sciendo
Control and Cybernetics
vol. 51 (2022) No. 4
pages: 501-522
DOI: 10.2478/candc-2022-0027

# Estimating and testing relations and trees on the basis of pairwise comparisons*

by

**Leszek Klukowski**

Systems Research Institute, Polish Academy of Sciences, Newelska 6, 01-447 Warsaw, Poland
Leszek.Klukowski@ibspan.waw.pl

**Abstract:** The paper presents an ample account on author's results, concerning the methods of estimating and testing of relations (equivalence, tolerance, preference) and trees on the basis of multiple pairwise comparisons with random errors, which are presented in a new book, Klukowski (2021c). The book contains also main results presented earlier in Klukowski (2011). The new content pertains mainly to: the preference relation with incomparable elements, trees and tests for verification of estimates of the relations. The estimators and tests have good statistical properties; estimators - consistency with good convergence, tests - known distributions of test statistics (exact, limiting or approximate). Some properties of estimates (precision and convergence) have been determined using the simulation approach. The results here reported gain in significance in view of the renewed interest in practical application of pairwise comparisons in various application domains, which give rise to various concrete data analysis problems.

**Keywords:** pairwise comparisons with random errors, estimation and testing of relations and trees, nearest adjoining order idea

## 1. Introduction

The present paper constitutes succinct account on the recent results, obtained by the author in the area of estimating and testing of relations (the equivalence, tolerance and preference relations) and of trees (the non-directed and directed trees). The empirical basis for the respective considerations and results is here constituted by the multiple pairwise comparisons with random errors. The totality of these results is provided in an extended form in the author's new

*Submitted: October 2022; Accepted: November 2022.

book (Klukowski, 2021c). The book referred to contains many results that are entirely new to the domain, especially concerning the preference relation with incomparable elements and the trees.

Moreover, new, efficient tests for the verification of estimates of the relations are also presented, while similar new tests can also be constructed for the trees. The estimators developed and analysed are based on the idea of the nearest adjoining order (NAO), i.e. the structure, minimizing the difference between a relation or a tree on the one hand, and the results of the pairwise comparisons, with potential errors, on the other. The generic concept of NAO was first introduced by Slater (1961) (see also David, 1988, or Bradley, 1976). The estimates are obtained as optimal solutions of appropriately formulated discrete programming problems (for the sets of elements of small or moderate cardinality) or from similarly oriented, in terms of their objective, heuristic algorithms (for larger sets).

In the case of large sets (when application of heuristic algorithms is suggested) it is also possible to obtain the estimates having good precision. This can be done with the use of statistical procedures, providing for a significant decrease of probabilities of comparison errors (see Klukowski, 2017). The respective precision can be also evaluated on the basis of the criterion function of the corresponding mathematical programming problem – its minimal value is equal to zero.

The pairwise comparisons, constituting the input data, are assumed to take on two basic forms – binary (e.g. direction of preference) and multivalent (e.g. expressing the difference of ranks). The developed estimators and tests have good statistical properties; the estimators display consistency with good convergence, while tests – known distributions of test statistics (exact, limiting or approximate). Some of the properties of estimates (precision and speed of convergence) have been established also with the use of the simulation approach (for the preference relation). The proposed general approach is fully formalized and can be computerized, with moderate computational costs. The majority of the results, presented in the book here mainly referred to, had been initially presented in the articles and conference papers of the author, see, e.g., Klukowski (2017, 2021a, b).

The paper consists of four sections. The second section, following the present one, introduces the theoretical basis of the estimation problem, namely the assumptions about the distributions of errors of pairwise comparisons, as well as the form of the estimators and their statistical properties. In the next, third section the original tests are formulated for the estimates, based on the properties of the true relation, and the tests for the assumptions about comparison errors and for the existence of a relation or a tree in the data set considered. Last section of the paper summarizes the results.

## 2. The estimation problem, assumptions about pairwise comparisons, form of estimators and their properties

### 2.1. General formulation of the estimation problem

The estimation problem can be formulated, in its general form, as follows. We are given a finite set of elements $\mathbf{X} = \{x_1, \ldots, x_m\}$, $(3 \leq m < \infty)$. There exists in the set $\mathbf{X}$ one of the following structures of data:

- the equivalence relation $\mathbf{R}^{(e)}$ (reflexive, transitive, symmetric),

- the tolerance relation $\mathbf{R}^{(\tau)}$ (reflexive, symmetric),

- the preference relation $\mathbf{R}^{(p)}$ (alternative of the equivalence relation and strict preference relation),

- the preference relation $\mathbf{R}^{(i)}$ with incomparable elements,

- the non-directed tree or the directed tree.

Each complete relation (i.e. without incomparable elements) generates some family of subsets $\chi_1^{(h)*}, \ldots,\ \chi_n^{(h)*}\ \ (h \in \{e,\ \tau,\ p\}\,;\ n \geq 2)\,.$

The <u>equivalence relation</u> $\mathbf{R}^{(e)}$ generates the family $\chi_1^{(e)*}, \ldots,\ \chi_n^{(e)*}$ having the following properties:

$$\bigcup_{q=1}^{n} \chi_q^{(e)*} = \mathbf{X}, \tag{1}$$

$$\chi_r^{*(e)} \cap \chi_s^{(e)*} = \{\mathbf{0}\}, \tag{2}$$

where:

$\mathbf{0}$ – the empty set,

$$x_i,\ x_j \in \chi_r^{(e)*} \equiv x_i,\ x_j \text{ - equivalent elements,} \tag{3}$$

$$(x_i \in \chi_r^{(e)*}) \wedge (x_j \in \chi_s^{(e)*}) \equiv x_i, x_j \text{ - non-equivalent elements for } i \neq j,\ r \neq s; \tag{4}$$

The <u>tolerance relation</u> $\mathbf{R}^{(\tau)}$ generates the family $\chi_1^{(\tau)*} \ldots,\ \chi_n^{(\tau)*}$ with the property (1), i.e. $\bigcup_{q=1}^{n} \chi_q^{(\tau)*} = \mathbf{X}$, and the properties:

$\exists\, r,\ s,\ (r \neq s)$ such that $\chi_r^{(\tau)*} \cap \chi_s^{(\tau)*} \neq \{\mathbf{0}\}$,

$$x_i,\ x_j \in \chi_r^{(\tau)*} \equiv x_i,\ x_j - \text{equivalent elements,} \tag{5}$$

$$(x_i \in \chi_r^{(\tau)*}) \wedge (x_j \in \chi_s^{(\tau)*}) \equiv x_i, x_j \text{ – non-equivalent elements for}$$
$$i \neq j \text{ and } (x_i,\, x_j) \notin \chi_r^{(\tau)*} \cap \chi_s^{(\tau)*}, \tag{6}$$

each subset $\chi_r^{(\tau)*}$ $(1 \leq r \leq n)$ includes an element $x_i$ such that

$$x_i \notin \chi_s^{(\tau)*}\ (s \neq r). \tag{7}$$

The preference relation $\mathbf{R}^{(p)}$ generates the family $\chi_1^{(p)*} \ldots,\ \chi_n^{(p)*}$ with the properties (1), (2) and the property:

$$(x_i \in \chi_r^{(p)*}) \wedge (x_j \in \chi_s^{(p)*}) \equiv x_i \text{ is preferred to } x_j \text{ for } r < s. \tag{8}$$

