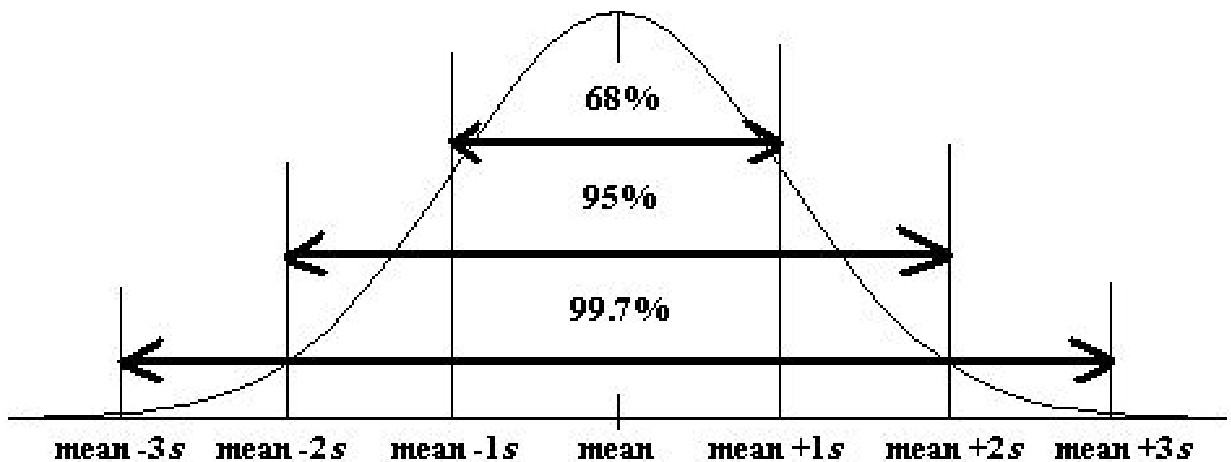
Objectives

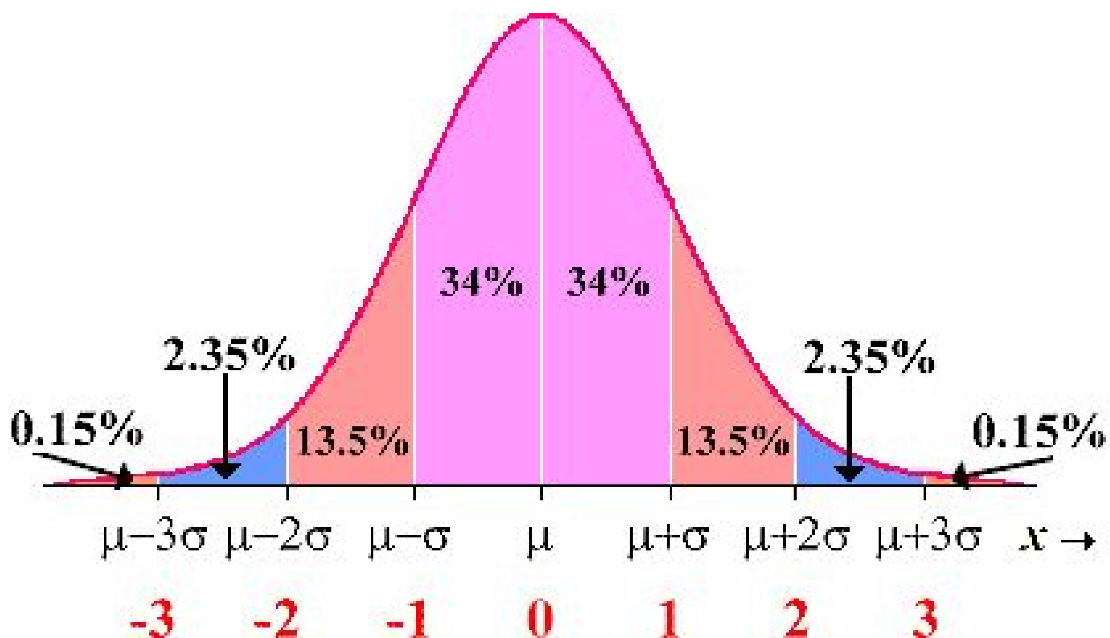
The student will be able to:

use empirical rule to calculate percentages

find the z-**scores** of a data set

The Normal Curve & Percentages





The Normal Curve

- Shows where the mean is
- Shows 1, 2, and 3 standard deviations above and below the mean
- Can easily get percentages between THOSE data points
- But what if we're talking about a data point that is not exactly 1, 2, or 3 s.d. away?

z-scores

When a set of data values are normally distributed, we can standardize each score by converting it into a **z-score**.

z-scores make it easier to compare data values measured on different scales.

z-scores

A z-score reflects how many standard deviations above or below the mean a raw score is.

