

## **Статистические подходы к обработке и интерпретации результатов экспериментов по выявлению способностей человека к экстрасенсорному восприятию**

*А.Г.Ли, А.Г.Чуновкина*

Показаны ограничения в использовании общепринятых подходов к статистической обработке результатов измерений экстрасенсорного восприятия (ЭСВ) по схеме Бернулли, не учитывающие специфическую природу явления: способность к ЭСВ непостоянна, плохо контролируется человеком и проявляется на короткий промежуток времени. Предложен новый подход к статистической обработке результатов измерений, основанный на анализе последовательности правильных ответов в серии, использовании последовательных статистических процедур (критериев), максимально учитывающий специфику ЭСВ и особенности проведения измерений. Для предложенного метода проведена оценка ошибки первого рода.

Традиционный путь получения статистически достоверных результатов - многократное повторение одних и тех же экспериментов с отдельными операторами, проведение многочасовых и многодневных испытаний. Однако такой подход не позволил до настоящего времени получить надежно воспроизводимые результаты: результаты получаются как небольшими по величине, что может быть обусловлено неадекватностью использованных методов исследования изучаемому явлению, так и плохо воспроизводимыми, что обусловлено несовершенством методики экспериментального исследования. Известные до сих пор способы контроля ЭСВ используют методики, предложенные 25-50 лет назад, основанные на перенесенных из области точных наук способах статистической обработки результатов измерений. Однако, как это часто бывает, методы исследования точных наук оказываются неэффективными при изучении биологических объектов и, особенно в области психофизиологии, психики человека, так как не учитывают специфические особенности процесса восприятия.

В общем виде известный способ измерения ЭСВ соответствует классической схеме испытаний Бернулли, а схема контроля ЭСВ сводится к многократному повторению одного и того же задания. Формализовано это может быть сформулировано следующим образом:  $N$  раз (10-20) ставится одно задание, на которое может быть дан один из  $m$  (например: двух-трех-четырёх-пяти) ответов. Если предположить, что испытуемый случайным образом пытается угадать правильный ответ, то эта схема соответствует испытаниям Бернулли с вероятностью успеха  $p=1/m$  и вероятностью неудачи  $q=(m-1)/m$ . Например, берут  $N$  конвертов, коробков, карт и т.п. и в них раскладывают по одному из  $m$  предметов (однотипных предметов различного цвета, формы, материала - железа, дерева, пластмассы и т.п.). Далее испытуемому предлагают определить, что и в каком конверте находится и фиксируют число правильных ответов. В Фонде парапсихологии им.Л.Л.Васильева используют одну из разновидностей этой методики: например, необходимо отобрать 10 конвертов или боксов, содержащих искомый предмет из 20, отделить 10 красных карточек от 10 синих (при исследовании так называемого "кожно-оптического зрения" (КОЗ)), определить в закрытых боксах различные предметы: металлы, пластмассы, дерево, части растений, мелких насекомых и животных (тест "биолокация"). Опыты по изучению проскопии проводят с использованием генератора псевдослучайных чисел. Исследовательские программы в ряде случаев были выполнены в игровой форме. Во всех случаях требуется предугадать один из 5 или 4 возможных вариантов, который выдает генератор случайных чисел (число, карту Зенера (тест Райна), направление движения самолета и т.д.) Всего за одну серию проводили 20 проб.

Число правильных ответов  $N_{\text{прав}}$  в серии из  $N$  вопросов при случайном угадывании распределено по биномиальному закону с параметром  $p$ :

$$P_k = P\{N_{\text{прав}}=k\} = C_n^k p^k q^{N-k} = C_n^k (1/m)^k ((m-1)/m)^{N-k}$$

Для любого значения  $k$  могут быть вычислены соответствующие вероятности по приведенной формуле и таким образом определены наиболее вероятные (правдоподобные) значения  $N_{\text{прав}}$ .

В основу любой модели должны быть положены некие предположения, которые в дальнейшем используются в рамках данной модели. Эти предположения формируются либо на основании предварительных исследований, либо принимаются на веру, исходя из здравого смысла. Выдвигается следующее предположение: ответы человека, обладающего способностью к ЭСВ от объектов, недоступных непосредственному чувственному восприятию соответствуют схеме испытаний Бернулли с вероятностью успеха  $p > 1/m$ .

