Секция 2. ГЕОМЕТРИЯ И АНАЛИЗ

УДК 519.23

Минимальный эффект цикла в диапазоне возможных значений коэффициента ранговой корреляции

С.В. Дронов, Е.В. Семенов
АлтГУ, г. Барнаул

Довольно часто в приложениях возникает необходимость изучения взаимосвязи дискретных или даже нечисловых показателей (см. [1]). Такие задачи возникают в медицине [2], при сравнении экспертных оценок [3], в общественных науках [4] и др. При этом практические ученые-статистики, как отмечено, например, в [5], продолжают пользоваться коэффициентом корреляции Пирсона, что, конечно же, следует считать в гораздо меньшей степени оправданным, чем при применении его в классических предположениях. С этой точки зрения актуальной представляется задача точного построения распределений коэффициента корреляции между дискретными, особенно ранговыми показателями, в случае небольшого числа наблюдений. Именно разработке подходов к определению точного распределения коэффициента корреляции между рангами и посвящена настоящая работа.

Возьмем n связанных наблюдений над двумя показателями

$$X : x_1;...;x_n, \quad Y : \ y_1;...;y_n, \tag{1}$$

где каждая пара $(x_i, y_i), i = 1,...,n$ – результат одновременного наблюдения этих показателей. Пусть каждый из рядов X и Y представляет собой перестановку чисел $1, 2, ..., n$.

При изучении значений коэффициента корреляции между двумя перестановками чисел 1, 2, …, n без ограничения общности можно считать, что первая из них (X) является тождественной. Пронумеруем возможные значения коэффициента корреляции $R(X,Y)$ по их убыванию. Множество всех этих значений назовем n-диапазоном. Ясно, что одно и то же значение из n-диапазона может получаться при разных перестановках Y. То количество раз, которое k-е значение встретится в n-диапазоне при последовательном выборе в качестве второй перестановки всех n! перестановок чисел 1, 2, …, n, будем называть его повторностью и обозначим $s_k(n)$.

Следующие два утверждения проверяются непосредственно.

Лемма 1. *Равноудаленные от концов n-диапазона значения R(X,Y) имеют равные повторности.*

Теорема 1. *n-диапазон рангового коэффициента корреляции содержится в множестве*

$$\mathbf{R}(n) = \left\{ R_k = 1 - kh, \quad k = 0,1,\dots,\frac{n(n^2-1)}{6} \right\}, \quad h = \frac{12}{n(n^2-1)}. \quad (2)$$

Эффектом перестановки Y назовем $Eff(Y) = \sum_{j=1}^{n} j(y_j - j).$ Тогда, оказывается,

$$R(X,Y) = 1 - Eff(Y) \cdot h, \quad (3)$$

т.е. эффект Y определяет положение *R(X,Y)* в его n-диапазоне (2): $R(X,Y) = R_{Eff(Y)}$.

Согласно известной теореме алгебры, произвольная перестановка разлагается в произведение независимых циклов. Более того, практически очевидно, что эффекты независимых (т.е. тех, в которых не участвуют одинаковые элементы) перестановок при их перемножении складываются. Поэтому займемся изучением эффектов циклов.

Пусть a – наименьшее из различных натуральных чисел $a_1,\dots,a_s$. Заметим, что эффект цикла, составленного из рассматриваемых чисел, равен

$$Eff((a_1,\dots,a_s)) = \sum_{j=1}^{s} a_j^2 - a_1 a_s - \sum_{j=1}^{s-1} a_j a_{j+1}. \quad (4)$$

Введем в рассмотрение набор $x_j,\ j = 1,\dots,s$ с помощью соотношений $a_j = a + x_j, \quad j = 1,\dots,s.$ Преобразуем (4) к следующему виду.

$$Eff(a_1,\dots,a_s)) = \sum_{j=1}^{s} x_j^2 - x_1 x_s - \sum_{j=2}^{s-1} x_j x_{j+1}. \quad (5)$$

Эта формула допускает циклические перестановки $x_j,\ j = 1,\dots,s$, а также смену порядка следования ее элементов на противоположный, когда $x_j \leftrightarrow x_{s+1-j}, \quad j = 1,\dots,s$ без изменения эффекта.