The relations, which are defined by the conditions (1) - (8) can be expressed, alternatively, by the values $T_c^{(h)}((x_i,\, x_j)\ ((x_i,\, x_j) \in \mathbf{X} \times \mathbf{X})$; $h \in \{p, e, \tau\}$, $c \in \{b, \mu\})$, where symbols $b$, $\mu$ correspond to – respectively – the binary and multivalent pairwise comparisons. These values are defined, respectively, as follows:

$$T_b^{(e)}(x_i,\, x_j) = \left\{ \begin{array}{l} 0 \ \text{ if there exists } r \text{ such that } (x_i, x_j) \in \chi_r^{(e)*}, \\ 1 \ \text{ otherwise;} \end{array} \right\} \tag{9}$$

• the values $T_b^{(e)}(x_i\, x_j)$, describing the equivalence relation, being binary, express the fact that a pair $(x_i, x_j)$ either belongs to a common subset or not;

$$T_b^{(\tau)}(x_i,\, x_j) =$$
$$\begin{cases} 0 \ \text{ if there exist } r, s\ (r = s \text{ not excluded}) \text{ such that } (x_i, x_j) \in \chi_r^{(\tau)*} \cap \chi_s^{(\tau)*}, \\ 1 \ \text{ otherwise;} \end{cases} \tag{10}$$

• the values $T_b^{(\tau)}(x_i,\, x_j)$, describing the tolerance relation, being again binary, express the fact that a pair $(x_i, x_j)$ either belongs to any conjunction of subsets (also to the same subset) or not; condition (8) guarantees the uniqueness of the description;

$$T_\mu^{(\tau)}(x_i,\, x_j) = \#\{\Omega_i^{(*)} \cap \Omega_j^{(*)}\}, \tag{11}$$

where: $\Omega_l^{(*)}$ – the set of the form $\Omega_l^{(*)} = \{s | x_l \in \chi_s^{(\tau)*}\}$,
$\#\{\Xi\}$ – the number of elements of the set $\Xi$;

• the values $T_\mu^{(\tau)}(x_i,\, x_j)$, describing the tolerance relation, being multivalent, express the number of subsets of conjunction including both elements; condition (8) guarantees the uniqueness of the description;

$$T_b^{(p)}(x_i,\, x_j) = \left\{ \begin{array}{l} 0 \text{ if there exists } r \text{ such that } (x_i, x_j) \in \chi_r^{(p)*}, \\ -1 \text{ if } x_i \in \chi_r^{(p)*},\ \ x_j \in \chi_s^{(p)*} \text{ and } r < s, \\ 1 \text{ if } x_i \in \chi_r^{(p)*},\ \ x_j \in \chi_s^{(p)*} \text{ and } r > s; \end{array} \right\} \tag{12}$$

• the values $T_b^{(p)}(x_i, x_j)$, describing the preference relation, express the direction of preference in a pair or the equivalence of its elements;

$$T_\mu^{(p)}(x_i, x_j) = d_{ij} \Leftrightarrow x_i \in \chi_r^{(p)*} \, x_j \in \chi_s^{(p)*},$$
$$d_{ij} = r - s(-(m-1) \leq d_{ij} \leq m-1); \quad d_{ij} = 0 \Leftrightarrow (x_i, x_j) \in \chi_r^{(p)*}; \tag{13}$$

• the values $T_\mu^{(p)}(x_i, x_j)$, describing the preference relation, being multivalent, express the difference of ranks of elements $x_i$ and $x_j$; the case of $d_{ij} = 0$ corresponds to the weak form of the relation, while the case of $-1 \geq d_{ij} \geq 1$ corresponds to the strict form of the relation.

The preference relation, including incomparable elements, and the trees (non-directed and directed) cannot be expressed by a family of subsets and are determined only with the use of values $T_b^{(\cdot)}(x_i, x_j)$ and $T_\mu^{(\cdot)}(x_i, x_j)$.

The preference relation including incomparable elements:

$$T_b^{(i)}(x_i, x_j) = \left\{ \begin{array}{l} -1 \text{ if } x_i \text{ precedes } x_j, \\ 1 \text{ if } x_j \text{ precedes } x_i, \\ 2 \text{ if } x_i \text{ and } x_j \text{ incomparable}; \end{array} \right\} \tag{14}$$

• the values $T_b^{(i)}(x_i, x_j)$, describing the preference relation (without equivalent elements) including incomparable elements, express the direction of preference in a pair $(x_i, x_j)$ or incomparability of the respective elements;

$$T_\mu^{(i)}(x_i, x_j) = \left\{ \begin{array}{l} d_{ij} \text{ if elements } x_i \text{ and } x_j \text{ are comparable}, \\ m \text{ if elements} x_i \text{ and } x_j \text{ are incomparable}; \end{array} \right\} \tag{15}$$

• the values $T_\mu^{(i)}(x_i, x_j)$ express the difference of ranks of comparable elements $(-m < d_{ij} < m, \ m \neq 0)$ or incomparability of elements $(d_{ij} = m)$; the difference of ranks of comparable elements can be presented through the digraph – the respective digraph corresponds to the number of edges connecting the elements of a pair.

The non-directed tree:

$$T_b^{(n)}(x_i, x_j) = \left\{ \begin{array}{l} 1 \text{ if elements } x_i \text{ and } x_j \text{ are conected by an edge}, \\ 0 \text{ if elements } x_i \text{ and } x_j \text{ are not connected by an edge}; \end{array} \right\} \tag{16}$$

• the values $T_b^{(n)}(x_i, x_j)$ express either the fact that two nodes $x_i$ and $x_j$ of a tree are connected with an edge or that they are not.

The directed tree:

$$T_b^{(d)}(x_i, x_j) = \begin{cases} -1 \text{ if there exists a path from a node } x_i \text{ to a node } x_j, \\ 1 \text{ if there exists a path from a node } x_j \text{ to a node } x_i, \\ 2 \text{ if there does not exists a path between nodes } x_i \text{ and } x_j; \end{cases} \tag{17}$$

• the values $T_b^{(d)}(x_i, x_j)$ express either the fact that two nodes $x_i$ and $x_j$ of a tree are connected with a path or that they are not, and show the direction of the path.

The values $T_c^{(h)}(x_i,\, x_j)\,(c \in \{b, \mu\}\,,\, h \in \{e, \tau, p, i, n, d\})$ define the relations and trees under consideration in a unique manner.