Высказанное предположение представляется естественным. Рассмотрим, как с его использованием может быть сконструирован критерий, на основании которого делается вывод о наличии или отсутствии искомых способностей у испытуемого. Слабые стороны высказанного предположения будут обсуждены ниже.

Выдвигаются две конкурирующие гипотезы:

$H_0$ : испытуемый обладает способностями к ЭСВ, т.е.  $p > 1/m$

$H_1$ : искомые способности у испытуемого отсутствуют, т.е.  $p = 1/m$

Объективная информация содержится в результатах эксперимента, наиболее естественным представляется анализировать число правильных ответов  $N_{\text{прав}}$ . Для того, чтобы сформулировать критерий на основании данного показателя (статистики) необходимо задать такое число  $N_{\text{крит}}$ , что при  $N_{\text{прав}} \geq N_{\text{крит}}$  будет приниматься гипотеза  $H_0$ , а при  $N_{\text{прав}} < N_{\text{крит}}$  гипотеза  $H_1$ . Любой статистический критерий в полной мере характеризуется вероятностью ошибок I и II рода, другими словами уровнем значимости  $\alpha$  и мощностью  $1 - \beta$ .

Обычно уровень значимости задается исследователем и тогда различные критерии сопоставляются по своим мощностям. Уровень значимости  $\alpha$  - это вероятность ошибочно отвергнуть нулевую гипотезу, т.е. вероятность события  $\{N_{\text{прав}} < N_{\text{крит}}\}$ , при условии, что справедлива гипотеза  $H_0$ :

$$\alpha = P\{N_{\text{прав}} < N_{\text{крит}} \mid H_0\} = P\{N_{\text{прав}} < N_{\text{крит}} \mid p = 1/m\}$$

Задавшись уровнем значимости  $\alpha$  можно определить значение  $N_{\text{крит}}$ . Гипотеза  $H_0$  является сложной и для вычисления  $N_{\text{крит}}$  может быть заменена рядом простых:

$$H_{0i}: p = p_i > 1/m; \quad p_1 < p_2 < \dots < 1$$

Тогда для заданного уровня значимости  $\alpha$  при выбранном или предварительно оцененном значении параметра  $p_i$  может быть определено значение  $N_{\text{крит}}$  в результате решения уравнения:

$$\alpha = P\{N_{\text{прав}} < N_{\text{крит}}\} = \sum_{n=0}^{N_{\text{крит}}-1} \binom{N}{n} p_i^n (1-p_i)^{N-n}$$

Мощность критерия тем выше, чем меньше вероятность ошибочно принять гипотезу  $H_1$ , т.е. чем меньше значение вероятности ошибки второго рода, вычисляемое по формуле:

$$\beta = P\{N_{\text{прав}} \geq N_{\text{крит}} \mid p = 1/m\} = \sum_{n=N_{\text{крит}}}^N C_N^n (1/m)^n ((m-1)/m)^{N-n} = (1/m)^N \sum_{n=N_{\text{крит}}}^N C_N^n (m-1)^{N-n}$$

В таблицах 1-3 для уровня значимости  $\alpha = 0,1$  и  $N = 20$  приведены значения  $N_{\text{крит}}$  и вероятности ошибок второго рода  $\beta$  для значений  $m = 2, 3$  и  $5$ . Видно, что с увеличением  $m$  резко возрастает мощность критерия.

