О ПРИМЕНИМОСТИ ОЦЕНКО МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ ПРИ СРАВНЕНИИ ИНТЕРВАЛЬНЫХ АЛЬТЕРНАТИВ

Стернин М.Ю., Шепелев Г.И.

***Аннотация**: В практических задачах сравнения альтернатив с интервальными, из-за неопределенности, показателями качества нередко применяются точечные (одно-числовые) оценки, предположительно эквивалентные исходным интервальным. В работе показано, что использование в таких задачах точечных оценок, в частности оценок математического ожидания, без квантификации и дополнительного анализа величины шансов истинности проверяемой гипотезы о предпочтительности одной из альтернатив в их паре неадекватно проблеме сравнения интервальных оценок. Предложен метод вычисления шансов истинности проверяемой гипотезы и риска, связанного с возможной истинностью гипотезы, противоположной анализируемой. Для ряда примеров сравнения интервальных альтернатив проведено сопоставление результатов сравнения по предложенному методу и по средним оценкам. Продемонстрировано, что упомянутые шансы далеко не во всех случаях могут быть представлены как функции математических ожиданий распределений на сравниваемых интервалах. Показано, что корректный анализ задачи сравнения интервальных альтернатив требует расчета указанных шансов и сопоставления их величины с предпочтениями лиц, принимающих решения.*

***Keywords**: comparing interval alternatives, method for estimation of chances of preferability compared alternatives, analysis of adequacy of average point estimates to comparing process*

***ACM Classification Keywords**: H.1.2 Human information processing. G3 Distribution functions. I.2.3 Uncertainty, “fuzzy”, and probabilistic reasoning.*

Введение

Сравнение интервальных альтернатив, - альтернатив, показатели качества которых имеют, из-за неопределенности, интервальное представление, достаточно распространенная на практике задача. Возникновение неопределенности часто связано здесь с необходимостью прогнозирования будущих значений анализируемых показателей. Цель такого сравнения дать заключение о предпочтительности какой-либо альтернативы или об эквивалентности альтернатив на момент прогнозирования, т.е. еще в условиях неопределенности. Обычно предполагается, что в интервале [*L*, *R*] (*L* и *R* левая и правая граница интервальной оценки соответственно) лежат все возможные точечные (одно-числовые) значения анализируемого показателя, но лишь единственное из них реализуется в будущем, когда исходная неопределенность будет устранена.

При сравнении интервальных оценок показателей качества альтернатив существует единственная конфигурация расположения пары сравниваемых интервалов, которая не вызывает трудностей. Это конфигурация, при которой левая граница одного, пусть второго, интервала не менее, чем правая граница первого. Тогда один из интервалов строго предпочтительнее другого. Именно, второй интервал «лучше» первого, если большему значению показателя качества соответствует более предпочтительное состояние, и «хуже» первого, если меньшему значению показателя качества соответствует более предпочтительное состояние.

Для конфигураций общего положения, когда сравниваемые интервальные оценки имеют ненулевое пересечение, в принципе нельзя с определенностью сделать вывод о предпочтительности какой-либо интервальной альтернативы в их паре, - любая из них может оказаться таковой в будущем. Поэтому на момент сравнения можно судить лишь о шансах на то, что одна альтернатива окажется предпочтительнее другой. При этом всегда существует риск того, что в действительности предпочтительной окажется другая альтернатива.

Таким образом, задача сравнения интервальных оценок не может быть решена чисто математическими методами и превращается в задачу принятия решений, поскольку для вывода о том, достаточна ли оцененная тем или иным способом величина указанных шансов для заключения о предпочтительности одного из интервалов, приходиться привлекать лицо, принимающее решение (ЛПР), или эксперта. Именно их оценка приемлемости упомянутых шансов в конечном счете позволяет принять гипотезу о предпочтительности одной из альтернатив или отклонить ее.

Из всех способов квантификации шансов на предпочтительность той или иной интервальной оценки в настоящей работе выбран аппарат функций распределения теории вероятностей, в наибольшей степени привычный, по нашему мнению, для экспертов, что существенно, поскольку параметры распределения должны быть заданы специалистами. Конечно, для чисто интервальных альтернатив точное распределение шансов на них неизвестно. Но мы можем предположить, что эти распределения унимодальные, и для многих типов распределений можно приближенно аппроксимировать эти неизвестные (но унимодальные) распределения треугольными распределениями. Указанная гипотеза о правомочности такой приближенной замены интервальных альтернатив интервально-вероятностными принята в настоящей работе. Некоторые соображения, связанные с принятием этой гипотезы, обсуждаются ниже.

В наших предыдущих публикациях предложен общий подход к сравнению интервально-вероятностных альтернатив [Стернин, 2011; 2012; Shepelyov, 2011]. Он состоит в следующем.

Пусть для определенности рассматривается ситуация, когда бóльшие значения показателя качества отвечают более предпочтительному состоянию. Пусть проверяется гипотеза о том, что вторая интервальная альтернатива I_2 предпочтительнее («больше») первой I_1. Если шансы появления различных точечных значений в сравниваемой паре интервалов заданы как распределения вероятностей, численное моделирование процесса реализации точечных оценок может служить основой для оценки шансов предпочтительности интервальных величин и их сравнения. 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 не благоприятствует ей.

