The Averaging of Interval Expert Evaluations

N. A. Gaidamakin, S. V. Leont’ev, and A. A. Yalpaev

Received May 14, 2012

Abstract—The features of the expert evaluation of intractable properties (parameters) in the form of interval values on number scales are analyzed. To find a consistent evaluation, two methods for averaging the evaluations in interval form are considered. The first is based on the simple (arithmetical mean) averaging of the interval boundaries and the second is concerned with weighted averaging. It is proposed that the “weighing” of the interval boundaries be implemented according to the interval width according to the following principle: the lower the width of the evaluation interval is, the more qualified the expert evaluation of the property under investigation is and the higher the weight of the boundaries of the corresponding interval under averaging is. When averaging two intervals, an increase in the qualification of a consistent expert evaluation during weighted averaging occurs.

Keywords: expert evaluations, measurement scales, quasinumerical scales, interval evaluations, interval scale, averaging of intervals, weighted averaging of intervals

DOI: 10.3103/S0005105512040061

In many cases, numerical (quantitative) scales are used for the expert evaluations of properties (parameters) of a difficult nature. Individual expert evaluations are formed as the values in corresponding scales (so-called point evaluations), and consistent (group) evaluation is determined by averaging the individual expert evaluations.

At the same time, in estimating intractable properties of objects, systems, and processes, including ones of a qualitative character, experts generally tend to evaluate using ordinal scales (“high,” “average,” and “low”), or assess the value of the analyzed property in the form of some interval on a numerical scale (the properties/parameters vary from $s$ to $2s$), when they are unable to state an exact evaluation on a numerical scale. Consequently, the analysis of the features of expert evaluations as intervals on numerical scales, including the so-called quasinumerical (point) scales, which are widely used for expert evaluations, is a topical problem.

The quasinumerical scales involve point scales, whose base $B$ (a range) is more than or equal to 100. In contrast to small range point scales (3-point, 5-point, or 10-point scales) which are varieties of index scales, experts who make evaluations on 100 point and larger scales are guided by not only the relationships of domination between the objects being estimated by a corresponding property (parameter), but also analyze the values (distances) and relationships between evaluations. Therefore, the evaluations of parameters on such point scales approach the estimates on numerical scales (including absolute ones) in terms of their properties. Since the measurement unit “point” is a universal abstraction of the quantitative evidence of some properties (parameters), such point scales are called quasinumerical.

The important parameters of evaluations in the form of the interval $[s_i, 2s_i]$ are the following:

$$\bar{s}_i = (s_i + 2s_i)/2$$

is the center of the interval;

$$\Delta_i = 2s_i - s_i$$

is the width of the interval.\(^2\)

The interval scale $S_{int}$ is said to be the set of all the intervals $(s_i, 2s_i)$ in some numerical scale $S(\leq, -, B, 0)$, on which the operations of domination ($\leq$), difference ($-$), and powering ($/$), as well as the base $B$ (range) and the zero of the scale, are determined.\(^3\)

The simplest logic for determining the operations ($\leq$), ($-$), and ($/$) in the interval scale involves the corresponding operators out of the centers of the intervals. In this case,

$$[s_i, 2s_i] \leq [s_j, 2s_j], \text{ when } \bar{s}_i \leq \bar{s}_j;$$

$$[s_i, 2s_i] - [s_j, 2s_j] = |\bar{s}_i - \bar{s}_j|;$$

$$[s_i, 2s_i]/[s_j, 2s_j] = \bar{s}_i/\bar{s}_j.$$}

The target aspect of the expert evaluation procedure is the formation of a so-called group (consistent)

\(^2\)The index $i$ indicates the individual evaluation of the $i$th expert.

\(^3\)The term “interval scale” is more correct, but it denotes the type of numerical scale for a point evaluation that conserves the relationship of the intervals between two couples of point evaluations, in literature sources [1].
evaluation [2, 3], which is the result of the aggregation of individual expert evaluations.

The aggregation process is based on the minimization criterion for the sum of the distances of individual evaluations from a group evaluation:

$$
\sum_{n=1}^{N} d(\hat{S}_n, \hat{S}_{\text{group}}) \rightarrow \min.
$$

(1)

Where $\hat{S}_{\text{group}}$ is the group (consistent) expert evaluation; $\hat{S}_n$ is the individual evaluation that is made by the $n$th expert; $d(\hat{S}_n, \hat{S}_{\text{group}})$ is the measure of the distance between the $n$th expert individual and group evaluations; and $N$ is the number of aggregated expert evaluations.

