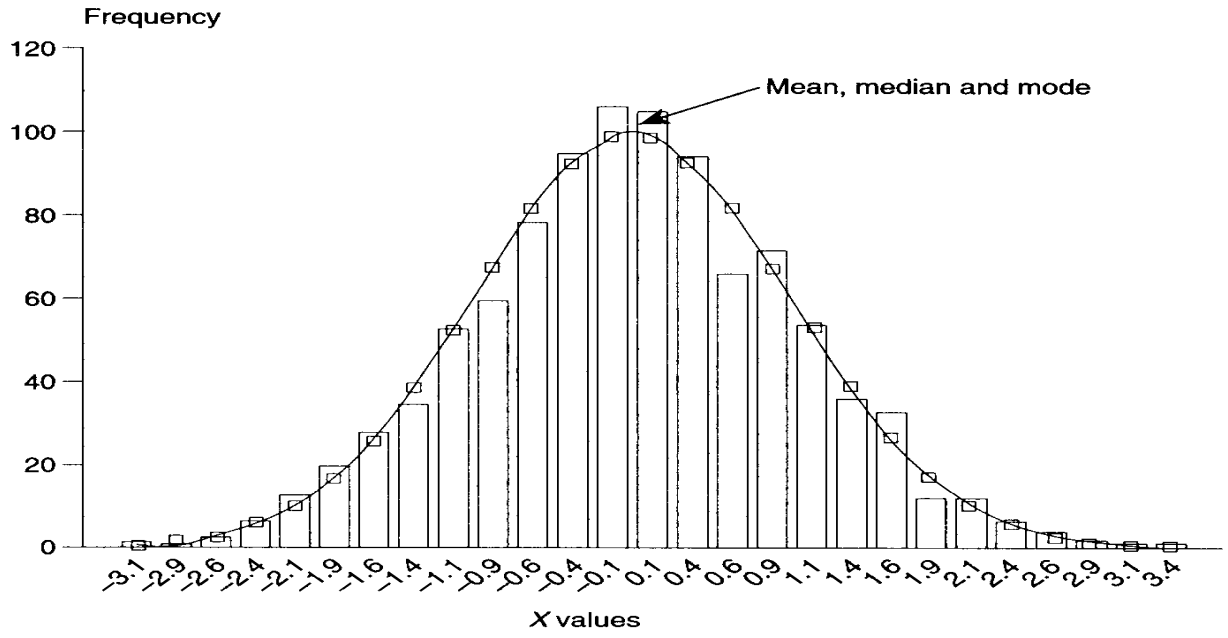
Normal Distribution

Normal or Gaussian distribution is a symmetric distribution. Normal distribution has two parameters: mean (μ) the location parameter and standard deviation (σ), the spread parameter.
If a X is a continuous random variable with mean μ and standard deviation σ, then it is written as, $X \sim N(\mu , \sigma)$. Note that, the bars in the histogram below must be connected to each other, not separate.



Source: emeraldinsight.com

Properties of Normal distribution

1. This is a distribution for continuous random variable. Therefore probability is computed by measuring the area under the curve rather than the curve height or frequency or count.

