

## МАТЕМАТИЧЕСКИЕ МЕТОДЫ АНАЛИЗА ЭКСПЕРТНЫХ ОЦЕНОК

© 2010 А.И. Чегодаев\*

**Ключевые слова:** экспертные оценки, результат экспертизы, теория измерений, метод медиан, метод средних арифметических, коэффициент ранговой корреляции, коэффициент конкордации, элементарные предпочтения, попарное сравнение.

Изучаются математические методы обработки и анализа результатов экспертизы, полученных экспертных оценок. Анализ заключений экспертов, выраженный в качественном виде, проводится с помощью теории измерений. Показывается применение этой теории к ранжировке объектов экспертизы, к вычислению средних баллов методом медиан и методом средних арифметических. Рассматривается использование результатов ранжирования посредством коэффициентов ранговой корреляции Спирмена и Кендалла, а также коэффициента конкордации Кендалла. Изучается вопрос обработки и анализа попарного сравнения и балльного оценивания объектов (выраженного в количественном и качественном виде) с целью выявления элементарных предпочтений (оценок), при этом используются методы математической статистики.

Научный подход к обоснованию принимаемых решений является гарантией успехов в управлении, если проводится глубокий анализ поставленных проблем на основе объективных данных с использованием различных научных методов принятия решений. Лицо, принимающее решения (ЛПР), использует не только личные знания, личный опыт и интуицию, но и чужой опыт, советы специалистов, экспертов при анализе данных и возникшей ситуации. Однако решение сложных, комплексных проблем в условиях неопределенностей и неполноты информации, в сложных ситуациях происходит посредством логико-эвристического анализа на основе экспертного оценивания, которое осуществляют компетентные в данной области специалисты (эксперты), используя специальные математические процедуры и математические методы анализа экспертных оценок, т.е. выводов, получаемых экспертами и носящих субъективный характер. Экспертные оценки не дают решения проблемы, они представляют информацию, необходимую для обоснования решения, для разработки окончательного решения, принимаемого ЛПР. Эксперты, по существу, моделируют ситуацию, связанную с проблемой, задачей, синтезируют какие-то объекты, оценивают их характеристики, их взаимосвязи, свойства, находят аргументы, доказательства, помогающие ЛПР принять решение, устраняющее возникшую проблему<sup>1</sup>.

Обработка полученной экспертной информации проводится с целью приведения ее к виду, удобному для подготовки предложений ЛПР и придания ей характера знаний, а также с целью выделения значения наиболее важных характеристик полученных новых данных. Способ обработки результатов экспертизы зависит от природы исследуемых факторов и типа шкалы результатов, требуемых точности и оперативности получения необходимых характеристик, при этом комплексно используются как строгие количественные, так и различные качественные способы обработки. Чтобы оценить качество того или иного метода обработки и анализа результатов экспертизы по сравнению с другими методами, его преимущества и недостатки, необходимо выявить его эффективность. Под эффективностью метода обработки или анализа обычно понимают какую-либо меру степени действенности этого или иного метода в отношении намеченных на экспертизу целей. В качестве меры действенности можно рассматривать, например, степень достоверности полученной информации, которая используется для выработки рационального решения, оценку оперативности и гибкости способа обработки или анализа, стоимость обработки и анализа результата или ценность получаемых в ходе таких операций новых интерпретаций или знаний и т.п.

Основой теории экспертных оценок, прежде всего, той ее части, которая связана

\* Чегодаев Анатолий Иванович, кандидат физико-математических наук, доцент Ярославского высшего зенитного ракетного училища противовоздушной обороны (военного института). E-mail: vestnik@sseu.ru.

