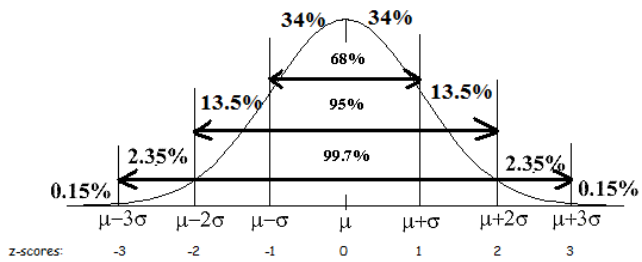
Secondary Math 3 Honors Note Guide 8.2
Drawing Conclusions from Data Unit

Assessing Normality

In statistics, density curves are used to describe data distributions. One particularly important class of density curves are Normal curves. They play a large role in statistics and are rather special but not necessarily common.

A Normal distribution is described by a Normal density curve. Any particular Normal distribution is completely specified by two numbers: its mean μ and standard deviation σ. The mean of a Normal distribution is at the center of the symmetric Normal curve. The standard deviation is the distance from the center to the change of curvature points on either side. We abbreviate the Normal distribution with mean μ and standard deviation σ as $N(\mu, \sigma)$.

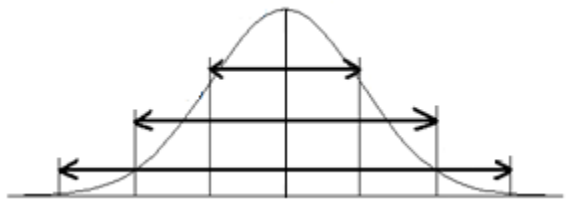
68-95-99.7 Rule



Recall:

- The mean μ is the sum of the data values divided by the number of values.
- The standard deviation σ measures the typical distance of the values in a distribution from the mean.

Example: The distribution of weights of 9-ounce bags of a particular brand of potato chips is approximately Normal with mean $\mu = 9.12$ ounces and standard deviation $\sigma = 0.05$ ounce. On the sketch below label the mean, as well as the points 1, 2, and 3 standard deviations away from the mean on the horizontal axis.



8.97 9.02 9.07 9.12 9.17 9.22 9.27

a- Identify the interval(s) that contain the given approximate areas under the curve.

95%	47.5%	16%	97.35%
9.02 to 9.22	9.02 to 9.12	below 9.07	8.97 to 9.22
	9.12 to 9.22	above 9.17	9.02 to 9.27

b- What percentage of potato chip bags weigh less than 9.02 ounces?

2.5%

c- What percentage of potato chip bags weigh more than 9.07 ounces?

84%

d- What percentage of potato chip bags weigh between 8.97 and 9.22 ounces?

97.35%

To use many of the basic inference procedures of statistics, it is necessary to make sure that the data is approximately normal. Consequently, we need to develop a strategy for assessing Normality. If a graph of the data is clearly skewed or has multiple peaks, or isn't bell-shaped, that's evidence that the distribution is not Normal. However, just because a plot of the data looks normal, we can't say that the distribution is Normal. The 68-95-99.7 rule can give additional evidence in favor of or against normality.

Example: Are the lengths of great white sharks normally distributed? The data from a random sample of 44 great white sharks is given below from smallest to largest for your convenience. The mean is 15.59 feet and a standard deviation of 2.55 feet.

9.4	12.4	13.2	13.6	14.7	15.7	16.1	16.4	16.8	18.2	18.7
12.1	12.6	13.2	13.8	14.9	15.7	16.2	16.6	17.6	18.3	19.1
12.2	13.2	13.5	14.3	15.2	15.8	16.2	16.7	17.8	18.6	19.7
12.3	13.2	13.6	14.6	15.3	15.8	16.4	16.8	17.8	18.7	22.8

A graph shows no extreme skewness or outliers and is single peaked so test 68-95-99.7 rule.
Proportion between 1 standard deviation (13.04 to 18.14): $30/44 \approx 0.6818$
Proportion between 2 standard deviation (10.49 to 20.69): $42/44 \approx 0.9545$
Proportion between 3 standard deviation (7.94 to 23.24): $44/44 = 1.0$
Follows the rule very closely so the data can be assumed to be normally distributed.

Normal curves are used in quality control of products.

A *control chart* is a graph used to study how a process changes over time. Data are plotted in time order. A *control chart* always has a central line for the average, an upper line for the upper *control* limit (3 standard deviations above the mean) and a lower line for the lower *control* limit (3 standard deviations below the mean).

A process is said to be out of control if a value falls outside the control limits or a run of 9 points of data is on one side of center. At this point the process would be stopped to look for a cause and fix it.