2. Normal distribution curve is bell-shaped, symmetric around its mean.

3. The normal distribution has a single mode. That is why this is also known as a unimodal distribution.

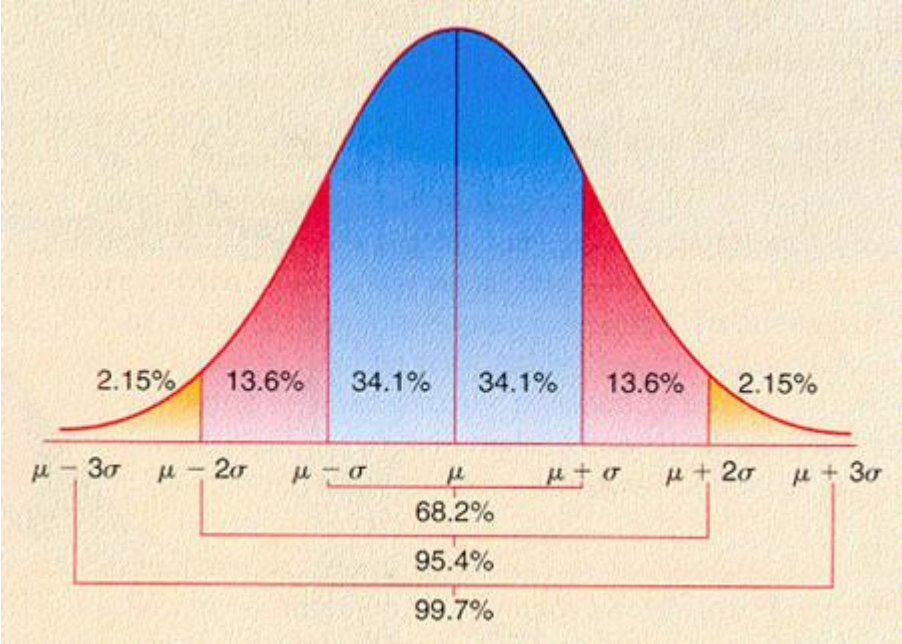
4. For normal distribution, mean = median = mode.

5. The probability distribution is highest exactly at the mean (in other words, mean has the highest frequency, it is the most likely value for the random variable).