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. Будем считать, что параметр $K_g = N_g/N$ служит 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. Видно, что $K_g + K_r = 1$. Таким образом, K_g – оценка вероятности того, что второй интервал предпочтительнее первого ($I_2 \succ I_1$), а K_r того, что $I_1 \succ I_2$.

Введем коэффициент уверенности K_{as} как разность между K_g и K_r: $K_{as}(I_2 \succ I_1) = K_g - K_r$. Ясно, что $K_{as}(I_2 \succ I_1) = P(I_2 \succ I_1) - P(I_1 \succ I_2)$. Так как $P(I_2 \succ I_1) + P(I_1 \succ I_2) = 1$, то $Kas(I_2 \succ I_1) = 2P(I_2 \succ I_1) - 1 = 1 -$

$2P(I_1 \succ I_2)$. $P(I_2 \succ I_1) = [1 + Kas(I_2 \succ I_1)]/2$, $P(I_1 \succ I_2) = [1 + Kas(I_2 \succ I_1)]/2$. Какой из формул следует пользоваться при расчете коэффициента уверенности, зависит, как мы увидим, от специфики конкретной конфигурации пар сравниваемых интервальных альтернатив.

Коэффициент уверенности показывает, насколько шансы реализации зеленой зоны превышают шансы реализации красной зоны. Другими словами, насколько шансы истинности проверяемой гипотезы о предпочтительности одной из альтернатив превышают шансы истинности противоположной гипотезы. 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 вычисленное 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. Можно, например, обсудить с экспертом целесообразность изменения распределения вероятностей, используемого для описания неопределенности, или порогового значения коэффициента уверенности K_{th}. Если значение K_{as} отрицательно, необходимо проверить противоположную гипотезу. Отметим, что привлекательные для ЛПР значения коэффициента уверенности начинаются, по нашему мнению, с $K_{as} = 0.4$. При этом шансы на предпочтительность второй альтернативы при сравнении с первой оказываются равными 0.7, что в два с лишним раза выше, чем шансы противоположной гипотезы. Конечно, приемлемый пороговый уровень K_{th} свой для каждого ЛПР и зависит от его склонности к риску.

В работе [Shepelyov, 2011] изучены некоторые свойства коэффициента уверенности. Пусть, как и раньше, $K_{as}(I_2 \succ I_1)$ - коэффициент уверенности, полученный при проверке гипотезы о том, что альтернатива I_2 более предпочтительна, чем I_1. Возможные значения коэффициента уверенности $K_{as}(I_2 \succ I_1)$ лежат в диапазоне [- 1, 1]: $-1 \le K_{as}(I_2 \succ I_1) \le 1$. Отметим, что условие $K_{as}(I_2 \succ I_1) = 1$ соответствует ситуации, когда I_2 доминирует I_1 (интервалы не пересекаются), а $K_{as}(I_2 \succ I_1) = -1$ отвечает противоположной ситуации. Условие $K_{as}(I_2 \succ I_1) = 0$ соответствует ситуации равнозначности интервальных альтернатив по предпочтению. Как функция двух переменных $K_{as}(I_2 \succ I_1)$ является антисимметричной функцией: $K_{as}(I_2 \succ I_1) = -K_{as}(I_1 \succ I_2)$. Кроме того, определение K_{as} согласуется с требованием транзитивности отношения предпочтительности. Действительно, если $K_{as}(I_2 \succ I_1) > 0$, тогда (теоретически) $I_2 \succ I_1$; если $K_{as}(I_3 \succ I_2) > 0$, то $I_3 \succ I_2$, и, поскольку, в силу транзитивности, $I_3 \succ I_1$, то $K_{as}(I_3, \succ I_1) > 0$.

Напомним, что в процессе сравнения предполагалось, что неизвестные, но унимодальные, распределения на интервалах могут быть аппроксимированы треугольными распределениями с различными положениями мод. При этом мода может быть значительно смещена вправо для I_1 (оптимистическая гипотеза) и влево для I_2 (пессимистическая гипотеза). 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.

Более полную, по сравнению со значением коэффициента уверенности, информацию для принятия решений содержит распределение вероятностей разности D случайных величин I_2 и I_1, получаемое в процессе численного моделирования совокупности величин $i_2 - i_1$. Ясно, что $K_{as} = P(d > 0) - P(d < 0)$, где $P(d)$ – функция распределения случайной величины D. Подход, основанный на соотношениях для функций распределения вероятностей разности сравниваемых интервально-вероятностных величин, развит нами в предыдущих работах [Стернин, 2011; 2012; Shepelyov, 2011] и реализован аналитически для случая, когда на обоих сопоставляемых интервалах заданы равномерные распределения. Определенным недостатком этого подхода является известная его громоздкость, приводящая к большому количеству формул, каждая из которых применима в своей области изменения переменной D. Это связано как с вариабельностью формы области интегрирования при расчете плотности распределения вероятностей разности, так и полиморфизмом выражений для некоторых функций распределения, задаваемых на сравниваемых интервалах, например, для треугольных распределений. Поэтому такой, реализуемый в принципе подход, приводит к соотношениям, уместным скорее в справочниках, чем в журнальных публикациях. В связи с этим нами предложен другой подход к получению аналитических выражений для коэффициентов уверенности, изложенный в следующем разделе статьи.

