Нормальное двумерное распределение

Нормальное одномерное распределение

Прежде чем рассмотреть, что представляет собой двумерное нормальное распределение, вспомним, что представляет собой его одномерный вариант. Главное, что нужно понимать: **нормальное распределение – это распределение вероятностей, которое в одномерном случае задаётся функцией плотности вероятности, совпадающей с функцией Гаусса**.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad [1]$$

В данной формуле:

- σ – среднеквадратичное отклонение (квадрат этой величины является диспресией)
- μ – математическое ожидание
- x – значение случайное величины

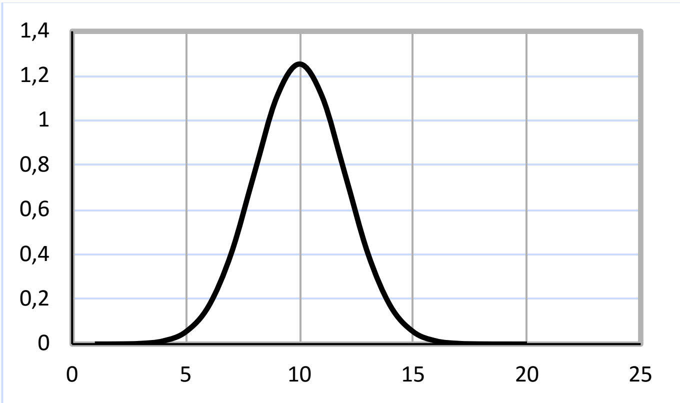

Рисунок 1

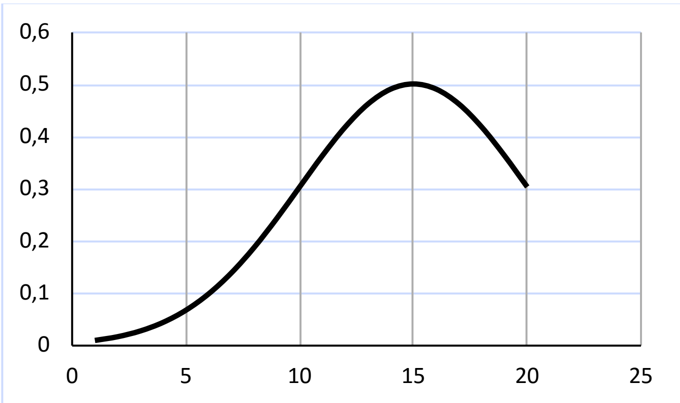

Рисунок 2

На рисунке 1 приведен пример нормального распределения случайной величины, при котором среднеквадратичное отклонение равно 2, математическое ожидание равно 10.

На рисунке 2 тоже равномерное нормальное распределение, но среднеквадратичное отклонение для него равно 4, математическое ожидание равно 15. Нетрудно заметить, что увеличение математического ожидания сказалось на смещении «вершины» графика вправо, а увеличение среднеквадратичного отклонения сделало график более пологим.

Нормальное распределение используется при оценке надежности изделий, на которые воздействует ряд случайных факторов, каждый из которых незначительно влияет на результирующий эффект (нет доминирующих факторов). Доказывается, что сумма достаточно большого числа независимых (или слабо зависимых) случайных величин, подчиненных каким угодно законам распределения (при соблюдении некоторых нежестких ограничений), приближенно подчиняется нормальному закону, и это выполняется тем точнее, чем большее количество случайных величин суммируется. Основное ограничение, налагаемое на суммируемые случайные величины, состоит в том, чтобы они все равномерно играли в общей сумме относительно малую роль. Если это условие не выполняется и, например, одна из случайных величин окажется по своему влиянию на сумму резко превалирующей над всеми другими, то закон распределения этой превалирующей случайной величины наложит свое влияние на сумму и определит в основных чертах ее закон распределения.

