Raport Badawczy

Research Report

RB/70/2008

Estimation of the preference relation on the basis of multiple pairwise comparisons in the form of differences of ranks

L. Klukowski

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

Control and Cybernetics

vol. **37** (2008) No. 3

Estimation of the preference relation on the basis of multiple pairwise comparisons in the form of differences of ranks*

by

Leszek Klukowski

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

Abstract: The problem of estimation of the preference relation in a finite set on the basis of pairwise comparisons, in the form of differences of ranks with random errors, with the use of nearest adjoining order idea (NAO), is investigated in the paper. The results presented are extension and correction of the earlier works of the author; especially the case of multiple independent comparisons of each pair is examined. The comparisons of each pair are aggregated through the average or the median of comparisons. The estimated form of the relation is obtained in both cases on the basis of discrete programming tasks. The properties of the estimators are obtained under weak assumptions about distributions of comparison errors, in particular, the distributions may be unknown.

Keywords: multiple pairwise comparisons, nearest adjoining order method, difference of ranks.

1. Introduction

The paper presents extensions of the method of ranking elements from a finite set on the basis of pairwise comparisons, in the form of differences of ranks with random errors, presented in Klukowski (2000). The results discussed in Klukowski (2000) relate to the case of one comparison of each pair and require some correction (see Section 4). The extension examines the case of multiple comparisons; the comparisons of each pair are aggregated in two ways, first by simply averaging of (each pair) comparisons; then through median from the comparisons. In both cases the idea of nearest adjoining order (NAO) is applied (see Slater, 1961; David, 1988, section 2.2). The results obtained are based on weak assumptions about distributions of comparison errors, their distributions

*Submitted: January 2008; Accepted: October 2008.

may, in particular, be unknown. The properties of distributions of comparison errors assumed in the paper may be verified with the use of statistical tests (for unimodality, mode, median, symmetry). The basis for the properties of estimators are the well known probabilistic inequalities: Hoeffding's inequality for sums of bounded independent random variables (Hoeffding, 1963) and the Tshebyshev inequality for expected value; some properties of the order statistics (David, 1970) are also used. In the case of averaging approach the probability of the event that some random variable (defined in Section 3) corresponding to the true relation (the errorless estimation result) is lower than the variable corresponding to any other relation, converges exponentially to one. Some asymptotic properties are obtained also for the median approach.

The empirical problems having such structure often appear in practice, e.g., an assessment "student X exceeds student Y, according to substance matter, for three classes" has such form. In statistics, estimation based on some function (difference) of two random variables is a typical problem.

Let us note that the comparisons in the form of differences of ranks with the properties assumed can be obtained also on the basis of rankings (estimates) resulting from comparisons in the form of direction of preference (Klukowski, 1994). It allows for constructing two-stage estimators: the first step — to obtain estimates of the relation form with the use of comparisons indicating direction of preference and to determine differences of ranks for each estimate; the second step — to apply the algorithm based on differences of ranks (Sections 4, 5). The assumptions on comparison errors may be verified wit the use of statistical tests. It seems that the two-stage approach can be more efficient than the approach from Klukowski (1994), Section 5. Therefore, examination of the estimator based on differences of ranks is needed.

The empirical results, based on actual data and initial simulation experiments, are promising — also for "inconvenient" forms of distributions of comparison errors (asymmetric, with non-zero expected value).

The literature on ranking problems is quite extensive; for example the probabilistic approach is presented in David (1988), Marden (1995), the learning approach in Hastie et al. (2001), Chapter 14, Kamishima and Akaho (2006), the fuzzy approach in Yager (2007).

2. Problem formulation

The problem considered is an extension of the one stated in Klukowski (2000) for the case of $N > 1$ independent comparisons of each pair.

It is assumed that in a finite set of elements $\mathbf{X} = \{x_1, ..., x_m\}$ $(m \geqslant 3)$ there exists an unknown complete weak preference relation $\mathbf{R}$ of the form:

$$\mathbf{R} = \mathbf{I} \cup \mathbf{P}, \tag{1}$$

where:

I — the equivalence relation (reflexive, transitive, symmetric),
P — the strict preference relation (transitive, asymmetric).

The preference relation **R** generates from elements of the set **X** the family (sequence) of subsets $\chi_1^*, \ldots, \chi_n^*$ $(n \leqslant m)$, such that each element $x_i \in \chi_r^*$ is preferred to any element $x_j \in \chi_s^*$ $(r < s)$ and each subset χ_q^* $(1 \leqslant q \leqslant n)$ comprises equivalent elements only.

Relation **R** can be characterised by the function $T : \mathbf{X} \times \mathbf{X} \rightarrow D_T$, $D_T = \{-(n-1), \ldots, 0, \ldots, n-1\}$, defined as follows:

$$T(x_i, x_j) = d_{ij} \quad \Leftrightarrow \quad x_i \in \chi_r^*, \quad x_j \in \chi_s^*, \quad d_{ij} = r - s. \tag{2}$$

The value of the function $T(x_i, x_j)$ expresses the difference of ranks of the elements x_i and x_j in the relation **R**. In the case $T(x_i, x_j) < 0$, $(T(x_i, x_j) > 0)$ the element x_i precedes element x_j (element x_j precedes x_i), for d_{ij} positions. The value $T(x_i, x_j) = 0$ means the equivalence of both elements (they belong to the same subset χ_q^*, $1 \leqslant q \leqslant n$). It is obvious that $T(x_i, x_j) = -T(x_j, x_i)$ for $T(\cdot) \neq 0$.

The relation form is to be determined (estimated) on the basis of pairwise comparisons of elements of the set **X**, disturbed by random errors. Each pair $(x_i, x_j) \in \mathbf{X}$ is compared independently (in stochastic sense) N times; the result of k-th comparison $(k = 1, \ldots, N;\ N > 1)$ is the value of the function:

$$g_k : \mathbf{X} \times \mathbf{X} \rightarrow D, \quad D = \{-(m-1), \ldots, m-1\}; \tag{3}$$

the result $g_k(x_i, x_j) = c_{ijk}$ is an assessment of the difference of ranks in the pair (x_i, x_j), in k-th comparison. The set D can include values from the range: $-(m-1), \ldots, m-1$ because the number of subsets n is assumed unknown.

It is assumed in the paper that each comparison $g_k(x_i, x_j)$ $(1 \leqslant k \leqslant N)$ can be disturbed by a random error; it means that the difference $T(x_i, x_j) - g_k(x_i, x_j)$ may assume values different from zero — with some probabilities. The comparisons $g_k(x_i, x_j)$ and $g_l(x_r, x_s)$ are assumed independent, i.e.:

$$\begin{aligned} &P((g_k(x_i, x_j) = c_{ijk}) \cap (g_l(x_r, x_s) = c_{rsl})) = \\ &= P(g_k(x_i, x_j) = c_{ijk}) P(g_l(x_r, x_s) = c_{rsl}) \quad (k \neq l). \end{aligned} \tag{4}$$

The probabilities, which characterize each distribution of comparison errors will be denoted with the symbols $\alpha_{ijk}(l)$, $\beta_{ijk}(l)$, $\gamma_{ijk}(l)$; the probabilities are defined as follows:

$$\begin{aligned} \alpha_{ijk}(l) = P(T(x_i, x_j) - g_k(x_i, x_j) = l;\ T(x_i, x_j) = 0) \\ (-(m-1) \leqslant l \leqslant (m-1)), \end{aligned} \tag{5}$$

$$\begin{aligned} \beta_{ijk}(l) = P(T(x_i, x_j) - g_k(x_i, x_j) = l;\ T(x_i, x_j) < 0) \\ (-2(m-1) \leqslant l \leqslant 2(m-1)), \end{aligned} \tag{6}$$

$$\begin{aligned} \gamma_{ijk}(l) = P(T(x_i, x_j) - g_k(x_i, x_j) = l;\ T(x_i, x_j) > 0) \\ (-2(m-1) \leqslant l \leqslant 2(m-1)). \end{aligned} \tag{7}$$

