

Методы и модели системного анализа

Сравнение интервальных альтернатив*

М. Ю. СТЕРНИН, Г. И. ШЕПЕЛЕВ

Аннотация. В статье предложен подход к сравнению пары альтернатив, описанных различными модификациями интервальных оценок. Рассмотрены как моно-, так и обобщенные интервальные оценки. Вначале обсуждается наглядный эмпирический метод «прямого» сравнения пары моноинтервалов. После этого предлагается его математическое обоснование. Рассматривается возможность применения полученных результатов для сравнения альтернатив, представленных в виде обобщенных интервальных оценок.

Ключевые слова: моно- и полиинтервальные оценки; обобщенные интервальные оценки, сравнение альтернатив.

Введение

Многие практические задачи исследуются в условиях неопределенности. В первую очередь это задачи прогнозирования, в которых необходимо предвидеть будущие значения анализируемых показателей. Во многих случаях эти показатели измеряются в количественных шкалах и из-за неопределенности получают интервальные, часто экспертные, оценки. В задачах выбора в условиях неопределенности индикаторы эффективности сравниваемых альтернатив также имеют интервальное представление.

В рамках интервальной парадигмы можно выделить два основных направления: моноинтервальный и полиинтервальный подходы. В первом из них анализируемые показатели задаются как одноинтер-

вальные оценки. Вторым подход является развитием моноподхода, при котором первичный объект моноподхода — точечная оценка — заменяется интервальной оценкой, а интервал как инструмент описания неопределенности заменяется набором интервальных оценок.

Потребность в полиинтервальном подходе вызвана тем фактом, что поскольку неопределенность знаний эксперта о параметре может быть связана как с длиной, так и с положением интервальной оценки параметра, эксперт может счесть неадекватным его представлениям задание значений какого-либо параметра в виде единственного интервального числа. Фактически знаниям об исходном параметре D может отвечать оценка в виде совокупности интервалов, простирающейся от наименее широкого (мини-интервала) $[D_{lu}, D_{ru}]$ до наиболее широкого (базового) $[D_{ld}, D_{rd}]$. При этом предполагается, что все промежуточные (по длине) интервалы также входят в указанную совокупность. Графическим представлением указанной структуры на диаграмме (D, h) , где ось ординат h является чаще всего, но не обязательно, осью меток интервалов совокупности, служит криволинейная трапеция, называемая полиинтер-

* Работа выполнена при финансовой поддержке программы Президиума РАН «Интеллектуальные информационные технологии, математическое моделирование, системный анализ и автоматизация» и программы ОНИТ РАН «Информационные технологии и методы анализа сложных систем», Российским фондом фундаментальных исследований (проекты № 09-07-00009, № 10-07-00242, № 11-07-00230, № 11-07-00398).

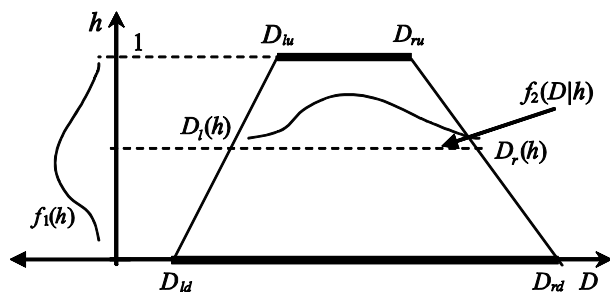


Рис. 1. Поли- и обобщенная интервальная оценка параметра D

вальной оценкой (ПИО). Задание на ПИО совместной функции распределения вероятностей с плотностью $f(h, D) = f_1(h)f_2(D|h)$ определяет обобщенную интервальную оценку (ОИО) анализируемого параметра [2, 5]. Простейший пример поли- и обобщенной интервальной оценки для случая вложенных интервалов показан на рис. 1.

Здесь эксперт определил как начальное приближение интервальной оценки параметра D верхнее основание трапеции, затем, принимая во внимание возможную неопределенность, «размыл» этот интервал в обе стороны, не обязательно симметрично, образовав другую «реперную» интервальную оценку ПИО. Каждый интервал в группе интервалов ПИО можно интерпретировать как сценарий возможной реализации интервальной оценки параметра D , и функция $f_1(h)$ показывает шансы этой реализации.

Одна из возможностей использования полиподхода в задачах принятия решений состоит в переходе от поли- к монокартине [3, 6] за счет усреднения $f(h, D)$ с получением маргинальной функции распределения $f_{av}(D)$, определенной на базовом интервале трапеции, как это показано в соотношении (1).

$$f_{av}(D) = \begin{cases} \int_0^{h_l(D)} f_1(h) f_2(D|h) dh, & D \in [D_{ld}; D_{lu}] \\ \int_0^1 f_1(h) f_2(D|h) dh, & D \in [D_{lu}; D_{ru}] \\ \int_0^{h_r(D)} f_1(h) f_2(D|h) dh, & D \in [D_{ru}; D_{rd}] \end{cases} \quad (1)$$

Здесь

$$h_l(D) = (D - D_{ld}) / (D_{lu} - D_{ld}), \\ h_r(D) = (D_{rd} - D) / (D_{rd} - D_{ru}).$$

