

## УПРАВЛЕНИЕ ДИЗАЙНОМ ИССЛЕДОВАНИЯ В ЗАДАЧАХ ОБРАБОТКИ СИГНАЛОВ МЕДИКО-БИОЛОГИЧЕСКОЙ ПРИРОДЫ<sup>72</sup>

Туровский Я.А.<sup>\*,\*\*</sup>, Борзунов С.В.<sup>\*\*</sup>, Вахтин А.А.<sup>\*\*</sup>

<sup>\*</sup>Институт проблем управления им. В.А. Трапезникова РАН,  
Россия, г. Москва, ул. Профсоюзная, д.65

<sup>\*\*</sup>ФГБОУ ВО Воронежский государственный университет,  
Россия, г. Воронеж, Университетская пл., 1

Yaroslav\_turovsk@mail.ru

*Аннотация:* Рассматривается задача управления дизайном исследования на основе оценки плотности распределения отклонения нулевой гипотезы. Предложен алгоритм оценки значимости различий, который был апробирован в задаче сравнения данных с парных каналов электроэнцефалографа.

Ключевые слова: множественные сравнения, парные сравнения, оценка значимости, электроэнцефалограмма, дизайн исследования.

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов №№ 19-29-01156\_мк, 19-07-01037 А

### Введение

Одной из традиционных задач, связанных с обработкой информации, полученной в ходе медико-биологических исследований, является статистическая обработка результатов, включающая использование определённого, достаточно небольшого набора критериев, применяемых согласно разработанному исследователем дизайну эксперимента. Однако, в подавляющем большинстве современных экспериментальных планов происходит более чем одно сравнение выборок. В этом случае проявляет себя эффект множественных сравнений, заключающийся в росте риска ложного отклонения нулевой гипотезы. Рассмотрим ситуацию поиска нервных центров, связанных с тем или иным видом деятельности, подразумевая наличие межполушарной асимметрии [1]. Взяв стандартный 21 канальный электроэнцефалограф, мы получим 8 пар сравнений (по 8 электродов на каждом полушарии). Пусть различия существуют только для одной пары электродов, т.к. только на один из них проецируется система нейронов, отвечающая за исследуемый феномен. Очевидно, согласно формуле Бернулли вероятность того, что при альфа-уровне в 5% будет хотя бы одно ложноположительное решение, составляет  $1 - (1 - \alpha)^8 = 0,3366$  что, очевидно, достаточно велико, чтоб им пренебречь. Как известно, многократное применение критерия парных сравнений приводит к значительному возрастанию вероятности ошибочного установления различий в тех случаях, когда они статистически незначимы [2]. Для борьбы с этим эффектом был разработан целый ряд подходов. Тем не менее, значительная часть из них, базирующаяся на FWER FDP имеет ряд недостатков, связанных не только с возрастанием риска ошибки второго рода, но и с практическим применением полученных в ходе статистической обработки результатов. Действительно, для FDP, зная долю ложноположительных решений, исследователь, тем не менее не может, строго следуя данному подходу, в большинстве случаев высказать предположение о том, какие конкретно результаты сравнений следует отнести к более «высокорисковым» в плане ошибки. Аналогично, для FWER требуется, например в следующей процедуре ранжировать результаты по показателю  $p$ , при этом он и только он служит маркером того, что данный результат был статистически значим.

Обозначим через  $P_{(1)}, \dots, P_{(m)}$  уровни значимости, упорядоченные по неубыванию,  $H_{(1)}, \dots, H_{(m)}$  — соответствующие  $p_{(i)}$  гипотезы.

1) Если  $P_{(1)} \geq \alpha/m$ , то принять гипотезы  $H_{(1)}, \dots, H_{(m)}$  и остановиться. Иначе, если  $P_{(1)} < \alpha/m$ , то отвергнуть гипотезу  $H_{(1)}$  и продолжить проверку оставшихся гипотез на уровне значимости  $\alpha/(m-1)$ .

2) Если  $P_{(2)} \geq \alpha/(m-1)$ , то принять гипотезы  $H_{(2)}, \dots, H_{(m)}$  и остановиться. Иначе, если  $P_{(2)} < \alpha/(m-1)$ , то отвергнуть гипотезу  $H_{(2)}$  и продолжить проверку оставшихся гипотез на уровне значимости  $\alpha/(m-1)$ .

<sup>72</sup>Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов №№ 19-29-01156\_мк, 19-07-01037 А

Однако, в этих подходах не учитывается широко известный факт: в многофакторных статистических исследованиях одна и та же выборка содержит в себе результирующие воздействия разных факторов. В этом случае, графически подобный экспериментальный план можно представить как совокупность осей в  $n$ -мерном пространстве, каждая из которых является той или иной переменной. Тогда очевидно, что если выявленные статистически значимые различия являются ложноположительным результатом, то в указанном пространстве они должны быть распределены достаточно равномерно, в то время как истинно положительные значения, в ряде случаев могут формировать группы различий в которых была отклонена нулевая гипотеза  $H_0$ .

