RISK BEHAVIOUR IN A SET OF INTERVAL ALTERNATIVES

Gennady Shepelev

Abstract*: Problem comparing alternatives under interval uncertainty is studied. It is assumed that the compared alternatives have indicators of quality in the form of interval estimates. It is shown that mentioned problem cannot be unambiguously resolved by purely mathematical methods and requires using of decision maker's preferences. From two possible situations of comparing, situation of unique choice and situation of repeated one, the first situation, which is typical for problems of forecasting, is analyzed. Quantitative measure of plausibility of implementation for tested hypothesis about preference of an alternative in comparison with others in their set and measure of risk connected with possible falsity of such hypothesis are introduced. It is shown that this risk is increased with increasing number of compared alternatives. Some methods to calculate risks as well as procedure of decision-making in the framework of a set of alternatives are proposed.*

Keywords*: comparing interval alternatives, dependence of the risk on number of alternatives, probability logic in comparing interval alternatives, methods and procedures of decision-making for the problem.*

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

Introduction and motivation

Many important problems are analyzed under uncertainty. First of all these are problems of forecasting when experts have to deal with estimates of the future values of the problem parameters and indicators. In most practical cases these quantities are measured in numerical scales and have, due to uncertainty, interval representations. Such interval estimates are given quite often by experts or are received as resultants of some models with interval input data. Certain interval indicators may play a role of comparison criteria for choice problems if it is necessary to choose some objects (alternatives) from their set. We will call alternatives with interval values of comparison criteria (or, synonymous, interval quality indicators) interval alternatives.

These choice problems cannot be exhaustively solved by purely mathematical methods. Indeed, if there is a non-zero intersection of compared intervals generally impossible definitively to conclude in a choice process on superiority of an interval alternative over the others in their set. Any alternative may be “better” in the future, at the time of “removal” of uncertainty, when the interval estimates are replaced by exact (point) values of comparison criteria. Therefore at the time of the comparison can be judged only on the chances that one alternative will be preferable to others. It means that the problem comparing interval alternatives is a decision-making problem demanding including preferences of decision makers (DM) or experts in the process of decision-making. It should be emphasized that even after the selection of an alternative, which seems preferable at the moment of decision-making; always there is a risk that in the future any other alternative will be actually better. This is an essential feature of such problems. Therefore formal methods of comparison serve in this case only as a means of information-analytical support for the decision making process and cannot guarantee choice of truly the best alternative. Thus comparison criteria become in fact measure of preference.

So we suppose that all alternatives are comparable in preference (system of alternatives is full) and at the moment of choice a disjunction containing alternatives is not a strict disjunction (XOR), when the only alternative is choose as preferable. The choice of a preferred alternative depends now on the chances of its preference among the members of the disjunction. Such preference relations can be called relations involve risks or relations based on the degree of assurance in the truth of a testable hypothesis of preference. Apparatus of distribution functions has been selected here for quantification of preference chances for compared interval alternatives and the associated risks. Then the problem may be studied in the framework of probability logic [Nilsson, 1986] when besides truth and falsity analyzed logic expressions (opinions, assertions) may have in-between values of truth interpreted as chances of truth of the expressions.

It seems that the tool of distribution functions is the most familiar to experts-practitioners. This is important because expert analysis of practical problems is most productive if it is carried out on professional language that is familiar to experts, with using methods and terminology that are conventional for them. Thus knowledge concerning uncertainty is expressed in the framework of distribution functions approach just as in probability theory (but, that is typical for the many problems of expert analysis, beyond the frequency concept of probability).

To permit experts describe their knowledge concerning uncertainty of values within the interval estimates arbitrary distribution functions will be used. This is contrary to the view of some researchers [Diligensky, 2004] who believe that only uniform distributions are permitted on the “true” intervals. But if one take this concept, operations of interval analysis, to which is necessary to resort when working with models, become in fact impossible. Indeed, it was shown in [Sternin, 2011] that the distribution of the difference (sum) of two uniform distributions would be trapezoidal distribution. Thus applying already the simplest arithmetic operations to the initial intervals with uniform distributions on them does not allow recognizing the results of such operations as “true” interval if we take the requirement of equality of chances of implementation for all values in it (Gibbs – Jaynes’ principle) as mandatory. If an expert would like nonetheless use only uniform distributions but express own knowledge more accurately he/she can switch to a class of generalized interval estimates [Chugunov et al, 2008] and resort to generalized uniform distributions [Sternin, 2010].

There are two main types of decision-making problems: problems of the unique and problems of the repeated choice. Each of these types of problems has specific comparison criteria. Note that situation of unique choice is typical for problems of forecasting.

Two approaches are usually used in the evaluating preference of interval alternatives and associated risks. In the framework of the first approach compared alternatives are considered as isolated, unconnected objects. Value of preference criterion is calculated for each of these alternatives and then, regardless of this indicator, one or the other risk indicator is calculated. To compare alternatives and choose preferred object the alternatives are evaluated on these two criteria. In spite of the fact that many problems of interval comparing belong to the class of unique choice as indicators of preference in this approach are often used averages of corresponding chance distributions (mathematical expectations), which are rather adequate to problems of repeated choice. Such indicators as variance, left and right semi variances, the mean absolute semi deviation and others [Ogryczak, 1999; Baker, 2015] are used as risk indicators in this approach. We draw attention to the fact that the values of the comparison criteria in this approach does not depend on the number of alternatives in their given set.

In the framework of the second approach compared alternatives are considered as an interconnected integrity. It is seems that this approach is more in line with features of the problems of unique choice. Risk selecting an alternative as preferred one in their set depends here firstly on the relative location of the compared intervals (configuration of compared alternatives) and then on their amount in the set. The presence of the group of mutually influencing alternatives increases the risk of making the wrong decision when choosing a preferred alternative. This is due to a “collective effect” just as it happens, for example, in condensed matter physics when properties of condensed matter systems composed of interacting components significantly differ from properties of more or less independent parts [Halperin, 2010].