Имеются, однако, и другие, более простые на первый взгляд методы сравнения интервальных величин. Они основаны на замещении интервальных оценок точечными, предположительно эквивалентными при сравнении исходным интервальным. Эти методы подкупают не только своей простотой, но и тем фактом, что, как отмечают эксперты, руководители-практики привыкли работать и предпочитают иметь дело с точечными оценками, в том числе для будущих значений сравниваемых показателей.

Простота и наглядность точечных оценок скрывает некоторые их недостатки. Складывается впечатление, что на их основе могут быть приняты исчерпывающие, при всех условиях верные решения о предпочтительности альтернатив, что, как мы видели, не всегда так. Не упоминается, что в ряде случаев эти оценки являются лишь границами некоторых суженных, по сравнению с исходными, интервалов. Кроме того, эти оценки не сопровождаются оценками шансов истинности проверяемой гипотезы о предпочтительности и риска истинности противоположной гипотезы.

При сравнении интервально-вероятностных альтернатив, которые главным образом рассматриваются в настоящей статье, наиболее употребителен на практике критерий математического ожидания показателей качества альтернатив [Смоляк, 1996]. Поэтому анализ адекватности критерия математического ожидания задачам сравнения интервально-вероятностных альтернатив и сопоставление результатов сравнения по предложенному в статье методу и по средним оценкам – также одна из целей настоящей работы.

Аналитический метод расчета коэффициентов уверенности

Для некоторых распределений вероятностей на сравниваемых интервалах могут быть найдены аналитические выражения для коэффициентов уверенности в истинности проверяемой гипотезы о предпочтительности одного из интервалов в их паре, что позволяет выполнить более общий, по сравнению с численным моделированием, анализ задачи. Это может быть сделано, в частности, при задании равномерных и треугольных распределений на интервалах в различных комбинациях.

Рассматриваются конфигурации: A) $L_2 > R_1$; B) $L_2 < R_1 < R_2$; C) $L_1 < L_2$, $R_2 < R_1$; D) $L_1 = L_2 = L$, $R_1 = R_2 = R$, которыми, с точностью до перестановки, исчерпываются все возможные конфигурации сравниваемых пар интервальных оценок. Таким образом, в конфигурации A) интервалы не пересекаются, случай B) – это конфигурация сдвинутых относительно друг друга интервалов с пересечением, C) – вложенные интервалы и D) – совпадающие интервалы.

Идея аналитического метода расчета коэффициентов уверенности $Kas(I_2 \succ I_1)$ фактически повторяет идею численного метода: для случайно выбранного фиксированного значения $i_{2f} \in I_2$ рассчитывается вероятность $P(i_{2f}>i_1)$, $i_1 \in I_1$. Затем такие вероятности для всех возможных i_2 суммируются с учетом шансов на реализацию каждого такого i_2. Другими словами, должен быть вычислен интеграл от $P(i_2>i_1)$ по случайной величине i_2 с учетом плотности распределения ее вероятностей $f(i_2)$.

Рассмотрим вначале случай равномерных распределений на обоих интервалах для всех нетривиальных (*D*, *C*, *B*) конфигураций пар сравниваемых интервалов.

Конфигурация *D*: $P(i_{2f}>i_1) = (i_{2f} - L)/S$, где $S = R - L$. $P(I_2 \succ I_1) = \frac{1}{S^2}\int_L^R dz(z-L) = \frac{1}{2}.$

Таким образом, в этом случае $Kas(I_2 \succ I_1) = 0$.

Конфигурация *C*: $P(i_{2f}>i_1) = P(L_1< i_1 < L_2) + P(L_2 < i_1 < i_{2f}) = (i_{2f} - L_1)/S_1$, где $S_1 = R_1 - L_1$. Поэтому ($S_2 = R_2 - L_2$)

$$P(I_2 \succ I_1) = \frac{1}{S_1 S_2}\int_{L_2}^{R_2} dz(z-L_1) = \frac{R_2+L_2-2L_1}{2S_1}.$$

Отсюда $Kas(I_2 \succ I_1) = (R_2 - R_1 + L_2 - L_1)/S_1 = 2(\langle I_2\rangle - \langle I_1\rangle)/S_1$, где $\langle I_2\rangle$ и $\langle I_1\rangle$ математические ожидания случайных величин на интервалах I_2 и I_1 соответственно: $\langle I_i\rangle = (R_i + L_i)/2$. Таким образом, в случае вложенных интервалов при равномерных распределениях вероятностей на них нормированная разность математических ожиданий (при правильно выбранной нормировке) совпадает с шансами предпочтительности проверяемой гипотезы.