Лемма 2. *Уменьшим на* 1 *все* $x_i, i = k,\dots,t\ (k > 1, t < s)$. *Если до уменьшения было справедливо* $x_{k-1} < x_k, \quad x_t > x_{t+1}$, *то эффект цикла в результате такого преобразования уменьшится.*

Доказательство. Пусть E – эффект первоначального цикла, E^- – эффект после преобразования. По формуле (5) имеем

$$E^{-}=\sum_{i=k}^{t}(x_i-1)^2-x_{k-1}(x_k-1)-\sum_{j=k}^{t-1}(x_j-1)(x_{j+1}-1)-(x_t-1)x_{t+1}+Q,$$

где Q уже не зависит ни от одного из уменьшенных x_i. Раскрывая здесь скобки, приходим к соотношению

$$E^{-}=E+(x_{k-1}-x_k)+(x_{t+1}-x_t)+1.$$

Поскольку все числа x_i целые и различные, то добавка к E в последней формуле отрицательна. Лемма доказана.

Упорядочим элементы цикла по возрастанию. Через $a_{(i)}$ будем обозначать i-й по величине его элемент. Число $d=a_{(s)}-a_{(1)}$ назовем размахом цикла.

Лемма 3. *Среди всех циклов фиксированной длины наименьший эффект достигается тогда и только тогда, когда цикл имеет один из видов*

$$(a_{(1)},a_{(3)},\ldots,a_{(4)},a_{(2)})\ \textit{или}\ (a_{(2)},a_{(4)},\ldots,a_{(3)},a_{(1)}). \tag{6}$$

При этом уменьшение размаха цикла влечет уменьшение наименьшего его возможного эффекта.

Доказательство леммы. Из (4) вытекает, что при любой перестановке между собой элементов цикла на величину эффекта может влиять только величина выражения

$$Q=a_1a_s+\sum_{j=1}^{s-1}a_ja_{j+1},$$

причем, чем эта величина больше, тем меньше эффект. Ясно таким образом, что для достижения минимального эффекта необходимо и достаточно, чтобы наибольшие элементы цикла имели бы наибольших по величине соседей в этом цикле. Следовательно, соседями $a_{(s)}$ в цикле (с разных сторон) обязаны быть $a_{(s-1)},a_{(s-2)}$. При этом у большего из этих двух чисел ($a_{(s-1)}$) вторым из соседей должно оказаться наибольшее из оставшихся – $a_{(s-3)}$. Продолжая такие рассуждения далее, с той стороны от наибольшего элемента, где находилось число $a_{(s-1)}$, мы дойдем либо до $a_{(1)}$ (при четном s), либо до $a_{(2)}$ (при нечетном). С другой же стороны последним обязано оказаться второе из этих двух чисел.

Таким образом, в том цикле, который обладает наименьшим эффектом, все большие по величине элементы образуют группу его центральных элементов и при фиксации некоторых границ k и t, окружающих среднюю часть такого цикла, участок между этими границами будет содержать элементы, большие, чем вне их. Взяв первоначально произвольный цикл с заданной длиной и размахом, сначала перейдем от него к виду (6), тем самым, уменьшив его эффект. После этого начнем работать уже с циклом вида (6).