Таблица 1

Критические значения параметра  $K_{\text{крит}}$  и вероятности ошибки второго рода  $\beta$  для разных вероятностей правильного распознавания  $p$  ( $n=20, m=2, \alpha=0,1$ )

$p$	$N_{\text{крит}}$	$\beta$
0,72	12	0,13
0,76	13	0,09
0,80	14	0,06
0,86	15	0,006
0,90	16	0,001
0,92	17	0,0002
0,96	18	0,00002

Таблица 2

Критические значения параметра  $K_{\text{крит}}$  и вероятности ошибки второго рода  $\beta$  для разных вероятностей правильного распознавания  $p$  ( $n=20, m=3, \alpha=0,1$ )

$p$	$N_{\text{крит}}$	$\beta$
0,63	10	0,09
0,67	11	0,01
0,72	12	0,004
0,77	13	0,0008
0,83	14	0,0001

Таблица 3

Критические значения параметра  $K_{\text{крит}}$  и вероятности ошибки второго рода  $\beta$  для разных вероятностей правильного распознавания  $p$  ( $n=20, m=5, \alpha=0,1$ )

$p$	$N_{\text{крит}}$	$\beta$
0,40	6	0,2
0,45	7	0,08
0,50	8	0,03
0,60	9	0,003
0,62	10	<0,003
0,65	11	<0,003
0,70	12	$\approx 0$

Остановимся подробнее на оценивании параметров  $p_i$ , используемых при определении значения  $N_{\text{крит}}$ . Естественно трактовать параметры  $p_i$  как параметры, количественно характеризующие степень проявления способностей к ЭСВ у испытуемого. На основании значений этих параметров, оцененных по результатам эксперимента можно обоснованно сопоставлять степень проявления способностей к ЭСВ различных людей по используемому тесту. Традиционной статистической оценкой вероятности успеха по выборке является:

$$\hat{p} = N_{\text{прав}} / N$$

Таким образом, после оценивания параметра  $p_i$  по результатам эксперимента, можно воспользоваться таблицами 1-3 для задания значения  $N_{\text{крит}}$  и принятия решения.

Вернемся к обсуждению недостатков выдвинутого предположения.

**Математическое описание эксперимента с использованием общепринятого подхода, который предполагает использование такой абстракции как испытания Бернулли неадекватно изучаемому явлению и вызывает серьезное возражение, так как априори предполагается, что способность оператора характеризуется некоторой величиной - вероятностью правильного ответа  $p$ , которая остается неизменной в ходе проведения тестирования. Последнее условие обеспечивает статистическую воспроизводимость, необходимую для обоснованного применения статистических методов, является очень жестким и закономерно вызывает возражения.**

Об этом говорят, в частности, результаты полученные нами при проведении массового обследования населения в Фонде парапсихологии. В таблице 4 приведены сводные данные по использованным методикам. По каждому виду тестов обследовано от 800 до 1500 человек. Погрешность определения 3 - 5% для разных тестов [1-3].

Таблица 4.

Средние значения количества правильных ответов по тестам на ЭСВ.

Серия	I	II	III
1 серия	5.36±0.05	5.56±0.06	4.22±0.03
2 серия	5.21±0.07	5.29±0.08	3.95±0.05
3 серия	5.10±0.1	5.14±0.12	-
математическое ожидание	5.00	5.00	4.00

В таблице в столбцах обозначены следующие опыты:

I – N=20, m=2 (экстрасенсорное восприятие в опытах по КОЗ и биолокации);

II – N=20, m=4 (игровой вариант исследовательской программы - предсказание направления движения самолета);

III – N=20, m=5 (Тест Райна - предсказание карты Зенера).

Каждый раз испытуемому предъявляли новый тест. При сопоставлении данных, сгруппированных из результатов, полученных разными операторами в 1 серии их опытов, с результатами, полученными во 2 и последующих сериях, следует отметить увеличение числа правильных ответов в 1 серии опытов по всем видам решаемых заданий (Таблица 4). В последующих сериях распределение числа правильных ответов приближается к случайному. Обращает на себя внимание асимметрия в распределении результатов 1 серии и смещение среднего значения распределения в сторону увеличения числа правильных ответов. Лучшие результаты получаются при использовании игровых программ, как у детей, так и у взрослых. Внутри каждой серии правильные ответы, как правило, следуют группами, время набора которых не превышает 30-45 секунд. Следует отметить, что даже у наиболее результативных операторов уже после второй серии распределение результатов приближается к случайному. При абсолютной случайности распределения ответов во 2 и последующих сериях результаты должны были бы колебаться около кривой, соответствующей случайному распределению со средним значением, равным математическому ожиданию. Однако происходит одностороннее приближение результатов последующих серий к случайному распределению.