It is obvious that the probabilities (5)–(7) have to fulfil the conditions:

$$\sum_{l=-(m-1)}^{(m-1)} \alpha_{ijk}(l) = 1, \qquad \sum_{l=-2(m-1)}^{2(m-1)} \beta_{ijk}(l) = 1, \qquad \sum_{l=-2(m-1)}^{2(m-1)} \gamma_{ijk}(l) = 1. \quad (8)$$

Moreover, it is assumed that the following assumptions hold:

$$\sum_{l\leqslant 0} P(T(x_i,x_j) - g_k(x_i,x_j) = l) > 1/2, \quad (9)$$

$$\sum_{l\geqslant 0} P(T(x_i,x_j) - g_k(x_i,x_j) = l) > 1/2, \quad (10)$$

$$P(T(x_i,x_j) - g_k(x_i,x_j) = l) \geqslant P(T(x_i,x_j) - g_k(x_i,x_j) = l+1); \quad l \geqslant 0, \quad (11)$$

$$P(T(x_i,x_j) - g_k(x_i,x_j) = l) \geqslant P(T(x_i,x_j) - g_k(x_i,x_j) = l-1); \quad l \leqslant 0. \quad (12)$$

The conditions (9)–(12) guarantee, that: • zero is the median of each distribution, • each probability function is unimodal and • assumes maximum in zero. The expected value of any comparison error can differ from zero (especially, for $T(x_i,x_j) = \pm(n-1)$, the expected value of the comparison is typically different from zero, because usually $P(T(x_i,x_j) - g_k(x_i,x_j) = 0) \neq 1$).

The probabilities $\alpha_{ijk}(0)$, $\beta_{ijk}(0)$ and $\gamma_{ijk}(0)$ may be lower than $\frac{1}{2}$; hence, the assumptions about errorless comparison are more general than those in zero-one approach (see Klukowski, 1994).

For simplification, it is assumed that distribution of any comparison error $T(x_i,x_j) - g_k(x_i,x_j)$ ($(x_i,x_j) \in \mathbf{X}\times\mathbf{X}$) is the same for each k, $1 \leqslant k \leqslant N$ (as a result the comparisons of each pair $g_1(x_i,x_j),\dots,g_N(x_i,x_j)$ are iid random variables). Therefore, the index k will be omitted in symbols: $\alpha_{ijk}(l)$, $\beta_{ijk}(l)$, $\gamma_{ijk}(l)$. The relaxation of the assumption about identical distributions is not difficult.

The probabilities $\alpha_{ij}(l)$ $(-(m-1) \leqslant l \leqslant m-1)$ determine the probability function of comparison errors for equivalent elements x_i and x_j (because $T(x_i,x_j) = 0$). The probability $\alpha_{ij}(l)$ means that a result of comparison assumes value l, when both elements are equivalent; especially $\alpha_{ij}(0)$ denotes the probability of errorless comparison (because $T(x_i,x_j) = g_k(x_i,x_j) = 0$). In the case of known number of the relation subsets (index n) the interval of integers $[-(m-1), m-1]$ ("support" of comparisons $g_k(\cdot)$) ought to be replaced with the interval $[-(n-1), n-1]$. The interpretation of the probabilities $\beta_{ij}(l)$ and $\gamma_{ij}(l)$ $(-2(m-1) \leqslant l \leqslant 2(m-1))$ is similar, with the difference that they both determine distributions of errors for non-equivalent elements.

The problem of estimation of the preference relation can be stated formally as follows:

To determine the relation $\mathbf{R}$ (or, equivalently, the sequence of subsets $\chi_1^*,\dots,\chi_n^*$) on the basis of the comparisons $g_k(x_i,x_j)$ $(k = 1,\dots,N)$, made for each pair $(x_i,x_j) \in \mathbf{X}\times\mathbf{X}$.

Let us emphasize that it is not assumed that the probability functions of comparisons errors (probabilities $\alpha_{ij}(l)$, $\beta_{ij}(l)$, $\gamma_{ij}(l)$) and the number of subsets n are known.

3. Basic definitions and notation

The following notation is introduced for further considerations:

• $t(x_i, x_j)$ — the function, which determines the difference of ranks for each pair $(x_i, x_j) \in \mathbf{X} \times \mathbf{X}$ in any preference relation in the set $\mathbf{X}$, i.e.:

$$t(x_i, x_j) = d_{ij} \Leftrightarrow x_i \in \chi_k, \quad x_j \in \chi_l; \quad d_{ij} = k - l. \tag{13}$$

• $I(\chi_1, \ldots, \chi_r)$, $P_1(\chi_1, \ldots, \chi_r)$, $P_2(\chi_1, \ldots, \chi_r)$ — the sets of pairs of indices $< i, j >$ generating a relation $(\chi_1, \ldots, \chi_r)$, i.e.:

$$I(\chi_1, \ldots, \chi_r) = \{< i, j > \mid t(x_i, x_j) = 0; \quad j > i\}, \tag{14}$$

$$P_1(\chi_1, \ldots, \chi_r) = \{< i, j > \mid t(x_i, x_j) < 0; \quad j > i\}, \tag{15}$$

$$P_2(\chi_1, \ldots, \chi_r) = \{< i, j > \mid t(x_i, x_j) > 0; \quad j > i\}; \tag{16}$$

• $R_m = I(\chi_1, \ldots, \chi_r) \cup P_1(\chi_1, \ldots, \chi_r) \cup P_2(\chi_1, \ldots, \chi_r) =$

$$\{< i, j > \mid 1 \leqslant i, j \leqslant m; \, j > i\}; \tag{17}$$

•$M = m(m-1)/2 = \#(R_m)$ (18)

(the symbol $\#(\Xi)$ denotes the number of elements of the set Ξ).

The properties of the estimators examined in the paper are based on random variables $U_{ij}^{(k)}(\chi_1, \ldots, \chi_r)$, $V_{ij}^{(k)}(\chi_1, \ldots, \chi_r)$, $Z_{ij}^{(k)}(\chi_1, \ldots, \chi_r)$, $W^{(k)}(\chi_1, \ldots, \chi_r)$ defined as follows:

$$U_{ij}^{(k)}(\chi_1, \ldots, \chi_r) = |\, g_k(x_i, x_j)\,|; \quad t(x_i, x_j) = 0, \tag{19}$$

$$V_{ij}^{(k)}(\chi_1, \ldots, \chi_r) = |\, t(x_i, x_j) - g_k(x_i, x_j)\,|; \quad t(x_i, x_j) < 0, \tag{20}$$

$$Z_{ij}^{(k)}(\chi_1, \ldots, \chi_r) = |\, t(x_i, x_j) - g_k(x_i, x_j)\,|; \quad t(x_i, x_j) > 0, \tag{21}$$

$$W^{(k)}(\cdot) = \sum_{<i,j>\in I(\cdot)} U_{ij}^{(k)}(\cdot) + \sum_{<i,j>\in P_1(\cdot)} V_{ij}^{(k)}(\cdot) + \sum_{<i,j>\in P_2(\cdot)} Z_{ij}^{(k)}(\cdot). \tag{22}$$

Random variables and other symbols corresponding to the relation $\mathbf{R}$ (errorless result of the estimation problem) will be marked with asterisks, i.e.: $U_{ij}^{(k)*}$, $V_{ij}^{(k)*}$, $Z_{ij}^{(k)*}$, I^*, P_1^*, P_2^*, $W^{(k)*}$, while variables and symbols corresponding to any other relation $\widetilde{X}_1, \ldots, \widetilde{X}_r$, different than the errorless one, will be denoted: $\widetilde{U}_{ij}^{(k)}$, $\widetilde{V}_{ij}^{(k)}$, $\widetilde{Z}_{ij}^{(k)}$, $\widetilde{I}$, $\widetilde{P}_1$, $\widetilde{P}_2$, $\widetilde{W}^{(k)}$. For fixed k $(1 \leqslant k \leqslant N)$ the difference $W^{(k)*} - \widetilde{W}^{(k)}$ can be written in the form[1]:

[1] The sum (23) comprises, in Klukowski (2000), six components only; it does not comprise the variables $Q_{ij}^{(k5)}(< i, j > \in S_5)$ and $Q_{ij}^{(k8)}(< i, j > \in S_8)$. Therefore, the evaluation (33) from Klukowski (2000) also requires correction (see formulas (46) and (63) in this paper).

$$W^{(k)*} - \widetilde{W}^{(k)} = \sum_{I^*\cap(\widetilde{P}_1-P_1^*)} (U_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}) + \sum_{I^*\cap(\widetilde{P}_2-P_2^*)} (U_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)}) +$$
$$+ \sum_{P_1^*\cap(\widetilde{I}-I^*)} (V_{ij}^{(k)*} - \widetilde{U}_{ij}^{(k)}) + \sum_{P_1^*\cap(\widetilde{P}_2-P_2^*)} (V_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)}) +$$
$$+ \sum_{(P_1^*\cap\widetilde{P}_1)\cap(T(\cdot)\neq\widetilde{t}(\cdot))} (V_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}) + \sum_{P_2^*\cap(\widetilde{I}-I^*)} (Z_{ij}^{(k)*} - \widetilde{U}_{ij}^{(k)})$$
$$+ \sum_{P_2^*\cap(\widetilde{P}_1-P_1^*)} (Z_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}) + \sum_{(P_2^*\cap\widetilde{P}_2)\cap(T(\cdot)\neq\widetilde{t}(\cdot))} (Z_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)}) \tag{23}$$

or shortly in the form

$$W^{(k)*} - \widetilde{W}(k) = \sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)},$$

where:

$$Q_{ij}^{(k,1)} = U_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}, \quad S_1 = \{< i,j > \mid < i,j > \in I^* \cap (\widetilde{P}_1 - P_1^*)\}, \tag{24}$$
$$Q_{ij}^{(k,2)} = U_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)}, \quad S_2 = \{< i,j > \mid < i,j > \in I^* \cap (\widetilde{P}_2 - P_2^*)\}, \tag{25}$$
$$Q_{ij}^{(k,3)} = V_{ij}^{(k)*} - \widetilde{U}_{ij}^{(k)}, \quad S_3 = \{< i,j > \mid < i,j > \in P_1^* \cap (\widetilde{I} - I^*)\}, \tag{26}$$
$$Q_{ij}^{(k,4)} = V_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)}, \quad S_4 = \{< i,j > \mid < i,j > \in P_1^* \cap (\widetilde{P}_2 - P_2^*)\}, \tag{27}$$
$$Q_{ij}^{(k,5)} = V_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)},$$
$$S_5 = \{< i,j > \mid < i,j > \in (P_1^* \cap \widetilde{P}_1) \cap (T(x_i,x_j) \neq \widetilde{t}(x_i,x_j))\}, \tag{28}$$
$$Q_{ij}^{(k,6)} = Z_{ij}^{(k)*} - \widetilde{U}_{ij}^{(k)}, \quad S_6 = \{< i,j > \mid < i,j > \in P_2^* \cap (\widetilde{I} - I^*)\}, \tag{29}$$
$$Q_{ij}^{(k,7)} = Z_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}, \quad S_7 = \{< i,j > \mid < i,j > \in P_2^* \cap (\widetilde{P}_1 - P_1^*)\}, \tag{30}$$
$$Q_{ij}^{(k,8)} = Z_{ij}^{(k)*} - \widetilde{Z}_{ij}^{(k)},$$
$$S_8 = \{< i,j > \mid < i,j > \in (P_2^* \cap \widetilde{P}_2) \cap (T(x_i,x_j) \neq \widetilde{t}(x_i,x_j))\}. \tag{31}$$

The following properties of the random variables $Q_{ij}^{(k\nu)}$ are necessary for further considerations:

LEMMA 1. *The expected value of each random variable* $Q_{ij}^{(k\nu)}$ ($1 \leqslant k \leqslant N$*;* $< i,j > \in S_\nu$*;* $\nu = 1,\ldots,8$) *satisfies the condition:*

$$E(Q_{ij}^{(k\nu)}) < 0. \tag{32}$$

Proof is similar to the proof of Theorem 1 in Klukowski (2007).

4. The case of averaged comparisons

The estimator presented in this section is based on averages from individual random variables $U_{ij}^{(k)}(\cdot)$, $V_{ij}^{(k)}(\cdot)$, $Z_{ij}^{(k)}(\cdot)$, i.e. the variables: $\bar{U}_{ij}(\cdot)$, $\bar{V}_{ij}(\cdot)$ and $\bar{Z}_{ij}(\cdot)$ defined in the following way:

$$\bar{U}_{ij}(\cdot) = \frac{1}{N}\sum_{k=1}^{N} U_{ij}^{(k)}(\cdot), \tag{33}$$

$$\bar{V}_{ij}(\cdot) = \frac{1}{N}\sum_{k=1}^{N} Vij^{(k)}(\cdot), \tag{34}$$

$$\bar{Z}_{ij}(\cdot) = \frac{1}{N}\sum_{k=1}^{N} Z_{ij}^{(k)}(\cdot). \tag{35}$$

Similarly, random variable $\bar{W}(\cdot)$ is defined as follows:

$$\bar{W}(\cdot) = \sum_{<i,j>\in I(\cdot)} \bar{U}_{ij}(\cdot) + \sum_{<i,j>\in P_1(\cdot)} \bar{V}_{ij}(\cdot) + \sum_{<i,j>\in P_2(\cdot)} \bar{Z}_{ij}(\cdot). \tag{36}$$

The symbols corresponding to the relation $\chi_1^*, \ldots, \chi_n^*$ will be also denoted with asterisks, i.e. $\bar{U}_{ij}^*(\cdot)$, $\bar{V}_{ij}^*(\cdot)$, $\bar{Z}_{ij}^*(\cdot)$, $\bar{W}^*$, while the symbols corresponding to any other relation $\widetilde{\chi}_1, \ldots, \widetilde{\chi}_r$ will be denoted by tildas, i.e. $\widetilde{\bar{U}}_{ij}(\cdot)$, $\widetilde{\bar{V}}_{ij}(\cdot)$, $\widetilde{\bar{Z}}_{ij}(\cdot)$, $\widetilde{\bar{W}}$.

Note that variables $\bar{U}_{ij}(\cdot)$, $\bar{V}_{ij}(\cdot)$ and $\bar{Z}_{ij}(\cdot)$ satisfy, under the assumption of identity of distribution functions $\alpha_{ijk}(l)$, $\beta_{ijk}(l)$, $\gamma_{ijk}(l)$ $(1 \leqslant k \leqslant N)$, the conditions:

$$E(\bar{U}_{ij}(\cdot)) = E(U_{ij}^{(k)}(\cdot)), \tag{37}$$

$$E(\bar{V}_{ij}(\cdot)) = E(V_{ij}^{(k)}(\cdot)), \tag{38}$$

$$E(\bar{Z}_{ij}(\cdot)) = E(Z_{ij}^{(k)}(\cdot)), \tag{39}$$

$$Var\,(\bar{U}_{ij}(\cdot)) = \tfrac{1}{N}\, Var\,(U_{ij}^{(k)}(\cdot)), \tag{40}$$

$$Var\,(\bar{V}_{ij}(\cdot)) = \tfrac{1}{N}\, Var\,(V_{ij}^{(k)}(\cdot)), \tag{41}$$

$$Var\,(\bar{Z}_{ij}(\cdot)) = \tfrac{1}{N}\, Var\,(Z_{ij}^{(k)}(\cdot)). \tag{42}$$

