

87/2011

**Raport Badawczy**  
**Research Report**

**RB/33/2011**

**Własności estymatorów relacji  
porządku opartych  
na różnicach rang  
– badanie symulacyjne**

**L. Klukowski**

**Instytut Badań Systemowych**  
**Polska Akademia Nauk**

**Systems Research Institute**  
**Polish Academy of Sciences**



# **POLSKA AKADEMIA NAUK**

## **Instytut Badań Systemowych**

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:  
Prof. zw. dr hab. inż. Zbigniew Nahorski

Warszawa 2011

Leszek Klukowski

Instytut Badań Systemowych PAN

## WŁASNOŚCI ESTYMATORÓW RELACJI PORZĄDKU OPARTYCH NA RÓŻNICACH RANG - BADANIE SYMULACYJNE

### Wprowadzenie

Własności estymatorów relacji porządku na podstawie wielokrotnych porównań parami w postaci różnicy rang, określone w pracach Klukowskiego (2007, 2008a, b), dotyczą głównie własności asymptotycznych - nie obejmują rozkładów błędów estymacji. Uzyskanie tych rozkładów w sposób analityczny nie wydaje się możliwe, natomiast ich znajomość jest niezbędna w zastosowaniach. Celem niniejszej pracy jest przedstawienie wyników badań symulacyjnych w tym zakresie – podano rozkłady błędów estymacji obu rozważanych estymatorów dla wybranych rozkładów błędów porównań parami. Przedmiotem rozważań był zbiór 9 – elementowy  $\mathbf{X} = \{x_1, \dots, x_9\}$ , z trzema postaciami relacji porządku:

$\{x_1\}, \dots, \{x_9\}$ ;  $\{x_1\}, \{x_2, x_3\}, \{x_4\}, \{x_5, x_6\}, \{x_7\}, \{x_8, x_9\}$ ;  $\{x_1, x_2\}$ ,  
 $\{x_3, x_4, x_5\}$ ,  $\{x_6, x_7, x_8, x_9\}$  oraz trzema rozkładami błędów porównań (łącznie 90 przypadków). Zadania optymalizacyjne stanowiące podstawę dla wyznaczenia ocen relacji rozwiązywano stosując przegląd pełny: liczba symulacji każdego rozważanego przypadku wynosiła 200. Wnioski z badań umożliwiają, w szczególności, wyznaczenie liczby porównań każdej pary zapewniającej ocenę relacji bliską bezbłędną lub bezbłędną oraz porównanie efektywności obu estymatorów.

Praca składa się z 4 części. W części pierwszej podano niezbędne definicje i oznaczenia, w drugiej parametry przeprowadzonego badania symulacyjnego. Wyniki badania, ich interpretację oraz ilustrację graficzną podano w części trzeciej; w części czwartej podsumowano wyniki.

### 1. Definicje oraz oznaczenia

Problem estymacji relacji porządku, na podstawie porównań parami, można sformułować w następujący sposób. Dany jest zbiór  $\mathbf{X} = \{x_1, \dots, x_m\}$  ( $m \geq 3$ ), w którym istnieje relacja słabego porządku  $R$  o postaci:

$$R = I \cup P, \quad (1)$$

gdzie:

$I$  – relacja równoważności (zwrotna, przechodnia, symetryczna).

$P$  – relacja porządku (przechodnia antysymetryczna).

Relacja  $R$  generuje rodzinę podzbiorów  $\chi_1^*, \dots, \chi_n^*$  ( $n \geq 3$ ) o następujących własnościach:

$$\bigcup_{q=1}^n \chi_q^* = \mathbf{X}, \quad (2)$$

$$\chi_r^* \cap \chi_s^* = \{\mathbf{0}\}, \quad (3)$$

$(x_r \in \chi_r^{(p)r}) \cap (x_s \in \chi_s^{(p)s}) \equiv$  element  $x_r$  jest preferowany w stosunku do  $x_s$ ,  
w przypadku  $r < s$ . (4)

gdzie:  $\{\mathbf{0}\}$  – zbiór pusty,

Relacja  $\chi_1^*, \dots, \chi_n^*$  może być wyrażona przy użyciu funkcji  $T_{(x_r, x_s)}$

$((x_r, x_s) \in \mathbf{X} \times \mathbf{X})$ , o postaci:

$$T_{(x_r, x_s)} = r - s. \quad (5)$$

przy czym:  $x_r \in \mathcal{X}_r^*$  oraz  $x_l \in \mathcal{X}_s^*$ .

