**Naive Bayes Classifier**

The Bayes Theorem is used to construct a set of classification algorithms known as Naive Bayes classifiers. It is a family of algorithms that share a similar idea, namely that each pair of features being classified is independent of the others.

In most real-world cases, the conclusions taken by Naive Bayes are incorrect. In truth, the presumption of freedom is never true, but it often works well in practice.

Naïve Bayes assumption is that each feature is:

* Independent
* Equal

We presume that no two functions are dependent. As a result, the characteristics are considered to be distinct.

Each function is evenly weighted. None of the features are unimportant, and they are all considered to have an equal effect on the result.

**Bayes’ Theorem**

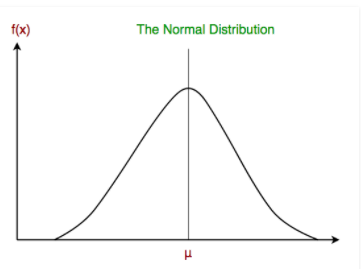
The Bayes Theorem measures the probability of an occurrence happening given the probability of a preceding event. The following equation expresses Bayes' theorem mathematically:

where A and B are events

* we're searching for the possibility of event A if event B is real. Proof is often referred to as Case B.
* The priori of A is P(A) (the prior probability, i.e. Probability of event before evidence is seen). The proof is a value attributed to an undefined instance's attribute (here, it is event B).
* P(A|B) is the a posteriori likelihood of B, or the probability of an occurrence after seeing proof.

**Gaussian Naive Bayes classifier**

Continuous values associated with each function are believed to be distributed according to a Gaussian distribution in Gaussian Naive Bayes. A regular distribution is also known as a Gaussian distribution. It generates a bell-shaped curve that is symmetric around the mean of the function values when plotted, as seen below:



**Source**: geeksforgeeks

Popular Naive Bayes classifiers are:

**Multinomial Naive Bayes**: The frequency with which certain events were induced by a multinomial distribution is represented by feature vectors. This is the most popular event model for record classification.

**Bernoulli Naive Bayes**: Features are independent Booleans (binary variables) that define inputs in the multivariate Bernoulli event model. This model, like the multinomial model, is common for document classification tasks where binary term occurrence (i.e., whether or not a word appears in a document) features are used instead of term frequencies (i.e. frequency of a word in the document)

Sentiment analysis, spam filtering, recommendation systems, and other applications use naive Bayes algorithms.