Dimensionless chances of truth of tested by expert hypothesis concerning preference of an analyzed alternative relative to others are comparison criteria within this approach. Chances of truth of the opposite hypothesis, which complement the first chances up to unity, serve as a measure of risk. In this approach point implementations of different alternatives from analyzed set are considered as independent and priori all the alternatives have equal rights with respect to the choice.

Slightly specify that will be understood further at risk. In accordance with the standard (ISO/FDIS 31000:2009) risk is defined as “the effect of uncertainty on the achieved goals”. More narrowly risk is a characteristic of decision-making situation, which has the uncertainty of the outcomes and what is more the presence of adverse effects is the obligatory condition. Thus the concept of risk is a combination of chances and consequences of adverse events. However consequences are specific to each particular decision-making problem and on the type and on the size but the chances is one of the universal characteristics of risk. As rule consequences are dimensional values, chances are dimensionless.

For these reasons chances of preference as the selection criterion will be used in the paper. Corresponding risks will be estimated on based of these chances.

The start step in the realization of this approach is pairwise comparing alternatives, when the number of objects to be compared and its impact on risk do not take into account [Shepelev, 2013; 2014]. Criterion of comparison of interval alternatives with arbitrary distributions of chances, which was called “assurance factor”, was proposed on this way. It is equal to the difference between chances of the truth of tested hypothesis on preference of an alternative in their set and the chances of the truth of the opposite hypothesis. Numerical (for arbitrary distributions of chance) and analytical (for uniform and triangular distributions) methods calculating the assurance factor as well as decision-making procedure based on this criterion were proposed. The assurance factor and chances of preference are equivalent as comparison criteria. The first of them is more convenient in some cases. For example, for some simple distributions of chances one may establish a relation between such criteria as the difference of the averages for two compared alternatives and assurance factor [Shepelev, 2013; 2014] and their dissimilarity in other cases. Questions about depending of the risk of making right and wrong decisions on the number of comparable alternatives and calculation of corresponding chances remained however not studied. Earlier this subject was touched in paper [Diligensky, 2004]; here we look at it in more detail.

General statements for estimating of risk in a group of interval alternatives

Suppose that there are K alternatives I_i, i = 1, 2,…, K with the same interval quality indicators. Let dimensionless quantity $C(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},…, I_K))$ is the chances, in other words degree of assurance, in the truth of a testable hypothesis of preference, that the alternative I_i is more preferable than all at once alternatives ($I_1, I_2,.., I_{i-1}, I_{i+1},…, I_K$) from initially given their set (I_i is “better” of others “on the whole”). The term “all at once” means here that

$$I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K) \equiv (I_i \succ I_1) \wedge (I_i \succ I_2) \wedge (I_i \succ I_3) \wedge \dots \wedge (I_i \succ I_{i+1}) \wedge \dots \wedge (I_i \succ I_K).$$

where $\equiv$ and $\wedge$ are symbols of equivalence and conjunction respectively. It is clear on sense that

$$0 \le C(I_i \succ (I_1, I_2,\dots, I_{i-1}, I_{i+1},\dots, I_K)) \le 1.$$

Risk that I_i would not in reality preferred will be measured by means of dimensionless quantity $R_s(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ complementing previous chances to unity so that

$$Rs(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)) = 1 - C(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)).$$

As can be seen $R_s(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ is degree of assurance in the truth of a hypothesis, which is opposite to the testable hypothesis of preference.

One may see that the following statement is true (T):

$$(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)) \vee \neg (I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)) = T,$$

where $\neg$ is symbol of negation. Then corresponding chances

$$C(I_1 \succ (I_2, I_3,\dots, I_K)) + C(\neg (I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))) = 1$$

and $Rs(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)) = C(\neg (I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)))$.

The statement $\neg I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ means that at least one alternative from their compared set would be preferable than I_i. Let illustrate the meaning of introduced in this way the measure of risk for case of three alternatives. Here we have the following possible preferences and chains of disjunctions:

$((I_1 \succ I_2 \succ I_3) \vee (I_1 \succ I_3 \succ I_2)) \vee ((I_2 \succ I_1 \succ I_3) \vee (I_2 \succ I_3 \succ I_1)) \vee ((I_3 \succ I_1 \succ I_2) \vee ((I_3 \succ I_1 \succ I_2)) \equiv (I_1 \succ (I_2, I_3)) \vee (I_2 \succ (I_1, I_3)) \vee (I_3 \succ (I_1, I_2))$ or $Rs(I_1 \succ (I_2, I_3)) = C(I_2 \succ (I_1, I_3)) + C(I_3 \succ (I_1, I_2))$.

Analogically for K alternatives

$C(I_1 \succ (I_2, I_3,\dots, I_K)) + C(I_2 \succ (I_1, I_3,\dots, I_K)) + C(I_3 \succ (I_1, I_2, I_4,\dots, I_K)) + \dots + C(I_K \succ (I_1, I_2,\dots, I_{K-1})) = 1$

Hence

$Rs(I_1 \succ (I_2, I_3,\dots, I_K)) = C(I_2 \succ (I_1, I_3,\dots, I_K)) + C(I_3 \succ (I_1, I_2, I_4,\dots, I_K)) + \dots + C(I_K \succ (I_1, I_2,\dots, I_{K-1}))$.