The form of a relation or a tree is a priori unknown and has to be estimated on the basis of available multiple pairwise comparisons, in binary or multivalent form, given with random errors. Any binary comparison, considered as datum, assumes the values from the set {-1, 0, 1}; it can determine (for instance) the direction of preference in the case of preference relation. Multivalent comparisons assume the values from a broader set of values and can express (for instance) the difference of ranks (the preference relation).

### 2.2. Assumptions concerning pairwise comparisons

Each relation and tree under consideration here is to be estimated on the basis of $N\,(N \geq 1)$ comparisons of each pair $(x_i,\, x_j)$ from the Cartesian product $\mathbf{X} \times \mathbf{X}$; any comparison

$$g_{ck}^{(h)}(x_i,\, x_j)\ (c \in \{b, \mu\}\,,\ h \in \{e, \tau, p, i, n, d\}\,,\ k = 1, \ldots, N)$$

assesses the actual value of $T_c^{(h)}(x_i,\, x_j)$ and can be disturbed by a random error. The following assumptions are made with respect to the comparison errors of the complete relations and non-directed trees (the assumptions concerning the remaining structures of data are slightly different).

**A1**. The relation type (equivalence or tolerance or preference or preference with incomparable elements) or the type of the tree (non-directed or directed) is known, the number of subsets of each complete relation $n$ is unknown (in the case of strict preference relation $n = m$).

**A2**. Any comparison $g_{ck}^{(h)}(x_i, x_j)$ is an evaluation of the value $T_c^{(h)}(x_i, x_j)$, disturbed by a random error. The probabilities of errors: $g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j)$,

$h \in \{e, \tau, p, n\}$ have to satisfy the following assumptions:

$$P(g_{bk}^{(h)}(x_i, x_j) - T_b^{(h)}(x_i, x_j) = 0 \,|\, T_b^{(h)}(x_i, x_j) = k_{bij}^{(h)}) \geq 1 - \delta$$
$$(k_{bij}^{(h)} \in \{-1, 0, 1\}\,, \;\; \delta \in (0, \frac{1}{2})), \tag{18}$$

$$\sum_{r \leq 0} P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}) > 1/2,$$
$$(k_{cij}^{(h)} \in \{0, \pm 1, \ldots, \pm (m-1)\}, \; c \in \{b, \mu\}), \; r - \text{ zero or integer number}, \tag{19}$$

$$\sum_{r \geq 0} P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}) > 1/2,$$
$$(k_{cij}^{(h)} \in \{0, \pm 1, \ldots, \pm (m-1)\}\,, \; c \in \{b, \mu\}) \; r - \text{zero or integer number}, \tag{20}$$

$$P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}) \geq$$
$$P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r + 1 \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}),$$
$$(k_{cij}^{(h)} \in \{0, \pm 1, \ldots, \pm m\} \;\; r > 0), \tag{21}$$

$$P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}) \geq$$
$$P(g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j) = r - 1 \,|\, T_c^{(h)}(x_i, x_j) = k_{cij}^{(h)}),$$
$$(k_{cij}^{(h)} \in \{0, \pm 1, \ldots, m\} \; r < 0), \tag{22}$$

**A3**. The (results of) comparisons,

$$g_{ck}^{(h)}(x_i, x_j) \quad (c \in \{b, \mu\}\,, \, h \in \{e, \tau, p, i, n, d\}\,, \, k = 1, \ldots, N)$$

are independent random variables.

The assumptions A1 - A3 make it possible to determine the distributions of estimation errors of the estimators defined in the paper. However, determination of the exact distributions of multidimensional errors, in an analytic way, is, in fact, unrealizable.

The assumptions A2 – A3 reflect the following properties of distributions of comparisons errors:

• the probability of correct comparison is greater than of the incorrect one - in the case of binary comparisons (inequality (18));

• zero is the median of each distribution of comparison error (inequalities (19) – (20)), in a “sharp” way,

• zero is the mode of each distribution of comparison error (inequalities (18) – (22));

• the set of all comparisons comprises the realizations of independent random variables (the assumption A3);

• the expected value of any comparison error can differ from zero (except a one value distribution); this fact is obvious for binary comparisons.

The assumptions, concerning the distributions of comparison errors are not restrictive. Especially, any error can have non-zero expected value and the probability of errorless result has to satisfy only the mode and median condition. Such assumptions are generally satisfied by the results of statistical tests and of other decision procedures with random errors. The main properties of the estimators are valid under weaker assumptions, and so, especially, the condition stipulating independence of all comparisons can be relaxed so that comparisons of pairs including different elements are independent and comparisons of the same pair (for $N > 1$) are independent.

### 2.3. The essential idea of estimation – minimization of absolute differences with comparisons

The main idea of the NAO estimators, i.e. minimization of (absolute) differences between the values $T_c^{(h)}(x_i, x_j)$, expressing the model of data considered, and the results of the pairwise comparisons $g_{ck}^{(h)}(x_i, x_j)$ with random errors, refers to a well-known statistical principle. However, in the case under consideration, it does not indicate directly the analytical properties, because it is not based on such criteria as maximization of the likelihood function or minimization of the sum of error squares or equality of moments of distributions, etc. (some likelihood properties of the NAO estimates are presented in Thompson and Remage, 1964, for strict preference relation and multiple independent comparisons). In the case under consideration here, the properties of the estimators have been obtained on the basis of differences between the properties of the errorless estimate (i.e. the actual form of the relation or tree) and the estimates, which are different from the errorless one. The properties have been demonstrated by the author on the basis of: the well-known probabilistic inequalities (see Hoeffding, 1963, Chebyshev – for expected value and variance), properties of order statistics (David, 1970) and the convergence of variances of the considered random variables to zero.

The theoretical properties have been also complemented through the simulation survey, because speed of convergence of estimators depends on many features of the problem, mainly on the number of elements considered, structure of data (relation or tree), number of comparisons $N$, and variance of distribution of comparison errors.

Two forms of estimators are examined, namely those based on the total sum of absolute differences between comparisons $g_{ck}^{(h)}(x_i,\, x_j)$ and values $T_c^{(h)}(x_i,\, x_j)$, and the sum of absolute differences between medians $\breve{g}_c^{(h)}(x_i,\, x_j)$ from multiple comparisons of each pair and values $T_c^{(h)}(x_i,\, x_j)$. The second form allows for the reduction of the number of variables of the discrete programming problem, which has to be solved for obtaining of estimates; its efficiency is, however, slightly lower.

The estimates based on the total sum of differences, denoted

$$\hat{T}_c^{(h)}(x_i,\, x_j)\,(h \in \{e, \tau, p, i, n, d\}\,,\, < i, j > \in R_m)$$

or

$$\hat{\chi}_1^{(h)} \ldots,\; \hat{\chi}_{\hat{n}}^{(h)}\;\; (h \in \{e, \tau, p\})$$

(for complete relations), are obtained as the optimal solutions of the minimization problem:

$$\min_{F_{\mathbf{X}}^{(h)}} = \left\{ \sum_{<i,j>\in R_m} \sum_{k=1}^{N} \left| g_{ck}^{(h)}(x_i,\, x_j) - t_c^{(h)}(x_i,\, x_j) \right| \right\} \tag{23}$$

where:

$F_{\mathbf{X}}^{(h)}$ – the feasible set, i.e. the set of all models of data (relations or trees) of the $h^{th}$ type in the set $\mathbf{X}$,

$t_c^{(h)}(x_i,\, x_j)$ - the values describing any relation or tree of the $h^{th}$ type,

$R_m$ – the set of the form $R_m = \{< i, j > | 1 \leq i, j \leq m,\, j > i\}$,

$g_{ck}^{(h)}(x_i,\, x_j)$ – $k^{th}$ pairwise comparison of the pair $(x_i,\, x_j)$.

