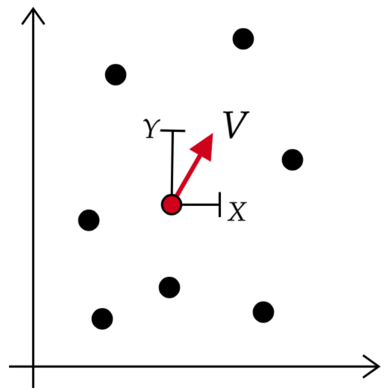
Конспектировала: Мария Щеголева.

1. Нормальное распределение

Путь через теорему Хершела-Максвелла: *Если распределение случайного вектора с независимыми компонентами инвариантно к повороту, то компоненты вектора распределены нормально* .



Аксиома 1. Закон распределения вектора скорости $V = \begin{pmatrix} X \\ Y \end{pmatrix}$ не зависит от направления, а только от длины: $f(x, y) = h(x^2 + y^2)$

Аксиома 2. Компоненты скорости X и Y независимы: $f(x, y) = g(x^2) \cdot g(y^2)$

Аксиома 3. Нормировка: $\mathrm{Var}(X) = 1$

Свойства:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \xrightarrow{\circlearrowleft 180^\circ} \begin{pmatrix} -X \\ -Y \end{pmatrix} \quad \Rightarrow \quad \mathrm{E}(X) = 0$$

$$\begin{pmatrix} X \\ Y \end{pmatrix} \xrightarrow{\circlearrowleft 90^\circ} \begin{pmatrix} Y \\ -X \end{pmatrix} \quad \Rightarrow \quad \mathrm{E}(X) = \mathrm{E}(Y) = 0; \mathrm{Var}(Y) = \mathrm{Var}(X) = 1$$

Определение. Вектор $V = \begin{pmatrix} X \\ Y \end{pmatrix}$ имеет *стандартное нормальное распределение*, если верны Акс.1-3.

Доказательство.

$$h(x^2 + y^2) = g(x^2) \cdot g(y^2)$$

$$\begin{cases} h'(x^2 + y^2) = g'(x^2) \cdot g(y^2) \\ h(x^2 + y^2) = g(x^2) \cdot g(y^2) \end{cases} \xrightarrow{x=0} \begin{cases} h'(y^2) = g'(0) \cdot g(y^2) \\ h(y^2) = g(0) \cdot g(y^2) \end{cases}$$

$$h'(t) = \frac{g'(0)}{g(0)} h(t) \quad \Rightarrow \quad h'(t) = k \cdot h(t) \quad \Rightarrow \quad h(t) = c \cdot \exp(kt)$$

$$f(x, y) = c \cdot e^{k(x^2+y^2)} = \sqrt{c} \cdot e^{x^2} \cdot \sqrt{c} \cdot e^{y^2} = g(x) \cdot g(y)$$

- $k-?$

$$\mathrm{E}(X^2) = \int\limits_{-\infty}^{+\infty} x^2\sqrt{c}\cdot e^{kx^2}dx = 1$$

$$\int\limits_{-\infty}^{+\infty} x\sqrt{c}\cdot x\cdot e^{kx^2}dx = x\cdot\sqrt{c}\cdot e^{kx^2}\cdot\frac{k}{2}\Big|_{-\infty}^{+\infty} - \int\limits_{-\infty}^{+\infty}\sqrt{c}\cdot e^{kx^2}\cdot\frac{1}{2k}dx = -\frac{1}{2k}\int\limits_{-\infty}^{+\infty}\sqrt{c}\cdot e^{kx^2} = -\frac{1}{2k}$$

$$-\frac{1}{2k} = 1 \quad\Rightarrow\quad k = -\frac{1}{2}$$

- $c-?$

$$f(x,y) = c\cdot e^{-\frac{1}{2}(x^2+y^2)}$$

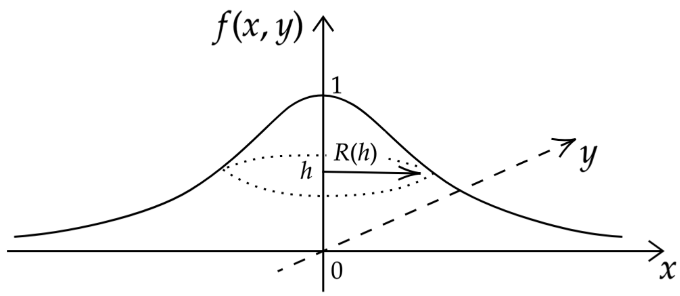
$$I = \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} c\cdot e^{-\frac{1}{2}(x^2+y^2)}dxdy = 1 \quad \text{(функция распределения)}$$