Если на обеих осях h и D распределения равномерные, то $f_1(h) = 1$, $f_2(D|h) = 1/[D_r(h) - D_l(h)]$, $D_l(h) = hD_{lu} + (1-h)D_{ld}$, $D_r(h) = hD_{ru} + (1-h)D_{rd}$. Вводя $L = D_{rd} - D_{ld}$, $U = D_{ru} - D_{lu}$, в качестве маргинального распределения получаем обобщенное равномерное распределение $f_{un}(D)$:

$$f_{un}(D) = \frac{1}{L-U} \begin{cases} L \frac{L(D_{lu} - D_{ld})}{L(D_{lu} - D) + U(D - D_{ld})}, & D \in [D_{ld}, D_{lu}] \\ L \frac{L}{L-U}, & D \in (D_{ld}, D_{ru}) \\ L \frac{L(D_{ru} - D_{rd})}{L(D_{ru} - D) + U(D - D_{rd})}, & D \in [D_{rd}, D_{ru}] \end{cases} \quad (2)$$

Соответствующие формулы для других форм ПИО получены в работе [4].

В моноинтервальном случае иерархия неопределенностей содержит точечные оценки (неопределенность отсутствует), интервально-вероятностные, интервальные оценки (наибольшая неопределенность). В рамках подхода ОИО существует следующая иерархия неопределенностей: точечные оценки; интервально-вероятностные оценки второго рода, когда известны функции $f_1(h)$ и $f_2(D|h)$; интервально-вероятностные оценки первого рода, когда известна лишь функция $f_1(h)$; полиинтервальные оценки.

Для того чтобы интервальные оценки можно было использовать при принятии решений, необходимо уметь сравнивать пары альтернатив с интервальными показателями качества. Существуют около 20 сочетаний различных интервальных оценок при сравнении пар альтернатив. Например, в монослучае мы имеем следующие возможные комбинации пар оценок сравниваемых альтернатив: интервально-интервальная пара, интервальная — интервально-вероятностная пара, интервальная — точечная комбинация и так далее. В случае полиподхода мы имеем комбинации ПИО — ПИО, ПИО — интервально-вероятностные оценки первого рода, ПИО — интервально-вероятностные оценки второго рода, ПИО — точечная оценка и так далее и, наконец, все перекрестные комбинации (моно- — ОИО). Чтобы обеспечить сопоставимость результатов этих сравнений, целесообразно построить единый критерий, позволяющий сравнивать любую возможную комбинацию в паре альтернатив. Заранее ясно, что эта проблема не может быть решена чисто математическими методами, а должна рассматриваться как задача принятия решений, т. е. с учетом предпочтений человека, эксперта или лица, принимающего решения (ЛПР), и

его/ее отношения к различным подходам к учету неопределенности. Такая единая концепция сравнения альтернатив предлагается в этой статье.

1. Общий подход к сравнению пары интервальных оценок

Будем считать, что проверяется гипотеза о том, что вторая интервальная альтернатива I_2 предпочтительнее¹ («больше») первой I_1 . Предполагается, что каждая интервальная оценка содержит неизвестное реализуемое точечное значение анализируемого параметра. Если известны шансы появления точечных значений в сравниваемой паре интервалов, численное моделирование процесса реализации точечных оценок может служить основой для сравнения интервальных величин. Будем говорить, что пара точечных реализаций (i_1, i_2) интервальных оценок I_1 и I_2 принадлежит зеленой зоне, если i_2 не меньше, чем i_1 , и принадлежит красной зоне в противоположном случае. Таким образом, зеленая зона — это зона, благоприятствующая истинности проверяемой гипотезы, а красная зона наоборот.

Для сравнения интервалов методом численного моделирования необходимо провести достаточно много (скажем N) испытаний, при каждом из которых соответствующий интервал заменяется точечной реализацией, и отметить количество попаданий в зеленую N_g и красную зоны N_r соответственно. При этом необходимо знать или постулировать и задать распределения вероятностей на сравниваемых интервалах. Можно видеть, что $K_g = N_g/N$ может служить мерой шансов реализации зеленой зоны (шансов, что альтернатива I_2 окажется предпочтительнее альтернативы I_1), а $K_r = N_r/N$ — аналогичной мерой для красной зоны.

Введем коэффициент уверенности K_{as} как разность между K_g и K_r : $K_{as} = K_g - K_r$. Коэффициент уверенности показывает, насколько шансы реализации зеленой зоны превышают шансы реализации красной зоны. Именно этот показатель предлагается в качестве критерия сравнения пар интервальных альтернатив для всех их модификаций. Будем говорить, что альтернатива I_2 теоретически предпочтительнее, чем I_1 , если значение K_{as} положительно. Теоретически, поскольку предпочтения ЛПР еще не учтены. Эти предпочтения могут быть выражены введением назначаемого ЛПР порогового значения коэффициента уверенности K_{th} . Именно, если вы-

численное для данной пары интервальных оценок значение K_{as} не меньше, чем K_{th} , то I_2 следует признать более предпочтительной альтернативой на уровне уверенности K_{as} с учетом предпочтений ЛПР. Если порог, назначенный ЛПР, не позволяет признать вторую альтернативу предпочтительной, необходимо анализировать ситуацию заново. Если K_{as} отрицательно, необходимо проверить противоположную гипотезу.