Relację  $\chi_1^*, \dots, \chi_n^*$  należy estymować na podstawie  $N$  – krotnej ( $N \geq 2$ ) oceny  $g_k(x_i, x_j)$  wartości  $T(x_i, x_j)$  ( $(x_i, x_j) \in \mathbf{X} \times \mathbf{X}$ ), zakłóconej błędem losowym. Błąd losowy oznacza, że każda różnica  $T(x_i, x_j) - g_k(x_i, x_j)$  jest zmienną losową o wartościach ze zbioru  $\{-(m-1), \dots, 0, \dots, m-1\}$ .

Przyjmuje się następujące założenia nt. błędów porównań:

Z1. Rozkłady błędów  $T(x_i, x_j) - g_k(x_i, x_j)$  ( $(x_i, x_j) \in \mathbf{X} \times \mathbf{X}$ ) są jednomodalne z wartością modalną i medianą równą zero; mediana spełnia warunki:

$$\sum_{l \geq 0} P((T(\cdot) - g_k(\cdot) = l) > 1/2, \sum_{l \geq 0} P((T(\cdot) - g_k(\cdot) = l) < 1/2.$$

Z2. Porównania  $g_k(x_i, x_j)$  oraz  $g_l(x_r, x_s)$  ( $k \neq l$ ) są niezależne. tzn.:

$$P((T(x_i, x_j) - g_k(x_i, x_j) = c_{kl}) \cap (T(x_r, x_s) - g_l(x_r, x_s) = c_{lr})) =$$

$$P((T(x_i, x_j) - g_k(x_i, x_j) = c_{kl}) P(T(x_r, x_s) - g_l(x_r, x_s) = c_{lr})).$$

Dwa rodzaje estymatorów rozważono we wcześniejszych pracach (Klukowski 1994, 2007b, 2008); pierwszy z nich opiera się na sumie różnic między postacią relacji i porównaniami, drugi – między postacią relacji i medianą z porównań każdej pary. Oceny otrzymuje się na podstawie rozwiązań zadań optymalizacji dyskretnej:

$$\min_{\chi_1^{(1)}, \dots, \chi_n^{(1)} \in F_N} \left\{ \sum_{\langle i, j \rangle \in R_m} \sum_{k=1}^N |t^{(1)}(x_i, x_j) - g_k(x_i, x_j)| \right\}, \quad (7)$$

$$\min_{\chi_1^{(1)}, \dots, \chi_n^{(1)} \in F_N} \left\{ \sum_{\langle i, j \rangle \in R_m} |t^{(1)}(x_i, x_j) - g^{(me)}(x_i, x_j)| \right\}, \quad (8)$$

gdzie:

$F_N$  - zbiór rozwiązań dopuszczalnych (rodzina wszystkich relacji porządku w zbiorze  $\mathbf{X}$ ).

$\chi_1^{(i)}, \dots, \chi_r^{(i)}$  -  $i$ -ty element zbioru  $F_N$ ,

$R_m$  - zbiór par indeksów o postaci:  $R_m = \{ \langle i, j \rangle \mid 1 \leq i, j \leq m; j > i \}$ ,

$f^{(i)}(x_i, x_j)$  - funkcja, określająca relację  $\chi_1^{(i)}, \dots, \chi_r^{(i)}$ , zdefiniowana analogicznie do  $T(x_i, x_j)$ ,

$g^{i^{med}}(x_i, x_j)$  - mediana w zbiorze  $\{g_1(x_i, x_j), \dots, g_N(x_i, x_j)\}$ .

