

## Task 4.2

Для каждой пары моделей нужно проверить нулевую гипотезу о равенстве средних распределений против двусторонней альтернативы, поскольку мы не можем заранее утверждать, что одна модель работает не хуже другой. Проверим принадлежность данных нормальному распределению методом Шапиро-Уилка. Достижимые уровни значимости критерия: Делаем вывод, что данные не противоречат гипотезе нормальности.

	$a_1$	$a_2$	$a_3$	$a_4$
p	0.26	0.98	0.28	0.83

Поэтому для задачи сравнения будет целесообразно использовать t-критерий Стьюдента для случая зависимых выборок (так как модели обучаются на одних и тех же наборах данных).

Ошибка первого рода в данной задаче - отвержение гипотезы о равенстве средних при условии ее истинности, то есть совершая такую ошибку, мы скажем, что одна модель может быть лучше другой, хотя это на самом деле не так. Чтобы совершить как можно меньше ошибок первого рода, сделаем поправку на множественную проверку гипотез методом Холма.

<b>выборки:</b>	$\mathbf{X} = \{X_1^6, \dots, X_4^6\},$
<b>нулевая гипотезы <math>H_0^{ij}</math>:</b>	$\mu_i = \mu_j, \quad i, j \in 1, 4$
<b>альтернативы <math>H_1^{ij}</math>:</b>	$\mu_i < \neq > \mu_j$
<b>статистика:</b>	$T(X_i^6, X_j^6) = \frac{\bar{x}_i - \bar{x}_j}{S/\sqrt{6}}$ $S = \sqrt{\frac{1}{5} \sum_{i=1}^6 (D_i - \bar{D})^2}$ $D_k = X_{ik} - X_{jk}, \bar{D} = \frac{1}{6} \sum_k D_k$
<b>нулевое распределение:</b>	$St(5)$

Достижимый уровень значимости считается для каждой пары классификаторов как

$$p_{ij} = 2(1 - F_{St(5)}(T(X_i^6, X_j^6)))$$

Далее составляется вариационный ряд из этих значений:

$$p_{i_1 j_1} \leq p_{i_2 j_2} \leq \dots \leq p_{i_6 j_6}$$

Далее принимаем решения, пользуясь методом Холма. Модифицируем достижимые уровни значимости:

$$\tilde{p}_{(i_k j_k)} = \min(1, \max((m - i + 1)p_{(i_k j_k)}, \tilde{p}_{(i_{k-1} j_{k-1})}))$$

и будем отвергать гипотезу  $H_0^{ij}$  при условии  $\tilde{p}_{ij} < \alpha = 0.05$ . Для нашей задачи результаты этих процедур вынесены в таблицу:

	$i = 1, j = 2$	$i = 1, j = 3$	$i = 1, j = 4$	$i = 2, j = 3$	$i = 2, j = 4$	$i = 3, j = 4$
$p$	0.47	0.3	0.34	0.66	0.04	0.04
$\tilde{p}$	1.0	1.0	1.0	1.0	0.04	0.19

Таким образом, данные позволяют отклонить нулевую гипотезу только для пары классификаторов  $a2, a4$ . До внесения поправки Холма мы могли также отклонить нулевую гипотезу для классификаторов  $a3, a4$ , но, возможно, совершили бы ошибку первого рода.