с анализом заключений экспертов, выраженных в качественном (а не в количественном) виде, является репрезентативная теория измерений. Эта теория связана с представлением отношений между реальными объектами в виде отношений между числами и является одной из составных частей эконометрики. Теория измерений нужна на практике для экспертного оценивания, например, в связи с агрегированием мнений экспертов, построением обобщенных показателей, называемых рейтингами. Экспертам часто предлагают осуществить упорядочение (ранжировку) объектов экспертизы, т.е. расположить их в порядке убывания интенсивности какой-то характеристики, которая интересует организаторов экспертизы. Ранг - это номер объекта экспертизы в упорядоченной совокупности. Формально ранги выражаются числами  $1, 2, 3, \dots, n, \dots$  натурального ряда, однако с этими числами нельзя производить обычные арифметические операции в связи с тем, что эти числа выражают интенсивность изучаемой характеристики в определенном смысле, на качественном уровне.

Ранжирование может быть строгим и нестрогим. При строгом ранжировании эксперт должен расположить оцениваемые элементы в порядке возрастания (убывания) их предпочтительности и приписать каждому из них ранг. Наиболее предпочтительный элемент имеет ранг  $1$ , а наименее предпочтительный - ранг  $k$ . При нестрогом ранжировании одинаковым по предпочтительности элементам присваиваются одинаковые ранги. Однако сумма рангов так же, как при прямом ранжировании, должна равняться сумме возможных мест, число которых равно числу элементов. Поэтому одинаковым по предпочтительности элементам присваивают стандартизованные ранги, которые представляют собой среднее арифметическое номеров соответствующих элементов. Заметим, что процедура ранжирования дает наиболее надежные результаты, когда число оцениваемых объектов меньше или равно  $10$ , а предельное рекомендуемое количество предъявленных элементов равно  $20$ .

Далее рассмотрим другой пример применения результатов теории измерений, связанных со средними величинами. При проведении различных опросов опрашиваемых про-

сят выставить баллы изделиям, предприятиям, проектам, политикам и т.п. Затем рассчитывают средние баллы и рассматривают их как обобщенные (итоговые, интегральные) оценки, выставленные коллективом опрошенных экспертов. Из многих существующих средних величин обычно применяют среднее арифметическое, однако такой способ некорректен с точки зрения специалистов по теории измерений, так как баллы обычно измерены в порядковой шкале и не дают обоснования важности характеристик рассматриваемых объектов, не позволяют произвести упорядочение объектов в зависимости от интенсивности этих характеристик. Обоснованным является использование медиан в качестве средних баллов. Напомним, что медианой  $M_j$  называют значение количественного признака  $X$ , приходящегося на середину ранжированного ряда наблюдений. В случае, если приведено нечетное число наблюдений  $n = 2k - 1$  и результаты их проранжированы, выписаны в последовательность, то значение  $x_k$ , расположенное в середине ряда, равно медиане:  $M_j = x_k$ . В случае, если приведено четное число наблюдений  $n = 2k$ , то на середину ранжированного ряда приходится два значения  $x_k, x_{k+1}$ . В этом случае за медиану принимают среднюю арифметическую наблюдений  $x_k, x_{k+1}$ :  $M_j = \frac{x_k + x_{k+1}}{2}$ .

В связи с тем, что полное игнорирование средних арифметических нецелесообразно из-за их привычности и распространенности, представляется рациональным одновременное использование метода медиан и метода средних арифметических, что находится в согласии с общенаучной концепцией устойчивости, рекомендующей применять различные методы для обработки одних и тех же данных с целью выделить выводы, получаемые одновременно при всех методах. Такие выводы соответствуют реальной действительности. Заключение, меняющиеся от метода к методу, зависят от субъективизма исследователя, выбирающего метод обработки исходных экспертных оценок.