The estimate, based on $\breve{g}_c^{(h)}(x_i, x_j)$, i.e. the medians from the comparisons $g_{c,1}^{(h)}(x_i, x_j), \ldots, g_{cN}^{(h)}(x_i, x_j)(N = 2\iota + 1;\, \iota \geq 1)$, $\iota$ being an integer number, denoted $\breve{T}_c^{(h)}(x_i,\, x_j)$ or $\breve{\chi}_1^{(e)}, \ldots,\, \breve{\chi}_{\breve{n}}^{(e)}$ $(h \in \{e, \tau, p\})$ (for complete relations), is obtained as the optimal solution of the following minimization problem:

$$\min_{F_{\mathbf{X}}^{(h)}} = \left\{ \sum_{<i,j>\in R_m} \left| \breve{g}_c^{(h)}(x_i,\, x_j) - t_c^{(h)}(x_i,\, x_j) \right| \right\}, \tag{24}$$

where:

$\breve{g}_c^{(h)}(x_i, x_j)$ - the median of comparisons $g_{c,1}^{(h)}(x_i, x_j), \ldots, g_{cN}^{(h)}(x_i, x_j)$.

The solutions of these problems can be obtained with the use of the discrete mathematical programming procedures (see, e.g., Garfinkel and Nemhauser, 1972; Hansen, Jaumard and Sanlaville, 1994) or using appropriate heuristic algorithms.

### 2.4. Properties of estimators - consistency

The analytical properties of the estimators have mainly asymptotic character, i.e. for $N \to \infty$ or/also for $m \to \infty$. The properties guarantee the basic feature of the estimators - consistency. It is clear that errorless estimates can be also obtained for finite $N$, with probability close to one, because of the exponential type of convergence of the estimators to the actual form of relation or tree. The precision level is not the same for both estimators considered; in general, the approach based on medians from comparisons is less efficient, but requires lower computational effort of optimization algorithms. The simulation approach allows to determine the minimal number of $N$, providing the necessary precision of an estimate, for given relation form and distributions of comparison errors. It can be also useful for verification of some hypotheses, concerning the estimates.

The analytical properties of the estimators are based on probabilistic properties of differences (random variables) between pairwise comparisons $g_{ck}^{(h)}(x_i, x_j)$ and (actual) values $T_c^{(h)}(x_i, x_j)$. It has been proven that the variables corresponding to the actual relation or tree (i.e. $g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j)$) have different properties than the variables corresponding to any other relations or trees $\tilde{T}_c^{(h)}(x_i, x_j)$ ($\tilde{T}_c^{(h)}(x_i, x_j) \not\equiv T_c^{(h)}(x_i, x_j)$). The following main results have been obtained:

(i) the expected value of the sum of random variables $\left|g_{ck}^{(h)}(x_i, x_j) - T_c^{(h)}(x_i, x_j)\right|$ (for $k = 1, \ldots, N$ and $< i, j > \in R_m$), corresponding to actual relation or tree is lower than the expected values of the sum of variables $\left|g_{ck}^{(h)}(x_i, x_j) - \tilde{T}_c^{(h)}(x_i, x_j)\right|$ corresponding to any other (incorrect) relations or trees;

(ii) the variances of the random variables expressing sum of (absolute) differences between comparisons and the relation or tree, both actual and different than actual (in the form $T_c^{(h)}(x_i, x_j)$, $\tilde{T}_c^{(h)}(x_i, x_j)$), divided by the number of comparisons $N$, in the case of the estimator (23) (in the form of sum of differences), converge to zero for $N \to \infty$;

(iii) the probability of the event that the sum of random variables corresponding to actual relation or tree assumes a value lower than any variable corresponding to other than actual relation or tree converges to one as $N \to \infty$ (see Klukowski, 2011); the property guarantees the consistency of the estimator; the speed of convergence, resulting from Hoeffding (1963) inequalities, guarantees good efficiency of the estimates, e.g. for the estimator of the equivalence relation the inequality assumes the form:

$$P(\sum_{<i,j>\in R_m} \sum_{k=1}^{N} \left| g_{bk}^{(e)}(x_i, x_j) - T_b^{(e)}(x_i, x_j) \right| <$$

$$< \sum_{<i,j>\in R_m} \sum_{k=1}^{N} \left| g_{bk}^{(e)}(x_i, x_j) - \tilde{T}_b^{(e)}(x_i, x_j) \right|) \geq$$

$$\geq 1 - \exp\{-2N(\frac{1}{2} - \delta)^2\};$$

(iv) the estimator based on medians from multiple comparisons has similar properties, but has slightly lower evaluation of the corresponding probability:

$$P(\sum_{<i,j>\in R_m} \left| \breve{g}_b^{(e)}(x_i, x_j) - T_b^{(e)}(x_i, x_j) \right| <$$

$$< \sum_{<i,j>\in R_m} \left| \breve{g}_b^{(e)}(x_i, x_j) - \tilde{T}_b^{(e)}(x_i, x_j) \right|) \geq 1 - 2\exp\{-2N(\frac{1}{2} - \delta)^2\};$$

the difference with respect to the evaluation from point (iii) above is insignificant for large $N$ and probability $\delta$ close to zero.

The properties (i) - (iv) provide for the reasonability of the estimators. They have been further complemented with some properties valid for relaxed assumptions and results of simulation approach (see Klukowski, 2011). Especially, the property of consistency can be proven also for the case of $m \to \infty$, under some additional assumptions.

## 2.5. Some examples

This short section provides a couple of illustrative examples, making the considerations presented better understood.

• Binary comparisons $\{g_{bk}^{(p)}(x_i, x_j);\ (1 \leq i, j \leq 5;\ j \neq i)\ ,\ N = 1\}$ from the strict preference relation have the following form (discussed in David, 1988, Section 2.2):

| $g_{b,1}^{(p)}(x_i, x_j)$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ |
|---|---|---|---|---|---|
| $x_1$ | $\times$ | 1 | -1 | -1 | -1 |
| $x_2$ | -1 | $\times$ | -1 | -1 | 1 |
| $x_3$ | 1 | 1 | $\times$ | -1 | -1 |
| $x_4$ | 1 | 1 | 1 | $\times$ | -1 |
| $x_5$ | 1 | -1 | 1 | 1 | $\times$ |

The NAO solution (estimate) of the problem has the form:

$\hat{\chi}_1^{(p)} = \{x_1\}$, $\hat{\chi}_2^{(p)} = \{x_2\}$, $\hat{\chi}_3^{(p)} = \{x_3\}$, $\hat{\chi}_4^{(p)} = \{x_4\}$, $\hat{\chi}_5^{(p)} = \{x_5\}$,

and the value of the criterion function equal to 1. The approach based on other rules, especially scoring methods (see David, 1988, section 6.1), does not indicate such solution and has no properties of the NAO estimator; moreover, different scorings indicate different solutions.