Нормальное двумерное распределение

В случае, когда есть не одна случайная величина, а несколько, причем все они не зависят друг от друга, можно дать следующее определение многомерного нормального закона: **если каждая случайная величина подчиняется нормальному закону и они совместно независимы, то система подчиняется многомерному нормальному закону**. Из этого следует, что f(x, y) – плотность двумерного нормального закона, если ее интегрирование по всем возможным значениям одной случайной величины дает нормальный закон распределения другой.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2{\sigma_x}^2}} \qquad [2]$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx = \frac{1}{\sigma_y\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2{\sigma_y}^2}} \qquad [3]$$

Формула 4 описывает плотность двумерного нормального распределения. Следует заметить, что это – во-первых, каноническая форма записи закона распределения. Иную форму можно получить, если перемножить функции 2 и 3. Помимо этого, как уже было сказано, две случайные величины независимы, и, кроме того, центрированы – это значит, что математическое ожидание должно быть равным нулю. Формула 5 демонстрирует вывод формулы 4 из формул 2 и 3.

$$f(\xi_1, \xi_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{\left(-\frac{1}{2}\left(\frac{\xi_1^2}{\sigma_1^2}+\frac{\xi_2^2}{\sigma_2^2}\right)\right)} \qquad [4]$$

$$f(x,y) = f_X(x)\, f_Y(y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{(x-\mu)^2}{2{\sigma_x}^2}+\frac{(y-\mu)^2}{2{\sigma_y}^2}\right)} = \frac{1}{2\pi\sigma_x\sigma_y} e^{\left(-\frac{1}{2}\left(\frac{(x-\mu)^2}{{\sigma_x}^2}+\frac{(y-\mu)^2}{{\sigma_y}^2}\right)\right)} = \left|\begin{array}{c} \xi_1 = x-\mu \\ \xi_2 = y-\mu \\ \sigma_1 = \sigma_x \\ \sigma_2 = \sigma_y \end{array}\right|$$
$$= \frac{1}{2\pi\sigma_x\sigma_y} e^{\left(-\frac{1}{2}\left(\frac{{\xi_1}^2}{{\sigma_x}^2}+\frac{{\xi_1}^2}{{\sigma_y}^2}\right)\right)} \qquad [5]$$

Линии уровня плотности двумерного нормального распределения

Можно заметить, что приведенная ранее формула не зависит от значений случайных величин в двух случаях: когда эти величины одни и те же, либо, когда показатель степени при их изменении остается постоянным.

Первый случай не представляет интереса, второй в геометрической интерпретации касается понятия линий уровня в математике. **Линией уровня функции называется линия на плоскости, в каждой точке которой функция сохраняет постоянное значение**. Формула 6 задает линию уровня для плотности двумерного нормального распределения при каком-либо значении величины k.

$$\left(\frac{\xi_1^2}{\sigma_1^2}+\frac{\xi_2^2}{\sigma_2^2}\right) = k^2 = const \qquad [6]$$

Формула 7 представляет канонический вид уравнения эллипса, соответственно, становится очевидно, что линия уровня плотности нормального распределения имеет вид эллипса.

$$\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right) = 1 \qquad [7]$$

Теперь введем три важных термина, касающихся рассмотренных линий уровня:

Эллипс равной плотности – линия уровня плотности двумерного нормального распределения. Эллипс рассеивания – фигура, ограниченная эллипсом равной плотности. Центр рассеивания – центр эллипса равной плотности.