Рассмотрим некоторые свойства коэффициента уверенности. Пусть $K_{as}(I_2, I_1)$ — коэффициент уверенности, полученный при проверке гипотезы о том, что альтернатива I_2 более предпочтительна, чем I_1 . Возможные значения коэффициента уверенности $K_{as}(I_2, I_1)$ лежат в диапазоне $[-1, 1]$: $-1 \leq K_{as}(I_2, I_1) \leq 1$. Отметим, что условие $K_{as}(I_2, I_1) = 1$ соответствует ситуации, когда I_2 доминирует I_1 (интервалы не пересекаются), а $K_{as}(I_2, I_1) = -1$ отвечает противоположной ситуации. Условие $K_{as}(I_2, I_1) = 0$ соответствует ситуации равнозначности интервальных альтернатив по предпочтению. Как функция двух переменных K_{as} является антисимметричной функцией: $K_{as}(I_2, I_1) = -K_{as}(I_1, I_2)$. Кроме того, K_{as} согласуется с требованием транзитивности отношения предпочтительности. Действительно, если $K_{as}(I_2, I_1) > 0$, тогда (теоретически) $I_2 \succ I_1$; если $K_{as}(I_3, I_2) > 0$, то $I_3 \succ I_2$, и, поскольку, в силу транзитивности, $I_3 \succ I_1$, то $K_{as}(I_3, I_1) > 0$.

Подчеркнем, что коэффициент уверенности описывает лишь усредненные шансы предпочтительности одной интервальной оценки другой в их сравниваемой паре. Поэтому более полную информацию для принятия решений содержит гистограмма значений коэффициента уверенности, возникающих в процессе численного моделирования и выраженных через $i_2 - i_1$, а также другие, помимо среднего, статистические характеристики соответствующего распределения.

Поскольку коэффициент уверенности — новая характеристика, полезно сравнить значения этого фактора со значениями K_g и K_r . Эти данные собраны в табл. 1.

ЛПР может интересоваться также оценка шансов того, что интервал I_2 больше, чем I_1 , на величину $\delta > 0$. Теперь попадание пары реализаций точечных оценок в зеленую зону определяется условием $i_2 - i_1 > \delta$. Для коэффициента уверенности имеем: $K_{as}(\delta) = 2K_g(\delta) - 1$, где $K_g(\delta) = N_g(\delta)/N$, и $N_g(\delta)$ — количество реализаций, удовлетворяющих указанному выше условию.

Проиллюстрируем теперь предложенную процедуру сравнения интервальных оценок на примере пар интервалов с треугольными распределениями на них. Этот пример полезен также потому, что точное распределение шансов на сравниваемых интервалах часто неизвестно. Но мы можем предположить, что

¹ Будем использовать обозначения I_1 и I_2 как для альтернатив, так и для их интервальных оценок.

Таблица 1

Сравнение значения K_{as} , K_g и K_r .

K_g	$K_r = 1 - K_g$	$K_{as} = K_g - K_r$
0	1	-1
0,1	0,9	-0,8
0,2	0,8	-0,6
0,3	0,7	-0,4
0,4	0,6	-0,2
0,5	0,5	0,0
0,6	0,4	0,2
0,7	0,3	0,4
0,8	0,2	0,6
0,9	0,1	0,8
1	0	1

эти распределения унимодальные, и для многих типов распределений можно приближенно заменить эти неизвестные (но унимодальные) распределения треугольными распределениями.

Получим формулу, связывающую случайные числа для равномерного (N_u) и треугольного (N_t) распределений, которую будем использовать при сравнении интервальных оценок методом статистических испытаний. Если L , R и M — левая, правая границы интервала и мода распределения, соответственно, то плотность $f(z)$ треугольного распределения имеет вид:

$$f(z) = \frac{2}{R-L} \begin{cases} \frac{z-L}{M-L}, & L \leq z \leq M, \\ \frac{R-z}{R-M}, & M < z \leq R. \end{cases}$$

Используя метод обратной функции, в соответствии с которым $N_u = \int_L^{N_t} f(z) dz$, получим:

$$N_t = \begin{cases} L + [N_u(R-L)(M-L)]^{1/2}, & N_u \leq (M-L)/(R-L), \\ R - [(1-N_u)(R-M)(R-L)]^{1/2}, & N_u > (M-L)/(R-L). \end{cases}$$

Используя это соотношение, рассмотрим теперь пример сравнения интервальных оценок с треугольным распределением на них. Границы интервалов (L , R) и значения моды (M) распределения представлены в табл. 2. Результаты округлены до второго знака после запятой.

Количество испытаний в каждом варианте равно 100 000. Сравнение сделано на основании значений $K_{as}(I_2, I_1)$ для $\delta = 0$. Введено обозначение $(K_{as})_c$, подчеркивающее, что значение коэффициента уверенности зависит от числа испытаний.

Прокомментируем полученные результаты. Строки 1–3 табл. 2 соответствуют конфигурации правого сдвига для пары сравниваемых интервальных оценок. Заранее ясно, что эта конфигурация благоприятствует истинности гипотезы $I_2 \succ I_1$. Но можно ли безоговорочно принять это предположение? Значения коэффициента уверенности показывают, что для приведенной ситуации — да, поскольку даже в худшем случае (строка 1), когда моды распределений находятся в практически крайних позициях (слева для второго интервала и справа для первого — пессимистические для истинности гипотезы оценки), его значение превышает 0,7. Такое значение, по видимому, может удовлетворить большинство ЛПР.