• Medians from binary comparisons $\{\breve{g}_b^{(p)}(x_i, x_j);\ (1 \le i, j \le 4;\ j \ne i)\ ,\ N = 15\}$, from strict preference relation, have the following form (discussed in David, 1988, section 6.3):

| $\breve{g}_b^{(p)}(x_i, x_j)$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $x_1$ | $\times$ | 1 | 1 | 1 |
| $x_2$ | -1 | $\times$ | -1 | 1 |
| $x_3$ | -1 | 1 | $\times$ | 1 |
| $x_4$ | -1 | -1 | -1 | $\times$ |

The NAO solution (estimate) of the problem has the form:

$\breve{\chi}_1^{(p)} = \{x_4\}$, $\breve{\chi}_2^{(p)} = \{x_2\}$, $\breve{\chi}_3^{(p)} = \{x_3\}$, $\breve{\chi}_4^{(p)} = \{x_1\}$,

and the value of the criterion function (24) equal to zero. The NAO approach has properties corresponding to the evaluation of probability of the errorless estimate based on Hoeffding inequality (see (iv) above) for the medians from 15 comparisons of each pair.

• Medians from three binary comparisons (based on results of three statistical tests) $\{\breve{g}_b^{(\tau)}(x_i, x_j);\ (1 \le i, j \le 7;\ j \ne i)\ ,\ N = 3\}$, from the tolerance relation, have the form (see Klukowski, 2006):

| $g_{b,1}^{(p)}(x_i, x_j)$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ |
|---|---|---|---|---|---|---|---|
| $x_1$ | × | 1 | 0 | 1 | 1 | 0 | 0 |
| $x_2$ | | × | 0 | 1 | 0 | 1 | 0 |
| $x_3$ | | | × | 1 | 1 | 0 | 0 |
| $x_4$ | | | | × | 1 | 1 | 1 |
| $x_5$ | | | | | × | 1 | 0 |
| $x_6$ | | | | | | × | 1 |
| $x_7$ | | | | | | | × |

The NAO estimate of the problem has the form of three relations with the value of criterion function (23) equal two:

the first one $\check{\chi}_1^{(\tau,1)} = \{x_1, x_3, x_6, x_7\}$, $\check{\chi}_2^{(\tau,1)} = \{x_2, x_3, x_5, x_7\}$, $\check{\chi}_3^{(\tau,1)} = \{x_4\}$,

the second $\check{\chi}_1^{(\tau,2)} = \{x_1, x_3, x_6, x_7\}$, $\check{\chi}_2^{(\tau,2)} = \{x_2, x_5, x_7\}$, $\check{\chi}_3^{(\tau,2)} = \{x_4\}$,

the third: $\check{\chi}_1^{(\tau,3)} = \{x_1, x_3, x_6\}$, $\check{\chi}_2^{3} = \{x_2, x_3, x_5, x_7\}$, $\check{\chi}_3^{(\tau,3)} = \{x_4\}$.

The first estimate has the maximal number of elements, $x_3, x_7$, included in the intersection; the maximal intersection consists of two subsets, $\check{\chi}_1^{(\tau,l)} \cap \check{\chi}_2^{(\tau,l)}$ $(l = 1, 2, 3)$ .

## 3. Verification of estimates and assumptions concerning pairwise comparisons

### 3.1. General approach to verification of relations and trees

The estimates, obtained with the use of the NAO concept can be verified with the use of statistical tests, based on assumptions relative to comparison errors, i.e. A1 – A3 from Section 2. Typically, the null hypothesis states that the estimate is the same as the true relation or tree, its alternative – that it is different. The tests, meant for such purposes are presented for different structures of data, under different assumptions, in the literature of the subject, see, e.g., David (1988), or Gordon (1999).

All of the assumptions, A1, A2, A3, regarding the true data structure and the distributions of comparison errors, can be also verified. First, it is necessary to verify the assumptions concerning the distributions of comparison errors. This can be done with the use of known tests for: independence of random variables, unimodal distribution of these variables and values of mode and median. Verification of existence of relation type or tree can be done on the basis of properties of distributions of comparisons corresponding to the true relation. Test statistics have to be appropriate to the data – binary or multivalent, i.e. having usually binomial, multinomial or limiting: Gaussian, $t$-Student and *chi-square* distributions.

### 3.2. Verification of estimates of individual relation or tree

Let us illustrate these considerations by the examples of tests, constructed mainly for complete relations:

- tests for relation type – equivalence or tolerance – for binary comparisons,
- tests for verification of estimate of equivalence relation for binary comparisons,
- tests for verification of estimate of tolerance relation for binary comparisons,
- tests for verification of estimate of tolerance relation for multivalent comparisons,
- tests for verification of estimate of preference relation for binary comparisons,
- tests for verification of estimate of the preference relation for multivalent comparisons,
- tests for assumptions about comparisons and verification of existence of relation or tree.

The tests for: equivalence relation and relation type (equivalence or tolerance), for binary comparisons, are based on the binomial distribution; for this distribution there exists a randomized most powerful test.

The null hypothesis of the test for relation type states that equivalence (tolerance) relation is true, an alternative - that tolerance (equivalence) relation is true. The tests are based on the random variables

$$\gamma_k^{(\tau)}(x_i,x_j) = \left|g_{bk}^{(\bullet)}(x_i,x_j) - \hat{T}_{bN}^{(e)}(x_i,x_j)\right| - \left|g_{bk}^{(\bullet)}(x_i,x_j) - \hat{T}_{bN}^{(\tau)}(x_i,x_j)\right|$$

for

$$\{(x_i,x_j)|\hat{T}_{bN}^{(e)}(x_i,x_j) \neq \hat{T}_{bN}^{(\tau)}(x_i,x_j)\}.$$

The properties of such variables have been determined for the case of errorless estimate of relation (equivalence or tolerance). Under the null hypothesis that the tolerance relation is true, the variable $\gamma_k^{(\tau)}(x_i,x_j)$ assumes values from the set $\{-1,\,1\}$ with probabilities, respectively, $\delta$ and $1-\delta$, i.e. it has the binomial distribution. Under the alternative hypothesis, the probabilities have the reversed values; thus the test based on the binomial distribution with two simple hypotheses can be applied. In the case of estimates with errors (i.e. not errorless) the test may be not valid. Therefore, effective application of the tests requires the value of probability of errorless estimate to be close to one. The probability can be determined via the simulation approach or can be evaluated analytically. This probability is applied for the correction of the probabilities of the first type and the second type errors in binomial test. The corrected probabilities of the test are higher than resulting from the true relation form, i.e. the non-zero probability of incorrect estimation make the properties of the tests worse. The tests presented in the book, Klukowski (2021c), are not the same as those in the earlier approach, presented in Klukowski (2011).

