Лемма Фишера. Распределение Стьюдента.

Пусть $X_i \sim \mathcal{N}(a, \sigma^2)$. Если оба параметра a и σ^2 неизвестны, то приходит на ум идея подставить вместо неизвестных параметров ОМП и использовать в случае, когда требуется оценить математическое ожидание a функцию при неизвестном σ, функцию

$$g_1(\overline{X}, S, a) = \frac{\sqrt{n}(\overline{X} - a)}{S},$$

а в случае, когда требуется оценить дисперсию σ^2 при неизвестном математическом ожидании a, функцию

$$g_2(S, \sigma) = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

К сожалению, подстановка оценок меняет распределение функций g_1 и g_2. Однако их можно найти, для чего пользуются следующей *леммой Фишера*:

Лемма 1. *Если $X_i \sim \mathcal{N}(\mu, \sigma^2)$, то величины*

$$g_3(\overline{X}, \mu, \sigma) = \sqrt{n}\frac{\overline{X} - \mu}{\sigma}, \;\; g_2(S, \sigma) = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{\sigma^2}$$

являются независимыми, причем $g_3 \sim \mathcal{N}(0, 1)$, $g_2 \sim \chi^2_{n-1}$.

С помощью полученного факта можно построить доверительные интервалы для параметров среднего и дисперсии в нормальной модели.

Для построения интервала для математического ожидания, в силу того, что мы знаем распределение функции $g_3(s)$, которая отличается от функции $g_1(s)$, нам потребуется также одно известное распределение, которое определено ниже.

Определение. Величину, чье распределение то же, что и у величины $\frac{U}{\sqrt{W/n}}$, где $U \sim \mathcal{N}(0, 1)$, $W \sim \chi^2_n$, называют величиной с распределением Стьюдента с n степенями свободы. Это распределение обозначают t_n.

Замечание 1. Отметим, что распределение Стьюдента симметрично: $t_\alpha = -t_{1-\alpha}$.

Замечание 2. Плотность распределения Стьюдента равна следующему выражению:

$$\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n}\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{y^2}{n}\right)^{-(n+1)/2}.$$

При $n = 1$ это плотность распределения Коши.