Oceny relacji otrzymane na podstawie zadań (7) oraz (8) będą oznaczane symbolami – odpowiednio:  $\hat{\chi}_1, \dots, \hat{\chi}_n$  oraz  $\bar{\chi}_1, \dots, \bar{\chi}_n$ , a funkcje określające te relacje symbolami:  $\hat{T}(x_i, x_j)$  oraz  $\bar{T}(x_i, x_j)$ . Należy zaznaczyć, że każde z zadań (7), (8) może mieć rozwiązanie wielokrotne. Błędy obu estymatorów – oznaczane symbolami  $\hat{\Delta}$  oraz  $\bar{\Delta}$  - przyjęto w postaci:

$$\hat{\Delta} = \sum_{\langle i, j \rangle \in R_m} |T(x_i, x_j) - \hat{T}(x_i, x_j)|, \quad (9)$$

$$\bar{\Delta} = \sum_{\langle i, j \rangle \in R_m} |T(x_i, x_j) - \bar{T}(x_i, x_j)|. \quad (10)$$

## 2. Parametry badania symulacyjnego

Badanie symulacyjne rozkładów błędów estymacji zostało wykonane na podstawie trzech postaci relacji o następujących parametrach.

- Zbiór  $X$  zawierający 9 elementów;
- trzy postacie relacji:
  - dziewięć podzbiorów (porządek liniowy):  $\{x_1\}, \dots, \{x_9\}$  ( $9 \times 9$ ),
  - sześć podzbiorów:  $\{x_1\}, \{x_2, x_3\}, \{x_4\}, \{x_5, x_6\}, \{x_7\}, \{x_8, x_9\}$  ( $6 \times 9$ ),
  - trzy podzbiory:  $\{x_1, x_2\}, \{x_3, x_4, x_5\}, \{x_6, x_7, x_8, x_9\}$  ( $3 \times 9$ );
- funkcja prawdopodobieństwa błędów porównań o postaci:

$$P(g_k(x_i, x_j) = T(x_i, x_j)) = \alpha_{ij},$$

$$P(g_k(x_i, x_j) - T(x_i, x_j) = -l) = (1 - \alpha_{ij}) / 2L_{ij}^{(d,l)}$$

$$(L_{ij}^{(d,l)} = T(x_i, x_j) + (m-1); l = -1, \dots, -L_{ij}^{(d,l)}),$$

$$P(g_k(x_i, x_j) - T(x_i, x_j) = l) = (1 - \alpha_{ij}) / 2L_{ij}^{(u,l)}$$

$$(L_{ij}^{(u,l)} = m-1 - T(x_i, x_j); l = 1, \dots, L_{ij}^{(u,l)});$$

zbadano trzy wartości  $\alpha_{ij}$ : 0,3334; 0,4167; 0,5000 ( $i, j \in R_m$ );

• liczba  $N$  porównań każdej pary: 1, 3, 5, 7, 9.

Łącznie zbadano 90 przypadków, tj.: trzy relacje, trzy postacie rozkładu błędów porównań, pięć wartości liczby porównań, przy wykorzystaniu dwóch estymatorów. Każdy przypadek symulowano 200rotnie, przy zastosowaniu generatora liczb losowych rozkładu równomiernego.

### 3. Wyniki badania symulacyjnego

Wyniki badania symulacyjnego mają postać rozkładów częstości błędów (7) i (8). Częstości te podano w 10 przedziałach: pierwsza częstość odnosi się do oceny bezbłędnej (błąd równy zeru), ostatnia – do błędów o wartości powyżej 72, osiem „środkowych” – do przedziałów o postaci  $(8v, 8(v+1)]$  ( $v=0, \dots, 8$ ). Rozpiętość „środkowych” przedziałów określono w taki sposób, aby zawierały oceny o podobnej odległości od rzeczywistej (symulowanej) relacji. Zauważmy, że np. zastąpienie relacji  $\{x_1\}, \{x_2\}, \dots, \{x_9\}$  przez ocenę  $\{x_1, x_2\}, \dots, \{x_9\}$  generuje błąd równy siedem tzn. błąd estymacji polegający na przemieszczeniu jednego elementu generuje błąd  $\hat{\Delta}$  lub  $\hat{\Delta}$  o wartości znacznie większej niż jedność. Wyniki badania zamieszczono w tablicach 1 – 9; uwzględniają one wielokrotne rozwiązania optymalne, np. w przypadkach dwóch

rozwiązań o wartościach błędu zero i siedem (średni) błąd wynosi 3,5.  
Wybrane rozkłady częstości przedstawiono na rysunkach 1 - 8.