Результаты ранжирования данных, полученных разными экспертами, могут использоваться с целью решения следующих задач:

1) нахождения тесноты связи между ранжировками двух экспертов; 2) нахождения тесноты связи между произвольным числом ранжированных признаков; 3) оценки согласованности мнений экспертов в группе, содержащей более двух экспертов. При решении первых двух задач используются коэффициенты ранговой корреляции Спирмена или Кендалла, а при решении второй задачи - коэффициент конкордации Кендалла. Коэффициент ранговой корреляции Спирмена  $\rho$  может использоваться как при строгом, так и при нестрогом ранжировании. При строгом ранжировании расчет производится по формуле:

$$\rho = \frac{6 \sum_{i=1}^m (r_{1i} - r_{2i})^2}{m(m^2 - 1) - 0,5 \cdot (S_1 + S_2)},$$

где  $m$  - число элементов;  $r_{1i}$  - ранг, приписанный первым экспертом  $i$ -му элементу;  $r_{2i}$  - ранг, приписанный вторым экспертом  $i$ -му элементу. Коэффициент корреляции Спирмена может изменяться от  $-1$  до  $+1$ . Если коэффициент корреляции равен  $1$ , то ранжировки одинаковы; если  $\rho = -1$ , то ранжировки противоположны (обратны друг другу); если  $\rho = 0$ , то ранжировки линейно независимы (некоррелированы).

Заметим, что при экспертном оценивании погрешность оценки носит случайный характер. Поэтому коэффициент корреляции представляет собой случайную величину. Если  $m > 10$ , то распределение величины

$$T = \rho \sqrt{\frac{m-2}{1-\rho^2}}$$

является распределением Стюдента с  $m-2$  степенями свободы, а если  $m > 30$ , то величина  $\rho$  имеет нормальное распределение с параметрами  $\rho$  и  $\frac{1}{m-1}$ ,

$D(\rho) = \frac{1}{m-1}$ . Для оценки степени близости мнений двух экспертов при строгом ранжировании можно применять коэффициент  $\tau$  корреляции Кендалла. С этой целью в обеих ранжировках рассматривают все возможные пары каких-либо элементов и выясняют, согласованы они или нет. Понятие согласован-

ной пары  $(a, b)$  несовпадающих элементов вводится так: пара  $(a, b)$  называется согласованной, если в обеих ранжировках элемент  $a$  стоит раньше (или в обеих парах - позже) элемента  $b$ . Для вычисления коэффициента ранговой корреляции Кендалла сначала вычисляют разность  $S^+ - S^-$ , где  $S^+$  - число согласованных пар, а  $S^-$  - число несогласованных пар, а затем и коэффициент  $\tau$  по формуле  $\tau = \frac{S^+ - S^-}{m}$ , где  $m$  - число элементов. При нестрогом ранжировании коэффициент Спирмена вычисляют (если  $|\rho| \leq 1$ ) по

$$\rho = 1 - \frac{6 \sum_{j=1}^m (r_{1j} - r_{2j})^2}{m(m^2 - 1) - 0,5 \cdot (S_1 + S_2)},$$

формулам:  $\rho = 1 - \frac{6 \sum_{j=1}^m (r_{1j} - r_{2j})^2}{m(m^2 - 1) - 0,5 \cdot (S_1 + S_2)}$ , где  $S_i = \sum_{K=1}^{K_i} (t_{Ki}^3 - t_{Ki})$ , где  $K_i$  - число групп одинаковых рангов в ранжировке  $i$ -го эксперта;  $t_{Ki}$  - число одинаковых рангов в  $k$ -й группе нестрогих ранжировок  $i$ -го эксперта.

Заметим, что справедливо приближенное равенство  $\rho \approx 1,5\tau$ . Следовательно, для  $\tau$  оценку можно дать после быстрого вычисления  $\rho$ , и наоборот. Кроме того, коэффициент  $\rho$  дает более точный результат в смысле минимума дисперсии, а коэффициент  $\tau$  выгоднее тем, что позволяет вычислять частные корреляции для отдельных членов ряда, а также производить простой пересчет при добавлении дополнительных элементов во множество предъявления.

