Suppose you have a normally distributed [random variable](javascript:glosItem%20=%20'random_variables';void%20openURL('../help/glossary.html','glossary',640,480)) and would like to calculate the probability of its value occurring in the interval of the mean plus or minus .5 standard deviations. How would you go about calculating the probability? This can be achieved using standardization.

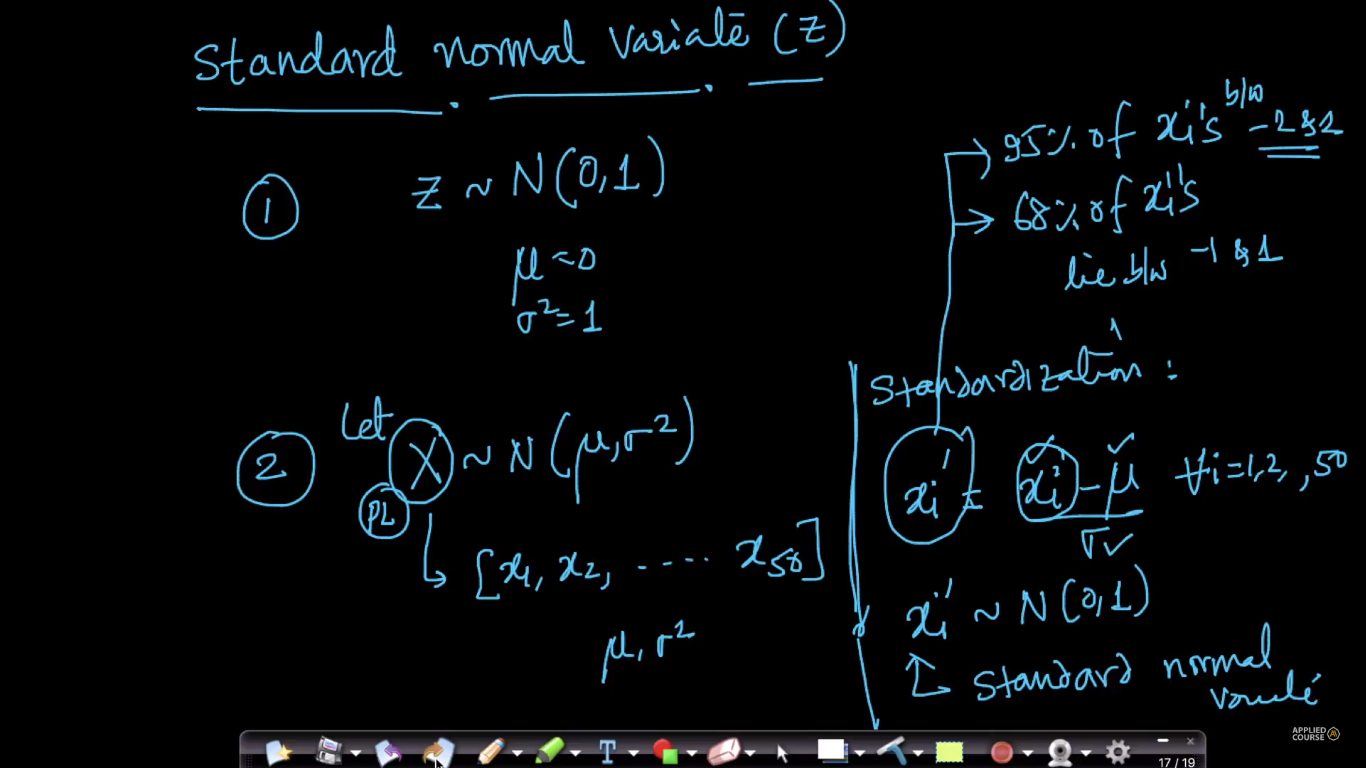
**Standardization:**

 Standardization comes from the term standard deviation and hence it is the task for mean centering meaning you subtract each term from the mean and divide by the standard deviation of the random variable.

If a given random variable follows normal distribution then, we can apply standardization on it to get a variable whose mean will be 0 and variance will be 1.