Оставим линии уровня плотности нормального распределения и рассмотрим альтернативный способ задания эллипса рассеивания (формула 8) – в виде операций над матрицами (в векторном виде): в этом случае рассматриваются две матрицы:

- $\bar{\xi}$ – матрица, состоящая из одного столбца (в математике называется вектором-столбцом), состоящего из двух случайных величин – в формуле также используется транспонированная к данной матрица
- D – диагональная матрица, состоящая из главных дисперсий, в формуле используется матрица, обратная к ней.

$$B_k = \{\bar{\xi}: \bar{\xi}\bar{\xi}^T D^{-1} = k^2\} \qquad [8]$$

Формула 9 демонстрирует вывод этого способа задания из приведенного ранее.

$$k^2 = \left(\frac{\xi_1^2}{\sigma_1^2} + \frac{\xi_2^2}{\sigma_2^2}\right) = \begin{pmatrix} \xi_1^2 \\ \xi_2^2 \end{pmatrix} * \begin{pmatrix} \sigma_1^{-2} & \sigma_2^{-2} \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} * \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}^{\mathrm{T}} * \begin{pmatrix} \sigma_1^{-2} & 0 \\ 0 & \sigma_2^{-2} \end{pmatrix} = \bar{\xi}\bar{\xi}^T D^{-1} \qquad [9]$$

Многомерный нормальный закон в канонической форме

Должно быть очевидно, что если есть одномерный нормальный закон и двумерный нормальный закон, то должен быть также и 3-мерный, и n-мерный. Получим формулу для плотности распределения много мерного нормального закона в канонической форме.

Первое, что следует сделать – это задать два вектора – столбца $\bar{x}$ и $\bar{\xi}$, которые связаны соотношением 10, в котором C – некая квадратная матрица размером n*n, где n – количество случайных величин. Первая матрица состоит из исходных случайных величин. Введение второй матрицы может оказаться полезным в том случае, если производится какое-либо преобразование координат, например, их поворот на некоторый угол α – этот пример описывается формулой 11.

$$\bar{\xi} = C\bar{x}$$

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \dots \\ x_n \end{bmatrix} \bar{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \dots \\ \xi_n \end{bmatrix} C = \begin{bmatrix} C_{11} & C_{12} & C_{13} & \dots & C_{1n} \\ C_{21} & C_{22} & C_{23} & \dots & C_{2n} \\ C_{31} & C_{32} & C_{33} & \dots & C_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ C_{n1} & C_{n2} & C_{n3} & \dots & C_{nn} \end{bmatrix} \qquad [10]$$

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \bar{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} C = \begin{bmatrix} cos\alpha & sin\alpha \\ -sin\alpha & cos\alpha \end{bmatrix}$$

$$\bar{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} * \begin{bmatrix} cos\alpha & sin\alpha \\ -sin\alpha & cos\alpha \end{bmatrix} = \begin{bmatrix} x_1 cos\alpha + x_2 sin\alpha \\ -x_1 sin\alpha + x_2 cos\alpha \end{bmatrix} \qquad [11]$$

Если имеют место вышеоговоренные условия, то матрицу D можно выразить формулой 12, где K – матрица корреляционных моментов, которая должна быть такой, чтобы выполнялось соотношение 13.

$$D = CKC^{\mathrm{T}} \qquad [12]$$

$$D_{ii} = {\sigma_i}^2 \qquad [13]$$

Возникает новый вопрос – чему равна матрица K? Выражения 14 демонстрируют вывод величин, определяющих матрицу K.

$$B_k = \left\{\bar{\xi}: \bar{\xi}\bar{\xi}^T D^{-1} = k^2\right\} = \{\bar{x}: (C\bar{x})(C\bar{x})^T D^{-1} = k^2\} = \{\bar{x}: \bar{x}\bar{x}^T C D^{-1} C^T = k^2\} = \qquad [14]$$

$$= \left\{ \bar{x} : \bar{x}\bar{x}^T \quad K^{-1} = k^2 \right\}$$

Итак, мы почти вплотную подобрались к определению вида многомерного нормального закона в канонической форме. Остается сделать еще несколько шагов. Формула 15 демонстрирует подстановку матрицы D в выражение 4.

$$f(\xi_1, \xi_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{\left(-\frac{1}{2}\left(\frac{\xi_1^2}{\sigma_1^2} + \frac{\xi_2^2}{\sigma_2^2}\right)\right)} = \frac{1}{2\pi\sqrt{detD}} e^{\left(-\frac{1}{2}\bar{\xi}\bar{\xi}^T D^{-1}\right)} = f(\bar{\xi}) \qquad [15]$$