Tabela 1

Częstości błędów estymacji relacja:  $9 \times 9$  oraz  $\alpha=0.3334$

Błąd estym.	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,140	0,780	0,545	0,930	0,800	0,995	0,955	0,995	0,980
(0, 8]	0,035	0,075	0,175	0,035	0,100	0	0,020	0	0,010
(8, 16]	0,055	0,045	0,125	0,025	0,050	0	0,005	0,005	0,010
(16, 24]	0,040	0,040	0,050	0	0,030	0	0,015	0	0
(24, 32]	0,080	0,035	0,055	0	0,010	0,005	0,005	0	0
(32, 40]	0,120	0,005	0,045	0,005	0	0	0	0	0
(40, 48]	0,060	0,015	0,020	0	0,005	0	0	0	0
(48, 56]	0,105	0,005	0,020	0	0,005	0	0	0	0
(56, 64]	0,105	0	0,010	0,005	0	0	0	0	0
(64, 72]	0,080	0	0,005	0	0	0	0	0	0
>72	0,180	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Tabela 2

Częstości błędów estymacji; relacja  $9 \times 9$  oraz  $\alpha=0,4167$

Bł.	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,350	0,890	0,770	0,990	0,965	1,0	0,985	1,0	1,0
(0,8]	0,090	0,065	0,075	0,005	0,015	0	0,010	0	0
(8,16]	0,085	0,010	0,050	0	0,015	0	0,005	0	0
(16,24]	0,100	0,015	0,040	0,005	0	0	0	0	0
(24,32]	0,065	0,015	0,030	0	0,005	0	0	0	0
(32,40]	0,035	0,005	0,025	0	0	0	0	0	0
(40,48]	0,085	0	0,010	0	0	0	0	0	0
(48,56]	0,050	0	0	0	0	0	0	0	0
(56,64]	0,045	0	0	0	0	0	0	0	0
(64,72]	0,035	0	0	0	0	0	0	0	0
>72	0,060	0	0	0	0	0	0	0	0

Źródło: obliczenia własne



Tabela 3

Częstości błędów estymacji; relacja  $9 \times 9$  oraz  $\alpha=0.500$ 

Błędy	N=1		N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.	
0	0,600	0,985	0,920	1,0	0,995	1,0	1,0	1,0	1,0	
(0,8]	0,095	0,010	0,040	0	0,005	0	0	0	0	
(8,16]	0,080	0,005	0,015	0	0	0	0	0	0	
(16,24]	0,025	0	0,010	0	0	0	0	0	0	
(24,32]	0,050	0	0,010	0	0	0	0	0	0	
(32,40]	0,040	0	0,005	0	0	0	0	0	0	
(40,48]	0,025	0	0	0	0	0	0	0	0	
(48,56]	0,035	0	0	0	0	0	0	0	0	
(56,64]	0,015	0	0	0	0	0	0	0	0	
(64,72]	0,015	0	0	0	0	0	0	0	0	
>72	0,025	0	0	0	0	0	0	0	0	