Конфигурация B: Здесь целесообразнее использовать соотношение $Kas(I_2 \succ I_1) = 1 - 2P(I_1 \succ I_2)$. $P(i_1>i_{2f}) =$ $= (R_1 - i_{2f})/S_1$, и

$$Kas(I_2 \succ I_1) = 1 - \frac{2}{S_1 S_2}\int_{L_2}^{R_1} dz(R_1 - z) = 1 - \frac{(R_1-L_2)^2}{S_1 S_2}.$$

Здесь шансы предпочтительности второго интервала не выражаются через средние величины. Для этого случая коэффициент уверенности всегда положителен (второй интервал теоретически предпочтительнее первого), поскольку его числитель равен $R_1(R_2 - R_1) + L_2(R_1 - L_2) - L_1(R_2 - L_2)$, его минимальное значение достигается при $L_1 = L_2$ и равно положительной для этой конфигурации величине $(R_2 - R_1)(R_1 - L_2)$.

Пусть теперь на первом интервале задано равномерное распределение, а на втором треугольное. Вспомним, что плотность треугольного распределения вероятностей имеет вид (*M* – мода распределения)

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

Для конфигурации *D*: $P(i_{2f}>i_1) = (i_{2f} - L)/S$,

$$P(I_2 \succ I_1) = \frac{2}{S^2}[\int_L^M dz(z-L)^2/(M-L) + \int_M^R dz(R-z)(z-L)/(R-M)] = \frac{M+R-2L}{3S}.$$

Тогда $Kas(I_2 \succ I_1) = (2M- L- R)/(3S) = 2(\langle I_2\rangle - \langle I_1\rangle)/S$, $\langle I_2\rangle = (L + M + R)/3$. Если определить положение моды соотношением $M = (1 - h)L + hR$, то $Kas(I_2 \succ I_1) = (2h - 1)/3$ и зависит только от положения моды треугольного распределения. Вновь величина искомых шансов может быть (при правильно выбранной нормировке) выражена через разность средних величин.

Конфигурация *C*: Здесь

$$P(I_2 \succ I_1) = 2[\int_{L_2}^M dz\frac{(z-L_2)(z-L_1)}{M-L_2} + \int_M^{R_2} dz\frac{(R_2-z)(z-L_1)}{R_2-M}]/(S_1 S_2).$$

Отсюда следует, что $Kas(I_2 \succ I_1) = 2[(M + R_2 + L_2)/3 - (R_1 + L_1)/2]/S_1 = 2(\langle I_2\rangle - \langle I_1\rangle)/S_1$.

Конфигурация *B*: Вновь, как и выше, $P(i_1>i_{2f}) = (R_1 - i_{2f})/S_1$. Необходимо различать случаи $M \ge R_1$ и $M < R_1$. В первом случае при интегрировании по всем возможным i_2 присутствует только левая ветвь плотности вероятности треугольного распределения, а во втором обе ветви. Тогда, при $M \ge R_1$

$$Kas(I_2 \succ I_1) = 1 - \frac{4}{S_1S_2}\int_{L_2}^{R_1} dz \frac{(R_1 - z)(z - L_2)}{M - L_2} = 1 - \frac{2(R_1 - L_2)^3}{3S_1S_2(M - L_2)}$$

А при $M < R_1$

$$P(I_2 \prec I_1) = \frac{2}{S_1S_2}[\int_{L_2}^{M} dz \frac{(R_1 - z)(z - L_2)}{M - L_2} + \int_{M}^{R_1} dz \frac{(R_1 - z)(R_2 - z)}{R_2 - M},$$

и, следовательно,

$$K(I_2 \succ I_1) = 1 - 2P(I_2 \prec I_1) = 1 - \frac{2}{3S_1S_2}[(M - L_2)(3R_1 - 2M - L_2) + \frac{(R_1 - M)^2}{R_2 - M}(3R_2 - 2M - R_1)].$$

Видно, что здесь величина коэффициента уверенности не равна с точностью до нормировки разности соответствующих средних.

Пусть теперь на обоих сравниваемых интервалах заданы треугольные распределения.

Конфигурация *D*. Здесь при сравнении фиксированных значений i_{2f} со значениями i_1 первые могут быть больше моды треугольного распределения M_1, заданного на I_1, или меньше его: $i_{2f} \geq M_1$ или $i_{2f} < M_1$. При $i_{2f} < M_1$ $P(i_{2f}>i_1) = (i_{2f} - L)/[S(M_1 - L)]$, а при $i_{2f} \geq M_1$ $P(i_{2f}>i_1) = 1 - (R - i_{2f})^2/[S(R - M_1)]$. Возможны два расположения мод распределений на сравниваемых интервалах: $M_1 \geq M_2$ или $M_1 < M_2$. Рассмотрим случай первого расположения мод. Проходя слева направо по I_2, i_2 обегает пути 1: $M_1 > M_2 > i_2$; 2: $M_1 > i_2 > M_2$ и, наконец, 3: $i_2 > M_1 > M_2$. Тогда для пути 1 имеем (напомним, что $M_1 \geq M_2$):

$$P(I_2 \succ I_1) = \frac{2}{S^2}\int_{L}^{M_2} dz \frac{(z - L)^3}{(M_1 - L)(M_2 - L)} = \frac{(M_2 - L)^3}{2S^2(M_1 - L)} = F_{1D}.$$