One can see now that $C(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ is monotonically non-increasing function of K, that is the chances that a certain alternative would be preferable in comparison with all the others do not increase with increasing number of the alternatives. Indeed, if the number of compared alternatives is increased the number of non-negative terms in the unit sum of corresponding chances is also increased. Therefore

$$C(I_i \succ (I_1, I_2,\dots, I_{i-1}, I_{i+1},\dots, I_K)) \le C(I_i \succ (I_1, I_2,\dots, I_{i-1}, I_{i+1},\dots, I_{K-1})), \tag{1A}$$

that proves monotonic non-increasing of chances. Then corresponding risk will be monotonically non-decreasing function of number of compared alternatives:

$$Rs(I_i \succ (I_1, I_2,\dots, I_{i-1}, I_{i+1},\dots, I_{K-1})) \le Rs(I_i \succ (I_1, I_2,\dots, I_{i-1}, I_{i+1},\dots, I_K)). \tag{1B}$$

By another words the more of number of the alternatives the more risk of wrong decision-making.

The fact of increasing overall risk with increasing number of alternatives is clearly demonstrated by the following equation:

$Rs(I_1 \succ (I_2, I_3,\dots, I_K)) + Rs(I_2 \succ (I_1, I_3,\dots, I_K)) + Rs(I_3 \succ (I_1, I_2, I_4,\dots, I_K)) + \dots + Rs(I_K \succ (I_1, I_2,\dots, I_{K-1})) = K - 1$.

The relationship (1A) takes place for chances of all possible exceptions of one interval estimation I_i from their set $(I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K)$. Therefore the chances $C(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ is not more than minimal chances, which may occur in the right side of (1A) and, respectively, $R_s(I_i \succ (I_1, I_2,.., I_{i-1}, I_{i+1},\dots, I_K))$ is not less than other risks, which may occur in the left side of (1B) after exception of some interval estimations.

Numerical method calculating chances and risks in a group of interval alternatives

Chances of preference for an alternative with arbitrary distributions on compared intervals can be calculated by the method of statistical testing. To be specific, situations of comparison will be analyzed when greater value of quality indicator corresponds to more preferred state and the hypothesis is tested that the first interval from their set with Q representatives is more preferable than others.

Let i_{lr} is point implementation of interval I_l in the r-th trial of Monte Carlo (l = 1, 2,.., Q; r = 1, 2,…, S). If S independent tests for each of the compared intervals was made and S_f is the number of tests for which i_{1r} > MAX(i_{2r},…, i_{Qr}), then S_f/S is an estimate for $C(I_1 \succ (I_2,\ldots, I_Q))$. This numerical method is applicable to any distributions on compared intervals when we cannot receive analytical formula as well as in the case of a large number of compared alternatives with simple distributions when obtained, in principle, analytical formula are difficult foreseeable.

Among simple distributions commonly used in practice should be noted uniform, triangular and trapezoidal distributions. Random numbers N_x for these distributions with density $f_x(z)$ used in the method of statistical testing can be obtained by the inverse function method from the standard random number N_u for uniform distribution defined on the interval [0, 1]. In accordance with this method $N_u = \int_L^{N_x} f_x(z)dz$. This integral can be taken for mentioned simple distributions.

We have for a uniform distribution on the interval [L, R]:

$$f_E(z) = 1/(R - L),\ N_E = (1 - N_u)L + N_uR.$$

For triangular distribution

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

where M is a mode of the distribution, and

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

For trapezoidal distribution

$$f_T(z) = \frac{2}{S}\begin{cases} \dfrac{z-L}{M_1-L}, \; L \le z \le M_1, \\ 1, \qquad M_1 < z < M_2, \\ \dfrac{R-z}{R-M_2}, \; M_2 \le z \le R \end{cases}$$

where $S = R + M_2 - M_1 - L$ and M_1 и M_2 are the left and right vertices of the distribution respectively. Then

$$N_T = \begin{cases} L + [N_u S(M_1 - L)]^{1/2}, \; N_u \le (M_1 - L)/S, \\ (N_u S + M_1 + L)/2, \; (M_1 - L)/S < N_u < (2M_2 - M_1 - L)/S, \\ R - [(1-N_u)(R-M_2)S]^{1/2}, \; N_u > (2M_2 - M_1 - L)/S. \end{cases}$$

Although numerical methods allow calculating the chances and risks involved in the specific problems considered here, only analytical methods make it possible to obtain more general conclusions. Let us consider firstly case of several coinciding interval alternatives and obtain analytical expressions for this case for the simplest distributions of chances.

Risk behavior in a set of coincided interval alternatives

Assume for simplicity that distributions in the compared interval alternatives are uniform distributions. Assume also that all alternatives at the number of K are represented by coinciding intervals. Illustrations of general statements, which were formulated above, will be in this case the brightest. It can be seen that for two coinciding intervals $C(I_1 \succ I_2) = \frac{1}{D}\int_L^R f_2(x_2)\int_{x_2}^R f_1(x_1)dx_1 dx_2$, where $f_i(x_i)$ are corresponding densities, $x_i \in I_i$ are current values in corresponding intervals and D is normalization factor. Then for uniform distributions we have: $C(I_1 \succ I_2) = \frac{1}{(R-L)^2}\int_L^R dx_2 \int_{x_2}^R dx_1$ = ½.

For three intervals

$$C(I_1 \succ (I_2, I_3)) = \frac{1}{(R-L)^3}[\int_L^R dx_3 \int_{x_3}^R dx_2 \int_{x_2}^R dx_1 + \int_L^R dx_2 \int_{x_2}^R dx_3 \int_{x_3}^R dx_1].$$

Here the first term corresponds to the current values of variables for which $x_2 > x_3$ and the second term to the situations when $x_3 > x_2$. This gives $C(I_1 \succ (I_2, I_3))$ = 1/3. For K intervals we have $(K-1)(K-2)\cdots 1$ of integral summands, which reflect all possible relations (<, >) between current values of variables. Ultimately $C(I_1 \succ (I_2, I_3, \dots, I_K)) = 1/K$ and risk associated with the adoption of the initial hypothesis is $1 - 1/K$.