Это выражение также будет описывать и плотность нормального распределения n-мерного вектора. Однако еще требуется перейти от $\bar{x}$ к $\bar{\xi}$:

$$f(\bar{\xi}) = \frac{1}{2\pi\sqrt{detD}} e^{\left(-\frac{1}{2}\bar{\xi}\bar{\xi}^T D^{-1}\right)} = \frac{1}{2\pi\sqrt{detK}} e^{\left(-\frac{1}{2}\bar{x}\bar{x}^T \quad K^{-1}\right)} = f(\bar{x}) \qquad [16]$$

Более общей, тем не менее, будет запись, учитывающая, что случайные величины могут быть не центрированы:

$$f(\bar{x}) = \frac{1}{2\pi\sqrt{detK}} e^{\left(-\frac{1}{2}(\bar{x}-\bar{m})(\bar{x}-\bar{m})^T \quad K^{-1}\right)} \qquad [17]$$

В общей форме многомерный закон выглядит как

$$f(\bar{x}) = \frac{1}{\sqrt{(2\pi)^k detK}} e^{\left(-\frac{1}{2}(\bar{x}-\bar{m})(\bar{x}-\bar{m})^T \quad K^{-1}\right)} \qquad [18]$$

Корреляционная матрица

Вероятно, что остались непонятыми некоторые моменты относительно того, что представляет собой корреляционная матрица и для чего она нужна. Первое, что надо сделать – это дать определение того, что является корреляционным моментом.

Корреляционный момент (ковариация, ковариационный момент) — в теории вероятностей и математической статистике мера линейной зависимости двух случайных величин. Она равна математическому ожиданию произведения разностей случайных величин и их математических ожиданий, иными словами:

$$cov(X, Y) = M[(X - MX)(Y - MY)] \qquad [19]$$

Если ковариация положительна, то с ростом значений одной случайной величины, значения второй имеют тенденцию возрастать, а если знак отрицательный — то убывать.

Однако только по абсолютному значению ковариации нельзя судить о том, насколько сильно величины взаимосвязаны, так как её масштаб зависит от их дисперсий. Масштаб можно отнормировать, поделив значение ковариации на произведение среднеквадратических отклонений (квадратных корней из дисперсий). При этом получается так называемый коэффициент корреляции Пирсона который всегда находится в интервале от −1 до 1:

$$r(X, Y) = \frac{cov(X, Y)}{\sigma_X \sigma_Y} \qquad [20]$$

Теперь переходим непосредственно к определению того, что является ковариационной матрицей:

Ковариацио́нная ма́трица (или ма́трица ковариа́ций) — это матрица, составленная из попарных ковариаций элементов одного или двух случайных векторов.

Ковариационная матрица случайного вектора — квадратная симметрическая неотрицательно определенная матрица, на диагонали которой располагаются дисперсии компонент вектора, а внедиагональные элементы — ковариации между компонентами.

Ковариационная матрица случайного вектора является многомерным аналогом дисперсии случайной величины для случайных векторов. Матрица ковариаций двух случайных векторов- многомерный аналог ковариации между двумя случайными величинами.

В случае нормально распределенного случайного вектора, ковариационная матрица вместе с математическим ожиданием этого вектора полностью определяют его распределение (по аналогии с тем, что математическое ожидание и дисперсия нормально распределенной случайной величины полностью определяют её распределение).