The difference $\bar{W}^*(\cdot) - \widetilde{\bar{W}}(\cdot)$ can be expressed in the form:

$$\bar{W}^*(\cdot) - \widetilde{\bar{W}}(\cdot) = \sum_{I^*\cap(\widetilde{P}_1 - P_1^*)} (\bar{U}_{ij}^* - \widetilde{\bar{V}}_{ij}) + \sum_{I^*\cap(\widetilde{P}_2 - P_2^*)} (\bar{U}_{ij}^* - \widetilde{\bar{Z}}_{ij}) +$$

$$+ \sum_{P_1^* \cap (\widetilde{I} - I^*)} (\bar{V}_{ij}^* - \widetilde{\bar{U}}_{ij}) + \sum_{P_1^* \cap (\widetilde{P}_2 - P_2^*)} (\bar{V}_{ij}^* - \widetilde{\bar{Z}}_{ij}) +$$

$$+ \sum_{(P_1^* \cap \widetilde{P}_1) \cap (T(\cdot) \neq \widetilde{t}(\cdot))} (\bar{V}_{ij}^* - \widetilde{\bar{V}}_{ij}) + \sum_{P_2^* \cap (\widetilde{I} - I^*)} (\bar{Z}_{ij}^* - \widetilde{\bar{U}}_{ij}) +$$

$$+ \sum_{P_2^* \cap (\widetilde{P}_1 - P_1^*)} (\bar{Z}_{ij}^* - \widetilde{\bar{V}}_{ij}) + \sum_{(P_2^* \cap \widetilde{P}_2) \cap (T(\cdot) \neq \widetilde{t}(\cdot))} (\bar{Z}_{ij}^* - \widetilde{\bar{Z}}_{ij})$$

$$= \sum_{\nu=1}^{8} \sum_{S_\nu} \bar{Q}_{ij}^{(\nu)}, \tag{43}$$

where:

$$\bar{Q}_{ij}^{(1)} = \bar{U}_{ij}^* - \widetilde{\bar{V}}_{ij}, \quad < i, j > \in S_1,$$
$$\bar{Q}_{ij}^{(2)} = \bar{U}_{ij}^* - \widetilde{\bar{Z}}_{ij}, \quad < i, j > \in S_2,$$
$$\bar{Q}_{ij}^{(3)} = \bar{V}_{ij}^* - \widetilde{\bar{U}}_{ij}, \quad < i, j > \in S_3,$$
$$\bar{Q}_{ij}^{(4)} = \bar{V}_{ij}^* - \widetilde{\bar{Z}}_{ij}, \quad < i, j > \in S_4,$$
$$\bar{Q}_{ij}^{(5)} = \bar{V}_{ij}^* - \widetilde{\bar{V}}_{ij}, \quad < i, j > \in S_5,$$
$$\bar{Q}_{ij}^{(6)} = \bar{Z}_{ij}^* - \widetilde{\bar{U}}_{ij}, \quad < i, j > \in S_6,$$
$$\bar{Q}_{ij}^{(7)} = \bar{Z}_{ij}^* - \widetilde{\bar{V}}_{ij}, \quad < i, j > \in S_7,$$
$$\bar{Q}_{ij}^{(8)} = \bar{Z}_{ij}^* - \widetilde{\bar{Z}}_{ij}, \quad < i, j > \in S_8,$$

(S_ν - same as in (24)–(31)).

It results from the lemma presented in Section 3 that:

$$E(\bar{W}^*(\cdot) - \widetilde{\bar{W}}(\cdot)) < 0. \tag{44}$$

Moreover, we can establish the evaluation of the probability $P(\bar{W}^* < \widetilde{\bar{W}})$. Hoeffding's inequality (see Hoeffding, 1963):

$$P\left(\sum_{k=1}^{N} Y_i - \sum_{k=1}^{N} E(Y_i) \geqslant Nt\right) \leqslant \exp\{-2Nt^2/(b-a)^2\}, \tag{45}$$

where:

Y_i, $(i = 1, \ldots, N)$ — independent random variables satisfying condition $P(a \leqslant Y_i \leqslant b) = 1$, $(-\infty < a < b < \infty)$;

t — positive constant,

will be used as the basis for evaluation.

THEOREM 1. *Probability $P(\bar{W}^* < \widetilde{\bar{W}})$ satisfies the inequality*

$$P(\bar{W}^* < \widetilde{\bar{W}}) \geqslant 1 - \exp\left\{ -2N \frac{(\sum\limits_{\nu=1}^{8} \sum\limits_{S_\nu} E(Q_{ij}^{(1,\nu)}))^2}{(2\vartheta(m-1))^2} \right\}, \tag{46}$$

where: ϑ — the number of elements of the set $\bigcup\limits_{\nu=1}^{8} S_\nu$.

Proof. Probability $P(\bar{W}^* < \widetilde{\bar{W}})$ can be expressed in the form:

$$P(\bar{W}^* < \widetilde{\bar{W}}) = 1 - P(\bar{W}^* - \widetilde{\bar{W}} \geqslant 0) \quad \text{and}$$

$$P(\bar{W}^* - \widetilde{\bar{W}} \geqslant 0) = P(\sum_{\nu=1}^{8} \sum_{S_\nu} \bar{Q}_{ij}^{(\nu)} \geqslant 0) =$$

$$= P(\sum_{\nu=1}^{8} \sum_{S_\nu} \tfrac{1}{N} \sum_{k=1}^{N} Q_{ij}^{(k\nu)} \geqslant 0) = P(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)}) \geqslant 0), \tag{47}$$

where:

$$Q_{ij}^{(k,1)} = U_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$Q_{ij}^{(k,8)} = Z_{ij}^{(k)*} - \widetilde{V}_{ij}^{(k)}.$$

Last inequality in (47) can be transformed in the following way:

$$P(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)}) \geqslant 0) =$$

$$= P(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)}) - E(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)})) \geqslant -E(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)}))) =$$

$$= P(\sum_{k=1}^{N} (\sum_{\nu=1}^{8} \sum_{S_\nu} Q_{ij}^{(k\nu)}) - N \sum_{\nu=1}^{8} \sum_{S_\nu} E(Q_{ij}^{(1,\nu)}) \geqslant -N \sum_{\nu=1}^{8} \sum_{S_\nu} E(Q_{ij}^{(1,\nu)})). \tag{48}$$

Equality (48) results from the assumption that for any k $(1 \leqslant k \leqslant N)$ the distributions of the random variables $U_{ij}^{(k)}(\cdot)$, $V_{ij}^{(k)}(\cdot)$ and $Z_{ij}^{(k)}(\cdot)$ are the same. Therefore, the expected values of the variables $Q_{ij}^{(k\nu)}$ $(1 \leqslant k \leqslant N)$ are also the same.

The probability (48) can be evaluated on the basis of Hoeffding inequality (45) in the following way:

$$P(\sum_{k=1}^{N}(\sum_{\nu=1}^{8}\sum_{S_\nu}Q_{ij}^{(k\nu)}) - N\sum_{\nu=1}^{8}\sum_{S_\nu}E(Q_{ij}^{(1,\nu)}) \geqslant -N\sum_{\nu=1}^{8}\sum_{S_\nu}E(Q_{ij}^{(1,\nu)})) \leqslant$$
$$\leqslant \exp\left\{-2N\frac{(\sum_{\nu=1}^{8}\sum_{S_\nu}E(Q_{ij}^{(1,\nu)}))^2}{(2\vartheta(m-1))^2}\right\}. \tag{49}$$

The evaluation (49) results from the following facts: the absolute value of each difference $|T(x_i,x_j)-g_k(x_i,x_j)|-|\widetilde{t}(x_i,x_j)-g_k(x_i,x_j)|$ $((x_i,x_j)\in \mathbf{X}\times\mathbf{X})$ cannot exceed the value $2(m-1)$, the number of components of the sum $\sum_{\nu=1}^{8}\sum_{S_\nu}E(Q_{ij}^{(1,\nu)})$ equals ϑ and each expected value $E(Q_{ij}^{(k\nu)})$ is negative (see Lemma in Section 3). The evaluation is equivalent to the proved inequality (46). ■