Для пути 2:

$$P(I_2 \succ I_1) = \frac{2}{S^2}\int_{M_2}^{M_1} dz \frac{(z - L)^2(R - z)}{(M_1 - L)(R - M_2)} = \frac{2}{S^2(M_1 - L)(R - M_2)}[\frac{(M_1^3 - M_2^3)(R + 2L)}{3} + L^2R(M_1 - M_2)$$
$$- \frac{(M_1^2 - M_2^2)(L^2 + 2RL)}{2} - \frac{M_1^4 - M_2^4}{4}] = F_{2D}.$$

Для пути 3:

$$P(I_2 \succ I_1) = \frac{2}{S}\int_{M_1}^{R} dz[1 - \frac{(R - z)^2}{S(R - M_1)}]\frac{R - z}{R - M_2} = \frac{(R - M_1)^2}{S(R - M_2)}[1 - \frac{R - M_1}{2S}] = F_{3D}.$$

Следовательно, $P(I_2 \succ I_1) = F_{1D} + F_{2D} + F_{3D}$, а $Kas(I_2 \succ I_1) = 2P(I_2 \succ I_1) - 1$.

Соответствующие соотношения для случая $M_1 < M_2$ следуют из того факта, что для этой конфигурации $Kas(I_2 \succ I_1|M_1 = A, M_2 = B) = -Kas(I_2 \succ I_1|M_1 = B, M_2 = A)$, $A \geq B$. И $P(I_2 \succ I_1|M_1 = A, M_2 = B) + P(I_2 \succ I_1|M_1 = B, M_2 = A) = 1$.

Конфигурация *C*. Здесь $P(I_2 \succ I_1) = P(L_1 < i_1 < L_2) + P(i_2 > i_1|i_1, i_2 \in [L_2, R_2])$. Необходимо различать случаи различного положения мод распределений. Пусть вначале $M_1 < L_2$. Тогда $P(L_1 < i_1 < L_2) = 1 - (R_1 - L_2)^2/[S_1(R_1 - M_1)]$. А на интервале $[L_2, R_2]$ $P(i_{2f}>i_1) = [(R_1 - L_2)^2 - (R_1 - i_{2f})^2]/[S_1(R_1 - M_1)]$. Это означает, что (при $i_{2f} = z$)

$$P(i_2 > i_1 \,|\, i_1, i_2 \in [L_2, R_2]) = \frac{2}{S_2(M_2 - L_2)}\int_{L_2}^{M_2} dz P(z > i_1)(z - L_2) + \frac{2}{S_2(R_2 - M_2)}\int_{M_2}^{R_2} dz P(z > i_1)(R_2 - z)$$

И, таким образом, при $M_1 < L_2$ имеем:

$$K(I_2 \succ I_1) = 1 - \frac{6R_1^2 + L_2^2 + R_2^2 + L_2R_2 - (M_2 + L_2 + R_2)(4R_1 - M_2)}{3S_1(R_1 - M_1)}.$$

При $M_1 > R_2$, действуя точно так же, получаем:

$$K(I_2 \succ I_1) = \frac{6L_1^2 + L_2^2 + R_2^2 + L_2R_2 - (M_2 + L_2 + R_2)(4L_1 - M_2)}{3S_1(M_1 - L_1)} - 1.$$

Пусть теперь $L_2 < M_1 < R_2$. Тогда $P(L_1 < i_1 < L_2) = (L_2 - L_1)^2/[S_1(M_1 - L_1)] = F_{1C}$. Следует различать случаи, когда $L_2 < M_1 < M_2$, и когда $M_2 < M_1$. При $L_2 < M_1 < M_2$ $P(i_{2f}>i_1|\ i_{2f} < M_1) = [(i_{2f} - L_1)^2 - (L_2 - L_1)^2]/[S_1(M_1 - L_1)] =$ $=G_1(i_{2f})$, $P(i_{2f}>i_1|\ i_{2f} > M_1) = 1 - (L_2 - L_1)^2]/[S_1(M_1 - L_1)] - (R_1 - i_{2f})^2]/[S_1(R_1 - M_1)] = G_2(i_{2f})$. Тогда

$$P(i_2 > i_1 \,|\, i_1, i_2 \in [L_2, R_2], M_1 > i_2) = \frac{2}{S_2(M_2 - L_2)} \int_{L_2}^{M_1} dz G_1(z)(z - L_2) = \frac{2}{S_1S_2(M_1 - L_1)(M_2 - L_2)} \circ$$
$$\circ [\frac{M_1^4 - L_2^4}{4} + \frac{(M_1^2 - L_2^2)(L_1^2 + 2L_1L_2) - (L_2 - L_1)^2(M_1 - L_2)^2}{2} - \frac{(M_1^3 - L_2^3)(L_2 + 2L_1)}{3} -$$
$$- L_1^2L_2(M_1 - L_2)] = F_{2C}.$$