One can see that if for the situation when two interval alternatives appear to be equivalent at the moment of comparison chances of right choice of preferred alternative equals risk of wrong choice then for a larger number of alternatives this risk increases rapidly. It can think that it is not a property of uniformity of distributions but property of equivalence of alternatives. To test this hypothesis we consider the case of triangular distributions.

The formula for calculating chances becomes more complicated in the case of a triangular distribution. The reason for this is rooted in the presence of two branches in the density of chances distribution for triangular distribution: left branch f_l, which lies on the left from value of mode, and right branch f_r on the right from mode. Let M_1, M_2 are modes of corresponding distributions and $M_2 \le M_1$ and we want to estimate chances $C(I_2 \succ I_1)$. These chances are the sum of four integrals C_i as may see on region of integration, which is represented here by square on the (X_1, X_2) plane:

$$C_1 = \int_L^{M_2} f_{1l}(x_1) \int_{x_1}^{M_2} f_{2l}(x_2)dx_1dx_2, C_2 = \int_L^{M_2} f_{1l}(x_1) \int_{M_2}^{R} f_{2r}(x_2)dx_1dx_2, C_3 = \int_{M_2}^{M_1} f_{1l}(x_1) \int_{x_1}^{R} f_{2r}(x_2)dx_1dx_2,$$

$$C_4 = \int_{M_1}^{R} f_{1r}(x_1) \int_{x_1}^{R} f_{2r}(x_2)dx_1dx_2.\ C(I_2 \succ I_1) = \frac{M_2 - L}{2(M_1 - L)} + \frac{(R-M_2)^3 - (R-M_1)^3}{6(M_1 - L)(R - M_2)(R - L)}.$$

.

Hence for $M_2 = M_1$ we have $Rs(I_1 \succ I_2) = C(I_2 \succ I_1) = ½$. The same result was obtained in the paper [Shepelev, 2014] by another method. Thus under equality of the two modes of triangular distributions defined on coinciding pair of intervals, when interval estimates are equivalent, risk $Rs(I_1 \succ I_2) = ½$ exactly as was for the case of uniform distributions. If the modes are not equal to each other risks (and chances) are more or less than ½ depending on location of modes. So compared alternatives are equivalent if their distributions of chances (not necessarily uniform) are the same and the same are their supports. Thus the same knowledges about interval alternatives resulting in their equivalence generate the highest risk during choice of the best alternative.

Chances and risks for such defined equivalent alternatives behave like for the uniform distributions i.e. the chances of preference fall hyperbolically with increasing of number K of alternatives and risk equals $1 - 1/K$. This is confirmed by numerical method calculations. Analytical relations for the chances of preference become rather cumbersome when the number of alternatives is more than two already for triangular distributions.

The case of other configurations

Besides the configuration of coinciding estimates for the two comparable alternatives there are, up to a permutation of alternatives in their pair, else two non-trivial configurations, i.e. configurations with non-zero intersections of intervals. It is configuration of the right shift when $L_2 < L_1 < R_2 < R_1$ and configuration of imbedded intervals when $L_1 < L_2 < R_2 < R_1$.

For these configurations and uniform distributions to receive relations for the preference chances and corresponding risks is convenient to use a method based on simple geometric considerations. For configuration of the right shift in the framework of complete system of events it is easier to distinguish events favoring the truth of the hypothesis $I_2 \succ I_1$ (size of risk for hypothesis $I_1 \succ I_2$). These are events when point implementations lie in the region $(i_1 \in [L_1, R_2]) \cap (i_2 \in [L_1, R_2])$, $i_1 \in I_1$, $i_2 \in I_2$. However some of these events at the same time are also favorable for the truth of hypothesis $I_1 \succ I_2$. Exactly half of the events favor each of these hypotheses in the case of uniform distributions on compared intervals as we can see above. This is not so in the case of other distributions. Then, for uniform distributions on the compared intervals,

$$C(I_2 \succ I_1) = Rs(I_1 \succ I_2) = (R_2 - L_1)^2/(2\Delta I_1 \Delta I_2),\ \Delta I_i = R_i - L_i,$$

and

$$C(I_1 \succ I_2) = 1 - \frac{(R_2 - L_1)^2}{2\Delta I_1 \Delta I_2}.$$

It's easy to see that the closer l_1 to l_2 and R_1 to R_2 the closer the risk to ½. For two intervals this value of the risk is maximal for configuration of the right shift. Indeed, the more R_1 the less the risk (at constant L_1 and L_2); the more L_1 the less the risk (at constant R_1 and R_2). Therefore this configuration is favorable for the preference of I_1.

Triangular distributions cases are somewhat more complicated because of the different possible positions of the modes. Let $\Delta I_1 > \Delta I_2$, $L_2 < L_1 < R_2 < R_1$ (right shift) and $M_2 < L_1$, $M_1 > R_2$ (the simplest configuration). The range of permissible point implementations in the (X_1, X_2) plane is a rectangle elongated to the right and its part, which corresponds area $X_2 > X_1$, lies above and on the left from line segment $X_2 = X_1$ with boundary points (L_1, L_1) and (R_2, R_2). Hence

$$C(I_2 \succ I_1) = Rs(I_1 \succ I_2) = \int_{L_1}^{R_2} f_{1l}(x_1) \int_{x_1}^{R_2} f_{2r}(x_2) dx_1 dx_2 = \frac{(R_2 - L_1)^4}{6\Delta_1 \Delta_2 (R_2 - M_2)(M_1 - L_1)}.$$