Групповое мнение экспертов можно формировать только тогда, когда индивидуальные предпочтения хорошо согласуются. Для оценки согласованности мнений в группе из более чем двух человек чаще всего используют коэффициент конкордации Кендалла, который имеет различный вид в зависимости от типа ранжирования. При строгом ранжировании коэффициент конкордации находится по формуле

$$C_K = \frac{12S_C}{n^2 m(m^2 - 1)},$$

где  $S_C = \sum_{j=1}^m \left( \sum_{i=1}^n r_{ij} - \frac{n(m+1)}{2} \right)^2$ ,  $n$  - число экс-

пертов,  $m$  - число оцениваемых параметров,  $r_{ij}$  - ранг  $j$ -го элемента, присвоенный  $i$ -м экспертом.

Если значение коэффициента  $S_K$  конкордации мало (велика вероятность того, что результаты ранжирования случайны), то следует изменить состав экспертной группы. Для этого нужно разбить экспертов на пары и найти для каждой пары коэффициент парной ранговой корреляции Спирмена. Затем нужно объединить в подгруппы тех экспертов, у которых значения коэффициентов парной корреляции близки. Наконец, для каждой подгруппы нужно определить коэффициент конкордации Кендалла.

Рассмотрим обработку, анализ попарного сравнения и балльного оценивания объектов, элементов. Попарное сравнение - наиболее выразительный способ выявления элементарных предпочтений. При сравнении двух произвольных элементов имеет место всегда один из трех альтернативных вариантов суждения: 1) элемент  $a_i$  предпочтительнее элемента  $a_j$ ; 2) оба элемента  $a_i$  и  $a_j$  одинаково предпочтительны; 3) элемент  $a_j$  предпочтительнее элемента  $a_i$ . Чаще всего при попарном сравнении ограничиваются простой констатацией того, что один из элементов предпочтительнее другого. В этом случае попарное сравнение есть измерение в номинальной шкале. Иногда удается выявить степень предпочтения, и тогда используют специальные шкалы, где каждой степени предпочтения присваивается определенная оценка (измерение в порядковой шкале). Чаще всего при попарном сравнении один или несколько экспертов (группа) проводят оценку предпочтительности для всех пар элементов по следующему правилу:

$$e_{jk} = \begin{cases} 1, & \text{если } a_j \succ a_k, \\ 0,5, & \text{если } a_j \sim a_k, \\ 0, & \text{если } a_k \succ a_j, \end{cases}$$

при этом  $j, k = 1, 2, \dots, m$ . Числа  $e_{jk}$  представляют собой элементы матрицы  $E$  результатов

попарных сравнений элементов каждым  $i$ -м экспертом. Полученные матрицы  $E$  обычно нормируют (усредняют) путем деления на

число экспертов:  $Z_{jk} = \frac{1}{n} e_{jk}$ . При балльном оценивании эксперт должен каждому элементу из множества предъявления приписать соответствующее число (балл), которое отражает субъективное мнение эксперта о предпочтительности, ценности, важности этого элемента. Указанные числа выбираются из специальной балльной шкалы. Самый известный пример балльных оценок - это школьные оценки. Балльные оценки бывают двух видов. Оценки первого вида производятся по объективному критерию, по общепринятому эталону. Как правило, таковы оценки в спортивном судействе или правила присвоения рабочих разрядов. Балльная оценка второго вида - это оценка, производимая, когда нет общепринятых эталонов и единственного объективного критерия, субъективными отражениями которого являются оценки. Метод попарного сравнения применяется потому, что считается более легким качественное сравнение двух объектов, чем выражение предпочтения в балльной (или ранговой) шкале. Иногда при попарных сравнениях эксперту предлагается баллами оценить интенсивность своего предпочтения для каждой пары объектов.

Для получения и обработки количественными методами качественной экспертной информации могут использоваться вербально-числовые шкалы, в состав которых входят содержательно описываемые наименования ее градаций и соответствующие им численные значения или диапазоны численных значений.

