MEASUREMENTS IN DECISION-MAKING

Oleg I. Larichev
Institute for Systems Analysis, Russian Academy of Sciences, 9, pr.60 let Octjabrja, Moscow, 117312, Russia

Keywords: decision making; elicitation of preferences; human measurement, expert, decision maker, qualitative measurement, verbal decision analysis.

Contents

1. Introduction
2. Five Different Methodologies in Decision-Making.
2.1. Cost-Benefit Analysis
2.2. Multi-Attribute Utility Theory (MAUT)
2.3. Outranking Relations
2.4. Analytical Hierarchy Approach
2.5. Verbal Decision Analysis
3. Different Operations of Measurement
3.1. Operations with Criteria as Items
3.2. Operations with Alternative Estimation by Criteria
3.3. Operations with alternatives
4. Verbal and Numerical Probabilities
4.1. In Which Form Do People Prefer to Give and Receive the Evaluations?
4.2. The Comparison of Pro and Contra
4.3. Experimental Comparison
4.4. Preference Reversal
4.5. Comparative Verbal Probabilities
5. Importance of Measurement in Practical Decision Tasks
6. Replacement of Qualitative Evaluations by Numbers
7. Sensitivity Check
8. The Correspondence of Measurement and Type of Problem
9. The Base for Qualitative Measurements
10. Case-Study: An Application of the Method ZAPROS for R&D Evaluation
11. Conclusion
Glossary
Bibliography
Biographical Sketch

Summary

Information of human beings-decision-makers and experts is needed for the choice of best variant of a decision. There are different operations of information elicitation from human beings that could be called as human measurement. The operations are estimated from the point of view of possibilities and limitations of human information processing system. The results of different psychological experiments are the base of such estimation. The most reliable operations are of qualitative nature. The qualitative (ordinal) approach to measurement is presented.

1. Introduction

Decision Making is the research discipline devoted to the development of methods helping people in the choice of the best variant (or variants) of problem solution. Follow this definition, one could say that a task of Decision Making exists in the cases where are:

- One or several Decision Makers-human beings responsible for the choice of best decision
- Several variants of decision alternatives

Usually people need decision support when the decision problem is complex: there are big uncertainty and several criteria for the evaluation of alternatives.

For the typical problems in natural sciences and in operations research, the information needed for a problem's solution is given in problem's statement. Contrary to it, additional subjective information—information of a Decision Maker (DM) and experts is required for the solution of a decision-making problem. That is why; any decision-making method has inside some procedures of information elicitation from DM and experts. For example, experts could measure the quality of decision' variants (alternatives) on criteria, DM could measure the relative importance of criteria and so on.

Human beings play the role of measurement devices in different procedures of information elicitation. Accepting technical analogy, one could say that such measurement devices give essential information needed for elimination of uncertainty and making a decision. In the connection with it, it is necessary to discuss the following questions:

- Which kind of measurement is required from people in different decision making methods?
- What are characteristics of a human being as a measurement device?
- To which extent one could rely upon the information given by such measurement device?
- How to develop new decision making methods taking into account the characteristics of such measurement device?

The questions above are of great importance. The research of many psychologists having origin in first bright papers of A. Tversky, P. Slovic, B. Fischhoff, and others demonstrated limited capacity of human information processing system. There is now a lot of evidence that people make errors and contradictions in some operations of information processing. That is why problems of measurement are the most important in Decision Making.

2. Five Different Methodologies in Decision-making

Among different approaches in the development of decision making methods it is possible to select five quite different methodologies.

2.1. Cost-Benefit Analysis

The basic idea of cost-benefit analysis is to evaluate in monetary terms all characteristics of decision problem. In principle, this is an ideal method of analysis, as everything is ideally represented by its economic worth. This approach is widely used. In simple problems where only expenses and profit are important, such approach is very appropriate. If, for example, a person wants to buy a truck to deliver some goods, he/she could estimate future profit and needed spending.