The z-score is positive if the data value lies above the mean and negative if the data value lies below the mean. What would the z-score be if it lies exactly at the mean?

z-score formula

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

Where x represents an element of the data set, the mean is represented by μ and standard deviation by σ.

Analyzing the data

Suppose SAT scores among college students are normally distributed with a mean of 500 and a standard deviation of 100. If a student scores a 700, what would be her z-score?

Answer Now

Analyzing the data

Suppose SAT scores among college students are normally distributed with a mean of 500 and a standard deviation of 100. If a student scores a 700, what would be her z-score?

$$z = \frac{700 - 500}{100} = 2$$

Her z-score would be 2 which means her score is two standard deviations above the mean.

Analyzing the data

- A set of math test scores has a mean of 70 and a standard deviation of 8.
- A set of English test scores has a mean of 74 and a standard deviation of 16.

For which test would a score of 78 have a higher standing?

Answer Now

Analyzing the data

A set of math test scores has a mean of 70 and a standard deviation of 8.
A set of English test scores has a mean of 74 and a standard deviation of 16.
For which test would a score of 78 have a higher standing?

To solve: Find the z-score for each test.

$$\text{math } z\text{-score} = \frac{78-70}{8} = 1$$

$$\text{English } z\text{-score} = \frac{78-74}{16} = .25$$

The math score would have the highest standing since it is 1 standard deviation above the mean while the English score is only .25 standard deviation above the mean.

Analyzing the data

What will be the miles per gallon for a Toyota Camry when the average mpg is 23, it has a z value of 1.5 and a standard deviation of 2?

Answer Now

Analyzing the data

What will be the miles per gallon for a Toyota Camry when the average mpg is 23, it has a z value of 1.5 and a standard deviation of 2?

Using the formula for z-scores: $z = \dfrac{x - \mu}{\sigma}$

$$1.5 = \frac{x - 23}{2} \qquad 3 = x - 23 \qquad x = 26$$

The Toyota Camry would be expected to use 26 mpg of gasoline.

Analyzing the data

A group of data with normal distribution has a mean of 45. If one element of the data is 60, will the z-score be positive or negative?

Answer Now

Analyzing the data

A group of data with normal distribution has a mean of 45. If one element of the data is 60, will the z-score be positive or negative?

The z-score must be positive since the element of the data set is above the mean.

Using Z-score to find Area

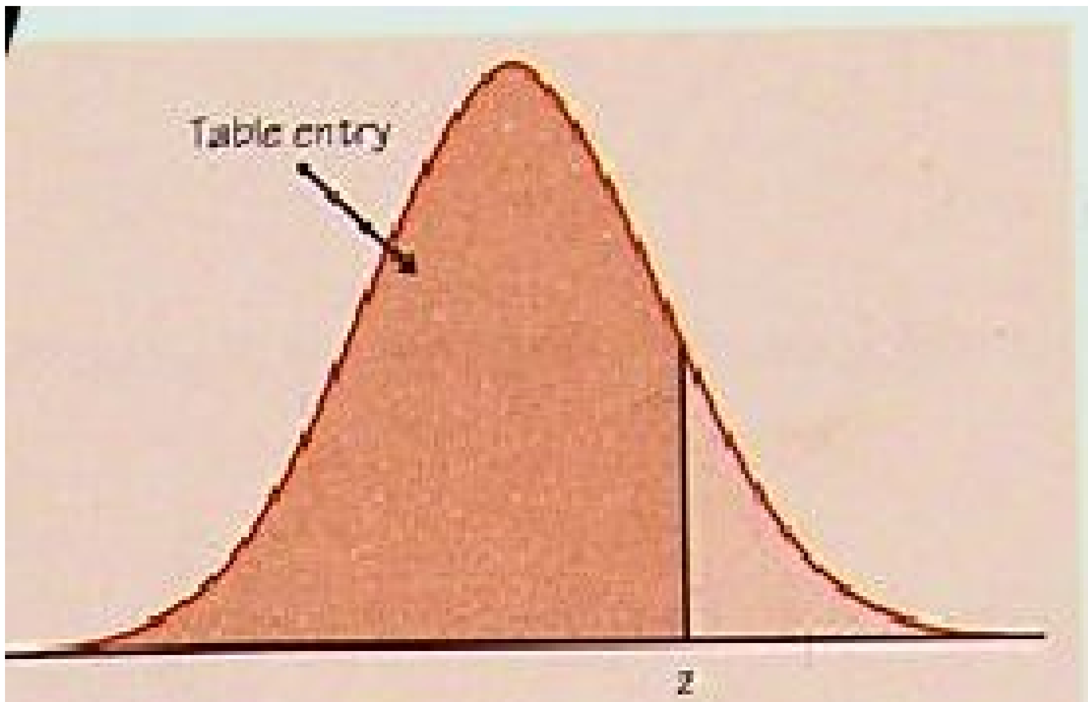
- Normal curves are useful for probability.
- The area under any section of the curve describes the probability that a value falls within a given range
- A table of z-scores gives us the area under the curve