На практике для выражения предпочтений часто оказывается полезным сначала произвести балльное оценивание элементов множества предъявления, а затем, ориентируясь на величины баллов, получить искомую ранжировку.

При количественном подходе балльные оценки  $x_{ij}$ , данные  $i$ -му объекту  $j$ -м экспертом, могут обрабатываться по-разному. Если эксперты равноправны, то простейшая груп-

повая оценка  $i$ -го объекта равна  $x_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$ ,

где  $n$  - число экспертов (т.е. средний балл). Если хотят учесть компетентность, объективность, информированность экспертов, то вводят весовые показатели компетентности: для

$j$ -го эксперта это  $q_j$ . Тогда  $x_i = \sum_{j=1}^n q_j x_{ij}$ . Час-

то оцениваемый показатель бывает сложным для непосредственной оценки. Например, таким показателем может быть показатель "качества изделия". Тогда такой показатель разбивают на более простые показатели, каждый из которых выражает некоторую сторону, часть исходного показателя. В этом случае  $x_{ijk}$  - оценка  $i$ -го изделия  $j$ -м экспертом по  $k$ -му признаку, а  $Z_{kj}$  - оценка весомости  $k$ -го признака  $j$ -м экспертом. Тогда полагают, что оценка весомости  $k$ -го признака рав-

на  $Z_k = \sum_{j=1}^n q_j Z_{kj}$ , а оценка  $i$ -го изделия равна

$$x_i = \sum_{j=1}^n \sum_{k=1}^m Z_{kj} q_j x_{ijk},$$

где  $m$  - число признаков.

Эта оценка  $x_i$  представима в виде:

$$x_i = \sum_{j=1}^n q_j \sum_{k=1}^m Z_{kj} x_{ijk} = \sum_{j=1}^n q_j x_{ij}.$$

Эксперименты и практика экспертизы показывают, что такие средние баллы  $x_i$  ведут себя достаточно устойчиво при изменении состава экспертной группы. Однако, вообще говоря, необходимо относиться с осторожностью к последнему утверждению, так как оно не всегда верно. Степень согласованности экспертов оценивается либо дисперсиями индивидуальных балльных оценок:

$$D_i = \frac{1}{n-1} \sum_{j=1}^m (x_{ij} - x_i)^2, \quad i = 1, 2, \dots, n, \text{ либо коэф-}$$

фициентами вариации, если балльные оценки положительны. Коэффициенты вариации

вычисляются следующим образом:  $V_i = \frac{\sqrt{D_i}}{x_i}$ ,

$i = 1, 2, \dots, n$ . При использовании коэффициентов вариации согласованность мнений экс-

пертов считается хорошей, если все  $V_i < 0,2$ , и удовлетворительной, если все  $V_i < 0,3$ .

Для определения интервала достоверных значений используют методы интервального оценивания. Для симметричного закона распределения случайной величины (значения оценки) доверительный интервал  $(x_{cp} - E, x_{cp} + E)$  можно определить, если использовать следующее отношение:

$$P(|x - x_{ucm}| < E) = L, \text{ где } x_{ucm} - \text{истинное значение оцениваемой характеристики; } E - \text{половина длины доверительного интервала; } L - \text{доверительная вероятность из интервала } (0,9; 0,99); x_{cp} - \text{среднее значение оценки } x.$$

Если число экспертов в группе  $n > 30$ , то считают распределение оценки нормальным, а поэтому для нахождения вероятности  $L$  используют таблицы нормального распределения. В противном случае доверительную вероятность находят с помощью таблиц распределения Стьюдента.