Способ 1: Перейти к полярным координатам

$$dx\wedge dy = d(r\cos\phi)\wedge d(r\sin\phi) = rd\phi\wedge d\phi$$

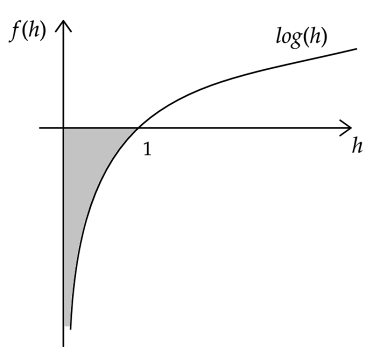
Способ 2:(почти устно) Найти площадь

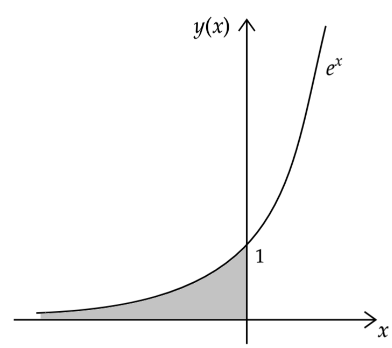
Принцип Кавальери:



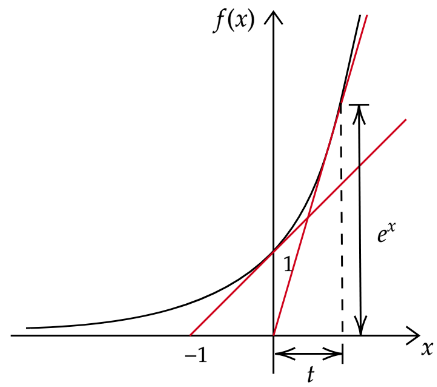
$$I = c\int\limits_0^1 \pi R^2(h)dh = \left[h = e^{-\frac{1}{2}R^2(h)} \;\Rightarrow\; R^2(h) = -2log(h)\right] = c\int\limits_0^1 -2\pi\log(h)dh$$

Замена переменной и пределов интегрирования: $I = -2\pi\cdot c\int\limits_{-\infty}^{0} e^x dx$





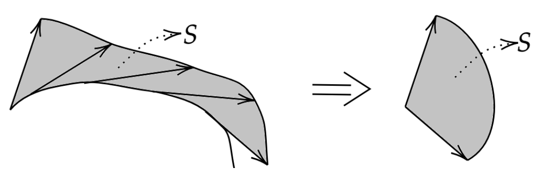
Геометрическое свойство экспоненты:



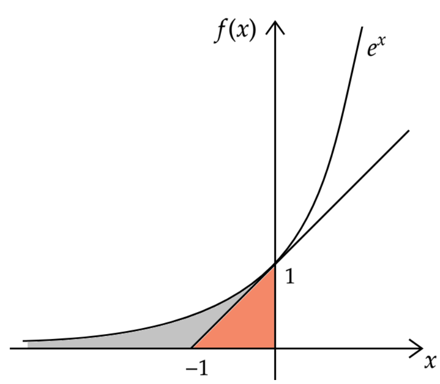
$$(e^x)' = e^x \;\Rightarrow\; \operatorname{tg}\alpha = e^x \;\Rightarrow\; \frac{e^x}{t} = e^x \;\Rightarrow\; t = 1$$

Принцип Мамикона:

Площадь S, закрашиваемая касательными при движении вдоль кривой, равна площади, закрашиваемой касательными, если только поворачивать их, но не перемещать вдоль кривой, независимо от формы исходной кривой.



Площадь фигуры, которую закрашивают касательные при движении по кривой e^x от $-\infty$ до 0, по принципу Мамикона равна $\frac{1}{2}$. Тогда $\int\limits_{-\infty}^{0} e^x dx$ равен сумме площади серой фигуры, закрашенной касательными, и площади треугольника ($S_\Delta = \frac{1}{2} \cdot 1 \cdot 1$):



$$I = -2\pi \cdot c \int\limits_{-\infty}^{0} e^x dx = -2\pi \cdot c \cdot \left(\frac{1}{2} + \frac{1}{2}\right) = 1 \;\Rightarrow\; c = -\frac{1}{2\pi}$$

$$f(x, y) = -\frac{1}{2\pi} \cdot e^{-\frac{1}{2}(x^2+y^2)} \;\Rightarrow\; X \sim \mathcal{N}(0,1), Y \sim \mathcal{N}(0,1)$$

Замечание: Без Аксиомы (3) X, Y имеют нормальное распределение $\sim \mathcal{N}(0, \sigma^2)$. □

2. χ^2-распределение

Определение. Если вектор $y = (y_1, \ldots, y_n)^T$ имеет стандартное нормальное распределение ($y_i \sim \mathcal{N}(0,1)$, y_i - независимы), V - k-мерное подпространство, $\hat{y}$ - проекция y на V и $q = \|\hat{y}\|^2$ - квадрат длины проекции, то q имеет χ^2-распределение с k степенями свободы.