Small, at first glance, changes of locations of distributions modes greatly complicate the formula for risk. So, when $L_1 < M_2 < M_1 < R_2$ (other things being equal), $Rs(I_1 \succ I_2) = C_1 + \ldots + C_4$,

$$C_1 = \int_{L_1}^{M_2} f_{1l}(x_1) \int_{x_1}^{M_2} f_{2l}(x_2) dx_1 dx_2, C_2 = \int_{L_1}^{M_2} f_{1l}(x_1) \int_{M_2}^{R_2} f_{2r}(x_2) dx_1 dx_2,$$
$$C_3 = \int_{M_2}^{M_1} f_{1l}(x_1) \int_{x_1}^{R_2} f_{2r}(x_2) dx_1 dx_2, C_4 = \int_{M_1}^{R_2} f_{1r}(x_1) \int_{x_1}^{R_2} f_{2r}(x_2) dx_1 dx_2.$$

Taking these integrals is not difficult; the resulting formulas allow us to establish the following. Choice of the type of distribution, which is distinct from uniform one, results in increasing of demands to the knowledge of experts. So in a situation of right shift under choice of triangular distributions expert should has in mind that for the same values of modes that lie at the area of intersection of interval estimates risk is practically unchanged with the displacement modes in this area. For example, for I_1= [10, 20], I_2 = [9, 19] $Rs(I_1 \succ I_2) = 0.42$; $C(I_1 \succ I_2) = 0.58$ for all equal values of modes ($M_1 = M_2$) from 11 to 18. Interestingly, that selecting uniform distributions practically does not change the results: $Rs(I_1 \succ I_2) = 0.4$; $C(I_1 \succ I_2) = 0.6$. (But the risk is slightly reduced). At the same time for M_1= 18, M_2 = 16 $Rs(I_1 \succ I_2) = 0.33$; $C(I_1 \succ I_2) = 0.67$. Thus the specification of chance distributions on compared alternatives requires high qualification of the expert and elaborating special procedures for working an expert with probability distributions on interval estimates.

If one from distributions is uniform it is easy to obtain formulas for the risk by the method used here or to use the relations obtained earlier in [Shepelev, 2014]. General conclusion in both cases is that, ceteris paribus, the more the value of mode in I_1 in general (and in comparison with the mode in I_2 for triangle distribution there) the less risk of making a wrong decision on the preference of the first alternative.

For case of embedded intervals events that are favorable to the truth of the hypothesis $I_1 \succ I_2$ are $\{(i_1 \in [R_2, R_1]) \cap (i_2 \in [L_2, R_2])\} \cup \{(i_1 \in [L_2, R_2]) \cap (i_2 \in [L_2, R_2])\}$. Hence, for uniform distributions,

$$R(I_1 \succ I_2) = 1 - \frac{R_1 - L_2}{\Delta I_1} + \frac{\Delta I_2}{2\Delta I_1},\ L_1 < L_2 < R_2 < R_1.$$

Further we will need the formula for the same configuration when $L_2 < L_1 < R_1 < R_2$:

$$R(I_1 \succ I_2) = 1 - \frac{L_1 - L_2}{\Delta I_2} - \frac{\Delta I_1}{2\Delta I_2}\ .$$

The case of triangular distributions is technically similar here to the case of right shift.

Set of possible configurations is considerably richer for three compared intervals. We will consider only one of them, for which $L_2 < L_1 < L_3 < R_2 < R_3 < R_1$, as an example. Subset of the complete system of events favoring the truth of the hypothesis $I_1 \succ (I_2, I_3)$, is as follows:

$\{(i_1 \in [R_3, R_1]) \cap (i_2 \in [L_2, R_2]) \cap (i_3 \in [L_3, R_3])\} \cup \{(i_1 \in [R_2, R_3]) \cap (i_2 \in [L_2, R_2]) \cap (i_3 \in [R_2, R_3])\} \cup \{(i_1 \in [R_2, R_3]) \cap (i_2 \in [L_2, R_2]) \cap (i_3 \in [L_3, R_2])\} \cup \{(i_1 \in [L_3, R_2]) \cap (i_2 \in [L_3, R_2]) \cap (i_3 \in [L_3, R_2])\} \cup \{(i_1 \in [L_3, R_2]) \cap (i_2 \in [L_2, L_3]) \cap (i_3 \in [L_3, R_2])\}$..

In the transition to the respective chances it should keep in mind that for uniform distributions at the intersection of the two sub-intervals in expression for chances ratio ½ appears, and at the intersection of three subintervals ratio 1/3. After some transformations we obtain:

$$C(I_1 \succ (I_2, I_3)) = \frac{R_1 - R_3}{\Delta I_1} + \frac{R_3 - R_2}{\Delta I_1 \Delta I_3}(\frac{R_3 - R_2}{2} + R_2 - L_3) + \frac{(R_2 - L_3)^2}{\Delta I_1 \Delta I_2 \Delta I_3}(\frac{L_3 - L_2}{2} + \frac{R_2 - L_3}{3}).$$

and $Rs(I_1 \succ (I_2, I_3)) = 1 - C(I_1 \succ (I_2, I_3))$.

Operating in a similar manner, we have:

$$C(I_2 \succ (I_1, I_3)) = \frac{(R_2 - L_3)^2}{\Delta I_1 \Delta I_2 \Delta I_3}(\frac{R_2 - L_3}{3} + \frac{L_3 - L_1}{2}) \text{ and}$$

$C(I_3 \succ (I_1, I_2)) = 1 - C(I_1 \succ (I_2, I_3)) - C(I_2 \succ (I_1, I_3)) = 1 - Rs(I_3 \succ (I_1, I_2))$.

Since the above expression depends only on the differences of the boundaries of intervals, then, in the case of uniform distributions, relations for chances of preference does not change when the borders are changed on the same number (translation invariance).