Źródło: obliczenia własne

Tabela 4

Częstości błędów estymacji; relacja  $6 \times 9$  oraz  $\alpha=0.3334$ 

Błędy	N=1		N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.	
0	0,025	0,420	0,155	0,775	0,485	0,915	0,635	0,990	0,825	
(0,8]	0,045	0,220	0,240	0,185	0,329	0,075	0,275	0,010	0,165	
(8,16]	0,100	0,165	0,245	0,015	0,125	0,005	0,070	0	0,005	
(16,24]	0,190	0,110	0,195	0,010	0,060	0,005	0,020	0	0,005	
(24,32]	0,195	0,040	0,055	0,015	0,005	0	0	0	0	
(32,40]	0,165	0,025	0,075	0	0,01	0	0	0	0	
(40,48]	0,090	0,015	0,020	0	0,005	0	0	0	0	
(48,56]	0,100	0	0,010	0	0	0	0	0	0	
(56,64]	0,025	0	0	0	0	0	0	0	0	
(64,72]	0,030	0,005	0	0	0	0	0	0	0	
>72	0,035	0	0	0	0	0	0	0	0	

Źródło: obliczenia własne

Tabela 5

Częstości błędów estymacji; relacja  $6 \times 9$  oraz  $\alpha=0,4167$ 

Błąd estym.	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,115	0,790	0,560	0,970	0,815	0,995	0,945	1,0	0,970
(0,8]	0,090	0,155	0,260	0,025	0,145	0,005	0,055	0	0,030
(8,16]	0,170	0,020	0,120	0	0,030	0	0	0	0
(16,24]	0,190	0,025	0,004	0,005	0,005	0	0	0	0
(24,32]	0,150	0,010	0	0	0,005	0	0	0	0
(32,40]	0,100	0	0,015	0	0,005	0	0	0	0
(40,48]	0,090	0	0,005	0	0	0	0	0	0
(48,56]	0,025	0	0	0	0	0	0	0	0
(56,64]	0,035	0	0	0	0	0	0	0	0
(64,72]	0,020	0	0	0	0	0	0	0	0
>72	0,015	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Tabela 6

Częstości błędów estymacji; relacja  $6 \times 9$  oraz  $\alpha=0,5000$ 

Błąd estym.	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,255	0,880	0,775	1,000	0,950	1,000	1,000	1,000	1,000
(0,8]	0,180	0,100	0,155	0	0,050	0	0		
(8,16]	0,220	0,015	0,045	0	0	0	0	0	0
(16,24]	0,115	0,005	0,015	0	0	0	0	0	0
(24,32]	0,060	0	0,010	0	0	0	0	0	0
(32,40]	0,075	0	0	0	0	0	0	0	0
(40,48]	0,050	0	0	0	0	0	0	0	0
(48,56]	0,020	0	0	0	0	0	0	0	0
(56,64]	0,015	0	0	0	0	0	0	0	0
(64,72]	0,010	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Tabela 7

Częstości błędów estymacji; relacja  $3 \times 9$  oraz  $\alpha=0,3334$

Błąd	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0	0,530	0,280	0,840	0,540	0,955	0,720	0,980	0,840
(0,8]	0	0,345	0,385	0,125	0,345	0,045	0,275	0,020	0,16
(8,16]	0,015	0,105	0,235	0,035	0,095	0	0,005	0	0
(16,24]	0,120	0,015	0,090	0	0,020	0	0	0	0
(24,32]	0,125	0,005	0,010	0	0	0	0	0	0
(32,40]	0,145	0	0	0	0	0	0	0	0
(40,48]	0,230	0	0	0	0	0	0	0	0
(48,56]	0,225	0	0	0	0	0	0	0	0
(56,64]	0,105	0	0	0	0	0	0	0	0
(64,72]	0,035	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Tabela 8

Częstości błędów estymacji; relacja  $3 \times 9$  oraz  $\alpha=0,4167$

Błąd	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,145	0,785	0,530	0,970	0,820	0,990	0,920	1,000	1,000
(0,8]	0,255	0,190	0,360	0,030	0,165	0,010	0,080	0	0
(8,16]	0,340	0,020	0,100	0	0,010	0	0	0	0
(16,24]	0,170	0,005	0,010	0	0,005	0	0	0	0
(24,32]	0,045	0	0	0	0	0	0	0	0
(32,40]	0,020	0	0	0	0	0	0	0	0
(40,48]	0,015	0	0	0	0	0	0	0	0
(48,56]	0,005	0	0	0	0	0	0	0	0
(56,64]	0,005	0	0	0	0	0	0	0	0
(64,72]	0	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Tabela 9