The inequality (46) shows that $P(\bar{W}^* < \widetilde{\bar{W}})$, i.e. the probability of the event that the value of the random variable $\bar{W}^*$ is lower than of any other variable $\widetilde{\bar{W}}$, converges exponentially to one, for $N \to \infty$. Moreover, each variance $Var\,(\bar{W}(\chi_1,\ldots,\chi_r))$ converges to zero, when $N\to\infty$. Therefore, any variable $\bar{W}(\chi_1,\ldots,\chi_r)$ converges stochastically to some constant $\bar{\omega}(\chi_1,\ldots,\chi_r)$; the constant $\bar{\omega}^*$, corresponding to the variable $\bar{W}^*$ (i.e. relation $\chi_1^*,\ldots,\chi_n^*$) assumes minimal value in the set $\{\bar{\omega}(\chi_1,\ldots,\chi_r)\,|\,\chi_1,\ldots,\chi_r\in F_X;\ F_X$ — the family of all preference relations in the set $\mathbf{X}\}$. This facts indicates the form of the estimator — to determine the relation $\widehat{\chi}_1,\ldots,\widehat{\chi}_{\hat{n}}$, which minimizes the value of the random variable $\bar{W}(\chi_1,\ldots,\chi_r)$ for given comparisons $g_k(x_i,x_j)$ $(k=1,\ldots,N;$ $(x_i,x_j)\in\mathbf{X}\times\mathbf{X})$. Let us notice that the value of the right-hand side of inequality (46) depends on the form of relation $\widetilde{\chi}_1,\ldots,\widetilde{\chi}_r$; an increase of "dissimilarity" between $\widetilde{\chi}_1,\ldots,\widetilde{\chi}_r$ and actual relation $\chi_1^*,\ldots,\chi_n^*$ increases the expected value $\sum_{\nu=1}^{8}\sum_{S_\nu}E(Q_{ij}^{(1,\nu)})$ and — finally — decreases the probability $P(\bar{W}^*\geqslant\widetilde{\bar{W}})$. In other words — a more dissimilar relation $\widetilde{\chi}_1,\ldots,\widetilde{\chi}_r$, in comparison to the relation $\chi_1^*,\ldots,\chi_n^*$, is less probable.

The optimization task for the case under examination assumes the form:

$$\min_{\chi_1^{(\iota)},\ldots,\chi_{r^{(\iota)}}^{(\iota)}\in F_X}\left\{\sum_{<i,j>\in R_m}\sum_{k=1}^{N}\left|t^{(\iota)}(x_i,x_j)-g_k(x_i,x_j)\right|\right\}, \tag{50}$$

where:

F_X – the feasible set of the problem, i.e. the family of all preference relations in the set $\mathbf{X}$,

$t^{(\iota)}(x_i, x_j)$ – the function describing the relation $\chi_1^{(\iota)}, \ldots, \chi_{r^{(\iota)}}^{(\iota)}$ from the feasible set F_X

(the factor $\frac{1}{N}$ is omitted, because it does not influence the form of the optimal solution).

There may be more than one solution to problem (50). In case of multiple solutions, the inequality (46) refers to whole set of solutions obtained. The unique form of the relation can be chosen randomly or with the use of an additional criterion, e.g. minimal value of the expression

$$\sum_{<i,j>\in I(\widehat{\chi}_1,\ldots,\widehat{\chi}_{\widehat{n}})} \sum_{k=1}^{N} \left|\widehat{t}(x_i, x_j) - g_k(x_i, x_j)\right|$$

the function $\widehat{t}(\cdot)$ describes the estimate $\widehat{\chi}_1, \ldots, \widehat{\chi}_{\widehat{n}}$.

The evaluation (46) is similar to those presented in Klukowski (1994) point 5.1, corresponding to the case, when comparisons indicate the direction of preferences (not difference of ranks). The right-hand side of probability (46) is better (assumes higher value), than the evaluation presented in Klukowski (1994) in the case, when:

$$(\sum_{\nu=1}^{8} \sum_{S_\nu} E(Q_{ij}^{1\nu}))^2/(2\vartheta(m-1))^2 > (1/2-\delta)^2,$$

where δ denotes maximum probability of error in comparisons expressing the direction of preference.

The numerical value of the right-hand side of inequality (46) can be determined in the case of known distributions of comparison errors and the form of relation $\chi_1^*, \ldots, \chi_n^*$. If not, they can be replaced with estimates or evaluations (see Klukowski, 2007). The estimation requires sufficient number of comparisons N, at least several.

Let us notice that the evaluation (46) is, in general, significantly underestimated; its negative feature is dependence on the number of elements m. Therefore, in the case of "reliable" estimation of the relation form (it is indicated by the minimal value of the function (50) close to zero), the value $m-1$ can be replaced by the estimate $\widehat{n}-1$. The estimate can be usually "reliable" for moderate N, e.g. several, because of the exponential form of the right-hand side of inequality (46).

It is also possible to consider the estimation problem with the quadratic function instead of the absolute value, e.g.:

$$U_k^2(x_i, x_j) = (t(x_i, x_j) - g_k(x_i, x_j))^2; \quad t(x_i, x_j) = 0; \tag{51}$$

$$\bar{W}^2(x_i,x_j) = \frac{1}{N}\sum_{k=1}^{N}\Big(\sum_{<i,j>\in I(\cdot)} U_k^2(x_i,x_j) + \sum_{<i,j>\in P_1(\cdot)} V_k^2(x_i,x_j) + \\ + \sum_{<i,j>\in P_2(\cdot)} Z_k^2(x_i,x_j)\Big), \tag{52}$$

or with the use of the average $\bar{g}(x_i,x_j)$ instead of individual $g_k(x_i,x_j)$, e.g.:

$$\check{U}(x_i,x_j) = |t(x_i,x_j) - \bar{g}(x_i,x_j)|\,; \quad t(x_i,x_j) = 0; \tag{53}$$

$$\check{W}(x_i,x_j) = \sum_{<i,j>\in I(\cdot)} \check{U}(x_i,x_j) + \sum_{<i,j>\in P_1(\cdot)} \check{V}(x_i,x_j) + \\ + \sum_{<i,j>\in P_2(\cdot)} \check{Z}(x_i,x_j), \tag{54}$$

where: $\bar{g}(x_i,x_j) = \frac{1}{N}\sum_{k=1}^{N} g_k(x_i,x_j)$, etc.

The minimization problems assume the following forms for these functions:

$$\min_{\chi_1^{(\iota)},\ldots,\chi_{r(\iota)}^{(\iota)}\in F_X} \Big\{ \sum_{<i,j>\in R_m} \sum_{k=1}^{N} (t^{(\iota)}(x_i,x_j) - g_k(x_i,x_j))^2 \Big\}, \tag{55a}$$

$$\min_{\chi_1^{(\iota)},\ldots,\chi_{r(\iota)}^{(\iota)}\in F_X} \Big\{ \sum_{<i,j>\in R_m} \Big| t^{(\iota)}(x_i,x_j) - \bar{g}(x_i,x_j) \Big| \Big\}. \tag{55b}$$

The properties of such estimators need further investigation.

