А.А. Венедиктов, доктор экономических наук, профессор
К.Г. Серебряков, кандидат военных наук

Характеризует ли коэффициент ранговой конкордации степень согласованности экспертных оценок?[1]

Нет.

При решении задач в области военного планирования и прогнозирования довольно широко используется метод экспертных оценок, когда то или иное управленческое решение принимается не на основе строгого научного доказательства его рациональности, а посредством обобщения субъективных оценок проблемы специалистами. Очевидно, что задачи, требующие количественного ответа, затруднительно решать экспертными методами. Без использования точных методов человек более или менее успешно может производить лишь качественное сравнение двух объектов (по схеме «лучше-хуже»). Если имеется более двух сравниваемых факторов (объектов), то может возникнуть потребность в упорядочения (ранжировании) их множества по некому принципу. Эта задача также доступна для решения экспертным путем, поскольку может быть сведена к попарному сравнению элементов.

Предполагается, что имеется m экспертов и n оцениваемых объектов. Задача экспертов – ранжировать объекты по тому или иному признаку (совокупности признаков), т. е. присвоить каждому из факторов порядковый номер от 1 до n, причем любой номер должен встречаться ровно один раз. В этом случае упорядоченные по неубыванию ранги выглядят так: 1; 2; 3; ...; n-1; n. Иногда методика оценки разрешает экспертам объединять объекты, которые, по их мнению, затруднительно дифференцировать, поскольку они имеют близкие параметры. В этом случае эксперт выставляет таким объектам одну и ту же оценку, равную среднему арифметическому рангов объединяемых факторов. Тогда последовательность упорядоченных по неубыванию рангов может выглядеть, например, так: 1; 2; 3,5; 3,5; 5; 7; 7; 7; 9; ...; n-1; n.

После сбора мнений экспертов, как правило, возникает необходимость оценить степень их согласованности. Для решения этой задачи нередко применяется так называемый коэффициент ранговой конкордации, иногда именуемый также коэффициентом конкордации (коэффициентом согласия) Кендалла (Кендэла[2]), следующего вида [1, 2]:

$$W=\frac{12}{m^2\left(n^3-n\right)}\sum_{j=1}^{n}\left(\sum_{i=1}^{m}C_{ij}-m\frac{n+1}{2}\right)^2, \tag{1}$$

где W – коэффициент ранговой конкордации, принимающий значения из диапазона $[0;1]$;

m – количество экспертов;

n – количество ранжируемых факторов;

C_{ij} – ранг j-го фактора по мнению i-го эксперта ($C_{ij}\in[1,n]$).

---

1 Статья подготовлена в рамках гранта РФФИ № 17-06-0052217.

2 Поисковые системы Интернет не позволили выяснить, кто же такой (или что такое) Кендалл (Кендэл). Удалось найти Н.Дж. Кендалл (англ. Nicole Jenner Kendall) — американская модель, участница телевизионного реалити-шоу «Семейство Кардашьян»; Э.Р. Кандел (англ. Eric Richard Kandel) — американский психиатр; а также города и провинции с таким названием в нескольких странах. Представляется, что все эти значения слова не имеют отношения к данному случаю.

При этом обычно предполагается, что значения W из диапазона $[0;0{,}3]$ свидетельствуют о низкой согласованности мнений экспертов, а из диапазона $[0{,}7;1]$ – о высокой [2].

Преобразуем формулу (1):

$$W=\frac{12}{n^3-n}\sum_{j=1}^{n}\frac{1}{m^2}\left(\sum_{i=1}^{m}C_{ij}-m\frac{n+1}{2}\right)^2,$$

$$W=\frac{12}{n^3-n}\sum_{j=1}^{n}\left(\frac{\sum_{i=1}^{m}C_{ij}}{m}-\frac{n+1}{2}\right)^2. \qquad (2)$$

Выражение $\frac{\sum_{i=1}^{m}C_{ij}}{m}$ представляет собой среднее арифметическое оценок, высказанных экспертами относительно ранга j-го фактора, обозначим его $\bar{C}_j$. Тогда формула (2) примет вид:

$$W=\frac{12}{n^3-n}\sum_{j=1}^{n}\left(\bar{C}_j-\frac{n+1}{2}\right)^2. \qquad (3)$$