Классическое определение. Случайная величина q имеет χ^2_k-распределение, если q представима в виде $q = z_1^2 + z_2^2 + \ldots + z_k^2$, где $z_i \sim \mathcal{N}(0,1)$ и независимы.

Эквивалентность: если $n = k$ и V - все пространство, тогда $y = \hat{y}$ и $q = \|\hat{y}\|^2 = y_1^2 + y_2^2 + \ldots + y_k^2$

Упражнение. $(y_1, y_2, y_3)^T$ - стандартное нормальное распределение; $V = \{y_1 + y_2 = y_3\}$ - двумерное подпространство.

а) $\hat{y}-?$; б) $q-?$

$$\hat{y} = Hy; \qquad H = (X^T X)^{-1} X^T y$$

Базисные векторы подпространства V: $\quad X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$

$$X^T X = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}; \quad (X^T X)^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \frac{1}{3} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

$$\hat{y} = \frac{1}{3} \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3}(2y_1 - y_2 + y_3) \\ \frac{1}{3}(-y_1 + 2y_2 + y_3) \\ \frac{1}{3}(y_1 + y_2 + 2y_3) \end{pmatrix}$$

$$q = \left(\frac{2y_1 - y_2 + y_3}{3}\right)^2 + \left(\frac{-y_1 + 2y_2 + y_3}{3}\right)^2 + \left(\frac{y_1 + y_2 + 2y_3}{3}\right)^2 \quad \Rightarrow \quad \text{по классич.опред. } q \sim \chi^2_2$$

Теорема. Если $y = X\beta + u; u_i \sim \mathcal{N}(0, \sigma^2)$ и независимы, то $\frac{RSS}{\sigma^2} \sim \chi^2_{n-k}$

Доказательство.

$\hat{y}$ - проекция y на пространство столбцов X; $\dim(col(X)) = k$
$\hat{u}$ - проекция u на $V = col^{\perp}(X)$ - ортогональное дополнение к $col(X)$; $\dim(V) = n - k$

$$u_i \sim \mathcal{N}(0, \sigma^2) \quad \Rightarrow \quad \frac{u_i}{\sigma} \sim \mathcal{N}(0,1)$$

$$(I - H)X\beta \text{ - проекция } X\beta \text{ на } col^{\perp}(X) = V; \quad X\beta \perp V \quad \Rightarrow \quad (I-H)X\beta = 0$$

$$\hat{u} = (I - H)y = (I - H)(X\beta + u) = (I - H)u$$

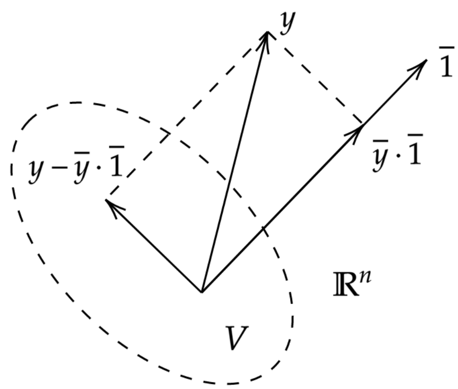
$$\frac{RSS}{\sigma^2} = \frac{\|(I-H)u\|^2}{\sigma^2} = \left\|(I-H)\frac{u}{\sigma}\right\|^2 \sim \chi^2_{n-k} \text{ по определению, т.к.}$$

$$\left\|(I-H)\frac{u}{\sigma}\right\|^2 \text{ - длина проекции} \left\{\frac{u_i}{\sigma} \sim \mathcal{N}(0,1)\right\}_n \text{ на подпространство } V;\ \dim(V) = n - k$$

$\square$

Упражнение. $y = u; u_i \sim \mathcal{N}(0, \sigma^2)$ и независимы. Оценивается модель $\hat{y} = X\hat{\beta}$ по МНК.