Строки 4–6 соответствуют конфигурации совпадающих интервальных оценок. Можем ли мы принять, что эти оценки эквивалентны при любом расположении мод распределений на сравниваемых интервалах? Вспомним, что для эквивалентных оценок

Таблица 2

Сравнение интервальных оценок для треугольного распределения

Номер примера	Интервал 1			Интервал 2			N_g	$(K_{as})_c$
	L_1	R_1	M_1	L_2	R_2	M_2		
1	50	100	90	70	170	80	85 055	0,70
2	50	100	60	70	170	160	99 244	0,98
3	50	100	75	70	170	120	97 832	0,95
4	50	100	75	50	100	75	50 100	0,00
5	50	100	90	50	100	60	25 672	-0,49
6	50	100	60	50	100	90	74 652	0,49
7	70	170	75	80	150	145	75 970	0,52
8	70	170	160	80	150	100	20 091	-0,59
9	70	170	120	80	150	115	42 648	-0,15

$K_{as} = 0$. Очевидно, что это условие выполнено только для симметричных треугольных распределений (строка 4). Это не так в других случаях (строки 5, 6). Однако для приведенной ситуации даже для самых благоприятных условий (строка 6), когда мода треугольного распределения для второй оценки резко сдвинута вправо, а мода первого распределения — влево, значение коэффициента $K_{as} = 0,49$ может оказаться недостаточным для того, чтобы рациональный ЛПР сделал заключение, что $I_2 \succ I_1$.

Результаты сравнения двух вложенных интервальных оценок показаны в строках 7–9. Видно, что значения K_{as} и, стало быть, выводы о предпочтительности существенно зависят от положения мод распределений.

Эти примеры показывают, что если информация о шансах на реализацию зеленых и красных зон известна, можно сказать, что проблема сравнения альтернатив с помощью введенного коэффициента уверенности в принципе решена. Во всяком случае, она решена для интервально-вероятностных пар и пар, заданных интервально-вероятностными и точечными оценками.

Но если хотя бы одна оценка чисто интервальная, мы не сможем осуществить эту программу. Что можно предложить в этой ситуации? Возможны два пути. Первый основан на принципе максимальной энтропии (принцип Гиббса—Джейнса), согласно которому все возможные состояния природы имеют равные шансы на реализацию и, следовательно, необходимо признать наличие на интервале равномерного распределения. Мы обсудим некоторые последствия этого предположения в разделах 2 и 3.

Второй путь основан на предположении о том, что неизвестное распределение может быть любым, включая смесь распределений. Эта ситуация уже обсуждалась на примерах сравнения оценок при треугольных распределениях, когда предполагалось, что неизвестные (но унимодальные) распределения могут быть аппроксимированы треугольными распределениями.

2. Эвристический метод прямого сравнения интервальных оценок

В этом разделе обсуждается эвристический метод прямого сравнения альтернатив, заданных интервальными оценками, в котором неявно предполагается наличие равномерных распределений на сравниваемых интервалах. Ниже будет показано, что значения коэффициентов уверенности, полученные по этому методу, оказываются нижними границами точных

значений. Поэтому решения, принятые на основе метода прямого сравнения, будут не менее надежны, чем решения на базе точного метода. Эвристический метод может быть полезен как быстрый и интуитивно ясный способ сравнения интервальных оценок в условиях предположения о равномерных распределениях на сравниваемых интервалах.

Рассмотрим два числовых интервала $I_i = [L_i, R_i]$, $i = 1, 2$, заданных их левой L_i и правой R_i границами, $L_i < R_i$. С точностью до перестановки возможны следующие четыре конфигурации взаимного расположения I_1 и I_2 : (a) $R_1 < L_2$ (интервалы не пересекаются); (b) $L_1 < L_2 < R_1 < R_2$ (конфигурация правого сдвига); (c) $L_1 < L_2 < R_2 < R_1$ (вложенные интервальные оценки); (d) $L_1 = L_2, R_1 = R_2$ (совпадающие оценки). Мы, как и ранее, будем тестировать гипотезу о том, что, с точки зрения ЛПР, альтернатива (интервал) I_2 более предпочтительна, чем I_1 ($I_2 \succ I_1$).

Конфигурация (a). Здесь нет затруднений: во всех возможных будущих реализациях i_1 и i_2 на интервалах I_1 и I_2 , соответственно, интервал I_1 не может оказаться предпочтительнее I_2 т. е. $I_2 \succ I_1$. В этой ситуации фактически нет неопределенности, и, стало быть, нет задачи принятия решений.

Конфигурация (d). Она характеризуется наибольшей неопределенностью, интервалы эквивалентны друг другу по предпочтительности. В этой ситуации ЛПР не может принять рационального решения в пользу какой-либо альтернативы без дополнительной информации. Необходимо или отказаться от выбора, или, если выбор неизбежен, а получить дополнительную информацию за требуемое время невозможно, придется использовать случайный выбор. В этой конфигурации коэффициент уверенности для «интуитивного» метода прямого сравнения $(K_{as})_{int}$ равен 0. Связь $(K_{as})_{int}$ с ранее введенным коэффициентом K_{as} обсуждается в следующем разделе. Этот пример показывает, что величину области пересечения $I_1 \cap I_2$ можно рассматривать как меру неопределенности в задаче сравнения интервалов. Область пересечения назовем желтой зоной.