Much more difficult task is to estimate the cost of non-monetary factors. The approach used in such cases could be illustrated by the following example. The municipal counsel of a big city takes the decision to construct the zoo. Such decision could satisfy the several goals: the pleasure of people, additional income for the city, creation of new working places and so on. Only some of the goals could be expressed in monetary terms (for example, the cost of working places). To obtain the general profit, the supposition is made about the possibility to assign the relative importance weights to different goals. The municipal counsel could decide that the pleasure of people is three times more important than creation of new working places. It is taken that the profit from achievement of first goal three times more than the one of achievement of third goal.

Such method is based on the supposition about the complete compensation of values for all goals and ability of human beings to measure the relative importance of goals. Such supposition is very questionable (see below). The severe critics of cost-benefit analysis are connected with great difficulties of measurement for non-monetary factors.

2.2. Multi-Attribute Utility Theory (MAUT)

Multi-attribute utility combines all criteria, but in terms of utility (or value) rather than in monetary terms. The utility of evaluations on the criteria and general utility of an alternative is taken as the measure of human preferences.

There is the strict mathematical theory of utility under multiple criteria. The theory is constructed axiomatically, the general axioms of connectivity and transitivity on a set of alternatives, etc., being complemented by the axioms (conditions) of independence. There exist many such conditions that conceptually define the possibility of comparing alternatives in some criteria, the estimates in other criteria being fixed (at different levels). In the case of certainty, the value function is used instead of the utility function. The multi-criteria utility theory is mainly vectored to the problems where many alternatives justify great efforts required to construct the utility (value) function.

There is big number of different multi-criteria methods based on MAUT. In majority of them, there are two main steps: 1. One criterion utility functions are constructed for each criterion. For this task DM is asked to compare lotteries having different outcomes with different probabilities. 2. The construction of general utility function. DM is to define equivalent points on the planes for the pairs of criteria, to compare the lotteries having as outcomes different combinations of criteria estimates and so on. On the base of such comparison, the relations between the weights of importance for all criteria are established. In the dependence from the results of the comparison, the expression for the general utility function could be obtained. In big majority of cases the general utility function is additive. All measurements are to be made in quantitative way.

In the cases when criteria have natural quantitative expression (for example, distances, expenses, sizes) such approach is appropriate. But in cases of qualitative human values, there are difficulties in measurements.

2.3. Outranking Relations

Another approach to comparing and estimating multi-criteria alternatives—outranking approach is oriented to choosing from a group of alternatives a subgroup of the best ones. The approach has the following two original characteristics: 1. Criteria are regarded as persons (jurors) voting for one or another choice, which explains the special attention paid to each of the criteria whose weights as if reflect the degree of influence of each juror. If an estimate by one criterion is low, then the alternative has a serious defect (negative opinion of one of the jurors). 2. The notion of incomparability of two alternatives is introduced. If the estimates of alternatives to a large measure are contradictory, that is, an alternative is superior in some criteria and inferior in the other criteria, then the contradictions do not compensate anyhow and cannot be compared. This notion is also of extreme practical importance because it identifies alternatives with “contrast” estimates that deserve special consideration.

The major ideas of the majority of methods of outranking approach can be described as follows. To each of the N criteria having numerical or qualitative scales, a number p characterizing its importance is assigned. For any pair of alternatives A and B, a binary relation is constructed according to which A is superior to B under certain values of the agreement and disagreement indices that are defined as follows. The agreement index (with A superior to B) is established from the fact of superiority of the weights of criteria in which A is superior or equal to B over the weights of criteria in which the estimate B is superior to the estimate A. The disagreement index is defined as a function of the most significant difference between the estimates B and A in the criteria where the alternative B is preferable.

In outranking methods, the binary superiority relation is defined in terms of levels of the agreement and disagreement indices. If the agreement index is above the given level and the disagreement one is below it, then the alternative A is declared to be superior to the alternative B. If for these levels alternatives cannot be compared, then they are declared to be incomparable.