Таким образом, целью работы была разработка и апробация метода управления дизайном исследования, основанном на оценке плотности распределения отклонения  $H_0$  в пространстве сравнения переменных.

## 1 Методы множественных сравнений

Как известно, многократное применение критерия парных сравнений приводит к значительному возрастанию вероятности ошибочного установления различий в тех случаях, когда они статистически незначимы [2, 4]. Среди известных методов работы с множественными сравнениями часто применяются следующие: поправка Бонферрони, критерий Ньюмена-Кейлса, критерий Тьюки [3, 4].

Метод, основанный на поправке Бонферрони, требует отклонить те гипотезы, для которых уровень значимости не удовлетворяет неравенству:

$$p' < p/n, \quad (1)$$

где  $p$  – уровень значимости, соответствующий применяемому попарному критерию. Из этого неравенства непосредственно следует особенность метода при достаточно большом числе сравнений:  $p'$  очень мало, и даже масштабные различия признаются статистически незначимыми, другими словами, метод является излишне «строгим», и шансы отклонить ошибочные гипотезы падают. Кроме того, не учитываются данные о возможной группировке серий сравнений в кластеры (т. е. группы сравнений, близких в некотором смысле согласно медико-биологической основе эксперимента). Из общих соображений ясно, что группировка в кластеры является существенным свидетельством в пользу зависимости выборок.

Более сложные методы, использующие критерии Ньюмена-Кейлса, Тьюки и др., также выявляют значимые различия без использования изначальных данных о кластеризации сравнений.

Обобщим метод парных сравнений, дополнив его влиянием расстояния в абстрактном пространстве попарных сравнений. Последовательность из независимых  $n$  сравнений, каждое из которых приводит к положительному результату с вероятностью  $p$ , имеет вероятностное распределение вида

$$P_B(k) = C_n^k p^k (1-p)^{n-k}, \quad (2)$$

то есть биномиальное распределение с вероятностью «успеха», равной  $p$ . Предположим, что в серии из  $n$  испытаний ровно  $s$  из них привели к успешному исходу. Согласно свойствам биномиального распределения, среднее значение случайной величины  $s$  в случае попарной независимости исходов равно математическому ожиданию  $s_0 = p \cdot n$ .

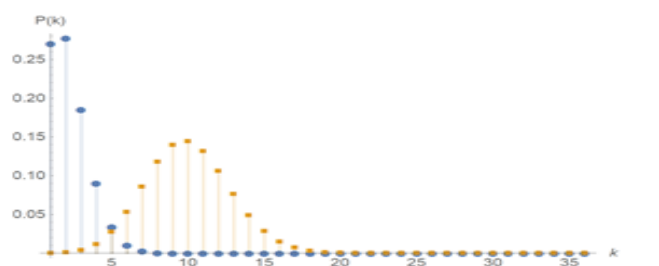


Рис. 1. Биномиальное распределение  $n=40$  испытаний с вероятностью успеха  $p_1=0,05$  (синие точки),  $p_2=0,25$  (оранжевые точки)

В терминах планирования эксперимента  $s_0$  определяет число ложноположительных сравнений. Приблизительно равенство значений  $s_0$  и  $s$  говорит о независимости результатов испытаний в ходе данного эксперимента. В свою очередь, неравенство  $s_0 > p \cdot n$  свидетельствует о наличии сравнений, превышающих математическое ожидание числа ложноположительных сравнений.

## 2 Алгоритм оценки значимости различий

В настоящем разделе предложим алгоритм, устанавливающий значимость множественных сравнений на основе выделения из множества всевозможных результатов эксперимента подмножеств, отвечающих общим деталям реализаций эксперимента в том или ином опыте. Такие подмножества  $U_i$  (где  $i=1,2,\dots,M$ ) универсального множества  $U$  будем называть «блоками» испытаний. Заметим, что множество  $U$  может быть рассмотрено как пространство исходов анализируемого эксперимента.

Первый шаг алгоритма заключается в оценке вероятности положительного исхода конкретного одиночного опыта  $p^*$  в предположении равномерного распределения положительных исходов в множестве согласно формуле  $p^* = s/n$ , где  $s$  — количество выявленных в эксперименте положительных исходов,  $n=|U|$  — мощность универсального множества. Далее определим порог  $\Pi$  количества ложноположительных решений, воспользовавшись неравенством

$$\sum_{i=0}^{\Pi} (p^*)^i (1-p^*)^{n-i} > 1-p_0, \quad (3)$$

где  $p_0$  — уровень значимости, соответствующий применяемому попарному критерию, например  $p_0=0,05$ . Значение числа положительных исходов, превышающее  $\Pi$ , трактуется как присутствие в результатах исследования ложноположительных результатов.