A numerical example

Consider a numerical example for this configuration, its parameters and the results of calculations based on the above analytical formulas are presented in Table 1. Data of the Table 1 show that $C(I_1 \succ (I_2, I_3))+C(I_2 \succ (I_1, I_3))+ + C(I_3 \succ (I_1, I_2)) = 1$, $Rs(I_1 \succ (I_2,I_3)) + Rs(I_2 \succ (I_1, I_3)) + Rs(I_3 \succ (I_1, I_2)) = 3 - 1 = 2$, the chances of preference for one interval alternative with respect to two others are less than the minimum chances of its preference in the pair-wise comparison (e.g. 0.602 < MIN(0.833, 0.625)). These findings are in full agreement with the statements of section 2 of this paper. With increasing the right (left) boundary of the second interval the chances of its preference are increased and the chances of the other two alternatives are reduced.

Previously we have seen that risks grow larger and chances of preference become small in the case of nearness of interval alternatives boundaries for a large number of compared objects. These indicators are also very sensitive to changes in the borders of the worst alternative. Suppose, for example, we have three alternatives $I_1 = [10, 18]$, $I_2 = [9, 16]$, $I_3 = [11, 17]$, then $C(I_1 \succ (I_2, I_3)) = =0.438$, $C(I_2 \succ (I_1, I_3)) = 0.161$, $C(I_3 \succ (I_1, I_2)) = 0.401$. Wherein $C(I_1 \succ I_2) = 0.679$, $C(I_1 \succ I_3) = =0.490$, $C(I_3 \succ I_2) = 0.702$. Small increasing the right boundary of the second alternative substantially changes the magnitude of the preference chances: if the first and third alternatives are unchanged and $I_2 = [9, 16.5]$ then $C(I_1 \succ (I_2, I_3)) = 0.423$, $C(I_2 \succ (I_1, I_3)) = 0.196$, $C(I_3 \succ (I_1, I_2)) = =0.381$. Wherein $C(I_1 \succ I_2) = 0.648$, $C(I_1 \succ I_3) = 0.490$, $C(I_3 \succ I_2) = 0.664$. It is evident that a change in the parameters of at least one alternative changes the magnitudes of all chances in the “collective” estimation unlike the pairwise comparison.

Table 1. Preference chances and risks for three compared intervals

Left and Right Borders of the Compared Intervals			
L_1	10	10	10
R_1	18	18	18
L_2	8	8	9
R_2	14	15	14
L_3	11	11	11
R_3	15	15	15
Preference Chances and Risks of Testing Hypothesis			
$C(I_1 \succ (I_2, I_3))$	0.602	0.577	0.597
$C(I_1 \succ I_2)$	0.833	0.777	0.800
$C(I_1 \succ I_3)$	0.625	0.625	0.625
$Rs(I_1 \succ (I_2, I_3))$	0.398	0.423	0.403
$C(I_2 \succ (I_1, I_3))$	0.070	0.131	0.084
$C(I_2 \succ I_1)$	0.167	0.223	0.200
$C(I_2 \succ I_3)$	0.188	0.286	0.225
$Rs(I_2 \succ (I_1, I_3))$	0.930	0.869	0.916
$C(I_3 \succ (I_1, I_2))$	0.328	0.292	0.319

$C(I_3 \succ I_1)$	0.375	0.375	0.375
$C(I_3 \succ I_2)$	0.813	0.714	0.775
$Rs(I_3 \succ (I_1, I_2))$	0.672	0.708	0.681

What will happen with the estimates of preferences when fourth alternative will be added to a set of three ones? For definiteness let I_1 = [10, 18], I_2 = [8, 14], I_3 = [11, 15], I_4 = [10.5, 16]. Calculations carried out by statistical testing method for four interval estimates and by the foregoing formula for three and two estimates give the following results:

$C(I_1 \succ (I_2, I_3, I_4))$ = 0.498; $C(I_2 \succ (I_1, I_3, I_4))$ = 0.034; $C(I_3 \succ (I_1, I_2, I_4))$ = 0.187; $C(I_4 \succ (I_1, I_2, I_3))$= = 0.281; $C(I_1 \succ (I_2, I_3))$ = 0.602; $C(I_1 \succ (I_2, I_4))$ = 0.567; $C(I_1 \succ (I_3, I_4))$ = 0.507; $C(I_1 \succ I_2)$ = =0.833; $C(I_1 \succ I_3)$ = 0.625; $C(I_1 \succ I_4)$ = 0.594. (We restricted ourselves to the hypothesis of preference of the first alternative when comparing estimates of the preference chances for “collective” and pairwise comparison). Again 0.498 < MIN(0.602, 0.567, 0.507), $C(I_1 \succ (I_2, I_3, I_4))$+ $C(I_2 \succ (I_1, I_3, I_4))$ + $C(I_3 \succ (I_1, I_2, I_4))$ + $C(I_4 \succ (I_1, I_2, I_3))$ = 1.

So the chances are reduced and the risks grow.

Hypothetical case of decision-making with account of “collective” effect

Let’s consider the hypothetical case of decision-making with account of “collective” effect for three interval alternatives described in the paper [Kononov, 2010] where three possible projects using the resources of the Kovykta gas condensate field are examined. Internal rate of return (IRR) of the project[1] is used in the cited paper as a criterion of comparison. Interval estimates for possible values of IRR in each project were obtained there and ranking of projects by preference and choice of the best one were performed by Hurwicz’s method while different values of “pessimism – optimism” coefficient λ were used for different projects. Data mentioned in [Kononov, 2010] are presented in Table 2.

[1] We leave aside the question of the validity finding preference of projects on the basis of such criteria as IRR (see in this regard [Vilensky, 2015]).