It is important to emphasize that for given estimates of alternatives the given levels of agreement and disagreement, where alternatives are comparable, provide an analytical tool to the consultant who can investigate the set of alternatives by defining the levels and gradually reducing the required level of the agreement coefficient and increasing that of the disagreement coefficient. For each given pair of levels, a kernel of non-dominated incomparable (or equivalent) elements is isolated. A smaller kernel can be extracted from it by varying the levels, and so forth. The analyst offers to the DM a whole range of possible solutions to a problem in form of different kernels. A single best alternative can be obtained eventually. The degrees of “violence” to the data characterize here the values of agreement and disagreement indices. The measurement operations performed in outranking methods are the nomination of quantitative criteria weights, definition of different comparability levels and so on.

2.4. Analytical Hierarchy Approach

The method of analytical hierarchy (AH) is based also on a multi-criteria description of the problem. The AH method is oriented to working with a given (usually small) group of multi-criteria alternatives from which best alternative must be isolated.
One can identify four main stages in the AH method:

- Structuring of the problem in form of level hierarchy—from objectives to criteria and from criteria to real alternatives. The elements of each level are listed exhaustively. Sometimes, intermediate levels (for example, of objectives) are introduced.
- The elements of each level are compared pair-wise in the degree of their preference to the DM. A comparison language with nine degrees of superiority ranging from equivalence through weak superiority and so on to very strong superiority is introduced, and a numerical scale ranging from 1 (equivalence) to 9 (very strong superiority) is assigned to it, that is, to each verbal description a certain number is assigned. The results of pair-wise comparisons are represented as a skew-symmetric matrix, and its latent vector, which components characterize "on the average" the degree of superiority of each of the compared elements over other elements, is computed.
- At the lowermost hierarchical level, the real alternatives are compared pair-wise in each criterion, that is, N comparison matrices, where N is the number of criteria, are constructed at this level.
- The index of value of each alternative is established using the method of weighted sums of estimates of criteria where the estimate (coefficient of superiority of the given criterion over other criteria) is multiplied by the weight (coefficient of superiority of the given alternative over the other alternatives in the i-th criterion).

The cause of popularity of the AH method lies not only in its simplicity, but also in that it enables the user to compare real alternatives separately in each criterion, which surely is of practical interest. In the framework of AH method, DM is to measure the degree of preference one item (objective, criterion, alternative) against different item. Although such comparisons are made in qualitative way, the language of comparisons is superimposed (small preference, big preference, great preference and so on). The results of such measurements are transferred into numbers without asking the real preferences of DM.

2.5. Verbal Decision Analysis

Verbal Decision Analysis (VDA) is specially developed for so-called unstructured problems. In such problems criteria are of pure qualitative, subjective nature, especially difficult for formalization and measurement (prestige of an organization, attractiveness of a dress, attitude towards reforms, etc.). The main steps for the methods of VDA are:

- Selection of one from qualitative type of questions for DM preferences elicitation; usually such question is tested in psychological experiments to demonstrate the possibility to receive enough reliable DM answers.
- The special procedure of checking the decision maker's information for the consistency is used. The possibility to make and correct errors in the process of gradual development of a decision rule is provided for DM.
- The qualitative information obtained from DM is used in the same form without any transformation into numbers.

A decision rule is developed on the base of logical transformation of qualitative DM information. The scientific criteria for the justification of a decision method from the family VDA methods are psychological criteria of "decision maker-method" interaction.

3. Different Operations of Measurement

There exist many different normative methods of decision-making, which belong to the groups mentioned above. They correspond to the spectrum of real life problems going from problems with objective mathematical models and reliable scales for quantitative measurement to the ones with subjective description in the terms of qualitative criteria (unstructured problems).

Normative decision methods present quite different requirements to their users such as "assign weights to criteria," "construct the probability distribution of this outcome," etc. To meet these requirements, an individual performs various operations of information processing which can be composite (incorporate other operations) or simple (elementary) operations that are not decomposable into elementary ones. Take, for example, the problem of constructing the utility function by a single criterion exercised within the framework of multi-attribute utility theory (MAUT). It involves a number of similar problems of finding a certainty equivalent for lotteries. The probability distributions are constructed on the base of such operations.