$$P(i_2 > i_1 \,|\, i_i \in [L_2, R_2], M_2 > i_2 > M_1) = \frac{2}{S_2(M_2 - L_2)} \int_{M_1}^{M_2} dz G_2(z)(z - L_2) = \{[1 - \frac{(L_2 - L_1)^2}{S_1(M_1 - L_1)}] \circ$$
$$\circ \frac{(M_2 - L_2)^2 - (M_1 - L_2)^2}{2} - \frac{1}{S_1(R_1 - M_1)}[\frac{M_2^4 - M_1^4}{4} + \frac{(M_2^2 - M_1^2)(R_1^2 + 2R_1L_2)}{2} +$$
$$+ \frac{(M_1^3 - M_2^3)(L_2 + 2R_1)}{3} - R_1^2L_2(M_2 - M_1)]\}\frac{2}{S_2(M_2 - L_2)} = F_{3C}.$$

$$P(i_2 > i_1 \,|\, M_2 < i_2) = \frac{2}{S_2(R_2 - M_2)} \int_{M_2}^{R_2} dz G_2(z)(R_2 - z) = \frac{2}{S_2(M_2 - L_2)}\{\frac{(R_2 - M_2)^2}{2}[1 -$$
$$- \frac{(L_2 - L_1)^2}{S_1(M_1 - L_1)}] - [\frac{M_2^4 - R_2^4}{4} - \frac{(R_2^2 - M_2^2)(R_1^2 + 2R_1R_2)}{2} + \frac{(R_2^3 - M_2^3)(R_2 + 2R_1)}{3} +$$
$$+ R_1^2R_2(R_2 - M_1)]\frac{1}{S_1(R_1 - M_1)}\} = F_{4C}.$$

$P(I_2 \succ I_1) = F_{1C} + F_{2C} + F_{3C} + F_{4C}$. $Kas(I_2 \succ I_1) = 2P(I_2 \succ I_1) - 1$.

При $L_2 < M_2 < M_1$, действуя точно так же, получаем:

$$P(i_2 > i_1 \,|\, i_1, i_2 \in [L_2, R_2], M_1 > i_2) = \frac{2}{S_2(M_2 - L_2)} \int_{L_2}^{M_2} dz G_1(z)(z - L_2) =$$
$$= \frac{2}{S_1S_2(M_1 - L_1)(M_2 - L_2)}[\frac{M_2^4 - L_2^4}{4} + \frac{(M_2^2 - L_2^2)(L_1^2 + 2L_1L_2) - (L_2 - L_1)^2(M_2 - L_2)^2}{2} -$$
$$- \frac{(M_2^3 - L_2^3)(L_2 + 2L_1)}{3} - L_1^2L_2(M_2 - L_2)] = F_{5C}.$$

$$P(i_2 > i_1 \,|\, i_1, i_2 \in [L_2, R_2], M_1 > i_2 > M_2) = \frac{2}{S_2(R_2 - M_2)} \int_{M_2}^{M_1} dz G_1(z)(R_2 - z) =$$
$$= [\frac{(R_2 - M_1)^2 - (R_2 - M_2)^2}{2}(L_2 - L_1)^2 - \frac{M_1^4 - M_2^4}{4} + \frac{(M_2^2 - M_1^2)(L_1^2 + 2R_2L_1)}{2} +$$
$$+ \frac{(M_1^3 - M_2^3)(R_2 + 2L_1)}{3} + L_1^2R_2(M_1 - M_2)]\frac{2}{S_1S_2(R_2 - M_2)(M_1 - L_1)} = F_{6C}.$$

$$P(i_2 > i_1 \mid i_1, i_2 \in [L_2, R_2], M_1 < i_2) = \frac{2}{S_2(R_2 - M_2)} \int_{M_1}^{R_2} dz G_2(z)(R_2 - z) =$$

$$= \frac{(R_2 - M_1)^2}{S_2(R_2 - M_2)}[1 - \frac{(L_2 - L_1)^2}{S_1(M_1 - L_1)}] + \frac{1}{S_1 S_2 (R_1 - M_1)(R_2 - M_2)}[(R_2^2 - M_1^2)(R_1^2 + 2R_1 R_2) -$$

$$- 2R_2 R_1^2 (R_2 - M_1) + \frac{R_2^4 - M_1^4}{2} - \frac{2(R_2^3 - M_1^3)(R_2 + 2R_1)}{3}] = F_{7C}.$$

$P(I_2 \succ I_1) = F_{1C} + F_{5C} + F_{6C} + F_{7C}$. $Kas(I_2 \succ I_1) = 2P(I_2 \succ I_1) - 1$.

Конфигурация *B*. Пусть $M_1 < L_2$, а $M_2 > R_1$. Вновь, как и выше, целесообразнее использовать соотношение $Kas(I_2 \succ I_1) = 1 - 2P(I_1 \succ I_2)$. $P(i_1 > i_{2f} \mid i_1, i_2 \in [L_2, R_1]) = (R_1 - i_{2f})^2/[S_1(R_1 - M_1)]$. Используя ту же технику, что и ранее, получаем:

$$P(I_2 \prec I_1) = \frac{R_1 - L_2}{S_1 S_2 (R_1 - M_1)(M_2 - L_2)}[(R_1 + L_2)(R_1^2 + 2R_1 L_2) + \frac{(R_1 + L_2)(R_1^2 + L_2^2)}{2} - 2R_1^2 L_2 -$$
$$- \frac{2(R_1^2 + L_2^2 + R_1 L_2)(2R_1 + L_2)}{3}].$$