Table 2. Interval estimates of IRR for three investment projects

Alternatives	Project Name	IRR (%)	λ
Alt_1	The gas supply to the Unified Gas Supply System of Russia	14.8 – 19.8	0.75
Alt_2	Export of liquefied natural gas to the Asia-Pacific region	11.7 – 23.6	0.5
Alt_3	Gas export to China	10.7 – 27.7	0.25

Recall that according to Hurwicz's method interval estimate $[L, R]$ is replaced by a point estimate $T(\lambda)$ by the formula $T(\lambda) = (1 - \lambda)L + \lambda R$, where $0 < \lambda < 1$ is "pessimism – optimism" coefficient. Note that the choice of different values of λ for different alternatives is not commonly accepted and the concrete choice of these values is difficult to justify. At the same time in some cases mentioned choice is the only way to reconcile knowledge/prediction of expert with the results of Hurwicz's method for right shift configuration of compared intervals when $L_2 < L_1 < R_2 < R_1$. For this configuration under the identical values of the "pessimism – optimism" coefficients the first interval is more preferable than the second one for any λ from the segment [0, 1]. Differing values of λ lead to different results that can fit in with the expectations of experts.

Using the table 2 values of λ we have for results of alternatives ordering after applying Hurwicz's method: $Alt_1 \succ Alt_2 \succ Alt_3$. Indeed, $T_1(0.75) = 18.55$; $T_2(0.5) = 17.65$; $T_3(0.25) = 14.95$. However, the order in the set of comparable alternatives is highly dependent on the choice of λ values if these values are the identical. Namely, for alternatives that are considered in table 2 for $\lambda < < 0.196$ $Alt_1 \succ Alt_2 \succ Alt_3$, for $0.196 < \lambda < 0.34$ $Alt_1 \succ Alt_3 \succ Alt_2$, for $0.34 < \lambda < 0.45$ $Alt_3 \succ Alt_1 \succ Alt_2$, at last for $\lambda > 0.45$ $Alt_3 \succ Alt_2 \succ Alt_1$. We draw attention to the fact that value of $\lambda = 1/3$ recommended in [Vilensky, 2015] for choosing similar alternatives is in this case just on the border of the two above selected bands of values λ. Therefore the difference of preferences determined by Hurwicz's method for the first and the third alternatives becomes insignificant.

It is advisable to analyze this problem now by means of the proposed in this paper approach. We estimate the same problem situation comparing alternatives "as a whole". To do so we need the relations for preference chances for the configuration $L_3 < L_2 < L_1 < R_1 < R_2 < R_3$, that is for configuration of embedded intervals. This is the configuration of interval estimates of IRR in the discussed case. Suppose as before that the distributions on compared intervals are uniform. These relations are shown below.

$$C(I_1 \succ (I_2, I_3)) = \frac{(\Delta I_1)^2}{3\Delta I_2 \Delta I_3} + \frac{(L_1 - L_3)(R_1 - L_2) + (L_1 - L_2)(R_1 - L_3)}{2\Delta I_2 \Delta I_3},$$

$$C(I_2 \succ (I_1, I_3)) = \frac{(\Delta I_1)^2}{3\Delta I_2 \Delta I_3} + \frac{(R_2 - R_1)(R_2 - 2L_3 + R_1) + \Delta I_1 (L_1 - L_3)}{2\Delta I_2 \Delta I_3}.$$

$$C(I_3 \succ (I_1, I_2)) = 1 - C(I_1 \succ (I_2, I_3)) - C(I_2 \succ (I_1, I_3)).$$

The results calculating the chances are shown in Table 3.

Table 3. Preference chances for gas utilization alternatives

Tested Hypothesis	Chances Values
$Alt_1 \succ (Alt_2, Alt_3)$	0.193
$Alt_1 \succ Alt_2$	0.471
$Alt_1 \succ Alt_3$	0.388
$Alt_2 \succ (Alt_1, Alt_3)$	0.298
$Alt_2 \succ Alt_1$	0.529
$Alt_2 \succ Alt_3$	0.409
$Alt_3 \succ (Alt_1, Alt_2)$	0.509

$Alt_3 \succ Alt_2$	0.591
$Alt_3 \succ Alt_1$	0.612

One can see that both in pair-wise comparison and in the comparison “as a whole” the third alternative preferred others. However the risk of wrong decision is not much less than preference chances of the third alternative. After removing the first alternative from the list of compared ones as an alternative with preference chances for the pairwise comparison less than half, the risk associated with choice of the third alternative is reduced to about 0.4. Note that comparison on base of value of the mathematical expectation leads to the similar choice: $Av(Alt_1) = 17.3$; $Av(Alt_2) = 17.65$; $Av(Alt_3) = 19.2$ (see [Shepelev, 2014]). However using the latter approach doesn’t permit estimate the risk associated with decision-making. Besides coincidence of the choice results for these two selection criteria takes place only for uniform distributions on compared intervals [Shepelev, 2013].

Conclusion

Thus the effect of the comparison of interval alternatives “as a whole” is manifested primarily in reducing value of preference chances for each alternative with respect to its chances under pair-wise comparison. This leads to a quantitative increasing risk value of selection as preferred alternative such one, which may not actually be per se later. The nature of this effect lies in the fact that in the presence of non-zero intersection for already two compared alternatives there is a non-zero risk of making the wrong decision. This risk is enhanced with increasing amounts of compared alternatives especially if some of these chances are not too different from each other. However, it should be borne in mind that the perception of risk is individual and can vary from one DM to another. Therefore the risk value resulting from the use of the proposed method is nothing more than a calculated risk, which can serve only as an estimate for the DMs.

What can be done to reduce the calculated risk? During deciding on preferred alternative choice or in the process of ordering alternatives by preference it’s useful to conduct a preliminary analysis of their initial set. Firstly, after selecting an alternative that preference is tested, one should select in the set of alternatives those, which do not have the intersection with analyzed alternative. If the left boundary of such intervals no less than the right boundary of the tested one the latter is certainly worse. If the right boundary of such intervals not greater than the left boundary of the tested one they can be excluded because they are certainly worse than the last interval. Secondly, one may try to unify some similar or complementary alternatives. By reducing the number of intervals in their initial set one may increase the calculated preference chances of analyzed alternative and decrease risks. At last, after calculating the preference chances of tested alternative during pairwise comparisons it is advisable to exclude those alternatives whose preference chances with respect to tested alternative is less than 0.5 and, respectively, the risk is more than 0.5.