На основе сведений о конкретных деталях реализации эксперимента выполним всевозможные группировки элементарных опытов в подмножества блоков  $U_i$ ,  $i=1,2,\dots,M$ . Элементы объединяются в блоки с учетом экспертной оценки экспериментатором меры «близости» параметров той или иной серии опытов. В частности, в  $U_i$  объединяем элементы, близкие в смысле аналогичности набора команд, соответствующих опытам.

Рассмотрим некоторый блок  $U_k$ , для которого вычислим количество положительных результатов  $s_k$  и вероятность в предположении равномерного распределения  $p_k^* = s_k / |n_k|$ , в знаменателе введено обозначение  $n_k = |U_k|$  для количества элементов в блоке. В качестве критерия неслучайного распределения исходов в подмножестве  $U_k$  выберем выполнение условия

$$C_{n_k}^{s_k} (p_k^*)^{s_k} (1-p_k^*)^{n_k-s_k} < p_0, \quad (4)$$

которое имеет смысл сравнения вероятности реализации ровно  $s_k$  исходов в серии  $n_k$  независимых испытаний с уровнем значимости критерия.

В результате получаем множество «блоков»

$$W = \{U_{k_1}, U_{k_2}, \dots, U_{k_h} : \text{for } U_{k_j} \text{ (4) is valid}\} \quad (5)$$

вероятности положительных исходов в которых значимо выше по сравнению с равномерным распределением. Набор полученных множеств позволяет сделать обоснованные выводы по организации дальнейших серий экспериментов.

Блок-схема алгоритма, содержащая формальное представление перечисленных шагов, приведена на рис. 2.

*Пример.* Предположим, что в результате серии экспериментов из  $n=100$  индивидуальных опытов получены вектор исходов

$$a = (a_1, a_2, \dots, a_n), \quad a_1=a_2=\dots=a_{17}=1, a_{18}=a_{19}=\dots=a_{100}=0,$$

где используется общепринятое бинарное обозначения результата (1 — положительный исход, 0 — отрицательный). Выполним шаги алгоритма, предположив, что выделены, например, два блока:  $U_1=(a_2, a_3, \dots, a_9)$  и  $U_2=(a_{16}, a_{17}, \dots, a_{23})$ . Как легко вычислить,  $\Pi=23$ . Для блока  $U_1$  неравенство (4) выполняется, поскольку  $p_1^*=17/100$ ,  $n_1=8$  и  $s_1=7$ , следовательно, результаты, сгруппированные в первом блоке, демонстрируют значимые различия. Напротив, для блока  $U_2$  условие (4) не выполняется, поскольку  $n_2=8$ ,  $s_2=2$ , следовательно, эти результаты не значимы.

