

4.2

Усманова Анастасия, 674 группа

28 марта 2020 г.

1 Формальное описание задачи

Хотим проверить гипотезу равенства средних для классификаторов. Если мы не отклоним эту гипотезу, то исследователь не может утверждать, что какая-то модель лучше.

Выборка: $X^i = \{X_j^i\}_{j=1}^6, i \in [1,4]$

Нулевая гипотеза: $H_0 : medX_i = medX_j, i \neq j$

Альтернатива: $H_1 : medX_i \neq medX_j, i \neq j$

2 Предлагаемый метод решения

Воспользуемся критерием знаковых рангов. Выборки связанные.

Для попарного сравнения двух моделей воспользуемся статистикой:

$$T(X^i, X^j) = \sum_{k=1}^6 rank(|X_k^i - X_k^j|) * sign(X_k^i - X_k^j)$$

Нулевое распределение: табличное. Стремится к нормальному при увеличении объема выборки для каждого из классификаторов.

3 Без поправки на множественность гипотез

```
1 p_values = []
2 for i in range(1, 5):
3     for j in range(i+1, 5):
4         t, p_val = st.wilcoxon(data.iloc[:, i].values, data.iloc[:, j].values)
5         p_values.append(p_val)
6         print('Классификаторы: ', data.columns[i], data.columns[j], ' P-value: ', p_val)
```

```
Классификаторы: a1 a2 P-value: 0.463071015014588
Классификаторы: a1 a3 P-value: 0.2071604489114608
Классификаторы: a1 a4 P-value: 0.463071015014588
Классификаторы: a2 a3 P-value: 0.7531523647659145
Классификаторы: a2 a4 P-value: 0.07473549830588248
Классификаторы: a3 a4 P-value: 0.027707849358079864
```

Гипотеза отклоняется только для классификаторов a3 и a4, т.к. $p\text{-val}(a3, a4) < \alpha = 0.05$

4 С поправкой на множественность гипотез

```
1 from statsmodels.stats.multitest import multipletests
2 for k in ['bonferroni', 'holm', 'fdr_bh', 'fdr_by', 'holm-sidak']:
3
4     result = multipletests(p_values, method = k)
5     print(k)
6     print(result[0])
7     print(result[1])
8     print('-'*30)
```

```
bonferroni
[False False False False False False]
[1.          1.          1.          1.          0.44841299 0.1662471 ]
-----
holm
[False False False False False False]
[1.          0.8286418  1.          1.          0.37367749 0.1662471 ]
-----
fdr_bh
[False False False False False False]
[0.55568522 0.4143209  0.55568522 0.75315236 0.22420649 0.1662471 ]
-----
fdr_by
[False False False False False False]
[1.          1.          1.          1.          0.54930591 0.40730539]
-----
holm-sidak
[False False False False False False]
[0.84520727 0.60486889 0.84520727 0.84520727 0.32184417 0.15514792]
-----
```

Мы не можем отклонить ни одну из гипотез на нашем уровне значимости ($\alpha = 0.05$), исходя из результатов применения поправок.

5 Вывод

Исследователь не может утверждать что какая-то из моделей лучше другой.