It is known that the different means (arithmetic mean, mean square, Kolmogorov mean, etc.), calculated in terms of corresponding measures (distances), satisfy criterion (1).

To choose the distance between two intervals $d$, it is obvious to use the module of the difference between the interval centers:

$$
d[1s_1, 2s_1], [1s_2, 2s_2] = |\bar{s}_1 - \bar{s}_2|.
$$

(2)

It is an easy matter to see that the magnitudes $d_{ij}$ satisfy the metric requirements (non-negativeness, “triangle rule,” and symmetry).

Thus, we can conclude that the simplest solution for determining the mean on a set of evaluations as intervals $[[s_1, 2s_1], [s_2, 2s_2], …, [s_N, 2s_N]]$, which expresses a group’s expert opinion, is the estimation of the boundaries of the interval of an aggregated (group) evaluation in the form of the mean square of the corresponding set of averaged intervals:

$$
[\hat{1s}_{\text{group}}, \hat{2s}_{\text{group}}] = \left(\frac{1}{N} \sum_{n=1}^{N} 1s_n, \frac{1}{N} \sum_{n=1}^{N} 2s_n\right).
$$

(3)

As noted above, the evaluations of values of analyzed properties in interval form are characterized by another aspect, viz., the interval width ($\Delta_i = 2s_i - 1s_i$). This can be indicated as some quality feature (“qualification”) of the expert evaluation. To put it differently, the narrower the interval of the individual expert evaluation is (the lower the scatter of the values of the estimated property), the more qualified its evaluation is.

It is an easy matter to see that the interval width is the mean square of the corresponding values of the interval width of individual expert evaluations upon obtaining a group interval evaluation using (3):

$$
\bar{\Delta}_{\text{group}} = \frac{1}{N} \sum_{n=1}^{N} \Delta_n.
$$

(4)

However, if the interval width is taken as an individual evaluation quality (the other variant is to consider it as the level of the evaluated qualification of an expert), the contribution of an individual evaluation to the formation of a group evaluation must increase with a decrease in the interval width.

This property of the averaging procedure (3) can be attained by replacing the equal averaging coefficients $(1/N)$ with normalized coefficients $\delta_n = (\sum_{n=1}^{N} \delta_n = 1)$, which grow with a decrease in the $l$ width of the corresponding interval. One such method is the calculation of weighting averaging coefficients $\delta_n$ using the following formula:

$$
\delta_n = \frac{\left(\sum_{n=1}^{N} \Delta_i\right) - \Delta_n}{(N - 1) \sum_{n=1}^{N} \Delta_i},
$$

(5)

while the procedure for obtaining the group (averaging) evaluation in interval form is determined by

$$
(\hat{1s}_{\text{group}}, \hat{2s}_{\text{group}}) = \left(\sum_{n=1}^{N} \delta_n 1s_n, \sum_{n=1}^{N} \delta_n 2s_n\right).
$$

(6)

It is an easy matter to see that the width of the averaging interval from (6) is estimated by

$$
\Delta_{\text{group, weighted}} = \sum_{n=1}^{N} \delta_n \Delta_n,
$$

(7)

where $\Delta_n$ is the width of the interval of the $n$th individual evaluation.

Figure 1 shows examples of simple averaging and averaging weighted by the width intervals of a set of three intervals in different variants: the intervals are approximately of the same width and do not cross each other; the intervals partly meet; the intervals sequentially include each other; and two intervals do not cross each other, but the third, which is narrower than the second, is included in the last one.

Analysis of the results of averaging conducted by (3) and (6) shows that group evaluation in an interval form, which is obtained on the basis of weighted averaging with respect to the interval width, is more correct in comparison with the results of the simple arithmetic mean averaging of the boundaries.

Specifically, the width of the interval corresponding to the consistent (group) evaluation obtained with respect to the width of the averaged intervals is equal to or less than the width of the group evaluation obtained according to the mean square averaging. In addition, the center of the interval of the consistent group evaluation based on the averaged interval weighted with

---

4 When the widths of all the averaged intervals are common.