Таким образом, изучаемые явления наиболее отчетливо проявляются в первых сериях и становятся случайными в последующих. При классическом подходе (статистическая обработка результатов одного оператора в ряде последующих серий) эффекты "исчезают" - становятся статистически недостоверными (значимо неразличимыми). Увеличение продолжительности опытов и количества серий не приводит к повышению достоверности результатов, а даже наоборот, т.к. в течение всей продолжительности опыта все феномены,

по-видимому, проявляются лишь на короткий промежуток времени, не более 30 - 45 секунд. Очевидно, для статистически достоверного выявления эффектов, необходимо анализировать первые серии опытов, полученные при решении одной и той же задачи большим количеством разных операторов. При таком подходе происходит как бы "деперсонализация" ЭСВ: удается доказать существование ЭСВ при исследовании больших групп людей, хотя статистически достоверное доказательство наличия ЭСВ у конкретного человека остается сложной проблемой.

Подобные результаты получены при проведении ЭЭГ исследований [2-4], в которых установлено, что длительность особых состояний сознания, во время которых реализуется ЭСВ, не превышает 5-30 секунд. В процессе выполнения профессиональной деятельности у экстрасенсов возможна кратковременная (продолжительностью порядка нескольких секунд) инверсия профиля функциональной асимметрии мозга с повышением активности правого полушария при любом исходном профиле функциональной асимметрии, что необходимо для проявления ЭСВ [2-4].

Таким образом, исследования показывают, что на результатах тестирования сказывается множество факторов, таких как соответствующий настрой, умение сконцентрироваться на выполнении задания, самочувствие, усталость, потеря интереса к выполнению задания и многое другое. В реальных условиях нельзя говорить о статистической воспроизводимости вообще, а только в течение некоторого интервала времени, который субъективен для каждого оператора и может быть очень небольшим. **Следовательно, основное возражение в интерпретации результатов тестирования с помощью испытаний Бернулли вызывает предположение о неизменном значении параметра  $p$  и вытекающий отсюда необоснованный выбор числа наблюдений (заданий) в эксперименте.**

Например, достаточно представить эксперимент из 20 опытов при двух вариантах ответов, когда испытуемый дал правильные ответы на первые десять опытов, а общее число правильных ответов оказалось меньше  $N_{\text{крит}}$ . В соответствии с сконструированным критерием в этом случае гипотеза  $H_0$  должна быть отвергнута. Однако серию из десяти правильных ответов трудно объяснить случайным угадыванием (вероятность такого события практически равна нулю).

**Возможная альтернатива - использование последовательных статистических процедур (критериев).**

Описанные примеры говорят о том, что **число правильных ответов не является достаточной статистикой, т.е. не содержит всю полезную информацию, содержащуюся в результатах эксперимента. Интерес представляет также максимальная длина серии из правильных ответов  $\tau(N)$ , а также общее число серий из правильных или неправильных ответов  $\delta(N)$ .** Эти статистики обычно используют для проверки случайного характера выборки и независимости наблюдений в выборке.

В нашем случае превышение этими статистиками критических значений говорит об отличии параметра  $p$  от  $1/2$  при рассмотрении результатов теста с двумя возможными ответами.

Критические значения статистик для уровня значимости  $\alpha=0,05$  соответственно равны:

$$\tau(N)_{\text{крит}}=3,3\log_{10}(N+1); \quad \delta(N)_{\text{крит}}=[1/2(N+1-1,96(N-1)^{1/2})]$$

Таким образом, если нарушается хотя бы одно из неравенств:

$$\tau(N) < \tau(N)_{\text{крит}}; \quad \delta(N) > \delta(N)_{\text{крит}}$$

то это говорит о неслучайном характере ответов.