Первый множитель в формуле (3) $\left(\frac{12}{n^3-n}\right)$ является константой для рассматриваемого набора объектов, поэтому как-то характеризовать степень согласованности мнений экспертов он не может. Второй множитель представляет собой сумму квадратов разницы между усредненной по всем экспертам ранговой оценкой j-го фактора и константой $\frac{n+1}{2}$. (Попутно отметим, что данная константа является средним арифметическим значений допустимых рангов анализируемых факторов: как для варианта с возможностью объединения объектов, так и без такового.) Таким образом, второй множитель также не зависит от степени согласованности мнений экспертов. Очевидно, что произведение величин, не зависящих от значения некого показателя, также от него не зависит.

Предварительный вывод: коэффициент ранговой конкордации никоим образом не характеризует степень согласованности (или несогласованности) мнений экспертов по конкретному вопросу. Его значение зависит лишь от среднего арифметического оценок экспертов каждого из анализируемых факторов.

Отметим, что формулу (1) иногда пытаются использовать для определения степени согласованности мнений экспертов в тех случаях, когда перед ними ставится не задача ранжирования (упорядочения) анализируемых объектов, а задаются иные вопросы, предполагающие высказывание мнения специалиста (например, поручается оценить вероятности развития тех или иных возможных сценариев; в этом случае C_{ij} – вероятность реализации j-го сценария по мнению i-го эксперта).

Однако подобное применение формулы (1) является некорректным, поскольку употребляемые в ней константы $\frac{12}{m^2(n^3-n)}$ и $m\frac{n+1}{2}$ обусловлены упомянутым в начале статьи ограничением на варианты ответов экспертов (присваивание каждому из анализируемых объектов ранга от 1 до m с возможностью объединения близких факторов либо без такового), т. е.

$$\forall i:\sum_{j=1}^{n}C_{ij}=\sum_{j=1}^{n}j=\frac{n(n+1)}{2}. \qquad (4)$$

Относительно второй константы $\frac{n+1}{2}$ мы уже отмечали выше, что она является средним арифметическим значений рангов анализируемых факторов, а выражение $\sum_{j=1}^{n}\left(\bar{C}_j-\frac{n+1}{2}\right)^2$, соответственно, имеет смысл суммы квадратов отклонений среднего арифметического рангов каждого из анализируемых параметров от среднего арифметического значения всех оценок (рангов), выставленных экспертами.

Константа $\frac{12}{m^2(n^3-n)}$ является нормирующим множителем, призванным привести значение коэффициента W к диапазону $[0;1]$. Ниже это будет показано.

$W=0$, когда каждое из неотрицательных слагаемых $\left(\bar{C}_j-\frac{n+1}{2}\right)^2$ равно нулю, т. е. при $\bar{C}_j=\frac{n+1}{2}$. Это достигается, в частности, в том случае, когда всем объектам эксперты присвоили одинаковое количество каждого рангового значения из допустимого диапазона. Пример такого распределения экспертных оценок приведен в таблице 1.

Таблица 1 – Пример распределения экспертных оценок, при котором коэффициент ранговой конкордации $W=0$

Номер вопроса	**Номер эксперта**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	1	2	3	1	2	3	1	2	3
2	2	3	1	2	3	1	2	3	1
3	3	1	2	3	1	2	3	1	2

Максимальная согласованность мнений экспертов, очевидно, достигается в том случае, когда все эксперты высказались единодушно по каждому оцениваемому объекту (таблица 2).

Таблица 2 – Пример распределения экспертных оценок, при котором коэффициент ранговой конкордации $W=1$

Номер анализируемого фактора	**Номер эксперта**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3