Analysis of different normative techniques described above, enables one to distinguish three groups of information processing operations such as operations with criteria, operations with estimates of alternatives by criteria, and operations with alternatives. An operation is called as elementary if it is not decomposable into simpler operations over to the objects of the same group that is, to criteria, alternatives, and alternative estimates by criteria.

It is possible to collect the results of psychological studies of the degree of human confidence and reliability in exercising one or another operation of information processing. If the data can be collected, then the psychological validity of a normative technique can be characterized in terms of psychological validity of the constituent elementary operations of information processing. The elementary operations would be defined as:

- Complex (C), if the psychological studies show that in performing such operations the decision maker displays many inconsistencies and makes use of simplified strategies, that is, eliminates a number of criteria

- Admissible (A), if the psychological studies show that the decision maker is capable of performing them with small inconsistencies and using complex strategies, that is, combinations of criteria estimates
- Admissible for small size (ASZ), if there are facts testifying that operations are performed rather reliably in case of few objects (criteria, outcomes, alternatives, and multi-criteria estimates), but become increasingly more difficult as their number increases
- Uncertain (U, UC, UA), if the psychological research of these operations is insufficient. Yet, a tentative conclusion on admissibility (UA) or complexity of the operation (UC) can be drawn by reasoning by analogy from already known facts

Table 1 contains the description of elementary operations and their estimates. Each elementary operation is described below in more detail.

No. of operation	**Name of elementary operation**	**Evaluation**
01	OPERATIONS WITH CRITERIA AS ITEMS	
011	Ordering in utility (value)	A
012	Assigning quantitative criteria weights	C
013	Decomposition of complex criterion into simpler ones	ASZ
02	OPERATIONS WITH SEPARATE ALTERNATIVE ASSESSMENTS BY CRITERIA	
021	Assigning a quantitative equivalent to qualitative estimate by a criterion	UC
022	Determination of quantitative equivalent of a lottery	C
023	Qualitative comparison of two estimates taken from two criteria scales	A
024	Determination of quantitative tradeoff value for two criteria estimates	UC
025	Determination of a satisfactory level by one criterion	UA
026	Nomination of probability for criteria estimate	C
03	OPERATIONS WITH ALTERNATIVES AS ITEMS	
031	Comparison of two alternatives viewed as a set of estimates by criteria and selection of the best one	ASZ
032	Comparison of two alternatives viewed as a whole, and selection of the best one	UA
033	Nomination of probabilistic estimates of alternatives	C
034	Attribution of alternatives to decision classes	ASZ
035	Quantitative estimation of utility	C
036	Decomposition of complex alternatives into simple ones	ASZ

037	Qualitative comparison of the probabilities of two alternatives	A

Table 1. The three groups of elementary operations and their estimates.

3.1. Operations with Criteria as Items

Operations 011 and 012 are measurements of comparative importance of criteria to the DM. Operation 011 is studied insufficiently, although there are some publications on subject’s consistency in ranking of criteria. In case of seven criteria with binary estimates, the subjects were shown to rank criteria rather consistently. The subjects rank consistently criteria, which are most important to them, although they tolerate permutations in the ranks of secondary criteria.

Several recent papers showing that the subjects make substantial errors in quantitative measurements of criteria importance. Indeed, the quantitative measurement of weights is an unusual operation for a person hardly realizing its consequences (e.g. whether an insignificant variation in the weight of a criterion can result in choosing another alternative). Although this operation is used by many normative techniques (and is sometimes regarded as “natural”) recent researches show that weights assigned by the subjects cannot be regarded as reliable and stable information.

Operation 013 was studied when constructing the criteria hierarchy frequently employed in MAUT. The results indicate that decomposition is not stable to DM's errors if the number of criteria is considerable. At the same time, it is probable that a complex criterion can be decomposed quite reliably into two or three sub-criteria that are obvious in terms of their meaning.

3.2. Operations with Alternative Estimation by Criteria

Operation 021 is a groundless assignment of arbitrary numbers to the qualitative notions on scales. This operation seems difficult for the decision maker. Just as in assigning quantitative weights, an insignificant variation in numbers can affect the relationship between alternatives. The reliability of preference measurement by lotteries was studied in detail with negative results.

