

Нормальное распределение (Normal Distribution)

Синонимы: Распределение Гаусса, Gauss Distribution

Разделы: [Метрики](#)

Loginom: [Статистика \(визуализатор\)](#).

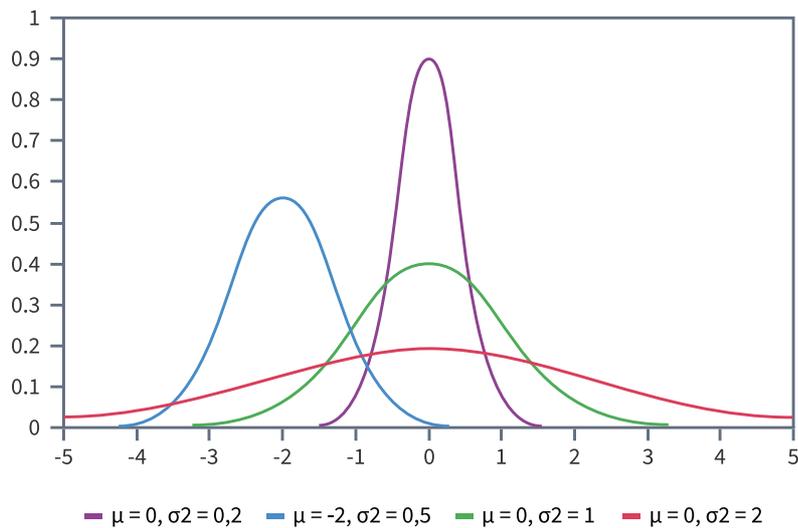
Нормальным называется распределение вероятностей, которое для одномерного случая задается функцией Гаусса.

Нормальное распределение играет важнейшую роль во многих областях знаний. Случайная величина подчиняется нормальному закону распределения, когда она подвержена влиянию большого числа случайных факторов, что является типичной ситуацией в анализе данных. Поэтому нормальное распределение служит хорошей моделью для многих реальных процессов.

Нормальное распределение зависит от 4-х параметров:

- математическое ожидание — «центр тяжести» распределения;
- дисперсия — степень разброса случайной величины относительно математического ожидания;
- коэффициент асимметрии — параметр формы распределения, определяющий его симметрию относительно математического ожидания;
- коэффициент эксцесса — параметр распределения, задающий «остроту» пика распределения.

Типичные формы нормального распределения для различных средних и дисперсии представлены на рисунке:



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

где μ — математическое ожидание, σ^2 — дисперсия.

Коэффициент асимметрии определяется следующим образом:

$$\gamma_1 = E\left[\left(\frac{x-\mu}{\sigma}\right)^3\right],$$

где E -знак усреднения. Если коэффициент асимметрии положителен, то правый «хвост» распределения длиннее левого, и отрицателен в противном случае. Если распределение симметрично относительно математического ожидания, то его коэффициент асимметрии равен нулю.

Коэффициент эксцесса вычисляется по формуле:

$$\gamma_2 = E\left[\left(\frac{x-\mu}{\sigma}\right)^4\right].$$

Возрастание коэффициента эксцесса увеличивают остроту пика распределения.

Смысл нормального распределения становится понятен из его формы. Наиболее вероятные значения случайной величины расположены вблизи его пика. По мере удаления от него, вероятность значений уменьшается и если значение расположено в «хвосте» распределения, то оно очень маловероятно.

Если случайная величина подчиняется нормальному закону распределения (о чем принимается соответствующая статистическая гипотеза, проверяемая по различным критериям), то статистический анализ описываемого ей процесса существенно упрощается.

В частности, предположение о нормальности распределения коэффициентов регрессии используется при построении регрессионных моделей, а в простом классификаторе Байеса — предположение о нормальности распределения непрерывных атрибутов, на основе которых производится классификация.