Анализ конфигураций (b) и (c) позволяет выделить две другие меры, играющие важную роль при сравнении интервалов, — величины зеленой и красной зон. Мы уже вводили эти понятия ранее, здесь это сделано с несколько иной точки зрения. Равнозначность этих двух подходов показана в следующем разделе.

Конфигурация (b). Если для I_1 текущая реализация $i_1 \in [L_1, L_2]$ или для I_2 текущая реализация $i_2 \in [R_1, R_2]$, то это говорит в пользу гипотезы о предпочтительности интервала I_2 . В соответствии с принятым определением это зеленая зона. Она состоит из двух областей: $[L_1, L_2]$ и $[R_1, R_2]$. Ее размах $S_g = L_2 - L_1 + R_2 - R_1$.

Размеры интервала $[L_2, R_1]$ характеризуют размах зоны неопределенности при сравнении I_1 и I_2 . Это желтая зона с размахом $S_y = R_1 - L_2$.

Конфигурация (с). Это конфигурация вложенных интервалов (второй интервал вложен в первый). Кроме зеленой и желтой зон здесь возникает неблагоприятная («красная») зона. Если мы проверяем гипотезу $I_2 \succ I_1$, размах красной зоны S_r в этой конфигурации равен $S_r = R_1 - R_2$. Размах зеленой зоны $S_g = L_2 - L_1$ и желтой — $S_y = R_2 - L_2$.

Пусть S — нормирующий множитель. Тогда для всех конфигураций приемлемо решающее правило, в соответствии с которым позитивное решение о предпочтении $I_2 \succ I_1$ может быть сделано, если коэффициент уверенности для интуитивного метода $(K_{as})_{int} = (S_g - S_r)/S > K_{th}$. Здесь K_{th} индивидуальное для каждого ЛПР пороговое значение $(K_{as})_{int}$, отвечающее его предпочтениям.

В соответствии с этим определением $(K_{as})_{int}$ при сравнении пары альтернатив можно не учитывать размах желтой зоны при вычислении значения коэффициента уверенности. Это, как будет показано в следующем разделе, прямое следствие предположения о равномерном распределении на интервалах. Мы уже видели, что для других распределений, в частности для треугольного распределения, это не так.

Коэффициент $(K_{as})_{int}$ имеет простой экономический смысл. Если мы рассматриваем конфигурацию (b) и в качестве предпочтительного выбран интервал I_2 , то максимальный выигрыш за счет принятия верного решения равен $R_2 - L_1$. Если бы интервал I_1 был предпочтен I_2 , то максимальный выигрыш был бы равен $R_1 - L_2$. Относительный прирост возможного максимального выигрыша за счет правильного выбора предпочтительного интервала равен как раз $(K_{as})_{int} = [(R_2 - L_1) - (R_1 - L_2)]/S$.

Нормирующий множитель S можно выбрать поразному. В частности, можно принять, что S равен размаху объединения сравниваемых интервалов, так что их пересечение учитывается один раз. Тогда $S = R_2 - L_1$ для конфигурации (b) и $S = R_1 - L_1$ для конфигурации (с). Этот выбор нормировки достаточно прост и нагляден. Однако при таком выборе нормировки в прямом методе сравнения вычисляемое значение коэффициента уверенности не согласуется для конфигурации (b), вообще говоря, с определением K_{as} , введенным в разделе 1. Это не препятствует, однако, использованию прямого метода: такой выбор S уменьшает значение коэффициента уверенности $(K_{as})_{int}$ по сравнению с его точным значением K_{as} для конфигурации (b) и равен ему для конфигурации (с). Иными словами, полученное прямым методом значение $(K_{as})_{int}$ позволяет оценить нижнюю границу

уверенности ЛПР в истинности гипотезы о предпочтительности второго интервала.

Сходные формулы для коэффициента уверенности, названного «показателем интервального неравенства», получены из других соображений в работе [1].

3. Точное сравнение интервальных оценок при равномерном распределении

Найдем распределение разности Z двух равномерных распределений X_2, X_1 , заданных на интервалах I_2, I_1 соответственно, и изучим его свойства. Определим кумулятивную функцию распределения $F(z)$ равенством:

$$F(z) = F(Z < z) = \iint_{\substack{x_2 - x_1 < z \\ L_1 < x_1 < R_1 \\ L_2 < x_2 < R_2}} f(x_1)f(x_2)dx_1dx_2,$$

$$x_1 \in I_1, x_2 \in I_2; f(x_1) = 1/(R_1 - L_1); f(x_2) = 1/(R_2 - L_2).$$

Здесь возникают два типа областей интегрирования на плоскости с декартовыми координатами ($X = X_1, Y = X_2$): прямоугольник, вытянутый вверх, $R_1 - L_1 < R_2 - L_2$ (область I), и прямоугольник, вытянутый вправо, $R_1 - L_1 > R_2 - L_2$ (область II).

В каждой области интегрирования возникают три подобласти переменной z .

Область I, подобласть 1: $L_2 - R_1 < z < L_2 - L_1$; подобласть 2: $L_2 - L_1 < z < R_2 - R_1$; подобласть 3: $R_2 - R_1 < z < R_2 - L_1$.

Область II, подобласть 1: $L_2 - R_1 < z < R_2 - R_1$; подобласть 2: $R_2 - R_1 < z < L_2 - L_1$; подобласть 3: $L_2 - L_1 < z < R_2 - L_1$.