6. Most of the statistical inference theory is based on the properties of the normal distribution.

7. Total area under the normal distribution curve is 1.

The Empirical Rule

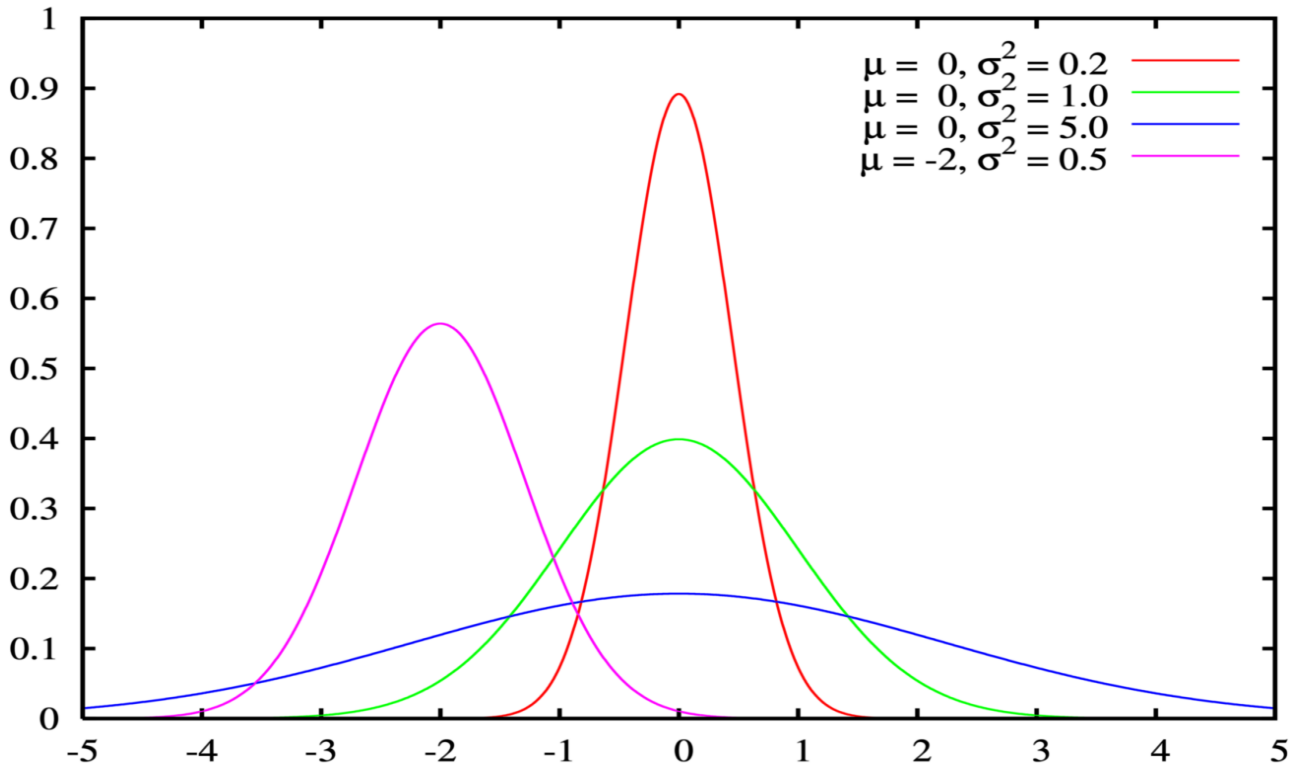


Source: pdnotebook.com

If the histogram of values in a data set can be reasonably approximated by a normal curve, then

a) approximately 68% of the observations are within 1 standard deviation of the mean;
b) approximately 95% of the observations are within 2 standard deviations of the mean;
c) approximately 99.7% of the observations are within 3 standard deviations of the mean.

Normal distribution has two parameters – mean (μ) the location parameter and standard deviation (σ), the spread parameter.



psychology.wikia.com

Changing the mean will cause the distribution to shift its location. Standard deviation controls the spread of the distribution; the larger it is the wider is the distribution shape.

Standard normal distribution

This is a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$.

If a variable x has any normal distribution with mean μ and standard deviation σ, then the standardized variable

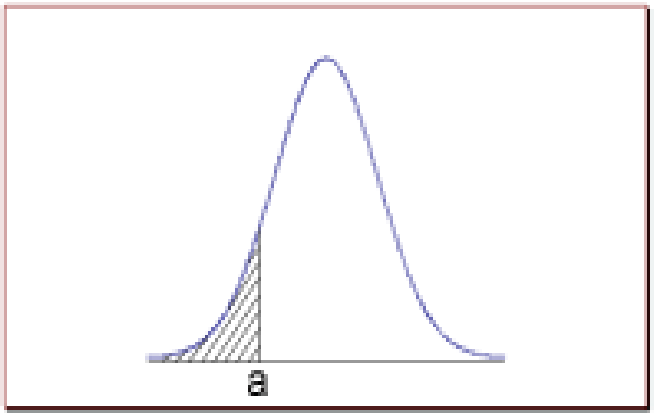
$$z = (x - \mu)/\sigma$$

has the standard normal distribution, $z \sim N(0,1)$.

Finding normal probabilities

Areas under a normal curve represent proportions of observations from that normal distribution. So they represent probabilities of randomly selecting an individual from that distribution.

Cumulative areas or probabilities under the standard normal curve are available in a table form. Standard normal curve is symmetric, the distribution can be divided into two equal parts at $\mu = 0$. So all the negative z values are on the left side and positive z values are on the right side of the curve. Probabilities on the left side of the curve are the same as the probabilities on the right side of the curve.



The shaded region in this graph represents the cumulative probability that z is less than or equal to a, that is $P(z \leq a)$.

Section 6.1

The random variable x has normal distribution with mean μ and variance σ^2. We can find probabilities for three different cases:

Case 1: $P(x \leq a)$. Find Z score for 'a' and then use the z table.

Case 2: $P(x \geq b) = 1 - P(x \leq b)$. Find Z score for 'b' and then use the z table.

Case 3: $P(a \leq x \leq b) = P(x \leq b) - P(x \leq a)$. Find Z scores for both 'a' and 'b' and then use the z table.

The z score formula is $z = (x - \mu)/\sigma$.