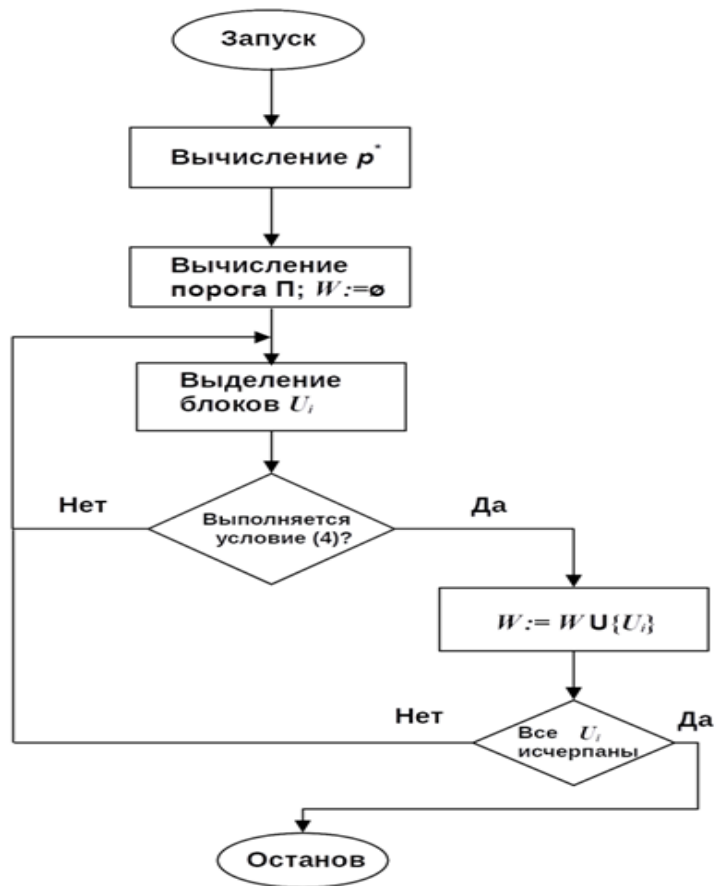


Рис.2. Блок схема алгоритма

Таблица 1. Результаты применения критерия Манна-Уитни

| Показатель 1       | Достижение цели | Команды | Пары ЭЭГ-каналов |       |       |
|--------------------|-----------------|---------|------------------|-------|-------|
|                    |                 |         | T3-T4            | F7-F8 | F3-F4 |
| Реальное движение  | попал           | вперед  | -                | -     | -     |
|                    |                 | право   | -                | -     | -     |
|                    |                 | лево    | -                | -     | -     |
|                    |                 | стоп    | -                | -     | -     |
|                    | промах          | вперед  | -                | -     | -     |
|                    |                 | право   | -                | -     | -     |
|                    |                 | лево    | по               | -     | -     |
|                    |                 | стоп    | -                | -     | -     |
| Мысленное движение | попал           | вперед  | -                | -     | -     |
|                    |                 | право   | +                | +     | -     |
|                    |                 | лево    | -                | -     | +     |
|                    |                 | стоп    | -                | -     | +     |
|                    | промах          | вперед  | -                | -     | -     |
|                    |                 | право   | -                | -     | -     |
|                    |                 | лево    | -                | -     | -     |
|                    |                 | стоп    | -                | -     | +     |
| Показатель 2       |                 |         | T3-T4            | F7-F8 | F3-F4 |
| Реальное движение  | попал           | вперед  | -                | -     | +     |
|                    |                 | право   | -                | -     | -     |
|                    |                 | лево    | -                | -     | +     |
|                    |                 | стоп    | -                | -     | +     |
|                    | промах          | вперед  | +                | -     | +     |
|                    |                 | право   | -                | -     | -     |

| Показатель 1       | Достижение цели | Команды | Пары ЭЭГ-каналов |       |       |
|--------------------|-----------------|---------|------------------|-------|-------|
|                    |                 |         | T3-T4            | F7-F8 | F3-F4 |
| Мысленное движение |                 | лево    | по               | -     | +     |
|                    |                 | стоп    | +                | -     | +     |
|                    | попал           | вперед  | +                | -     | -     |
|                    |                 | право   | +                | -     | -     |
|                    |                 | лево    | -                | -     | -     |
|                    |                 | стоп    | -                | -     | -     |
|                    | промах          | вперед  | -                | -     | -     |
|                    |                 | право   | -                | -     | -     |
|                    |                 | лево    | -                | -     | -     |
|                    |                 | стоп    | -                | -     | -     |

Примечание: «+» –  $p < 0.05$ , «-» –  $p > 0.05$ , «по» - нет данных.

В таблице 1 представлены результаты сравнения по двум показателям пары симметричных электродов в экспериментах по депрессии  $\mu$ -ритма в рамках интерфейса мозг-компьютер [5].

На основе данных таблицы можно сделать вывод, что, например, подмножество, где наиболее часто встречаются выявленные различия, наблюдается для пары электродов F3-F4 при выполнении реального движения (6 различий из 8 пар сравнений). Оценим, являются ли эти наблюдения результатом случайного совпадения. Общее количество сравнений составляет 96. Общее количество отклонений  $H_0$  составляет 14. Таким образом, используя формулу Бернулли, мы можем увидеть, что с вероятностью более 95% число ошибочных отклонений  $H_0$  будет составлять 9 и ниже. Приняв отношение вероятности  $p^* = 14/96$ , оценим, с какой вероятностью из 8 попыток 6 будут успешными в том смысле, что обеспечат ложноположительный результат. Полученное значение вероятности, равное  $2 \cdot 10^{-4}$ , говорит о том, что отклонения  $H_0$  сконцентрированы в этой области не случайно. Следовательно, мы можем рассматривать их как статистически-значимые результаты.

## Заключение

Разработан метод управления дизайном исследования основанном на оценке плотности распределения отклонения  $H_0$  в пространстве сравнения переменных. Метод апробирован на примере задачи планирования эксперимента на основе данных, полученных при сравнении потоков данных с парных каналов электроэнцефалограммы

## Литература

1. Гнездицкий В.В. Вызванные потенциалы мозга в клинической практике. М.: МЕДпресс-ин, 2003. –264 с.
2. Гланц С. Медико-биологическая статистика. – М.: Практика, 1998. – 459 с.
3. Рунцион Р. Справочник по непараметрической статистике: Современный подход. – М.: Финансы и статистика, 1982. – 198 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика. – 4-е изд. – М.: Высшая школа, 1972. – 368 с.
5. Туровский Я.А., Кургалин С.Д., Вахтин А.А., Борзунов С.В., Алексеев А.В. Информационные потоки при формировании дополнительного канала обратной связи для интерфейса мозг-компьютер // Технологии живых систем, 2016. – Т. 13, №7. – С. 34-40.