

ИСТОРИЯ И ПОЛИТОЛОГИЯ

Гаспарян Михаил Владимирович

канд. полит. наук, доцент

ФГАОУ ВО «КФУ им. В.И. Вернадского»

г. Симферополь, Республика Крым

НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ СТАТИСТИКИ: ПРИМЕНЕНИЕ В ПОЛИТИЧЕСКИХ ИССЛЕДОВАНИЯХ

Аннотация: в статье рассматриваются особенности применения непараметрических тестов в политических исследованиях. Представлены наиболее востребованные в практики непараметрические критерии. Иллюстрируется применение методов непараметрической статистики в рамках политических исследований. В частности, вычисление коэффициента W Кендалла (коэффициент конкордации), позволяющего измерить степень согласованности оценок в группе, при проведении определенного вида политических исследований, а именно: фокус-групп, экспертных опросов и т.п.

Ключевые слова: политические исследования, непараметрические тесты, коэффициент конкордации, независимые выборки, зависимые выборки.

Использование математических методов в политических исследованиях, к сожалению, не всегда сопровождается четкими представлениями о возможностях и границах их применения. Сложность, с которой мы зачастую сталкиваемся при применении статистических тестов, при анализе данных политических исследований обусловлены, с одной стороны, доминированием переменных, измеренных на уровне шкал низкого порядка, с другой стороны, наблюдаемые распределения могут значительно отличаться от нормального распределения. В такой ситуации исследователь использует тот или иной метод анализа данных не потому, что он наиболее целесообразен, а лишь в силу его доступности.

В результате этого значительная часть прикладных исследований социально-политических процессов довольствуется упрощенным, одномерным ана-

лизом данных, который существенно обедняет информацию, снижает диагностическую и прогностическую ценность получаемых выводов [1].

Использование более сложных видов анализа (многомерный, факторный, дисперсионный, регрессионный, и др.) сопряжено с рядом трудностей. Между тем, существуют статистические методы, использование которых позволяет проводить вполне корректную математическую обработку данных, не заставляя исследователя осваивать сложные разделы высшей математики. Мы имеем в виду методы непараметрической статистики.

Непараметрическая статистика – непараметрические методы статистики – часть математической статистики, комплекс методов обработки статистических данных, не требующих, чтобы распределение вероятностей было описано каким-либо законом распределения. Она опирается на более широкие и менее ограничительные свойства распределений вероятностей: статистическую независимость распределений (ошибок наблюдений), непрерывность этих распределений; часто – на ту или другую симметрию распределений и т.п. [2].

Название «непараметрические методы» подчеркивает их отличие от классических – параметрических – методов, в которых предполагается, что генеральное распределение известно с точностью до конечного числа параметров, и которые позволяют по результатам наблюдений оценивать неизвестные значения этих параметров и проверять гипотезы относительно их значений.

Непараметрические критерии – это такие приемы обработки экспериментальных данных, которые не рассматривают анализируемое статистическое распределение как функцию, их применение не предполагает предварительного вычисления параметров распределения. Эти критерии обрабатывают не само измеренное значение, а его ранг – положение внутри выборки.

На сегодня разработано значительное число непараметрических критериев, по существу, для каждого параметрического критерия имеется, по крайней мере, один непараметрический аналог. Эти критерии можно отнести к одной из следующих групп:

- критерии различия между независимыми выборками;

- критерии различия между зависимыми выборками;
- критерии связи между переменными.

Различия между независимыми выборками. Обычно, когда имеются две выборки (*например, мужчины и женщины*), которые мы хотим сравнить относительно среднего значения некоторой изучаемой переменной, то мы имеем возможность применить один из следующих критериев: критерий Уалда-Вольфовица, *U*-тест по методу Манна-Уитни, критерий Колмогорова-Смирнова.

Для сравнения более чем двух независимых выборок может быть использован *H*-тест по Краскелу-Уоллису, являющийся модификацией *U*-теста по методу Манна-Уитни (на случай для более двух независимых выборок) или Медианный тест.

Различия между зависимыми выборками. Если стоит задача сравнить две переменные, относящиеся к одной и той же выборке (*например, изменения в восприятии целевыми аудиториями образа политика или политической партии в начале и в конце PR или рекламной кампании*), то обычно используется: критерий Уилкоксона парных сравнений. Если рассматриваемые переменные по природе своей категориальны или являются категоризованными (т.е. представлены в виде частот, попавших в определенные категории), то подходящим будет критерий χ^2 Мак-Немара.

Если рассматривается более двух переменных, относящихся к одной и той же выборке, то обычно используется метод рангового дисперсионного анализа Фридмана или *Q* критерий Кохрена (последний применяется, например, если переменная измерена в номинальной шкале). *Q* критерий Кохрена используется также для оценки изменений частот (долей).

В случае, когда необходимо оценить степень согласованности между несколькими связанными выборками (*например, оценка участников фокус группового исследования некоторого количества предвыборных лозунгов*) используется коэффициент согласованности Кендалла.