Теперь несложно выполнить интегрирование и, вводя обозначения $S_1 = R_1 - L_1, S_2 = R_2 - L_2, S_q = S_1S_2$, получить для кумулятивной функции распределения $F(z)$ и ее плотности $f(z)$ следующие соотношения.

Если $z < L_2 - R_1$, то $f = 0, F = 0$, если $z > R_2 - L_1$, то $f = 0, F = 1$.

Для $R_1 - L_1 < R_2 - L_2$ (область I):

если $L_2 - R_1 < z < L_2 - L_1$, то

$$f(z) = \frac{z + R_1 - L_2}{S_q}, F(z) = \frac{(z + R_1 - L_2)^2}{2S_q};$$

если $L_2 - L_1 < z < R_2 - R_1$, то

$$f(z) = \frac{R_1 - L_1}{S_q}, F(z) = \frac{R_1 - L_1}{2S_q} [R_1 + L_1 + 2(z - L_2)];$$

если $R_2 - R_1 < z < R_2 - L_1$, то

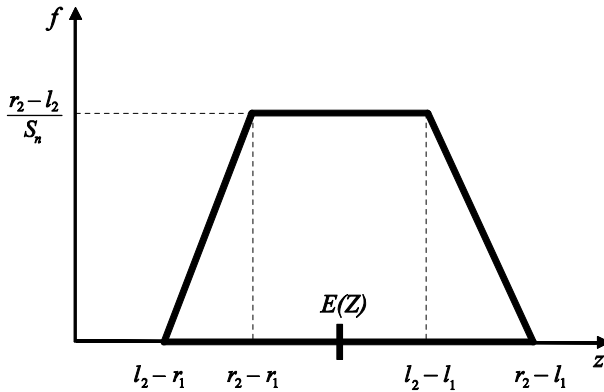


Рис. 2. Плотность вероятности разности двух равномерных распределений для области II

$$f(z) = \frac{R_2 - L_1 - z}{S_q}, F(z) = 1 - \frac{(R_2 - L_1 - z)^2}{2S_q}.$$

Для $R_1 - L_1 > R_2 - L_2$ (область II):
если $L_2 - R_1 < z < R_2 - R_1$, то

$$f(z) = \frac{z + R_1 - L_2}{S_q}, F(z) = \frac{(z + R_1 - L_2)^2}{2S_q};$$

если $R_2 - R_1 < z < L_2 - L_1$, то

$$f(z) = \frac{R_2 - L_2}{S_q}, F(z) = \frac{R_2 - L_2}{2S_q} [-R_2 - L_2 + 2(z + R_1)];$$

если $L_2 - L_1 < z < R_2 - L_1$, то

$$f(z) = \frac{R_2 - L_1 - z}{S_q}, F(z) = 1 - \frac{(R_2 - L_1 - z)^2}{2S_q}.$$

Таким образом, распределение разности двух равномерных распределений является трапециевидальным распределением (см. пример на рис. 2). Это распределение симметрично, его математическое ожидание $E(Z) = (R_2 + L_2 - R_1 - L_1)/2$, $F(E(Z)) = 1/2$.

В соответствии с общим подходом к расчету коэффициента уверенности в рамках точного метода при равномерных распределениях на сравниваемых интервальных оценках имеем для него: $K_{as} = F(z > 0) - F(z < 0)$.

Теперь можно проанализировать результаты использования этих формул для сравнения интервальных оценок. Отметим, что существует возможность «быстрого» сравнения, основанного на свойствах математического ожидания $E(Z)$ трапециевидального распределения. Именно, $E(Z)$ может быть положительным и отрицательным. Если $E(Z) < 0$, то можно заранее сказать, что I_2 не может быть предпочтительнее, чем I_1 , поскольку $F(Z < E(Z)) = 1/2$, $K_{as} = 0$

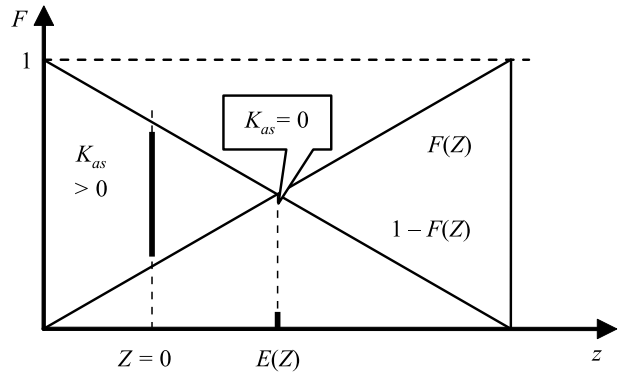


Рис. 3а. K_{as} для $E(Z) > 0$

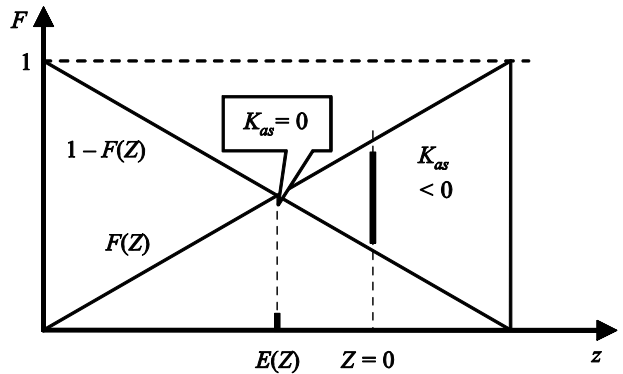


Рис. 3б. K_{as} для $E(Z) < 0$

для этой точки и, следовательно, переход в область $z > 0$ значений z приводит к $K_{as} < 0$. Эти ситуации показаны на рис. 3а ($E(Z) > 0$) и 3б ($E(Z) < 0$).