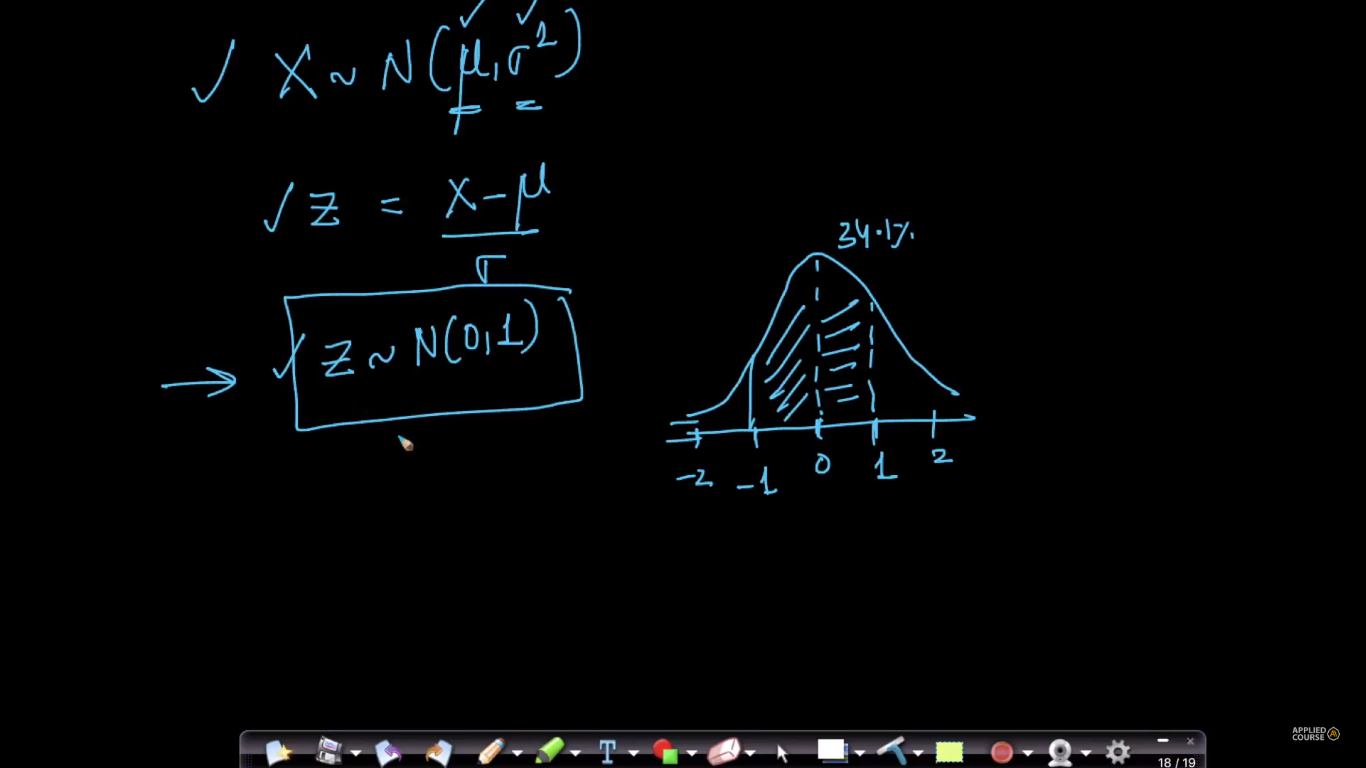
Standardization will convert each point of given random variable into standard form using formula given below.

The distribution(with mean 0 and SD 1) we get after standardization is called **standard variate.**



After doing standardization we can say directly that

* 68% of x’s lies between -1 & 1
* 95% of x’s lies between -2 & 2
* 99.7% of x’s lies between -3 & 3.



**Must read referecnce:** <http://ci.columbia.edu/ci/premba_test/c0331/s6/s6_4.html>

**Some Comments:**

Normalization and Standardization are two techniques of Feature Scaling. We generally perform Feature Scaling to get all the features on to the same scale and make the interpretation easier.

**Normalization**:  
Normalization of a feature is transforming all the values of a feature onto a scale with a minimum value of 0 and a maximum value of 1.

Let us assume we have  a feature 'f' in our dataset. Then if we want to normalize this feature, the formula is

**f\_i = (f\_i - f\_min)/(f\_max - f\_min)**

where

f\_i --> ith element in the column vector of 'f'  
f\_min --> the minimum value in the column vector of 'f'  
f\_max --> the maximum value in the column vector of 'f'  
  
  
**Standardization**:

Standardization is the process of transforming the features on to a scale with a mean of 0 and variance of 1. If we have 'f' as a feature, then the standardization of a feature is given by

**f\_i = (f\_i - f\_mean)/f\_sd**  
where  
f\_i --> ith value in the column vector of 'f'  
f\_mean --> the mean of all the values in the column vector of 'f'  
f\_sd --> the standard deviation of all the values in  the column vector of 'f'  
  
  
In sklearn, the normalization can be performed using **MinMaxScaler**module and the standardization can be performed using **StandardScaler**module.

**Note:** Standardization should be used only where you don’t want to preserve variance or on only that features which have capability of classifying them even after standardization.

Therefore places where you need to preserve variance then standardization should not be performed, Example: height of people, some are very tall and some are short and we need to keep this info as it is, because we don’t want to reduce variance which is classifying them.