Задание 4.2.
Курузов Илья, 676 группа

1. В таблице записаны четыре реализации для каждой случайной величины $A_1, A_2, A_3, A_4$. Положим, что исследователь, выбрал как лучшую модель $k$-ую. Проверим гипотезы $H_j : \mathbb{P}(A_k > A_j) = \frac{1}{2}$ при альтернативе $H'_j : \mathbb{P}(A_k > A_j) > \frac{1}{2}$ для $j \neq k$. Всего три гипотезы.

2. Воспользуемся для этой задачи критерием знаков. В таком случае статистика имеет вид:

$$T(\mathbf{x}_1, \mathbf{x}_2) = \sum_{k=1}^{n} \left| x_1^k > x_2^k \right| \tag{1}$$

В силу того, что мы проверяем множество гипотез воспользуемся методом Холма для модификации достигаемых уровней значимости.

3. Нулевое распределение $\mathrm{Bin}(n, \frac{1}{2})$

4. Достигаемый уровень значимости:

$$p(T) = 1 - F_{\mathrm{Bin}(n, \frac{1}{2})}(T)$$

Нулевая гипотеза отвергается, если $p(T) \leq \alpha$

5. Наилучшим по среднему (мат.ожиданию) является третий классификатор. Используя описанный метод, получаем следующую результаты (см. ).

| Гипотеза | $p_{\text{value}}$ | $\alpha_m$ | Решение |
|---|---|---|---|
| $A_3 > A_4$ | 0 | 0.02 | Отвергнуть |
| $A_3 > A_1$ | 0.02 | 0.025 | Отвергнуть |
| $A_3 > A_2$ | 0.34 | 0.05 | Принять |

Table 1: Результаты проверки гипотез

Согласно полученным данным, мы можем утверждать, что третий классификатор лучше первого и четвертого. Однако, мы не можем утвержать, что второй классификатор хуже первого.