The subsequent test, for the estimates of the tolerance relation, is proposed in the form:

$$H_0 : \hat{\chi}_1^{(\tau)}, \ldots, \hat{\chi}_{\hat{n}}^{(\tau)} \equiv \chi_1^{(\tau)*}, \ldots, \chi_n^{(\tau)*}$$

and

$$H_1 : \hat{\chi}_1^{(\tau)}, \ldots, \hat{\chi}_{\hat{n}}^{(\tau)} \not\equiv \chi_1^{(\tau)*}, \ldots, \chi_n^{(\tau)*}, \tag{25}$$

and verified by the set of partial hypotheses:

$$H_{01} : \hat{\chi}_1^{(\tau)} = \chi_1^{(\tau)*},\ H_{11} : \hat{\chi}_1^{(\tau)} \neq \chi_1^{(\tau)*}, \ldots,\ \mathrm{H}_{0,\hat{n}} : \hat{\chi}_{1,\hat{n}}^{(\tau)}$$
$$= \chi_{1,n}^{(\tau)*} \ldots,\ \mathrm{H}_{1,\hat{n}} : \hat{\chi}_{1,\hat{n}}^{(\tau)} \neq \chi_{1,n}^{(\tau)*}. \tag{26}$$

Each of the partial hypotheses is verified separately with the use of the statistic having binomial distribution with known parameters or the limiting Gaussian distribution. The proposed statistic has the form (for each partial hypothesis, i.e. for any subset $\hat{\chi}_q^{(\tau)}$ $(1 \leq q \leq \hat{n})$):

$$\eta_q = \frac{1}{N\nu_q} \sum_{<i,j> \in S_{\hat{q}}} \sum_{k=1}^{N} g_{bk}^{(\tau)}(x_i, x_j) + \frac{1}{N\upsilon_q} \sum_{(x_i \in \hat{\chi}_q^{(\tau)}) \wedge (x_j \notin \hat{\chi}_q^{(\tau)})} \sum_{k=1}^{N} (1 - g_{bk}^{(\tau)}(x_i, x_j)), \tag{27}$$

where:

$$S_{\hat{q}} = \{< i, j > | x_i \in \hat{\chi}_q^{(\tau)}, \quad x_j \in \hat{\chi}_q^{(\tau)} \quad (j > i)\},$$

$\nu_q$ – number of comparisons for pairs of elements included in the subset $\hat{\chi}_q^{(\tau)}$,
$\upsilon_q$ – number of pairs $(x_i, x_j)$ such, that $x_i \in \hat{\chi}_q^{(\tau)}$ and $\hat{T}_b^{(\tau)}(x_i x_j) = 1$.

The second sum in (27) corresponds to these pairs, in which the first element $x_i \in \hat{\chi}_q^{(\tau)}$, while the second $x_j \notin \hat{\chi}_q^{(\tau)}$. The expected value and variance of the statistic assume, under the null hypothesis, the form:

$E(\eta_q) = \delta, \quad Var(\eta_q) = \frac{\delta(1-\delta)}{N}(\frac{1}{\nu_q} + \frac{1}{\upsilon_q}).$

Under the alternative hypothesis the expected value is higher, that is, $E(\eta_q) > \delta$, while the variance is the same. The partial tests allow for rejecting the null hypothesis (25) and show the subset(s) incorrectly estimated (the significance level can be assumed separately for any partial hypothesis). The form of the partial hypotheses can be modified, e.g. they can show in a more detailed manner the source of estimation error, based on distribution of each sum in the statistic (27) separately.

The tests for the tolerance relation, in the case of multivalent comparisons, are based on multinomial distribution. The null and alternative hypotheses have similar general form, as in the case of binary comparisons, namely:

$$H_0 : \ \hat{\chi}_1^{(\tau)}, \ldots, \hat{\chi}_{\hat{n}}^{(\tau)} \equiv \chi_1^{(\tau)*}, \ldots, \chi_n^{(\tau)*} \ and \ H_1 : \hat{\chi}_1^{(\tau)}, \ldots, \hat{\chi}_{\hat{n}}^{(\tau)} \not\equiv \chi_1^{(\tau)*}, \ldots, \chi_n^{(\tau)*}. \tag{28}$$

This test is verified also by the set of partial hypotheses, in the form compatible with the multivalent comparisons:

$H_{00} : T_\mu^{(\tau)}(x_i, x_j) = 0$ for all pairs $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = 0 \ (j > i)$,

$H_{10} : T_\mu^{(\tau)}(x_i, x_j) \neq 0$ for at least one pair $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = 0$,

$H_{01} : T_\mu^{(\tau)}(x_i, x_j) = 1$ for all pairs $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = 1 \ (j > i)$,

$H_{11} : T_\mu^{(\tau)}(x_i, x_j) \neq 1$ for at least one pair $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = 1$,

$H_{02} : T_\mu^{(\tau)}(x_i, x_j) = 2$ for all pairs $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = 2 \ (j > i), H_{12} : T_\mu^{(\tau)}(x_i x_j) \neq 2$ for at least, one pair $(x_i x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i x_j) = 2$,

...

$H_{0\kappa}: T_\mu^{(\tau)}(x_i, x_j) = \kappa$ for all pairs $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = \kappa\,(j > i)$,

$H_{1\kappa}: T_\mu^{(\tau)}(x_i, x_j) \neq \kappa$ for at least one pair $(x_i, x_j)$ satisfying the condition $\hat{T}_\mu^{(\tau)}(x_i, x_j) = \kappa$, (29)

where: $\kappa\,(\kappa \leq \hat{n})-$ maximum number of subsets in the intersection occurring in an estimate $\hat{\chi}_1^{(\tau)}, \ldots, \hat{\chi}_{\hat{n}}^{(\tau)}$.

The verification of these hypotheses is based on two well-known tests, i.e. the tests for the mode (see Domański, 1990) and median of multinomial distribution. The test for mode requires the knowledge of distributions of comparison errors, but can be also applied, in an approximate way, in the case of unknown distributions. The actual distributions have to be replaced by an estimate (if possible) or by the so called quasi-uniform distributions, with maximal possible variance (see Klukowski, 2011). The test for median does not require the exact form of distributions. The significance levels of both tests can be determined separately for each partial hypothesis. The results of the partial tests confirm the obtained estimate or indicate its incorrect elements.

The tests for an estimate of preference relation, for binary comparisons, are constructed separately for the strict and weak forms. Moreover, a test has been also developed that allows for detection of the proper relation form – strict or weak.