Связь между переменными. Для того, чтобы оценить связь между двумя переменными, обычно вычисляют коэффициент корреляции. Непараметрическими

аналогами стандартного коэффициента корреляции Пирсона являются статистики Спирмена, τ Кендалла и коэффициент Гамма. Если две рассматриваемые переменные по природе своей категориальны, подходящими непараметрическими критериями для тестирования зависимости будут: χ^2 , ϕ -коэффициент, точный критерий Фишера.

Проиллюстрируем применение методов непараметрической статистики на примере вычисления коэффициента W Кендалла.

При проведении определенного вида политических исследований (фокус-групп, экспертных опросов и т.п.) мы весьма ограничены в обобщающих высказываниях, лишены возможности относительных оценок. Преодолеть эти сложности позволяет применение непараметрических тестов. В частности, коэффициент W Кендалла (коэффициент конкордации) [3] позволяющий измерить степень согласованности оценок в группе.

Рассмотрим результаты экспертного опроса, в рамках данного исследования 5-ти экспертам необходимо было проранжировать по важности 4 разных фактора, расставляя ранги от 1 до 4.

Формула для расчета Коэффициента W Кендалла:

$$W = \frac{12 \sum_{i=1}^n D_i^2}{m^2(n^3 - n)},$$

где m – число ранговых последовательностей (число экспертов);

n – число оцениваемых объектов (факторов);

$D_i = d_i - \bar{d}$ – отклонение суммы рангов i -го объекта $d_i = \sum_{j=1}^m R_{ij}$ от средней суммы рангов всех объектов $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$. Средняя сумма рангов всех объектов может быть вычислена по формуле $\bar{d} = \frac{1}{2} m(n + 1)$, которая используется для контроля.

Значения коэффициентов конкордации заключены на отрезке $[0; 1]$. Увеличение коэффициента от 0 к 1 означает проявление большей согласованности суждений. Коэффициент конкордации равен единице при полном совпадении всех

ранговых последовательностей. Если мнения экспертов (ранговые последовательности) полностью противоположны, коэффициент конкордации равен нулю.

В таблице 1 представлены результаты оценок экспертов.

Таблица 1

	Эксперт 1	Эксперт 2	Эксперт 3	Эксперт 4	Эксперт 5
Фактор 1	1	3	4	2	2
Фактор 2	3	2	3	3	4
Фактор 3	2	1	1	4	1
Фактор 4	4	4	2	1	3

Таким образом, в нашем примере $m = 5$, $n = 4$, остается найти значение D .

Для дальнейших вычислений нам нужно добавить в исходную таблицу 1 следующие значения, см. таблицу 2.

Таблица 2

	Эксперт 1	Эксперт 2	Эксперт 3	Эксперт 4	Эксперт 5	d_i	D_i	D_i^2
Фактор 1	1	3	4	2	2	12	-0,5	0,25
Фактор 2	3	2	3	3	4	15	2,5	6,25
Фактор 3	2	1	1	4	1	9	-3,5	12,25
Фактор 4	4	4	2	1	3	14	1,5	2,25
						50		21

Средняя сумма рангов всех объектов равна $\bar{d} = \frac{50}{4} = 12,5$.

В качестве контроля используем выражение $\bar{d} = \frac{1}{2} m(n + 1) = \frac{25}{2} = 12,5$.

Подставляем значения из таблицы 2 в формулу коэффициента конкордации Кендалла:

$$W = \frac{(12 \times 21)}{(25 \times (64 - 4))} = 0,168$$

Как видно из результатов вычисления коэффициента конкордации Кендалла, его значение равняется 0,168, а это говорит нам о слабой согласованности суждений.

Слабая согласованность обычно является следствием следующих причин:

– в рассматриваемой группе экспертов действительно отсутствует общность мнений;

– внутри группы существуют коалиции с высокой согласованностью мнений, однако, обобщенные мнения коалиций противоположны.

В заключение следует отметить, что развитие компьютерной техники и появление мощных программных пакетов для обработки статистических данных, позволяет, при всей относительной сложности современных методов статистического анализа, сосредоточить внимание исследователя не на технической стороне методов, а на особенностях их применения в каждом конкретном случае.

Список литературы

1. Джужа Н.Ф. Применение методов непараметрической статистики в психолого-педагогических исследованиях / Н.Ф. Джужа // Вопросы психологии. – 1987. – №4. – С. 145–151.

2. Тюрин Ю.Н., Шмерлинг Д.С. Непараметрические методы статистики / Ю.Н. Тюрин, Д.С. Шмерлинг // Социология: методология, методы, математические модели. – 2004. – №18. – С.154–166.

3. Ромашкина Г.Ф., Татарова Г.Г. Коэффициент конкордации в анализе социологических данных / Г.Ф. Ромашкина, Г.Г. Татарова // Социология: методология, методы, математические модели. – 2005. – №20. – С. 131–158.