Пример корреляционной матрицы для двумерного вектора:

$$K = \begin{bmatrix} \sigma_x^2 & r\sigma_x\sigma_y \\ r\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix} \qquad [21]$$

Эта матрица характеризуется определителем

$$detK = \sigma_x^2\sigma_y^2 - r^2\sigma_x^2\sigma_y^2 = (1-r^2)\sigma_x^2\sigma_y^2 \qquad [22]$$

Можем посчитать обратную к К матрицу и сделать подстановку в выражение 17:

$$K^{-1} = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\begin{bmatrix} \sigma_y^2 & -r\sigma_x\sigma_y \\ -r\sigma_x\sigma_y & \sigma_x^2 \end{bmatrix} \qquad [23]$$

$$f(\bar{x}) = \frac{1}{2\pi\sqrt{detK}} e^{\left(-\frac{1}{2}(\bar{x}-\bar{m})(\bar{x}-\bar{m})^T \ K^{-1}\right)} =$$

$$\frac{1}{2\pi\sigma_x \ \sigma_y \ \sqrt{1-r^2}} e^{\left(\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - 2r\frac{(x-m_x) \ (y-m_y)}{\sigma_x \ \sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]\right)} = f(x,y) \qquad [24]$$

Чтобы убедиться в том, что в выражении 24 нет ошибок, проинтегрируем полученный результат, скажем, по y – должны получить плотность нормального распределения случайной величины x.

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_x \ \sigma_y \ \sqrt{1-r^2}} e^{\left(-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - 2r\frac{(x-m_x) \ (y-m_y)}{\sigma_x \ \sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]\right)} dy \qquad [25]$$

$$= \frac{1}{2\pi\sigma_x \ \sigma_y \ \sqrt{1-r^2}} \int_{-\infty}^{\infty} e^{\left(-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - 2r\frac{(x-m_x) \ (y-m_y)}{\sigma_x \ \sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right]\right)} dy = \left| \begin{matrix} u = \dfrac{x-m_x}{\sigma_x} \\ v = \dfrac{y-m_y}{\sigma_y} \\ dy = \sigma_y dv \end{matrix} \right| = I$$

$$I = \frac{1}{2\pi\sigma_x \ \sqrt{1-r^2}} \int_{-\infty}^{\infty} e^{\left(-\frac{u^2-2ruv+v^2}{2(1-r^2)}\right)} dv \qquad [26]$$

$$\int_{-\infty}^{\infty} e^{\left(-\frac{u^2-2ruv+v^2}{2(1-r^2)}\right)} dv = \int_{-\infty}^{\infty} e^{\left(-\frac{v^2-2ruv+(ru)^2-(r^2-1)u^2}{2(1-r^2)}\right)} dv = e^{-\frac{u^2}{2}} \int_{-\infty}^{\infty} e^{\left(-\frac{(v \ -ru)^2}{2(1-r^2)}\right)} dv = \left| \begin{matrix} z = \dfrac{(v-ru)}{\sqrt{1-r^2}} \\ dv = \sqrt{1-r^2}dz \end{matrix} \right|$$

$$= e^{-\frac{u^2}{2}}\sqrt{1-r^2}\int_{-\infty}^{\infty} e^{\left(-\frac{z^2}{2}\right)} dz = e^{-\frac{u^2}{2}}\sqrt{1-r^2}\sqrt{2\pi}$$

$$I = \frac{1}{2\pi\sigma_x \ \sqrt{1-r^2}} e^{-\frac{1}{2}\left(\frac{(x-m_x)}{\sigma_x}\right)^2} \sqrt{1-r^2}\sqrt{2\pi} = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}}$$

Получили одномерный вариант нормального распределения

Пример задачи

Рассмотрим простую задачу. Условие выглядит следующим образом:

Случайные величины X и Y независимы и распределены по нормальному закону. Известно, что $m_x = a, m_y = b$, $\sigma_x = \sigma_y = \sigma$. Найти радиус R круга с центром в точке (a, b), вероятность попадания в который случайной точки (X, Y) равна 0,997.