For the strict form of the relation two tests have been developed:

- for verification of the null hypothesis that $\hat{\chi}_1^{(p)}, \ldots, \hat{\chi}_n^{(p)}$ is the same as $\chi_1^{(p)*}, \ldots, \chi_n^{(p)*}$ $(n = m)$, with $H_1$ stating that it is not the same,
- for verification of null hypothesis that difference of ranks of any two elements $(x_i, x_j)\,(j \neq i)$ is equal $r - i$ (assuming $x_i \in \hat{\chi}_i^{(p)}$ $(i = 1, \ldots, n)$).

The first hypothesis is verified with the use of random variables $g_k(x_i, x_j) + g_k(x_{m-i+1}, x_{m-j+1})\,(j \neq i; k = 1, \ldots, N))$ having multinomial distributions with the set of values $\{-2, 0, 2\}$. Three properties of these variables are examined: expected value equal to zero, median equal to zero and symmetry of the distribution. The test for expected value is based on parameters (expected value and variance) of multinomial or limiting Gaussian distribution. The tests for median and symmetry are non-parametric. Rejection of the null hypothesis indicates the incorrect elements of the obtained estimate, the opposite result confirms an estimate.

The test for difference of ranks is based on the statistic having binomial distribution, with parameters determined in Klukowski (2021c). Thus, exact binomial or limiting Gaussian distribution can be applied. The test statistic is the average of zero-one random variables, depending on the parameter $\delta$; the expected value and variance of the average are equal, respectively:

$(1-\delta)^2+\delta^2$ and $\frac{2}{(m-2)N}\delta(1-\delta)((1-\delta)^2+\delta^2)$.

It is rational to use the test for the fixed value of difference $\frac{m}{2}$ ($m$ even) and the sequential pairs of estimate $x_1, x_{m/2}, \ldots, x_{m/2}, x_m \quad (x_i \in \hat{\chi}_i)$; it verifies the whole form of the obtained estimate. Rejection of the null hypothesis indicates incorrect value of difference of rank, the opposite result confirms the verified value of differences.

The tests for an estimate of weak form of the preference relation are constructed on the basis of properties of the random variables (differences):

$$g_{bk}^{(p)}(x_i,x_j)-g_{bk}^{(p)}(x_r,x_j) \quad (j\neq i,r;\ k=1,\ldots,N). \tag{30}$$

The null hypothesis states that the estimate $\hat{\chi}_1^{(p)},\ldots,\hat{\chi}_{\hat{n}}^{(p)}$ is the same as the true relation $\chi_1^{(p)*},\ldots,\chi_n^{(p)*}$, the alternative – that it is not the same. In the case when the elements $x_i$ and $x_r$ belong to the same subset, i.e. $(x_i,x_r)\in\chi_q^{(p)*}$ $(1\leq q\leq n)$, the differences (30) have the following properties:

- multinomial distribution over the set $\{-2,-1,0,1,2\}$,
- expected value, mode and median equal to zero,
- symmetry of distribution around zero.

In the opposite situation (the two elements belong to different subsets), the properties are not true. Thus, verification of an estimate examines the properties. It can be done with the use of well-known tests based on the exact or limiting or non-parametric distributions: binomial, multinomial, limiting $t$-Student, Gaussian, *chi*-square. Acceptance of all null hypotheses validates of an estimate, rejection – shows the questionable features.

Tests for detection of weak or strict form of the preference relation are useful for the problem of selecting the best (single) element of the set. The idea of these tests is close to that of the ones applied to equivalence and tolerance relations in the case of binary comparisons. The test statistic is a function of differences between the comparisons $g_{bk}^{(p)}(x_i, x_j)$ and an estimate (based on sum of inconsistencies) of both types of the relations: strict $\hat{T}_b^{(ps)}(x_i, x_j)$ or weak $\hat{T}_b^{(pw)}(x_i, x_j)$. The results of comparisons $g_{bk}^{(p)}(x_i, x_j)$ are assumed to take on three values $\{-1, 0, 1\}$ (with equivalency of elements), but under the assumption that all probabilities of correct comparison are greater than $\frac{1}{2}$. The strict form of the estimate is obtained as the optimal solution with the feasible set satisfying the condition $n=m$. The weak form is obtained for $n<m$; therefore, the two estimates are not the same. The test for the null hypothesis: $\mathrm{H}_0$: the strict form is true vs $\mathrm{H}_1$: the weak form is true, is based on the random

variables:

$$\zeta_{ijk}^{(ps)} = \left\{ \begin{array}{l} 0 \text{ for } \hat{T}_b^{(ps)}\left(x_i, x_j\right) = g_{bk}^{(p)}\left(x_i, x_j\right) \text{ and } \hat{T}_b^{(ps)}\left(x_i, x_j\right) \neq \hat{T}_b^{(pw)}\left(x_i, x_j\right); \\ 1 \text{ for } \hat{T}_b^{(ps)}\left(x_i, x_j\right) \neq g_{bk}^{(p)}\left(x_i, x_j\right) \text{ and } \hat{T}_b^{(ps)}\left(x_i, x_j\right) \neq \hat{T}_b^{(pw)}\left(x_i, x_j\right). \end{array} \right\}$$

having zero-one distributions with known parameters in the set

$$\{(x_i, x_j) | \hat{T}_b^{(ps)}\left(x_i,\, x_j\right) \neq \hat{T}_b^{(pw)}\left(x_i,\, x_j\right)\}.$$

Approval of the null hypothesis indicates the strict form of the relation, while the opposite result – the weak form.

### 3.3. Tests for assumptions as to pairwise comparisons and existence of relation or tree

The tests, related to the assumptions, concerning distributions of pairwise comparisons and verification of existence of relation or tree provide the basis of rationality of the estimation. They allow for the rejection of these cases, when comparisons have been generated in a random way or when some other data structure characterizes them.

The following assumptions about comparison errors have to be verified:

- independence of the whole set of comparisons or its part (comparisons of different pairs of elements, i.e. $(x_i, x_j)$ and $(x_r, x_s)$ $(r \neq i, j;\ s \neq i, j)$),
- unimodal distributions of comparison errors,
- mode and median of the distributions of errors equal zero.

These properties can be verified with the use of well-known tests; some limitations concern only the independence of comparison errors in the case of multivalent comparisons with unknown distributions. However, such comparisons can be transformed into the binary form.

Verification of existence of the relation or tree in the set $\mathbf{X}$ has to be done after the positive results of the tests for verification of properties of distributions of comparisons errors. Now, the null hypothesis $\mathrm{H}_0$ assumes the following form: the relation or tree under consideration exists in the set $\mathbf{X}$, the alternative hypothesis $\mathrm{H}_1$ – the relation does not exist in the set $\mathbf{X}$. The non-existence of the relation can mean either randomness of comparisons or equivalence of all elements of the set or existence of some other data structure. The general approach to verification of such hypothesis can be based on the differences between the comparisons $g_{\mu k}^{(h)}(x_i,\, x_j)$ and the estimates $\hat{T}_b^{(h)}(x_i,\, x_j)$ and/or $\check{T}_b^{(h)}(x_i,\, x_j)$.