Аналогичные утверждения справедливы для произвольного, а не только для равномерного распределения, если заменить ожидаемое значение $E(Z)$ медианой Me , поскольку для нее $F(Z < Me) = 1/2$, $K_{as} = 0$.

Обсудим теперь конфигурации (b), (c) и (d) раздела 1 с точки зрения только что полученных результатов.

Конфигурация (d). Область интегрирования — квадрат, $E(Z) = 0$, $F(z < 0) = F(z > 0) = 1/2$, и в соответствии с определением коэффициента уверенности $K_{as} = 0$. В этом случае $K_{as} = (K_{as})_{int}$. Напомним, что $(K_{as})_{int}$ — это коэффициент уверенности эвристического метода.

Конфигурация (b). В этой конфигурации для обоих типов областей интегрирования (I и II) $F(z < 0) = (R_1 - L_2)^2 / (2S_q)$. Следовательно, $K_{as} = 1 - (R_1 - L_2)^2 / S_q = 1 - S_y^2 / S_q$. Можно показать, что $(K_{as})_{int} \leq K_{as}$ (равенство лишь при $L_1 = L_2$). Это означает, что если признать $I_2 \succ I_1$ с коэффициентом уверенности $(K_{as})_{int}$ по эвристическому методу, это заключение будет тем более верным в случае использования точного метода сравнения. Заметим также, что для этой конфигу-

рации значение $E(Z)$ всегда положительно. Итак, можно использовать $(K_{as})_{int}$ для экспресс-анализа и K_{as} для точных вычислений.

Конфигурация (с). Здесь $R_1 - L_1 > R_2 - L_2$. Это означает, что в этой конфигурации имеется только область интегрирования, вытянутая вправо, а точка $z = 0$ принадлежит подобласти 2, поскольку $R_2 - R_1 < 0$, $L_2 - L_1 > 0$. Тогда $F(Z < 0) = (R_2 - L_2)(2R_1 - R_2 - L_2)/(2S_q)$ и, следовательно, $K_{as} = (R_2 - L_2)(R_2 + L_2 - R_1 - L_1)/S_q$. Или $K_{as} = [S_g(= L_2 - L_1) - S_r(= R_1 - R_2)]/(R_1 - L_1) = (K_{as})_{int}$.

Заметим, что формулы этого раздела для $Z = I_2 - I_1$ преобразуются в формулы для $Z = I_1 - I_2$ при перестановке индексов $1 \leftrightarrow 2$ и областей интегрирования (прямоугольник, вытянутый вверх, заменяется на прямоугольник, вытянутый вправо, и наоборот).

4. Некоторые следствия использования метода точного сравнения

В разделе 1 мы отмечали, что ЛПР может интерпретироваться оценкой шансов того, что I_2 окажется больше, чем I_1 , на $\delta > 0$. Это означает, что необходимо вычислить величину коэффициента уверенности $K_{as}(\delta) = F(z > \delta) - F(z < \delta)$ и сравнить полученное значение с предпочтениями ЛПР. Если проверяется истинность гипотезы $I_2 \succ I_1$, то это имеет смысл, как это ясно из предыдущего (см. рис. 3а, 3б), только для $E(Z) > 0$ и $\delta < E(Z)$. Если мы имеем дело не с равномерным распределением операндов, то это утверждение имеет место для значений δ , для которых $F(z > \delta) > 1/2$.

Отметим в этой связи следующее обстоятельство. Представляется, что для сравнения двух стохастических переменных X_1 и X_2 с известными функциями распределения можно сравнить их значения x_1, x_2 , отвечающие выбранному ЛПР уровню вероятности $F_1(X_1 > x_1) = F_2(X_2 > x_2)$. Однако, как мы только что видели, при достаточно большом различии этих значений уверенность в реализуемости такой разницы может быть недостаточной.

Приведем простой пример, иллюстрирующий это утверждение.

Предположим, что имеются два равномерных распределения на интервалах $[L_1 = 5, R_1 = 18]$, $[L_2 = 10, R_2 = 20]$, и ЛПР выбрал для сравнения уровень вероятности, равный $2/3$. Тогда $x_1 = 28/3$ и $x_2 = 40/3$, т. е. $\delta = 4$. Поскольку $E(X_1) = 11,5$ и $E(X_2) = 15$, то $E(Z = X_2 - X_1) = 3,5 < 4$. Тогда заранее ясно, что $K_{as}(4) < 0$. Вычисления показывают, что $F(Z > 4) = 0,46$, $F(Z < 4) = 0,54$ и, таким образом, $K_{as}(4) = -0,08$. Предельно допустимое значение разности с нулевым коэффициентом уверенности равно в этом примере $3,5$, $K_{as}(3,5) = 0$.

Аналогичные соображения могут быть использованы при анализе результатов сравнения интервалов по критерию Гурвица.

Но что произойдет, если хотя бы одна оценка в сравниваемой паре окажется чисто интервальной, и мы не хотим следовать предположению о равномерности распределений? В этой ситуации нельзя получить точных результатов, однако можно получить некоторые приближенные оценки, исходящие из пессимистических предположений для I_2 и оптимистических для I_1 . Примером использования такого подхода служат сравнения пар интервалов, проведенные в табл. 2.