Operation 023 was studied methodically while developing one of the families VDA method-the ZAPROS method. The results indicate that the DM performs it steady and with a small number of inconsistencies.

Operation 024, that is, determination of the quantitative change in the estimate of one criterion that is equivalent to a change in the estimate of another criterion. Unfortunately, no methodic verification of reliability of this operation was carried out. Here again we deal with measurement that is uncustomary to people. As a number of studies shows, operation 025 is a routine human operation of translating criteria into constraints. It is typically exercised while seeking admissible values and admissible (but not optimal) decisions.

-
-
-

Bibliography

Budescu D. and Wallsten T. (1995). *Processing Linguistic Probabilities: General Principles And Empirical Evidence*, *The Psychology Of Learning And Motivation*, Vol. **32**, Academic Press. [This represents the results of the experiments comparing linguistic and numerical evaluation of probabilities.]

Carnap R. (1969) *Philosophical Foundation Of Physics*, London, Basic Books Inc. Publishers. [This is foundation of qualitative measurement in natural sciences.]

Huber B. and Huber O. (1987). *Development Of The Concept Of Comparative Subjective Probability*, *Journal Of Experimental Child Psychology*, **44**, pps. 304–316. [This is the presentation of theory for qualitative probabilities and the results of the experiments.]

Kahneman D., Slovic P., and Tversky A., eds. (1982). *Judgment Under Uncertainty: Heuristics And Biases*. Cambridge: Cambridge University Press. [This is summary of results on human errors, heuristics and biases.]

Keeney R. L. and Raiffa H. (1976). *Decisions With Multiple Objectives: Preferences And Value Tradeoffs*, New York: Wiley. [This is detail presentation of MAUT theory and the examples of its application.]

Larichev O. I., Olson D. L., Moshkovich H. M., and Mechitov A. I. (1995). Numerical Vs. Cardinal Measurements In Multi-attribute Decision Making: How Exact Is Exact Enough? *Organizational Behavior And Human Decision Processes*, **64**(1), 9–21. [This is the presentation of the experiment on the comparison of verbal and quantitative measurement in decision-making.]

Larichev O. I. and Moshkovich H. M. (1997). *Verbal Decision Analysis For Unstructured Problems.* Boston: Kluwer Academic Publishers. [This is the presentation of main ideas and methods of Verbal Decision Analysis.]

Roy B. (1996). *Multi-criteria Methodology for Decision Aiding*, Kluwer Academic Publisher, Dordrecht. [This is presentation of methodology and the methods of outranking approach.]

Saaty T. L. (1980) *The Analytic Hierarchy Process*, New York, McGraw-Hill. [This is presentation of Analytical Hierarchy approach.]

Von Neumann J. and Morgenstern O. (1947). *Theory Of Games And Economic Behavior*. Princeton University Press. [This is presentation of classical utility theory.]

Biographical Sketch

Professor Oleg I. Larichev received his Candidate Degree in Optimal Control Theory in 1965 and his Doctoral Degree in Decision Making in 1975 from Institute for Control and Management Problems, Moscow. His fields of interest are: Decision Making, Artificial Intelligence, Multi-criteria Mathematical Programming, and Psychological Problems of Decision Making. O.I. Larichev is Professor of Moscow Institute for Physics and Technology (Technical University). He is Head of Department in the Institute for Systems Analysis, Russian Academy of Sciences. Prof. O. Larichev published 8 books (two in English) and 180 papers. In the field of Decision Making, Prof. Larichev developed new methodology for the construction of decision methods and decision support systems having psychological and mathematical foundation-Verbal Decision Analysis. In the field of Artificial Intelligence, he proposed a new approach to the study and exact imitation of experts' knowledge and to the construction of Intelligent Tutoring Systems. In 1991 Prof. O. Larichev was elected as Corresponding Member of Russian Academy of Sciences and in 1997-as a Full Member of the Academy. Awarded the Gold Medal by the International Society on Multiple Criteria Decision Making.