The basis for the null hypothesis for binary comparisons is constituted by the sums of differences between comparisons and estimates, i.e.:

$$\hat{\varrho}_b^{(h)} = \sum_{<i,j>\in R_m} \sum_{k=1}^{N} \left| g_{bk}^{(h)}\left(x_i,\, x_j\right) - \hat{T}_b^{(h)}\left(x_i,\, x_j\right) \right| \tag{31}$$

or

$$\check{\varrho}_b^{(h)} = \sum_{<i,j>\in R_m} \left| g_{bk}^{(h)}\left(x_i,\, x_j\right) - \check{T}_b^{(h)}\left(x_i,\, x_j\right) \right| \tag{32}$$

In the case of known $\delta$ and $\delta_{me}$ ($\delta_{me}$ – the error, corresponding to the median from the comparisons of each pair), equal for each comparison, and the errorless estimate ($\hat{T}_b^{(h)}\left(x_i,\, x_j\right)$ or $\check{T}_b^{(h)}\left(x_i,\, x_j\right)$), the expected value and the variance of the sums (31), (32) assume the forms:

$$E\left(\hat{\varrho}_b^{(h)}\right) = \frac{1}{2}m(m-1)N\delta, \tag{33}$$

$$E\left(\check{\varrho}_b^{(h)}\right) = \frac{1}{2}m(m-1)\delta_{me}, \tag{34}$$

$$Var\left(\hat{\varrho}_b^{(h)}\right) = m(m-1)N\delta(1-\delta), \tag{35}$$

$$Var\left(\check{\varrho}_b^{(h)}\right) = m(m-1)\delta_{me}(1-\delta_{me}). \tag{36}$$

The equalities (31) or (32), with properties (33) – (36), can be used as the basis for the null hypothesis, confirming the assumed model of data; the test statistic has the binomial or limiting Gaussian distribution. The probabilities of errors in the test have to be corrected with the use of the probability of errorless estimate. Approval of the null hypothesis confirms the assumed model of data, while its rejection suggests the incorrect form.

In the case of individual relations, e.g. strict preference relation and multiple comparisons, some other tests can be also applied. Especially, the positive correlation of estimates of ranks of elements $x_i$ ($i = 1, \ldots, m$), obtained on the basis of comparisons $g_{\upsilon k}^{(p)}\left(x_i, x_j\right)$ and $g_{\upsilon l}^{(p)}\left(x_i, x_j\right)$ ($\upsilon \in \{b, \mu\}, k \neq l;\ < i, j >\in R_m$), can be verified; significant positive correlation confirms the form of relation. The basis for verification of this fact is the well-known Spearman test, with the null hypothesis stating that the correlation of rankings for any $k \neq l$ is equal zero and the alternative hypothesis stating that it is positive. Moreover, it is also possible to verify the hypothesis as to the positive correlation (concordance) of the whole matrix of all ranks (see Raghavachari, 2004).

The tests for the assumptions and the existence of relations or trees allow for a versatile and effective verification of the obtained estimates. In the case of negative result of verification it is possible to detect the sources of errors or to reject an incorrect element of the estimate.

## 4. Summary and conclusions

The estimates, based on the idea of the nearest adjoining order, have good statistical properties and can be efficiently determined. The estimates, which satisfy all the tests, used in the validation process, ought to be considered as trustworthy and reliable. Moreover, the assumptions, concerning the distributions of pairwise comparisons are weaker than those commonly used in the literature of the subject.

An important feature of the approach to estimation and verification is the simplicity of statistical tools, but with the necessary analytical basis; it broadens the range of potential users. The approach, presented in the book, Klukowski (2021c), and reported here, will be developed in the following directions: statistical learning, estimation of more complex structures of data (e.g. hierarchical), multidimensional (multi-criteria) pairwise comparisons, etc. An important field is also constituted by the application of the estimators and tests developed.

## References

Bradley, R.A. (1976) Science, statistics and paired comparisons. *Biometrics*, 32, 213–232.

David, H. A. (1988) *The Method of Paired Comparisons*, $2^{nd}$ ed. Ch. Griffin, London.

Domański, C. (1990) *Statistical tests* (in Polish). PWE, Warsaw.

Gordon, A.D. (1999) *Classification*, $2^{nd}$ ed. Chapman&Hall/CRC.

Garfinkel, R. S. and Nemhauser, G. L. (1972) *Integer Programming.* J Wiley and Son, Inc., New York, London, Sydney, Toronto.

Hansen, P., Jaumard, B. and Sanlaville, E. (1994) *Partitioning Problems in Cluster Analysis: A Review of Mathematical Programming Approaches. Studies in Classification, Data Analysis, and Knowledge Organization.* Springer-Verlag.

Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *JASA*, **58**, 13–30.

Klukowski, L. (2006) Tests for relation type – equivalence or tolerance – in a finite set of elements. *Control and Cybernetics* **35**, 2, 369–84.

Klukowski, L. (2011) *Methods of Estimation of Relations of: Equivalence, Tolerance, and Preference in a Finite Set.* IBS PAN, Series: *Systems Research*, Vol. 69, Warsaw.

Klukowski, L. (2017) Determining an estimate of an equivalence relation for moderate and large sized sets. *Operations Research and Decisions*, **27**, 2, 45–58.

Klukowski, L. (2021a) Pairwise comparisons in the form of difference of ranks – testing of estimates of preference relation. *Uncertainty and Imprecision* *in Decision Making and Decision Support – New Challenges, Solutions and Perspectives. Advances in Intelligent Systems and Computing* **1081**. Springer, 214–224.

KLUKOWSKI, L. (2021b) Statistical tests to verification of estimate of the preference relation resulting from pairwise comparisons. *Uncertainty and Imprecision in Decision Making and Decision Support – New Challenges, Solutions and Perspectives Advances in Intelligent Systems and Computing* **1081**. Springer, 225–238.

KLUKOWSKI, L. (2021c) *Estimation and Verification of Relations and Trees on the Basis of Multiple Binary and Multivalent Comparisons.* Systems Research Institute, Polish Academy of Sciences, series: *Systems Research*, Vol. 79, Warsaw.

RAGHAVACHARI, M. (2004) *Measures of Concordance for Assessing Agreement in Ratings and Rank Order Data.* In: N. Balakrishnan, N. Kannan and H. N. Nagaraja, eds., *Advances in Ranking and Selection, Multiple Comparisons, and Reliability. Methodology and Applications. Statistics for Industry and Technology.* Birkhäuser, Boston, Basel, Berlin.

SERFLING, R. J. (1980) *Approximation Theorems of Mathematical Statistics.* J. Wiley and Sons Inc.

SLATER, P. (1961) Inconsistencies in a schedule of paired comparisons. *Biometrika*, 1961, **48**, 303–312.

THOMPSON, W. A., JR. AND REMAGE, R., JR. (1964) Rankings from paired comparisons. *Ann. Math. Statist.* **35**, 739–47.