В случае $M_1 < L_2$, $M_2 < R_1$ имеем:

$$P(I_2 \prec I_1) = \frac{1}{S_1 S_2 (R_1 - M_1)}\{(M_2 + L_2)(R_1^2 + 2R_1 L_2) + \frac{(M_2 + L_2)(M_2^2 + L_2^2)}{2} - 2R_1^2 L_2 -$$
$$- \frac{2(M_2^2 + L_2^2 + M_2 L_2)(2R_1 + L_2)}{3} + \frac{R_1 - M_2}{R_2 - M_2}[2R_1^2 R_2 - (R_1 + M_2)(R_1^2 + 2R_1 R_2) - \frac{(R_1 + M_2)(R_1^2 + M_2^2)}{2} +$$
$$+ \frac{2(M_2^2 + R_1^2 + M_2 R_1)(2R_1 + R_2)}{3}]\}.$$

При $M_1 > L_2$, $M_2 > R_1$ $P(i_1 > i_{2f} \mid i_1, i_2 \in [L_2, R_1], i_{2f} < M_1) = 1 - (i_{2f} - L_1)^2/[S_1(M_1 - L_1)]$, $P(i_1 > i_{2f} \mid i_1, i_2 \in [L_2, R_1], i_{2f} > M_1) = (R_1 - i_{2f})^2/[S_1(R_1 - M_1)]$. Тогда

$$P(I_2 \prec I_1) = \frac{(M_1 - L_2)^2}{S_2(M_2 - L_2)} + \frac{1}{S_1 S_2 (M_2 - L_2)}\{(R_1 + M_1)(R_1^2 + 2R_1 L_2) + \frac{(R_1 + M_1)(R_1^2 + M_1^2)}{2} - 2R_1^2 L_2 -$$
$$- \frac{2(M_1^2 + R_1^2 + M_1 R_1)(2R_1 + L_2)}{3} + \frac{1}{M_1 - L_1}[2L_1^2 L_2 (M_1 - L_2) - (M_1^2 - L_2^2)(L_1^2 + 2L_1 L_2) - \frac{M_1^4 - L_2^4}{2} +$$
$$+ \frac{2(M_1^3 - L_2^3)(2L_1 + L_2)}{3}]\}.$$

Аналогичные формулы для случая $M_1 > L_2$, $M_1 < M_2 < R_1$ имеют вид:

$$P(I_2 \prec I_1) = \frac{(M_1 - L_2)^2}{S_2(M_2 - L_2)} + \frac{1}{S_1 S_2 (M_2 - L_2)}\{[2L_1^2 L_2 (M_1 - L_2) + \frac{2(M_1^3 - L_2^3)(2L_1 + L_2)}{3} - (M_1^2 - L_2^2)(2L_1 L$$
$$+ L_1^2) - \frac{M_1^4 - L_2^4}{2}]\frac{1}{M_1 - L_1} + [2R_1^2 L_2 (M_1 - M_2) - (M_1^2 - M_2^2)(R_1^2 + 2R_1 L_2) + \frac{M_2^4 - M_1^4}{2} +$$
$$+ \frac{2(M_1^3 - M_2^3)(2R_1 + L_2)}{3}]\frac{1}{R_1 - M_1}\} + \frac{1}{S_1 S_2 (R_1 - M_1)(R_2 - M_2)}[2R_1^2 R_2 (R_1 - M_2) + \frac{2(R_1^3 - M_2^3)(2R_1 + R}{3}$$
$$- (R_1^2 - M_2^2)(2R_1 R_2 + R_1^2) - \frac{R_1^4 - M_2^4}{2}]$$

Наконец, для случая $M_2 < M_1$ получаем:

$$P_1(I_2 \prec I_1) = 1 - \frac{(R_2 - M_1)^2}{S_2(R_2 - M_2)} - \frac{1}{S_1 S_2 (M_1 - L_1)}[\frac{(M_2 + L_2)(M_2^2 + L_2^2)}{2} -$$
$$-\frac{2(L_2 + 2L_1)(M_2^2 + L_2^2 + M_2 L_2)}{3} + (M_2 + L_2)(L_1^2 + 2L_1 L_2) - 2L_1^2 L_2] = F_{1B}.$$

$$P_2(I_2 \prec I_1) = \frac{1}{S_1 S_2 (M_1 - L_1)(R_2 - M_2)}[\frac{M_1^4 - M_2^4}{2} - \frac{2(M_1^3 - M_2^3)(R_2 + 2L_1)}{3} -$$
$$-2L_1^2 R_2 (M_1 - M_2) + (M_1^2 - M_2^2)(L_1^2 + 2L_1 R_2)] = F_{2B}.$$

$$P_3(I_2 \prec I_1) = \frac{1}{S_1 S_2 (R_1 - M_1)(R_2 - M_2)}[\frac{M_1^4 - R_1^4}{2} + \frac{2(R_1^3 - M_1^3)(R_2 + 2R_1)}{3} +$$
$$+2R_1^2 R_2 (R_1 - M_1) + (M_1^2 - R_1^2)(R_1^2 + 2R_1 R_2)] = F_{3B}.$$

И $Kas(I_2 \succ I_1) = 1 - 2P(I_1 \succ I_2)$.