Найдем значение выражения $\sum_{j=1}^{n}\left(\bar{C}_j-\frac{n+1}{2}\right)^2$ для этого предельного случая. Очевидно, что оно не зависит от того, какому именно из анализируемых факторов присвоен тот или иной ранг (важно только то, что все эксперты были единодушны в его определении). Поэтому для простоты будем считать, что экспертные оценки распределились именно так, как это указано в таблице 2, т. е. порядковый номер анализируемого фактора совпадает с оценкой (рангом), выставленной ему всеми экспертами. Тогда данная формула примет вид:

$$\sum_{j=1}^{n}\left(j-\frac{n+1}{2}\right)^2.$$

Преобразуем данное выражение:

$$\sum_{j=1}^{n}\left(j-\frac{n+1}{2}\right)^2=\sum_{j=1}^{n}\left(j^2-j(n+1)+\frac{(n+1)^2}{4}\right)=\sum_{j=1}^{n}j^2-(n+1)\sum_{j=1}^{n}j+n\frac{(n+1)^2}{4}. \qquad (5)$$

Применим известные формулы: $\sum_{j=1}^{n}j^2=\frac{n(n+1)(2n+1)}{6}$ и $\sum_{j=1}^{n}j=\frac{n(n+1)}{2}$. Тогда выражение (5) примет вид:

$$\frac{n(n+1)(2n+1)}{6}-\frac{n(n+1)^2}{2}+n\frac{(n+1)^2}{4}=n(n+1)\left(\frac{2n+1}{6}-\frac{n+1}{2}+\frac{n+1}{4}\right)=$$
$$=\frac{n(n+1)}{12}(4n+2-6n-6+3n+3)=\frac{n(n+1)(n-1)}{12}=\frac{n^3-n}{12}. \qquad (6)$$

Как видим, итоговое значение формулы (6) $\left(\frac{n^3-n}{12}\right)$ является обратным первому нормирующему множителю в формуле (3) $\left(\frac{12}{n^3-n}\right)$, т. е. в случае полного совпадения мнений всех экспертов коэффициент W равен единице.

Таким образом, мы показали, что константы в формуле (1) подобраны применительно к условиям ограничения (4) и ее использование в отсутствие соответствующего требования к оценкам C_{ij}, как минимум, требует иной интерпретации результата.

Вернемся к сделанному нами ранее предварительному выводу о том, что коэффициент ранговой конкордации не характеризует степень согласованности мнений экспертов по конкретному вопросу. Данное заключение может показаться неточным (односторонним) в связи с тем, что хотя, как и было показано выше, значение каждого отдельно взятого слагаемого в скобках в формуле (3) определяется лишь средним арифметическим оценок экспертов $\bar{C}_j$ по каждому из анализируемых объектов, однако сами значения $\bar{C}_j$ не являются независимыми, поскольку действует ограничение (4). Следовательно, можно предположить, что коэффициент ранговой конкордации все-таки позволяет судить о степени согласованности мнений экспертов, тем более, как видно из рассмотренных выше примеров (таблицы 1 и 2), для крайних случаев (полная несогласованность или полная согласованность) W принимает соответственно значения 0 или 1. Можно ожидать, что промежуточным вариантам распределения мнений экспертов будут соответствовать некие числа из интервала $(0;1)$.

Однако данное предположение является неверным. Чтобы подтвердить это, приведем несколько контрпримеров. Рассмотрим вариант распределения мнений экспертов (таблица 3).

Таблица 3 – Пример распределения экспертных оценок с совпадением рангов («2»), присвоенных экспертами одному из факторов, $W=0$

Номер анализируемого фактора	**Номер эксперта**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	1	3	1	3	1	3	1	3	1
2	2	2	2	2	2	2	2	2	2
3	3	1	3	1	3	1	3	1	3

Подчеркнем существенную разницу между вариантами, приведенными в таблицах 1 и 3. В таблице 1 все факторы получили в качестве «оценки» одинаковое количество всех возможных рангов, т. е. три «единицы», три «двойки» и три «тройки». Нулевое значение коэффициента ранговой конкордации в этом случае выглядит вполне логичным. В таблице 3 приведен принципиально иной вариант: по второму фактору эксперты высказались единодушно, присвоив ему ранг «2», однако рассчитанный по формуле (1) коэффициент конкордации, тем не менее, также будет равен нулю.

В таблице 4 приведен весьма похожий на рассмотренный в таблице 3 вариант распределения экспертных оценок. Единственное отличие: в этом случае эксперты были единодушны при присвоении ранга «1» (а не «2», как в таблице 3). Однако значение коэффициента конкордации, рассчитанное по формуле (1), при таком незначительном изменении распределения меняется принципиально: вместо $W=0$ (для таблицы 3) в последнем примере мы получаем значение $W=0{,}75$. То есть два весьма близких (с содержательной точки зрения) распределения экспертных оценок дают совершенно разные значения коэффициента согласия W: минимально возможное в первом случае и весьма высокое во втором.