5. The median approach

In the case of the median approach it is assumed, that N is odd, i.e. $N = 2\tau + 1$ ($\tau = 0, 1\ldots$). The form of estimator is based on the random variables: $U_{me,N}(x_i,x_j)$, $V_{me,N}(x_i,x_j)$, $Z_{me,N}(x_i,x_j)$, $W_{me,N}(\chi_1,\ldots,\chi_r)$ defined in the following way:

$$U_{me,N}(x_i,x_j) = |g_{me,N}(x_i,x_j)|; \quad t(x_i,x_j) = 0, \tag{56}$$

$$V_{me,N}(x_i,x_j) = |t(x_i,x_j) - g_{me,N}(x_i,x_j)|; \quad t(x_i,x_j) < 0, \tag{57}$$

$$Z_{me,N}(x_i,x_j) = |t(x_i,x_j) - g_{me,N}(x_i,x_j)|; \quad t(x_i,x_j) > 0, \tag{58}$$

$$W_{me,N}(\chi_1,\ldots,\chi_r) = \sum_{I(\cdot)} U_{me,N}(x_i,x_j) + \sum_{P1(\cdot)} V_{me,N}(\cdot) + \sum_{P2(\cdot)} Z_{me,N}(x_i,x_j), \tag{59}$$

where: $g_{me,N}(x_i,x_j)$ – the median from comparisons $g_k(x_i,x_j)$ ($k = 1,\ldots,N$), i.e. $g_{me,N}(x_i,x_j) = g_{((N+1)/2)}(x_i,x_j)$, while symbols $g_{(1)}(x_i,x_j),\ldots,g_{(N)}(x_i,x_j)$ denote the comparisons: $g_1(x_i,x_j),\ldots,g_N(x_i,x_j)$ ordered in non-decreasing manner.

The symbols corresponding to the actual relation $\chi_1^*, \ldots, \chi_n^*$ are marked with asterisks, i.e.: $U^*_{ij,me,N}$, $V^*_{ij,me,N}$, $Z^*_{ij,me,N}$, $W^*_{me,N}$, while corresponding to any other relation $\widetilde{\chi}_1, \ldots, \widetilde{\chi}_{\widetilde{n}}$ – with tildas, i.e.: $\widetilde{U}_{ij,me,N}$, $\widetilde{V}_{ij,me,N}$, $\widetilde{Z}_{ij,me,N}$, $\widetilde{W}_{me,N}$.

Using such notation, the difference $W^*_{me,N} - \widetilde{W}_{me,N}$, the basis for the properties of the estimator based on medians, assumes the form:

$$
\begin{aligned}
W^*_{me,N} - \widetilde{W}_{me,N} = &\sum_{I^*\cap(\widetilde{P}_1 - P_1^*)} (U^*_{ij,me,N} - \widetilde{V}_{ij,me,N}) + \\
+ \sum_{I^*\cap(\widetilde{P}_2 - P_2^*)} (U^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}) &+ \sum_{P_1^*\cap(\widetilde{I} - I^*)} (V^*_{ij,me,N} - \widetilde{U}_{ij,me,N}) + \\
+ \sum_{P_1^*\cap(\widetilde{P}_2 - P_2^*)} (V^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}) &+ \sum_{(P_1^*\cap\widetilde{P}_1)\cap(T(\cdot)\neq\widetilde{t}(\cdot))} (V^*_{ij,me,N} - \widetilde{V}_{ij,me,N}) + \\
\sum_{P_2^*\cap(\widetilde{I} - I^*)} (Z^*_{ij,me,N} - \widetilde{U}_{ij,me,N}) &+ \sum_{P_2^*\cap(\widetilde{P}_1 - P_1^*)} (Z^*_{ij,me,N} - \widetilde{V}_{ij,me,N}) + \\
+ \sum_{(P_2^*\cap\widetilde{P}_2)\cap(T(\cdot)\neq\widetilde{t}(\cdot))} &(Z^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}),
\end{aligned}
\tag{60}
$$

equivalent to:

$$
W^*_{me,N} - \widetilde{W}_{me,N} = \sum_{\nu=1}^{8} \sum_{S_\nu} Q^{(\nu)}_{ij,me,N}, \tag{61}
$$

where:

$$
\begin{aligned}
Q^{(1)}_{ij,me,N} &= U^*_{ij,me,N} - \widetilde{V}_{ij,me,N}, \quad < i,j > \in S_1, \\
Q^{(2)}_{ij,me,N} &= U^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}, \quad < i,j > \in S_2, \\
Q^{(3)}_{ij,me,N} &= V^*_{ij,me,N} - \widetilde{U}_{ij,me,N}, \quad < i,j > \in S_3, \\
Q^{(4)}_{ij,me,N} &= V^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}, \quad < i,j > \in S_4, \\
Q^{(5)}_{ij,me,N} &= V^*_{ij,me,N} - \widetilde{V}_{ij,me,N}, \quad < i,j > \in S_5, \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
Q^{(8)}_{ij,me,N} &= Z^*_{ij,me,N} - \widetilde{Z}_{ij,me,N}, \quad < i,j > \in S_8,
\end{aligned}
$$

(S_ν — defined in (24)–(31)).

The properties of the difference $W^*_{me,N} - \widetilde{W}_{me,N}$ are determined in the following

THEOREM 2. *The difference* $W^*_{me,N} - \widetilde{W}_{me,N}$ *satisfies the inequalities:*

$$E(W^*_{me,N} - \widetilde{W}_{me,N}) < 0, \tag{62}$$

$$P(W^*_{me,N} < \widetilde{W}_{me,N}) \geqslant \frac{E(\sum_{\nu=1}^{8} \sum_{S_\nu} Q^{(\nu)}_{ij,me,N})}{\lambda_1(m-1) + 2\lambda_2(m-1) + \lambda_3(m-2)}, \tag{63}$$

where: $\lambda_1 = \#(S_1 \cup S_2 \cup S_3 \cup S_6)$*;* $\lambda_2 = \#(S_4 \cup S_7)$*;* $\lambda_3 = \#(S_5 \cup S_8)$ ($\#(\Xi)$ *denotes the number of elements of* Ξ).

Proof of inequality (62). Inequality (62) is true for $N = 1$ (on the basis of the inequality (32) — see Lemma 1 in Section 3). For $N = 2\tau + 1$ ($\tau = 1, \ldots,$) the proof is similar to the case examined in Klukowski (2007); therefore its draft is presented only below. The probability function $P(T(x_i, x_j) - g_{me,N}(x_i, x_j)) = l)$ ($N = 2\tau+1$; $\tau = 0, 1, \ldots,$) satisfies for each pair $(x_i, x_j) \in \mathbf{X}\times\mathbf{X}$ the inequalities:

$$P(T(x_i, x_j) - g_{me,N+2}(x_i, x_j) = 0) > P(T(x_i, x_j) - g_{me,N}(x_i, x_j) = 0), \tag{64a}$$

$$P(T(x_i, x_j) - g_{me,N+2}(x_i, x_j) = l) < P(T(x_i, x_j) - g_{me,N}(x_i, x_j) = l) \quad (l \neq 0). \tag{64b}$$

Inequalities (64a, b) result from the following facts: probabilities $P(T(x_i, x_j) - g_{me,N}(x_i, x_j) = l)$ can be expressed in the form (see David, 1970, Section 2.4):

$$P(T(x_i, x_j) - g_{me,N}(x_i, x_j) = 0) =$$
$$= P(T(x_i, x_j) - g_{me,N}(x_i, x_j) \leqslant 0) - P(T(x_i, x_j) - g_{me,N}(x_i, x_j) \leqslant -1) =$$
$$= \frac{N!}{(((N-1)/2)!)^2} \int_{G(-1)}^{G(0)} t^{(N-1)/2}(1-t)^{(N-1)/2} dt, \tag{65a}$$

$$P(T(x_i, x_j) - g_{me,N}(x_i, x_j) = l) =$$
$$= P(T(x_i, x_j) - g_{me,N}(x_i, x_j) \leqslant l) - P(T(x_i, x_j) - g_{me,N}(x_i, x_j) \leqslant l - 1) =$$
$$= \frac{N!}{(((N-1)/2)!)^2} \int_{G(l-1)}^{G(l)} t^{(N-1)/2}(1-t)^{(N-1)/2} dt, \tag{65b}$$

where:

$$G(l) = P(T(x_i, x_j) - g_k(x_i, x_j) \leqslant l).$$

The expressions (65a) and (65b) are determined on the basis of beta distribution $B(p, q)$, with parameters $p = q = (N+1)/2$. The expected value and variance of the distribution assume the forms, respectively: $\frac{1}{2}$ and $((N+1)/2)^2/(N+1)^2(N+2) = \frac{1}{4(N+2)}$. The variance of the distribution converges to zero for $N \to \infty$ and the integrand in integrals (65a, b) is symmetric around $\frac{1}{2}$. These facts guarantee that: the distributions of the random variables $T(x_i, x_j) - g_{me,N}(x_i, x_j)$ ($(x_i, x_j) \in \mathbf{X} \times \mathbf{X}$) are for each N unimodal, their probability functions assume maximum in zero (i.e. for $T(x_i, x_j) - g_{me,N}(x_i, x_j) = 0$) and satisfy the inequalities (64a, b). Last two conditions are sufficient (see the assumptions (9)–(12) and Lemma 1 from Section 3) for the inequality (62).