Are there any other amendments to the results of the pairwise comparison of alternatives due to “collective” effect? Particularly important is the following question: is there difference of the alternatives order in their set defined by the “collective” preference chances and the order for pairwise comparison? The answer to this question is negative: the order established in the process of pairwise comparison is the same as the order for comparison “as a whole”.

Thus for this approach the “best” alternative will be the alternative with the highest chances at pairwise comparisons in the set of compared alternatives. However adequate the risk estimation of making wrong decision we obtain by comparing this alternative simultaneously with all the others, “as a whole”.

The large number of alternatives aimed at achieving the same goals is a characteristic of the upper hierarchical levels of decision-making. Perhaps that is a reason why for the upper levels of management the risk of making not quit correct decision is more likely (ceteris paribus). Thus meaningful choice of a part of initial set of alternatives to lower levels with the delegation of authority to estimate such objects is represented with point of view of the considered here approach as quite justified.

Difficulties in the work of the experts and decision-makers under interval uncertainty are associated with the complexity of representation of their knowledge and grounding decision-making. DMs often desire to overcome the uncertainty by replacement of interval estimations with point ones on base of their experience, preferences and intuition. Experts from their side desire express their knowledge by rather simple distributions. Of course, exact distributions of chances are unknown for interval alternatives. But one can assume that in many cases these distributions are unimodal ones, and for many types of distributions can be roughly approximated by triangular distributions.

In this regard let’s pay attention to following circumstance. Although the methods used in the proposed approach for comparison of alternatives by preference are quantitative because of approximate nature of expert information hardly makes sense to emphasize exactly how much calculated indicator of the quality of one alternative is over/under than for the other one. It seems that here are more appropriate judgments based on ordinal scales i.e. on stating that one of the alternatives is preferable others without quantifying the degree of the preference, such as has place in problems with not quantitative but with qualitative criteria [Larichev, 2006].

Since process of decision-making is quite difficult DMs and experts need means of analytical support. In particular, it's useful that DMs or experts had some ideas of the magnitude of the risk associated with their choice, which is defined not only by configurations of pairwise comparisons but also by a specific set of comparable alternatives. Some such methods are proposed in this paper, which may permit to DMs or experts check how their knowledge and largely intuitive choice is consistent with the formal results and adjust their decisions. Using in the process of alternatives comparing different methods increases the volume and variety of information that is useful to DM and may contribute to increasing adequacy of decision-making.

Acknowledgements

The paper is published with partial support by the project ITHEA XXI of the ITHEA ISS (www.ithea.org) and the ADUIS (www.aduis.com.ua).

References

[Baker, 2015] Baker H.K., Filbeck G. Investment risk management. NY: Oxford university press. 2015.

[Chugunov, 2008] Chugunov N., Shepelyov G., Sternin M. The generalised interval estimations in decision making under uncertainty. //Int. J. Technology, Policy and Management. Vol. 8, No. 3, pp. 298 – 321. 2008.

[Diligensky, 2004] Diligensky N.V., Dymova L.G., Sevastyanov P.V. Nechetkoe modelirovanie i mnogokriterialnaya optimizatsiya proizvodstvennykh sistem v usloviyakh neopredelennosti: tekhnologiya, ekonomika, ekologiya. M.: Izdatelstvo mashinostroenie -1. 2004- in Russian.

[Halperin, 2010] Halperin B., Sevrin A. (Eds.). Quantum theory of condensed matter. World Scientific. 2010.

[Kononov, 2010] Kononov Yu. D., Loktionov V.I., Stupin P.V. Uchet faktora neopredelennosti pri otsenke variantov ispolzovaniya kovyktinskogo gaza //Proc. of the Int. Symposium on “Energy of Russia in XXI Century: Development Strategy – Eastern Vector”. Irkutsk, Russia. 2010.– in Russian.

[ISO/FDIS 31000, 2009]. International Standard ISO/FDIS 31000, Risk management - Principles and Guidelines. 2009

[Larichev, 2006] Larichev O.I. Verbalnyy analiz resheniy. M.: Nauka. 2006. – in Russian.

[Nilsson, 1986] Nilsson N. J. Probabilistic logic. Artificial Intelligence. Vol. 28, No. 1, pp. 71-87. 1986.

[Ogryczak, 1999] Ogryczak W, Ruszczyński A. From stochastic dominance to mean-risk models: semideviations as risk measures // European journal of operational research. 1999. V. 116. P. 33 – 50.

[Sternin, 2010] Sternin M., Shepelev G. //Generalized Interval Expert Estimates in Decision Making. Doklady Mathematics. Vol. 81, pp. 485 - 486. 2010.

[Shepelyov, 2011] Shepelyov G., Sternin M. Methods for comparison of alternatives described by interval estimations. //Int. J. Business Continuity and Risk Management. Vol. 2, No. 1, pp. 56 – 69. 2011.

[Shepelyov, 2013] Shepelyov G., Sternin M. The applicability of mathematical expectation indicators in comparing interval alternatives. //Advances in Decision Technology and Intelligent Information Systems. Vol. 14, pp. 32-36. 2013.

[Shepelev, 2014] Shepelev G., Sternin M. The adequacy of point criteria during the evaluation and comparison of interval alternatives problems. //Scientific and technical information processing. Vol. 41, No. 6, pp. 404 – 412. 2014.

[Vilensky, 2015] Vilensky P.L., Livshits V.N., Smolyak S.A. Otsenka effektivnosti investitsionnykh proektov. Teoriya i praktika. M.: Poly Print Service. 2015. - in Russian

Author's Information

Gennady Shepelev *– Laboratory Head of Institute for Systems Analysis of Federal Research Center "Informatics and Control" of RAS; e-mail: gis@isa.ru*