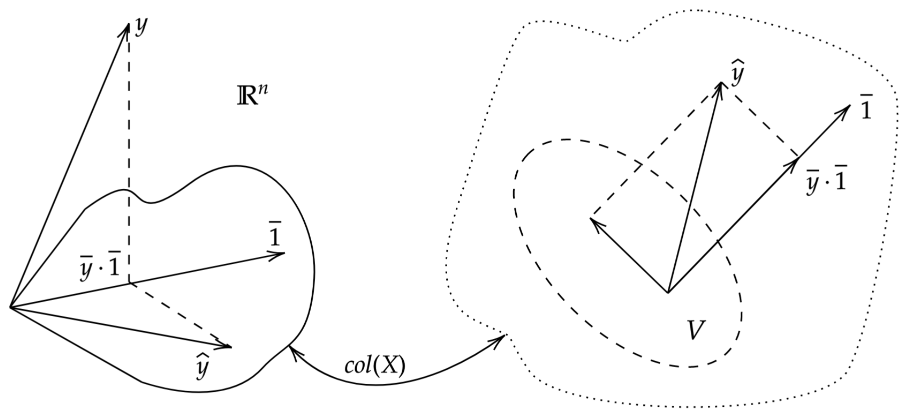
- $TSS = \sum(y_i - \bar{y})^2$



$TSS = \|y - \bar{y} \cdot \bar{1}\|^2$ - длина проекции y на подпространство $V = \text{Lin}^\perp\, \bar{1}-$ ортогональное дополнение к $\bar{1}$ из $col(X)$ в пространстве $\mathbb{R}^n \Rightarrow \dim(V) = n - 1$

$$\frac{TSS}{\sigma^2} \sim \chi^2_{n-1}$$

- $ESS = \sum(\hat{y}_i - \bar{y})^2$



$ESS = \|\hat{y} - \bar{y} \cdot \bar{1}\|^2$ - длина проекции y на подпространство $V = \text{Lin}^\perp\, \bar{1} \subset col(X)$ - ту часть $col(X)$, которая лежит в ортогональном дополнении к $\bar{1} \Rightarrow \dim(V) = k - 1$

$$\frac{ESS}{\sigma^2} \sim \chi^2_{k-1}$$

Удобное следствие. Аксиома (1) + Аксиома (2) = длины проекций на ортогональные подпространства независимы.

Доказательство.

$y = (y_1, \ldots, y_n)^T;\ y_i \sim \mathcal{N}(0, \sigma^2)$ и независимые.
W, V - подпространства в $\mathbb{R}^n$; $V \perp W$ (не обязательно покрывают все $\mathbb{R}^n$)

Повернем систему так, чтобы первые координаты отвечали за координаты в W:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{pmatrix} \xrightarrow{\circlearrowleft} \begin{pmatrix} \updownarrow \text{координаты в } W \\ \hline \vdots \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} \xrightarrow{\circlearrowleft} \begin{pmatrix} \updownarrow \text{координаты в } W \\ \hline \updownarrow \text{координаты в } V \\ \hline \vdots \end{pmatrix} : y^* = \begin{pmatrix} \begin{pmatrix} w_1 \\ \vdots \end{pmatrix} \\ \begin{pmatrix} v_1 \\ \vdots \end{pmatrix} \\ \begin{pmatrix} z_1 \\ \vdots \\ \vdots \end{pmatrix} \end{pmatrix}$$

y^* - координаты y в новом базисе

y_i независимые $\;\Rightarrow\;$ w и v независимые

□

3. t-распределение (Стьюдента)

Классическое определение. Случайная величина $T = \frac{Z}{\sqrt{\gamma_k / k}}$ имеет t-распределение с k степенями свободы, где $Z \sim \mathcal{N}(0,1)$, $\gamma_k \sim \chi^2_k$ и Z, γ_k независимы.

Геометрическое определение.

W - одномерное линейное подпространство в $\mathbb{R}^n$, задаваемое вектором единичной длины w.
$W = \mathrm{Lin}(w)$. $\quad W \perp V$. $\quad V = col(X) \;\Rightarrow\; \sqrt{dim(V)} = \sqrt{k}$.
$\hat{z}$ - проекция z на V; $\dim(V) = k \;\Rightarrow\; \|\hat{z}\|^2 \sim \chi^2_k \;\Rightarrow\; \gamma_k = \|\hat{z}\|^2$
Случайная величина $T = \operatorname{tg}\alpha \sqrt{dim(V)}$ имеет t-распределение с k степенями свободы, где α - угол между $\hat{y}$ (проекцией y на V) и $\tilde{y}$ (проекцией y на $W \oplus V$).