Proof of inequality (63). The inequality (63) is proved on the basis of Chebyshev inequality for expected value. For this purpose the left-hand side of the inequality is transformed to the form $P(W^*_{me,N} < \widetilde{W}_{me,N}) = 1 - P(W^*_{me,N} - \widetilde{W}_{me,N} \geqslant 0)$ and each random variable $Q^{(\nu)}_{ij,me,N}$ is transformed to the form, which provides non-negative expected value:

$$Q'^{(\nu)}_{ij,me,N} = Q^{(\nu)}_{ij,me,N} + (m-1) \quad (\nu = 1, 2, 3, 6), \tag{66}$$

$$Q'^{(\nu)}_{ij,me,N} = Q^{(\nu)}_{ij,me,N} + 2(m-1) \quad (\nu = 5, 8), \tag{67}$$

$$Q'^{(\nu)}_{ij,me,N} = Q^{(\nu)}_{ij,me,N} + (m-2) \quad (\nu = 4, 7). \tag{68}$$

The probability $P(W^*_{me,N} - \widetilde{W}_{me,N} \geqslant 0)$ can be evaluated in the following way:

$$\begin{aligned}
&P(W^*_{me,N} - \widetilde{W}_{me,N} \geqslant 0) = P(\sum_{\nu=1}^{8} \sum_{S_\nu} Q^{(\nu)}_{ij,me,N} \geqslant 0) = \\
&P(\sum_{\nu=1}^{8} \sum_{S_\nu} Q'^{(\nu)}_{ij,me,N} \geqslant \lambda_1(m-1) + 2\lambda_2(m-1) + \lambda_3(m-2)) \leqslant \\
&\leqslant \frac{E(\sum_{\nu=1}^{8} \sum_{S_\nu} Q'^{(\nu)}_{ij,me,N})}{\lambda_1(m-1) + 2\lambda_2(m-1) + \lambda_3(m-2)} = \\
&= 1 + \frac{E(\sum_{\nu=1}^{8} \sum_{S_\nu} Q^{(\nu)}_{ij,me,N})}{(\lambda_1 + 2\lambda_2)(m-1) + \lambda_3(m-2)}.
\end{aligned} \tag{69}$$

The inequality (69) is equivalent to proved inequality (63). ■

The right-hand side of the inequality (63) is included in the interval $(0, 1)$. Its numerical value can be determined in the case of known distributions of comparison errors $P(T(\cdot) - g_{me,N}(\cdot) = \iota)$. In the opposite case it can be estimated or approximated. An approximation procedure for this purpose, useful for moderate N (lower than several), based on the formulas (65a, b) and an assumption about symmetry of distribution tails, can be constructed in similar way, as in Klukowski (2007), for the tolerance relation. For N greater than several, unknown probability functions of comparison errors can be estimated.

The evaluation (63) based on the value m is usually underestimated (lower than actual probability) and — similarly as in the case of averaged comparisons — the values: $m-1$, $2(m-1)$ and $m-2$ can be replaced by the estimates based on $\widehat{n}$.

The minimisation task for the estimation of the preference relation is similar, as in the case of $N=1$ (see (50)). It assumes the form:

$$\min_{\chi_1^{(\iota)},\ldots,\chi_{r^{(\iota)}}^{(\iota)} \in F_X} \Big\{ \sum_{<i,j> \in R_m} \Big| t^{(\iota)}(x_i,x_j) - g_{me,N}(x_i,x_j) \Big| \Big\}. \tag{70}$$

There may be more than one solution of the task (70).

It should be emphasized that the evaluation (63) is based on rough probability inequality. However, it seems conceivable that for some types of distributions of comparison errors, the efficiency of the median approach is similar to that for the averaging approach.

The right-hand side of the inequality (63) does not converge exponentially to one. However, the estimator, which guarantees such convergence can be constructed for medians (from differences of ranks) on the basis of the approach presented in Klukowski (1994), point 5.2. The differences of ranks have to be transformed into comparisons indicating the direction of the preference, which satisfy the condition that probability of errorless comparison is higher than $\frac{1}{2}$. The idea of the transformation can be presented briefly in the following way. On the basis of the formulas (65a, b) we can determine the minimal value (integer) κ, $(\kappa \leqslant N)$, which guarantees, for each pair $(x_i,x_i) \in \mathbf{X} \times \mathbf{X}$, satisfaction of the condition:

$$P(T(\cdot) - g_{me(\kappa)}(\cdot) = 0) > \tfrac{1}{2} \tag{71}$$

where: $g_{me(\kappa)}(\cdot)$ is the median in the subset of κ consecutive comparisons, i.e. $\{g_1(\cdot),\ldots,g_\kappa(\cdot)\}$ or $\{g_{\kappa+1}(\cdot),\ldots,g_{2\kappa}(\cdot)\}$, etc.

Let us define the random variables $\mathcal{U}_{ij,\tau}(\chi_1,\ldots,\chi_r)$, $\mathcal{V}_{ij,\tau}(\chi_1,\ldots,\chi_r)$, $\mathcal{Z}_{ij,\tau}(\chi_1,\ldots,\chi_r)$ in the following way:

$$\mathcal{U}_{ij,\tau}(\chi_1,\ldots,\chi_r) = \begin{cases} 0; \ \mathfrak{g}_{me,\tau}(x_i,x_j) = t(x_i,x_j) \ \text{ for } <i,j> \in I(\chi_1,\ldots,\chi_r); \\ 1; \ \mathfrak{g}_{me,\tau}(x_i,x_j) \neq t(x_i,x_j) \ \text{ for } <i,j> \in I(\chi_1,\ldots,\chi_r), \end{cases} \tag{72a}$$

$$\mathcal{V}_{ij,\tau}(\chi_1,\ldots,\chi_r) = \begin{cases} 0; \ \mathfrak{g}_{me,\tau}(x_i,x_j) = t(x_i,x_j) \ \text{ for } <i,j> \in P_1(\chi_1,\ldots,\chi_r); \\ 1; \ \mathfrak{g}_{me,\tau}(x_i,x_j) \neq t(x_i,x_j) \ \text{ for } <i,j> \in P_1(\chi_1,\ldots,\chi_r), \end{cases} \tag{72b}$$

$$\mathcal{Z}_{ij,\tau}(\chi_1,\ldots,\chi_r) = \begin{cases} 0; \ \mathfrak{g}_{me,\tau}(x_i,x_j) = t(x_i,x_j) \ \text{ for } <i,j> \in P_2(\chi_1,\ldots,\chi_r); \\ 1; \ \mathfrak{g}_{me,\tau}(x_i,x_j) \neq t(x_i,x_j) \ \text{ for } <i,j> \in P_2(\chi_1,\ldots,\chi_r), \end{cases} \tag{72c}$$

where:

$\mathfrak{g}_{me,\tau}(x_i, x_j)$ – the median in the set of comparisons $\{g_{(\tau-1)\cdot\kappa+1}(\cdot), \ldots, g_{\tau\cdot\kappa}(\cdot)\}$ $(\tau = 1, \ldots, \vartheta)$; the expression $a \cdot b$ – (index in $g_{a\cdot b}(\cdot)$) means the product of a and b,

ϑ – integer part of the quotient N/κ (odd number), i.e. $\vartheta = \mathrm{ent}(\mathrm{N}/\kappa)$.