Попробуем уменьшить размах цикла, уменьшая на 1 $a_{(s)}$ (и одновременно наибольший из x_j). Это может оказаться невозможным сделать, только если $a_{(s-1)}$ ровно на 1 меньше наибольшего элемента. Если это так, перейдем к уменьшению $a_{(s-1)}$ (предыдущего по величине элемента) и т.д. Предположим, что наибольший из тех элементов цикла, который возможно уменьшить, находится в цикле вида (6) на позиции k слева от наибольшего элемента $a_{(s)}$ или на позиции t справа. Уменьшим на 1 этот допускающий уменьшение элемент и все элементы цикла, большие, чем он. Это возможно, поскольку, уменьшая меньший элемент, мы «освобождаем место» для уменьшения большего. Исходя из вида рассматриваемого цикла, все уменьшенные элементы будут заполнять целиком участок цикла между некоторыми позициями k и t, а, следовательно, после такого преобразования эффект цикла уменьшится согласно лемме 2.

Если же возможностей для уменьшения ни одного из элементов цикла не будет, то это означает, что все элементы цикла – последовательные натуральные числа, и его размах не может быть уменьшен без уменьшения его длины. Лемма доказана.

Из леммы 3 несложно выводится основной результат работы.

Теорема 2. *Для произвольного цикла длины s*

$$Eff((a_1,...,a_s)) \geq 2s-3.$$

Наименьший эффект достигается тогда и только тогда, когда цикл имеет один из видов (6), *а указанная нижняя граница достигается тогда и только тогда, когда* $x_{(j)} = j-1, \quad j=1,...,s$, *т.е. когда такой цикл составлен из последовательных натуральных чисел.*

Библиографический список

1. Boddy R., Smith G. Statistical methods in practice: for scientists and technologists. – Chichester, U.K.: Wiley, 2009. – pp. 95–96.

2. Шойхет Я.Н., Лукеренко Е.В., Ефремушкина А.А, Лебедев А.Б., Дронов С.В. Клиническая диагностика немассивной тромбоэмболии легочной артерии в условиях ограниченных инструментальных диагностических возможностей. // Клиническая медицина. – №9, 2009. – С. 38–43.

3. Гутовская О.М., Дронов С.В. Формирование экспертных групп методами кластерного анализа. // В сборнике: Ломоносовские чтения на Алтае: фундаментальные проблемы науки и образования. Сборник научных статей международной конференции. Ответственный редактор Е. Д. Родионов. 2017. – С. 574–578.

4. Bryukhanova E.A., Dronov S.V., Chekryzhova O.I. Spatial Approach to the Analysis of the Employment Data in Siberia Based on the 1897 Census (the Experience of the Multivariate Statistical Analysis of the Districts Data). // Journal of Siberian Federal University. Humanities & Social Sciences. – 2016. – № 7. – P. 1651–1660.

5. Dodge Y. The Concise Encyclopedia of Statistics. N.Y.: Springer-Verlag, 2010. – p. 502.

УДК 514.765

Исследование метрик Эйнштейна трехмерных групп Ли с векторным кручением с помощью универсальных математических пакетов

П.Н. Клепиков, О.П. Хромова

АлтГУ, г. Барнаул

Пусть *(M,g)* – (псевдо)риманово многообразие, на котором определена *метрическая связность* ∇ *с векторным кручением*, т.е. связность вида

$$\nabla_X Y = \nabla_X^g Y + g(X,Y)V - g(V,Y)X, \qquad (1)$$

где ∇^g – связность Леви-Чивита, V – некоторое фиксированное векторное поле, X,Y – произвольные векторные поля. Данная связность является одной из основных связностей, описанных Э.Картаном в [1], и также называется полусимметрической связностью. В случае двумерных поверхностей любая метрическая связность является связностью с векторным кручением. В произвольной размерности интересен тот факт, что (см. [2])

Теорема 1. *Риманово многообразие допускает метрическую связность с векторным кручением, тензор кривизны которой равен нулю, тогда и только тогда, когда оно является конформно плоским.*

Данное утверждение обобщается и на случай псевдоримановой метрики, т.к. доказательство не использует положительную определенность метрического тензора.

Определим тензор кривизны R и тензор Риччи r связности ∇ соответственно равенствами