Для  $N=20$  соответствующие значения равны:

$$\tau(20)_{\text{крит}}=5; \quad \delta(20)_{\text{крит}}=6$$

В данном случае не представляется возможным определить мощность критерия поскольку незатабулирован двумерный закон распределения статистики  $\{\tau(N), \delta(N)\}$  при  $p \neq 1/2$ . При значениях  $p \neq 1/2$  (т.е.  $m > 2$ ) можно предположить простой критерий, основанный на длине серии из правильных ответов в независимости от общего числа вопросов. Вероятность серии  $L$  правильных ответов при случайном угадывании равна:  $(1/m)^L$  и резко убывает с ростом  $L$  особенно при  $m > 2$ . Значения вероятностей для различных  $m$  и  $L$  приведены в таблице 5. Из таблицы видно, что для уровня значимости  $\alpha=0,01$  (выделено в таблице) длина серии правильных ответов, большая критического значения:  $q_{\text{крит}}(2)=7$  для  $m=2$ ,  $q_{\text{крит}}(3)=5$  для  $m=3$ ,  $q_{\text{крит}}(5)=3$  для  $m=5$  - говорит о неслучайном угадывании ответов.

Таблица 5

Вероятность серии правильных ответов  $\alpha$  в зависимости от длины серии  $L$  и числа возможных вариантов ответов  $m$ .

$m \backslash L$	1	2	3	4	5	6	7
2	0,5	0,25	0,12	0,06	0,03	0,02	<b>0,01</b>
3	0,33	0,1	0,06	0,02	<b>0,01</b>	$\approx 0$	$\approx 0$
5	0,2	0,04	<b>0,01</b>	$\approx 0$	$\approx 0$	$\approx 0$	$\approx 0$

При исследовании способностей к восприятию стимулов малой интенсивности применяют обычно не один, а несколько тестов. Обработка результатов такого объединенного эксперимента составляет отдельную задачу. Строго говоря, для ее решения необходимо построение многомерного закона распределения статистик, используемых в различных тестах, что само по себе представляет самостоятельную сложную задачу. Если ограничиться эвристическими соображениями при построении объединенного критерия, то процедура принятия решения может быть сформулирована подобно одному из нижеприводимых способов:

- $H_0$  принимается, если она принимается по всем частным критериям;
- $H_0$  принимается, если она принимается в более, чем половине случаев и т.п.

При вычислении уровня значимости такого критерия необходимо учитывать зависимость между результатами различных тестов при условии справедливости гипотезы  $H_0$  и естественно значение уровня значимости несколько возрастает. Однако мощность критерия значительно увеличится, поскольку значение  $\beta$  вычисляется в предположении справедливости гипотезы  $H_1$  (в этом случае результаты различных тестов независимы).

Кратко остановимся на вопросах организации тестирования. Принципиально возможны две схемы проведения эксперимента или их комбинации:

- без обучения. Например, испытуемому предлагается разложить конверты с различными цветами и он узнает об результатах только после выполнения всего задания,
- с обучением. Когда испытуемый знает о результатах предыдущего задания и может корректировать свои ответы в дальнейшем.

При обработке результатов во втором случае следует в большей мере ориентироваться на длину серий из правильных ответов. При постановке эксперимента важно исключить влияние субъективных факторов, в частности, такого рода: например, в тесте с конвертами

общее число красных и синих конвертов не должно быть известно испытуемому и должно быть случайным в каждом эксперименте.

В качестве примера опишем подробнее организацию эксперимента и интерпретацию его результатов для теста с определением цветowych вкладышей в конвертах (КОЗ) (аналогичные действия справедливы для теста Райна или тестов на биолокацию).

В ходе проведения теста оператору предлагают распознать  $m$  цветов. Необходимо приготовить достаточно большое количество конвертов с каждым цветом (приблизительно равное максимальному числу наблюдений). Исследователь предлагает оператору конверт случайным образом, не зная, какой цвет в конверте. После ответа оператора исследователь и оператор вместе проверяют правильность ответа и так далее.

Оператору предлагают ответить на вопрос и сразу же контролируют его ответ (правильный или нет). Если подряд следует достаточно длинная серия правильных ответов, то тестирование заканчивается с положительным результатом. Данный эксперимент может быть организован двояким образом: оператору сообщают о результате его ответа (тестирование с обратной связью) и оператору не сообщают о результате его отдельного ответа. В первом случае в ходе эксперимента происходит своеобразное обучение и в среднем следует ожидать более высоких результатов. Вообще говоря, следует стремиться, чтобы этап настройки на задачу (этап обучения) заканчивался до начала тестирования. Однако, при последовательной схеме эксперимента несколько первых неудачных ответов не испортят результата теста (как было бы в классическом случае ранее рассматриваемого критерия), так как ориентируются не на правильное число ответов из общего числа вопросов, а на длину серии правильных ответов.