Пример расчета коэффициента уверенности

В предыдущем разделе мы показали, что в ряде простых случаев (например, в случаях задания равномерных распределений на сопоставляемых интервалах, а в некоторых случаях и для комбинации равномерного и треугольного распределений) решения задачи сравнения интервальных альтернатив предлагаемым методом и на базе правильно нормированной разности средних для распределений вероятности совпадают. Ни в одном другом случае использование средних оценок не позволяет провести корректное сравнение интервальных альтернатив, поскольку при этом нельзя оценить риск принятия неверной гипотезы. Действительно, как видно из таблицы 1, сравнение по критерию разности математических ожиданий ΔAv правильно описывает тенденцию (знаки величин в столбцах ΔAv и $Kas(I_2 \succ I_1)$ совпадают), однако сделать заключение о шансах истинности проверяемой гипотезы $I_2 \succ I_1$ по этому критерию на наш взгляд невозможно. Не представляется возможным также подобрать единый нормирующий множитель для критерия ΔAv, который приводил бы к значениям искомых шансов. Нельзя отыскать и риск принятия ложной гипотезы, измеряемый вероятностью $P(I_1 \succ I_2)$, в то время как эта величина просто вычисляется по значениям коэффициента уверенности.

Таблица 1. Сравнение интервальных альтернатив для треугольных распределений

№ п/п	Интервал 1			Интервал 2				
	L_1	R_1	M_1	L_2	R_2	M_2	$Kas(I_2 \succ I_1)$	ΔAv
1	50	100	90	70	170	80	0.70	26.70
2	50	100	60	70	170	160	0.98	63.30
3	50	100	75	70	170	120	0.96	45.00
4	50	100	75	50	100	75	0.00	0.00
5	50	100	90	50	100	60	-0.49	-10.00
6	50	100	60	50	100	90	0.49	10.00
7	70	170	75	80	150	145	0.52	20.00
8	70	170	160	80	150	100	-0.59	-23.30
9	70	170	120	80	150	115	-0.15	-5.00

Заключение

При работе с интервальными оценками качества альтернатив лица, принимающие решения, в процессе решения задач выбора в условиях неопределенности часто предпочитают заменить неопределенную интервальную оценку детерминированной точечной. Помимо оценок математического ожидания при этом используются и другие точечные оценки, такие как коэффициенты «пессимизма – оптимизма» Гурвица [Hurwicz, 1951] и детерминированные эквиваленты теории ожидаемой полезности [Keeney, 1993]. Однако в силу неопределенности будущего исхода критериальная оценка не должна, как показано в настоящей работе, выражаться одним числом и, как следствие, всегда присутствует риск того, что принимаемое решение окажется неверным. Адекватный задаче метод сравнения должен позволять соизмерить шансы проверяемой гипотезы о предпочтительности и противоположной ей. В ряде отдельных простых случаев, как мы показали на примере оценок математического ожидания, разность точечных оценок при надлежащей нормировке действительно может быть использована для анализа шансов истинности гипотезы о предпочтительности одной из сравниваемых альтернатив. Так ли это для других типов точечных оценок еще предстоит проверить. Вместе с тем уже сейчас ясно, что в случае конфигураций пар сравниваемых альтернатив и распределений вероятностей на них общего вида необходимо использовать специализированные методы вычисления упомянутых шансов.

Благодарности

The paper is published with financial support by the project ITHEA XXI of the Institute of Information Theory and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS (www.aduis.com.ua).

Библиография

[Hurwicz, 1951] Hurwicz L. Optimality criteria for decision making under ignorance. 'Cowles Commission Discussion Paper', Statistics, # 370, New Haven. 1951.

[Keeney, 1993] Keeney R., Raiffa H. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. Cambridge University Press. 1993.

[Shepelyov, 2011] Shepelyov G., Sternin M. Methods for comparison of alternatives described by interval estimations// International Journal of Business Continuity and Risk Management. 2011. Vol. 2, No. 1. Pp.56-69.

[Смоляк, 2011] Смоляк С.А. О сравнении альтернатив со случайным эффектом. // Экономика и мат. методы. 1996. Т. 32. Вып. 4.

[Стернин, 2011] Стернин М.Ю, Шепелев Г.И. Сравнение интервальных альтернатив. // Труды Института системного анализа Российской академии наук. 2011. Т. 61, вып. 2. Сс. 7 – 11.

[Стернин, 2012] Стернин М.Ю., Шепелев Г.И. Оценка интервальных альтернатив: неопределенности и предпочтения. //International journal "Information models and analysis". 2012. V.1, No 4. Pp. 357 – 369.

Информация об Авторах

Михаил Стернин *– старший научный сотрудник Института системного анализа Российской академии наук, Россия, 117312, Москва, просп. 60-летия Октября, 9, ИСА РАН; e-mail: mister@isa.ru*

Геннадий Шепелёв *– заведующий лабораторией ИСА РАН; e-mail: gis@isa.ru*