Решение:

В условии сказано, что случайные величины независимы, поэтому плотность двумерного нормального распределения определена как

$$f(x,y) = f_X(x)\, f_Y(y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{(x-\mu)^2}{2\sigma_x{}^2}+\frac{(y-\mu)^2}{2\sigma_y{}^2}\right)} = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{(x-\mu)^2+(y-\mu)^2}{2\sigma^2}\right)} \tag{27}$$

Вероятность попадания точки в круг можно вычислить через двойной интеграл по области, ограниченной этим кругом:

$$\begin{gathered}
P(D) = \iint\limits_D f(x,y)dxdy = \iint\limits_{(x-a)^2+(x-b)^2\le R^2} f(x,y)dxdy = \frac{1}{2\pi\sigma^2} * \\
* \iint\limits_{(x-a)^2+(x-b)^2\le R^2} e^{-\left(\frac{(x-a)^2+(y-b)^2}{2\sigma^2}\right)}dxdy = \left|\begin{matrix} u = \dfrac{x-a}{\sigma} \\ v = \dfrac{y-b}{\sigma}\end{matrix}\right| = \frac{1}{2\pi} \iint\limits_{u^2+v^2\le\left(\frac{R}{\sigma}\right)^2} e^{-\left(\frac{u^2+v^2}{2}\right)}dudv = \left|\begin{matrix} u = r cos\varphi \\ v = r sin\varphi\end{matrix}\right| \\
= \frac{1}{2\pi}\int_0^{\frac{R}{\sigma}}\int_0^{2\pi} r e^{-\frac{r^2}{2}} d\varphi\, dr = \int_0^{\frac{R}{\sigma}} e^{-\frac{r^2}{2}} d\left(\frac{r^2}{2}\right) = \left| t = \frac{r^2}{2}\right| = \int_0^{\frac{R^2}{2\sigma^2}} e^{-t} d(t) = 1 - e^{-\frac{R^2}{2\sigma^2}}
\end{gathered} \tag{28}$$

Остается учесть данное в условии ограничение величины вероятности:

$$\begin{gathered}
1 - e^{-\frac{R^2}{2\sigma^2}} = 0{,}997 \\
-\frac{R^2}{2\sigma^2} = \ln 0{,}003 \\
R = \sqrt{-2\sigma^2 \ln 0{,}003} \approx 3{,}41\sigma
\end{gathered} \tag{29}$$

Ответ : $3{,}41\sigma$

Условные распределения нормального закона

Рассмотрим еще один небольшой раздел, который не хотелось бы обойти стороной, говоря о двумерном нормальном распределении. Применяя теорему об умножении плотностей получаем следующую формулу для условного распределения нормального закона:

$$\begin{gathered}
f_{Y|x}(y) = \frac{f(x,y)}{f_X(x)} = \frac{\sigma_x\sqrt{2\pi}}{2\pi\sigma_x\ \sigma_y\ \sqrt{1-r^2}} e^{\left(-\frac{1}{2(1-r^2)}\left[\frac{(x-m_x)^2}{\sigma_x^2} - 2r\frac{(x-m_x)\ \ (y-m_y)}{\sigma_x\ \sigma_y} + \frac{(y-m_y)^2}{\sigma_y^2}\right] * \left(\frac{(x-m_x)^2}{2\sigma_x^2}\right)^{-1}\right)} \\
= \frac{1}{\sqrt{2\pi}\sigma_y\ \sqrt{1-r^2}} e^{\left(-\frac{1}{2}\left(\frac{y-\left[m_y + r\frac{\sigma_y}{\sigma_x}(x-m_x)\right]}{\sigma_x\sqrt{1-r^2}}\right)^2\right)} = \frac{1}{\sqrt{2\pi}\sigma_{y/x}} e^{\left(-\frac{1}{2}\left(\frac{y-m_{y/x}}{\sigma_{y/x}}\right)^2\right)}
\end{gathered} \tag{30}$$

$$\sigma_{y/x} = \sigma_y\ \sqrt{1-r^2} \tag{31}$$

$$m_{y/x} = m_y + r\frac{\sigma_y}{\sigma_x}(x - m_x) \tag{32}$$