В описанной схеме проведения эксперимента остается открытым вопрос об его окончании, если он не был прерван после получения серии правильных ответов достаточной длины (положительный результат тестирования). Окончание эксперимента при отрицательном результате может произойти либо по предложению самого оператора (основание - усталость) либо по предложению исследователя (основание - объективные показания переутомления оператора либо ограниченное время проведения тестирования). В данном случае мы не стремимся провести строго обоснованной верхней границы времени проведения эксперимента, что повлекло бы необходимость дополнительных субъективных предположений. На наш взгляд этот вопрос следует решать эмпирическим путем.

Достаточную длину серии правильных ответов  $L$  определяют по таблице 5 для выбранного значения  $\alpha$  или по таблице 6 для двух значений  $\alpha = 0.05$  и  $\alpha = 0.01$ . Длину серии правильных ответов выбирают по минимальному  $L$ , удовлетворяющему выбранным требованиям. Обычно  $\alpha$  выбирают порядка 0,05 или 0,01.

Таблица 6.

Длина серии правильных ответов  $L$  для разных значений ошибок первого рода  $\alpha$  и числа определяемых предметов  $m$ .

$m$	$\alpha = 0,05$	$\alpha = 0,01$
2	5	7
3	3	5
5	2	3

**В заключении** хотелось бы еще раз подчеркнуть, что эффективность математических (в том числе статистических) методов обработки и интерпретации результатов экспериментов определяется правомерностью (корректностью) их применения. Предложенный подход к организации проведения тестирования и построению статистических критериев не отвергает полностью предыдущий. По-видимому, целесообразно совместно использовать оба подхода на начальных этапах и проведении предварительных "поисковых" исследований. Тест,

основанный на использовании биномиального распределения и интерпретации результатов в соответствии со схемой Бернулли является более жестким. Он требует не просто проявления экстрасенсорных способностей, но и демонстрации (воспроизводимости) в течение некоторого времени (числа заданий в тесте). Предпочтение какому-либо из них следует отдавать на основании конкретизации задачи исследования и накопления результатов их экспериментального сравнения с целью построения более адекватной модели исследуемого явления. В Фонде парапсихологии при проведении экспериментальных работ используют оба варианта обработки результатов: при предварительном просмотре на первых этапах и при оценочных экспериментах - отдают предпочтение второму подходу, при строгом обосновании выводов - первому.

## Литература

1. Ли А.Г. К вопросу о методике изучения некоторых необычных феноменов психики человека. - Парапсихология в СССР, 1991, N2, с.34-38.
2. Ли А.Г. Ясновидение. Формирование особых состояний сознания для раскрытия экстрасенсорных способностей человека. - М.: Изд-во Фонда парапсихологии им.Л.Л.Васильева, 1993. - 168с.
3. Ли А.Г. Разработка способов управления состояниями мозга для повышения эффективности восприятия стимулов малой интенсивности и создания систем управления в биологических и медицинских системах. - М.: Изд-во Фонда парапсихологии им.Л.Л.Васильева, 1993. - 112с.
4. Ли А.Г. Биофизические и нейрофизиологические основы восприятия подпороговых сигналов - М.: Изд-во Фонда парапсихологии им.Л.Л.Васильева, 1992. - 161с.

Ли Андрей Гендинович - докт.мед.наук, проф., канд.техн.наук  
Президент Фонда парапсихологии им.Л.Л.Васильева  
129281, Москва, Олонекский проезд 12 – 63; тел. +7 916 5253052  
E-mail: 777000777@mail.ru

Чуновкина Анна Гурьевна - канд.техн.наук, ст.науч.сотр.,  
Зав.лабораторией Центра Единства измерений (бывший Институт метрологии  
им.Д.И.Менделеева), г. С.-Петербург