Напомним, что в процессе сравнения предполагалось, что неизвестные, но унимодальные распределения на интервалах были аппроксимированы треугольными распределениями с различными положениями мод. При этом мода была значительно смещена вправо для I_1 и влево для I_2 . Сдвигая на одинаковые расстояния положения мод от их первоначальных позиций вправо для I_2 и влево для I_1 , можно проанализировать зависимость K_{as} от локализации мод и, вместе с ЛПР, найти их значения, которые согласуются с K_{th} . Результаты подобного моделирования могут помочь ЛПР осмыслить свои предпочтения при принятии решений.

5. О сравнении обобщенных интегральных оценок

Приведем вначале пример еще одной ситуации, в которой может возникнуть задача сравнения полиинтервальных и обобщенных интервальных оценок. Пусть имеются два инвестиционных проекта, эффективность которых сравнивается на основе значения их чистого дисконтированного дохода (ЧДД). Если, из-за неопределенности, цены и доходы каждого года в периоды реализации проектов получили интервальные оценки, то ЧДД также будут иметь интервальные представления. Если в процедуре оценивания цен и доходов участвуют по два эксперта, то для каждого ЧДД будут получены по две интервальные оценки. Число экспертов и, следовательно, количество сценариев может быть больше, чем два. Между заданными экспертами «реперными» сценариями могут находиться также промежуточные, не заданные явно, сценарии. Чтобы формализовать эту картину, на основании экспертных сценариев можно построить ПИО для каждого ЧДД и, чтобы сравнить эффективность проектов, следует сравнить ПИО их ЧДД.

Итак, рассматривается задача сравнения двух полиинтервальных оценок PI_1 и PI_2 . ПИО задаются их

угловыми точками (вершинами) D_{ild} , D_{ird} , D_{ilu} , D_{iru} , $i = 1, 2$ и боковыми границами криволинейных трапеций (см. рис 1). Как и ранее, в терминах коэффициента уверенности находятся условия, при которых P_2 может быть предпочтительнее P_1 .

Вначале обсудим задачу сравнения в предположении о равномерности распределений на обеих осях ПИО. В этом случае наиболее просто рассчитать коэффициент уверенности, используя для каждой ПИО формулу (2). Мы получаем при этом обобщенные равномерные распределения, определенные на базовых интервалах каждой ПИО. Тем самым задача сведена к моноинтервальной и может быть решена методами раздела 1 и стохастического моделирования. Наглядный геометрический подход, подобный эвристическому методу прямого сравнения монооценок, в случае полиоценок приводит к существенно более сложным соотношениям и требует дополнительного рассмотрения.

В общем случае сравнения ОИО с произвольными распределениями на осях ПИО следует использовать методы стохастического моделирования. Если необходимо получить экспресс-оценки, полезно привлечение метода прямого сравнения монослучая. При этом вначале интервалы-сценарии для сравниваемых ПИО выбираются методом Монте-Карло, а потом выбранные монооценки сравниваются методами прямого сравнения.

Заключение

В работе предложен общий подход к сравнению пары альтернатив, описанных различными модифи-

кациями интервальных оценок. Получены формулы для функции распределения вероятности разности двух равномерно распределенных величин. Предложенный аппарат является удобным инструментарием для анализа отношений на парах интервальных оценок. Этот метод будет развиваться нами для других комбинаций распределения вероятностей операндов, а не только для равномерных распределений.

Литература

1. Ащепков Л. Т., Давыдов Д. В. Показатель интервального неравенства: свойства и применение // Вычислительные технологии. 2006. Т. 11. № 4. С. 13–22.
2. Стернин М. Ю., Чугунов Н. В., Шепелев Г. И. Обобщенные интервальные оценки в моделях предметных областей систем поддержки экспертных решений // Труды ИСА РАН. 2005. Т. 12. С. 95–113.
3. Стернин М. Ю., Чугунов Н. В., Шепелев Г. И. Учет неопределенности экспертных знаний: синтез интервального и вероятностного подходов // Информационные технологии и вычислительные системы. 2005. № 4. С. 36–46.
4. Стернин М. Ю., Шепелев Г. И. Анализ сценариев в методе обобщенных интервальных оценок // Таврический вестник информатики и математики. 2008. № 2. С. 195–201.
5. Стернин М. Ю., Шепелев Г. И. Обобщенные интервальные экспертные оценки в принятии решений // Доклады академии наук. 2010. Т. 432. № 1. С. 33–34.
6. Chugunov N., Shepelyov G., Sternin M. The generalized interval estimations in decision making under uncertainty // International Journal on Technology, Policy and Management. 2008. Vol. 8. P. 298–321.

Стернин Михаил Юрьевич. С. н. с. ИСА РАН. Окончил МИРЭА в 1970 г. Количество печатных работ: более 60. Область научных интересов: математическое моделирование, системы поддержки принятия решений, системы, основанные на знаниях. E-mail: mister@isa.ru

Шепелев Геннадий Иванович. Заведующий лабораторией ИСА РАН. К. ф.-м. н., с. н. с. Окончил МГУ в 1965 г. Количество печатных работ: более 70. Область научных интересов: системный анализ, математическое моделирование, методы принятия решений. E-mail: gis@isa.ru