Example: A potato chip company checks the salt content of a random sample of potato chips every 15 minutes during the process. The salt content should have a mean of 2 mg and varies by 0.15 mg.

a) Give the upper and lower control limits.
   1.55 and 2.45 mg

b) Is the following data from an in control or out of control process? Explain.
   2.32, 1.93, 1.58, 1.87, 1.67, 1.99, 1.85, 1.62, 1.73, 1.77, 1.96, 2.01, 2.26
   Out of control, 10 points on one side of center

c) Is the following data from an in control or out of control process? Explain.
   2.39, 2.14, 2.01, 1.97, 2.11, 1.83, 1.94, 1.99, 2.12, 2.29, 1.97, 2.08, 2.41
   In control; no points outside limits and no series of 9 on one side of the mean

d) Is the following data from an in control or out of control process? Explain.
   2.43, 2.37, 2.17, 1.99, 2.03, 1.45, 1.72, 1.93, 2.07, 2.13, 1.87, 1.74, 1.97
   Out of control; 1.45 is smaller than 1.55 the lower limit

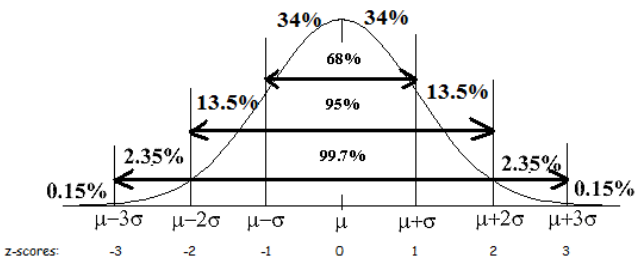
Secondary Math 3 Honors Note Guide 8.2
Drawing Conclusions from Data Unit

Assessing Normality

In statistics, density curves are used to describe data distributions. One particularly important class of density curves are Normal curves. They play a large role in statistics and are rather special but not necessarily common.

A _____ _____ is described by a Normal density curve. Any particular Normal distribution is _____ _____ by two numbers: its mean μ and standard deviation σ. The mean of a Normal distribution is at the _____ of the _____ Normal curve. The standard deviation is the _____ from the center to the change of _____ _____ on either side. We abbreviate the Normal distribution with mean μ and standard deviation σ as _____.

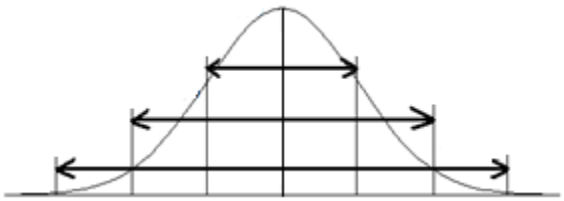
68-95-99.7 Rule



Recall:

- The _____, __, is the _____ of the data values _____ by the number of values.
- The _____ _____ , ___, measures the _____ _____ of the values in a distribution from the _____.

Example: The distribution of weights of 9-ounce bags of a particular brand of potato chips is approximately Normal with mean $\mu = 9.12$ ounces and standard deviation $\sigma = 0.05$ ounce. On the sketch below label the mean, as well as the points 1, 2, and 3 standard deviations away from the mean on the horizontal axis.



a- Identify the interval(s) that contain the given approximate areas under the curve.

95% 47.5% 16% 97.35%

b- What percentage of potato chip bags weigh less than 9.02 ounces?

c- What percentage of potato chip bags weigh more than 9.07 ounces?

d- What percentage of potato chip bags weigh between 8.97 and 9.22 ounces?

To use many of the basic inference procedures of statistics, it is necessary to make sure that the data is _____ _____. Consequently, we need to develop a strategy for _____ _____. If a graph of the data is clearly _____ or has _____ _____, or isn't _____, that's evidence that the distribution is _____ Normal. However, just because a plot of the data _____ _____, we can't say that the distribution is _____. The 68-95-99.7 rule can give _____ _____ in favor of or against normality.

Example: Are the lengths of great white sharks normally distributed? The data from a random sample of 44 great white sharks is given below from smallest to largest for your convenience. The mean is 15.59 feet and a standard deviation of 2.55 feet.

9.4	12.4	13.2	13.6	14.7	15.7	16.1	16.4	16.8	18.2	18.7
12.1	12.6	13.2	13.8	14.9	15.7	16.2	16.6	17.6	18.3	19.1
12.2	13.2	13.5	14.3	15.2	15.8	16.2	16.7	17.8	18.6	19.7
12.3	13.2	13.6	14.6	15.3	15.8	16.4	16.8	17.8	18.7	22.8

Normal curves are used in quality control of products.

A _____ _____ is a graph used to study how a process changes over _____. Data are plotted in time order. A *control chart* always has a _____ _____ for the _____, an _____ _____ for the upper *control* _____ (___ standard deviations above the mean) and a _____ _____ for the lower *control* _____ (___ standard deviations below the mean).

A process is said to be _____ ____ _____ if a value falls _____ the control limits or a run of _____ _____ of data is on one side of _____. At this point the process would be stopped to look for a cause and fix it.

Example: A potato chip company checks the salt content of a random sample of potato chips every 15 minutes during the process. The salt content should have a mean of 2 mg and varies by 0.15 mg.

a) Give the upper and lower control limits.

b) Is the following data from an in control or out of control process? Explain.
2.32, 1.93, 1.58, 1.87, 1.67, 1.99, 1.85, 1.62, 1.73, 1.77, 1.96, 2.01, 2.26

c) Is the following data from an in control or out of control process? Explain.
2.39, 2.14, 2.01, 1.97, 2.11, 1.83, 1.94, 1.99, 2.12, 2.29, 1.97, 2.08, 2.41

d) Is the following data from an in control or out of control process? Explain.
2.43, 2.37, 2.17, 1.99, 2.03, 1.45, 1.72, 1.93, 2.07, 2.13, 1.87, 1.74, 1.97