Компетентность эксперта есть степень его квалификации в определенной области знаний. Отбор экспертов в экспертную группу проводят на основе анализа доступной информации о профессиональной подготовке кандидатов, их ученой степени и ученом звании, стаже работы, кругозоре и эрудиции, участии в других экспертизах, уровне достижений в профессиональной деятельности. Однако можно воспользоваться другой возможностью определения степени компетентности экспертов. Для этого проводят контрольную экспертизу в предположении, что правильные ответы на поставленные вопросы заранее неизвестны. При этом используют подход, основанный на обработке нормированных балльных оценок.

Рассмотрим пример. Вычисления разбиваются на несколько этапов.

1. В контрольной экспертизе участвуют четыре эксперта, которые должны оценить проблему повышения рентабельности предприятия посредством выставления оценок по 10-балльной шкале за относительную важность покупательной способности, себестоимости  $a_2$  и затрат  $a_3$ .

Таблица 1

Оценки экспертов

Эксперт	Оцениваемые элементы			
	$a_1$	$a_2$	$a_3$	$a_1 + a_2 + a_3$
1	8	4	2	14
2	10	2	1	13
3	4	8	2	14
4	5	4	3	12

В табл. 1 указана сумма баллов, полученная каждым экспертом за оценку трех элементов  $a_1, a_2, a_3$ .

2. Затем составляется таблица нормированных балльных оценок для каждого эксперта путем деления каждого балла на суммарный балл для данного эксперта.

Таблица 2

Нормированные балльные оценки

Эксперт	Оцениваемые элементы		
	$a_1$	$a_2$	$a_3$
1	0,57	0,29	0,14
2	0,77	0,15	0,08
3	0,29	0,58	0,13
4	0,42	0,33	0,25
Средние баллы	0,51	0,34	0,15

В табл. 2 указаны средние баллы, полученные каждой из  $a_1, a_2, a_3$  (характеристик проблемы рентабельности предприятия).

3. Вычисляются взвешенные суммы относительных балльных оценок для каждого эксперта:

а) для первого эксперта:  $0,57 \cdot 0,51 + 0,29 \cdot 0,34 + 0,14 \cdot 0,15 = 0,41$ ;

б) для второго эксперта:  $0,77 \cdot 0,51 + 0,15 \cdot 0,34 + 0,08 \cdot 0,15 = 0,45$ ;

в) для третьего эксперта:  $0,29 \cdot 0,51 + 0,58 \cdot 0,34 + 0,13 \cdot 0,15 = 0,31$ ;

г) для четвертого эксперта:  $0,42 \cdot 0,51 + 0,33 \cdot 0,34 + 0,25 \cdot 0,15 = 0,36$ .

4. Вычисляется сумма полученных взвешенных оценок:

$$0,41 + 0,45 + 0,31 + 0,36 = 1,53.$$

5. Находят коэффициенты компетентности экспертов:

а) для первого эксперта:  $\frac{0,41}{1,53} = 0,26$ ;

б) для второго эксперта:  $\frac{0,45}{1,53} = 0,30$ ;

в) для третьего эксперта:  $\frac{0,31}{1,53} = 0,20$ ;

г) для четвертого эксперта:  $\frac{0,36}{1,53} = 0,24$ .

Средняя групповая компетентность составляет  $0,25$ . Коэффициенты компетентности у 1-го и 4-го эксперта ближе к средней групповой компетентности, поэтому можно полагать, что самыми компетентными экспертами являются первый и четвертый эксперты.

<sup>1</sup> См.: Орлов А.И. Теория принятия решений. М., 2006; Бешелев С.Д., Гурвич Ф.Г. Математико-статистические методы экспертных оценок. М., 1980; Евланов Л.Г., Кутузов В.А. Экспертные оценки в управлении. М., 1978; Чегодаев А.И. Математические методы и модели поддержки принятия решений в условиях неопределенностей: монография. Ярославль, 2007; Кувшинов М.С. Алгоритмизация расчетов рейтинговых оценок инвестиционного климата предприятий // Вестн. Самар. гос. ун-та. 2008. № 8 (46).

Потупила в редакцию 17.11.2009 г.