Example

A student received a score of 56 on a normal distributed standardized test, with a mean of 50 and a standard deviation of 5. What is the probability that a randomly selected student achieved a lower score?

$$z = \frac{x-\mu}{\sigma} = \frac{56-50}{5} = 1.2$$

Z-Score shows the area under the curve



Using the z-score table we find the area

TABLE A Standard normal probabilities *(continued)*

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441

So about 88.5% will score below

Example #3

The distribution of blood cholesterol levels in young males is roughly normal with a mean of 165 and a SD of 30.

a) What proportion of young males have a level above 120?

b) What proportion will have a level between 120 & 200

1. Calculate the z-scores

Mean = 165
SD = 30
X = 120

$$z = \frac{120 - 165}{30} = -1.5$$

Mean = 165
SD = 30
X = 200

$$z = \frac{200 - 165}{30} = 1.17$$

How many above 120?

TABLE A Standard normal probabilities

z	.00	.01	.02	.03	.04
−3.4	.0003	.0003	.0003	.0003	.0003
−3.3	.0005	.0005	.0005	.0004	.0004
−3.2	.0007	.0007	.0006	.0006	.0006
−3.1	.0010	.0009	.0009	.0009	.0008
−3.0	.0013	.0013	.0013	.0012	.0012
−2.9	.0019	.0018	.0018	.0017	.0016
−2.8	.0026	.0025	.0024	.0023	.0023
−2.7	.0035	.0034	.0033	.0032	.0031
−2.6	.0047	.0045	.0044	.0043	.0041
−2.5	.0062	.0060	.0059	.0057	.0055
−2.4	.0082	.0080	.0078	.0075	.0073
−2.3	.0107	.0104	.0102	.0099	.0096
−2.2	.0139	.0136	.0132	.0129	.0125
−2.1	.0179	.0174	.0170	.0166	.0162
−2.0	.0228	.0222	.0217	.0212	.0207
−1.9	.0287	.0281	.0274	.0268	.0262
−1.8	.0359	.0351	.0344	.0336	.0329
−1.7	.0446	.0436	.0427	.0418	.0409
−1.6	.0548	.0537	.0526	.0516	.0505
−1.5	.0668	.0655	.0643	.0630	.0618
−1.4	.0808	.0793	.0778	.0764	.0749

Remember the Z-score gives the amount below. So to find the amount of above we have to subtract from 100

100 - 6.68% = 93.32%

Amount between 120 and 200

z	.00	.01	.02	.03	.04	.05	.06	.07
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980

Below 200 is 87.90%

Between 120 and 200 would be

87.90 – 6.68=81.22%

Practice

Choose the sampling method identified by the following scenario:

1. A university polled 500 of its students, randomly selecting them proportional to the number of students enrolled in each degree program.
2. To do market research, a telemarketing firm randomly selected 1000 names from a store's database and contacted them.
3. A random starting point is chosen, followed by every 10^{th} individual.