Now, the majority approach, introduced in Klukowski (1994), point 5.2 (equivalent to the median in the set comprising zero-one random variables), can be applied to the random variables $\mathcal{U}_{ij,\tau}(\chi_1, \ldots, \chi_r)$, $\mathcal{V}_{ij,\tau}(\chi_1, \ldots, \chi_r)$, $\mathcal{Z}_{ij,\tau}(\chi_1, \ldots, \chi_r)$ $(\tau = 1, \ldots \vartheta$). As a result, one can obtain the variables $\mathcal{U}_{ij}^{(me)}$, $\mathcal{V}_{ij}^{(me)}$, $\mathcal{Z}_{ij}^{(me)}$ defined as follows:

$$\mathcal{U}_{ij}^{(me)}(\chi_1, \ldots, \chi_r) = \begin{cases} 0; \sum_{\tau=1}^{\vartheta} \mathcal{U}_{ij,\tau}(\chi_1, \ldots, \chi_r) < \vartheta/2 \text{ for } <i,j> \in I(\chi_1, \ldots, \chi_r); \\ 1; \sum_{\tau=1}^{\vartheta} \mathcal{U}_{ij,\tau}(\chi_1, \ldots, \chi_r) > \vartheta/2 \text{ for } <i,j> \in I(\chi_1, \ldots, \chi_r), \end{cases} \tag{73a}$$

$$\mathcal{V}_{ij}^{(me)}(\chi_1, \ldots, \chi_r) = \begin{cases} 0; \sum_{\tau=1}^{\vartheta} \mathcal{V}_{ij,\tau}(\chi_1, \ldots, \chi_r) < \vartheta/2 \text{ for } <i,j> \in P_1(\chi_1, \ldots, \chi_r); \\ 1; \sum_{\tau=1}^{\vartheta} \mathcal{V}_{ij,\tau}(\chi_1, \ldots, \chi_r) > \vartheta/2 \text{ for } <i,j> \in P_1(\chi_1, \ldots, \chi_r), \end{cases} \tag{73b}$$

$$\mathcal{Z}_{ij}^{(me)}(\chi_1, \ldots, \chi_r) = \begin{cases} 0; \sum_{\tau=1}^{\vartheta} \mathcal{Z}_{ij,\tau}(\chi_1, \ldots, \chi_r) < \vartheta/2 \text{ for } <i,j> \in P_2(\chi_1, \ldots, \chi_r); \\ 1; \sum_{\tau=1}^{\vartheta} \mathcal{Z}_{ij,\tau}(\chi_1, \ldots, \chi_r) > \vartheta/2 \text{ for } <i,j> \in P_2(\chi_1, \ldots, \chi_r). \end{cases} \tag{73c}$$

Let us apply the convention used in previous sections to the variables: $\mathcal{U}_{ij}^{(me)}(\cdot)$, $\mathcal{V}_{ij}^{(me)}(\cdot)$, $\mathcal{Z}_{ij}^{(me)}(\cdot)$, i.e. the symbols corresponding to the actual relation $\chi_1^*, \ldots, \chi_n^*$ will be marked with asterisks: $\mathcal{U}_{ij}^{(me)*}$, $\mathcal{V}_{ij}^{(me)*}$, $\mathcal{Z}_{ij}^{(me)*}$, while the symbols corresponding to any other relation $\widetilde{\chi}_1, \ldots, \widetilde{\chi}_r$ — with tildas: $\widetilde{\mathcal{U}}_{ij}^{(me)}$, $\widetilde{\mathcal{V}}_{ij}^{(me)}$, $\widetilde{\mathcal{Z}}_{ij}^{(me)}$.

Finally let us define the random variables $\mathcal{W}_\vartheta^*$ and $\widetilde{\mathcal{W}}_\vartheta$:

$$\mathcal{W}_\vartheta^* = \sum_{I^*} \mathcal{U}_{ij}^{(me)*} + \sum_{P_1^*} \mathcal{V}_{ij}^{(me)*} + \sum_{P_2^*} \mathcal{Z}_{ij}^{(me)*}, \tag{74}$$

$$\widetilde{\mathcal{W}}_\vartheta = \sum_{\widetilde{I}} \widetilde{\mathcal{U}}_{ij}^{(me)} + \sum_{\widetilde{P}_1} \widetilde{\mathcal{V}}_{ij}^{(me)} + \sum_{\widetilde{P}_2} \widetilde{\mathcal{Z}}_{ij}^{(me)}. \tag{75}$$

On the basis of the results presented in Klukowski (1994), point 5.2, it is clear that:

$$P(\mathcal{W}_\vartheta^* - \widetilde{\mathcal{W}}_\vartheta < 0) > 1 - 2\lambda_\vartheta, \tag{76}$$

where:

$$\lambda_\vartheta = \exp\{-2\vartheta(1/2 - \delta_{\max}^{(\kappa)})^2\} \tag{77}$$

and

$$\delta_{\max}^{(\kappa)} = \max_{(x_i,x_j)\in\mathbf{X}\times\mathbf{X}} \{P(T(x_i,x_j) \neq g_{me(\kappa)}(x_i,x_j))\}.$$

If $\kappa > 1$, then the convergence obtained as a result of the zero-one transformation is weaker, than in Klukowski (1994), because $\vartheta < N$ in the equality (77) (in other words, the exponent in the right-hand side of relationship (76) "decreases with the step κ"). The case $\kappa = 1$ is not excluded, in general, but it is satisfied only when $P(T(\cdot) - g_k(\cdot) = 0) > 1/2$ for each $(x_i, x_i) \in \mathbf{X}\times\mathbf{X}$.

It seems feasible to prove that the efficiency of the median approach in the case of difference of ranks is not worse than that based on the transformations (72a)–(73c); the problem needs further investigation.

6. Summary

The paper presents two approaches to estimation of the preference relation on the basis of multiple pairwise comparisons in the form of difference of ranks. The results are extensions and complements to the case $N = 1$ (one comparison for each pair), considered in Klukowski (2000); the extension is based on the ideas similar to those developed in Klukowski (1994) (for the case of comparisons indicating the direction of preference). The algorithms presented in the paper are based on weak assumptions about distributions of comparison errors. The properties of the averaging approach, especially exponential convergence of the probability $P(\bar{W}^* < \widetilde{\bar{W}})$ to one for $N \to \infty$, are meaningful. On the other hand, the optimisation problem corresponding to the median approach is easy to solve. The question about efficiency of the median approach, in comparison to averaging approach needs further investigations. It seems reasonable to investigate the properties of the estimators, difficult for analytic examination, with the use of simulation.

References

David, H.A. (1970) *Order Statistics.* J. Wiley, New York.

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

Hastie, T., Tibshirani, R. and Friedman, J. (2001) *The Elements of Statistical Learning.* Springer, New York-Berlin-Heidelberg.

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

Kamishima, T. and Akaho, S. (2006) Supervised Ordering by Regression combined Thurstone's Model. *Artificial Intelligence Review* **25**, 231–246.

KLUKOWSKI, L. (1994) Some probabilistic properties of the nearest adjoining order method and its extensions. *Annals of Operations Research* **51**, 241–261.

KLUKOWSKI, L. (2000) The nearest adjoining order method for pairwise comparisons in the form of difference of ranks. *Annals of Operations Research* **97**, 357–378.

KLUKOWSKI, L. (2007) Estimation of tolerance relation the basis of multiple pairwise comparisons with random errors. *Control and Cybernetics* **36**, 443–466.

MARDEN, J.I. (1995) *Analyzing and Modelling Rank Data. Monograph on Statistics and Applied Probability* **64**. Chapman & Hall.

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

YAGER, R.R. (2007) Aggregation of ordinal information. *Fuzzy Optim. Dec. Making* **6**, 199–219.