Таблица 4 – Пример распределения экспертных оценок с совпадением рангов («1»), присвоенных экспертами одному из факторов, $W=0{,}75$

Номер анализируемого фактора	**Номер эксперта**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	2	3	2	3	2	3	2	3	2
2	1	1	1	1	1	1	1	1	1
3	3	2	3	2	3	2	3	2	3

Рассмотрим аналогичные примеры для варианта, когда экспертами оцениваются четыре фактора. Принципиальные отличия между распределениями, приведенными в таблицах 5 и 6, также отсутствуют, однако в первом случае коэффициент ранговой конкордации равен 0,067, а во втором – 0,600, т. е. в 9 раз больше.

Таблица 5 – Пример распределения экспертных оценок для четырех анализируемых факторов с совпадением рангов («2»), присвоенных экспертами одному из факторов, $W=0{,}067$

Номер анализируемого фактора	**Номер эксперта**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	2	2	2	2	2	2	2	2	2
2	1	3	4	1	3	4	1	3	4
3	3	4	1	3	4	1	3	4	1
4	4	1	3	4	1	3	4	1	3

Наконец, в таблицах 7 и 8 приведены варианты распределения мнений экспертов для пяти анализируемых объектов. В каждом случае по одному из факторов специалисты высказались единодушно, а по остальным четырем половина экспертов выставила одну оценку, а вторая половину – другую. Разница между таблицами 7 и 8 лишь в конкретных рангах, присвоенных анализируемым факторам. Тем не менее, в первом случае коэффициент конкордации равен 0,9, а во втором – 0.

Таблица 6 – Пример распределения экспертных оценок для четырех анализируемых факторов с совпадением рангов («1»), присвоенных экспертами одному из факторов, $W=0{,}600$

Номер анализируемого фактора	Номер эксперта								
	1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1	1
2	2	3	4	2	3	4	2	3	4
3	3	4	2	3	4	2	3	4	2
4	4	2	3	4	2	3	4	2	3

Таблица 7 – Пример распределения экспертных оценок для пяти анализируемых факторов с совпадением рангов («1»), присвоенных экспертами одному из факторов, и разделившимися в отношении 50/50 значениями рангов для остальных анализируемых объектов, $W=0{,}9$

Номер анализируемого фактора	Номер эксперта									
	1	2	3	4	5	6	7	8	9	10
1	4	5	4	5	4	5	4	5	4	5
2	5	4	5	4	5	4	5	4	5	4
3	1	1	1	1	1	1	1	1	1	1
4	2	3	2	3	2	3	2	3	2	3
5	3	2	3	2	3	2	3	2	3	2

Таблица 8 – Пример распределения экспертных оценок для пяти анализируемых факторов с совпадением рангов («3»), присвоенных экспертами одному из факторов, и разделившимися в отношении 50/50 значениями рангов для остальных анализируемых объектов, $W=0$

Номер анализируемого фактора	Номер эксперта									
	1	2	3	4	5	6	7	8	9	10
1	1	5	1	5	1	5	1	5	1	5
2	5	1	5	1	5	1	5	1	5	1
3	3	3	3	3	3	3	3	3	3	3
4	2	4	2	4	2	4	2	4	2	4
5	4	2	4	2	4	2	4	2	4	2

Вывод: коэффициент ранговой конкордации (1) не может рассматриваться как адекватная характеристика степени согласованности мнений экспертов.

Вариант расчета коэффициента согласия ранговых экспертных оценок, свободный от перечисленных выше недостатков, будет рассмотрен в отдельной статье.

Список использованных источников

1. Жуков Г.П., Викулов С.Ф. Военно-экономический анализ и исследование операций: Учебник. – М.: Воениздат, 1987. – 440 с.

2. Григан А.М. Управленческая диагностика: теория и практика. – Ростов-на-Дону: Изд-во РСЭИ, 2009. – 316 с.