Częstości błędów estymacji: relacja  $3 \times 9$  oraz  $\alpha=0,500$ 

Błąd estym.	N=1	N=3		N=5		N=7		N=9	
	Sum.	Sum.	Med.	Sum.	Med.	Sum.	Med.	Sum.	Med.
0	0,310	0,915	0,795	0,995	0,940	1,000	0,975	1,000	1,000
(0,8]	0,325	0,085	0,190	0,005	0,060	0	0,025	0	0
(8,16]	0,215	0	0,015	0	0	0	0	0	0
(16,24]	0,130	0	0	0	0	0	0	0	0
(24,32]	0,015	0	0	0	0	0	0	0	0
(32,40]	0,005	0	0	0	0	0	0	0	0
(40,48]	0	0	0	0	0	0	0	0	0
(48,56]	0	0	0	0	0	0	0	0	0
(56,64]	0	0	0	0	0	0	0	0	0
(64,72]	0	0	0	0	0	0	0	0	0
>72	0	0	0	0	0	0	0	0	0

Źródło: obliczenia własne

Zadania (7), (8) stanowiące podstawę estymacji rozwiązywano metodą przeglądu quasi-pełnego, tzn. wykonywano przegląd dla liczby podzbiorów:  $n-1$ ,  $n$ ,  $n+1$ . Analiza wyników estymacji prowadzi do następujących wniosków.

1<sup>o</sup>. Estymator oparty sumie różnic (bezwzględnych) między funkcją  $T(x_i, x_j)$ , określającą postać relacji, i oceną, tj. na rozwiązaniu zadania (7), zapewnia, w przypadku wielokrotnych porównań każdej pary ( $N>1$ ), lepszą precyzję estymacji niż estymator oparty na medianach z porównań każdej pary, tj. na rozwiązaniu zadania (8). Określenie większa precyzja oznacza: wyższą częstość oceny bezbłędnej oraz większą koncentrację rozkładu częstości w otoczeniu błędu równego zero. Różnica precyzji jest niewielka w przypadku rozkładu częstości bliskiego jednopunktowemu.

tj. w przypadku, gdy częstość błędu równego zero estymatora opartego na medianach osiąga poziom nie mniejszy niż 0,95.

2<sup>0</sup>. Wzrost liczby porównań  $N$  powoduje szybkie zwiększanie precyzji estymacji. W przypadku estymatora opartego na sumie różnic oraz prawdopodobieństwa bezbłędneho porównania  $\alpha=0,5$ , trzy porównania ( $N=3$ ) zapewniają ocenę relacji bliską bezbłędnej; w przypadku estymatora opartego na medianach liczba porównań  $N$  musi być większa o dwa. W przypadku pozostałych wartości  $\alpha$ , tj. 0,3334 i 0,4167 oraz estymatora opartego na sumie różnic, liczba porównań, gwarantująca częstość bliską jedności, musi być nie mniejsza niż – odpowiednio: siedem i pięć. W przypadku estymatora opartego na medianach liczba ta musi być większa (co najmniej o dwa). Dokonanie 7 – 9 porównań każdej pary gwarantuje precyzję estymacji bliską lub równą 100%.

3<sup>0</sup>. Najlepsza precyzja estymacji ma miejsce w przypadku relacji mocnego porządku, tj.  $n=9$ ; pozostałe przypadki, tj.  $n=3$  oraz  $n=6$ , charakteryzują się nieco niższą – podobną precyzją. W przypadku relacji mocnego porządku akceptowalną precyzję (częstość bezbłędneho porównania większą niż 50%) zapewnia pojedyncze porównanie każdej pary  $N=1$  dla  $\alpha=0,5$ .

4<sup>0</sup>. Rozkłady błędów estymacji charakteryzują się dużą koncentracją w przypadku, gdy częstość bezbłędnej oceny relacji przekracza 0,5; częstość błędu nie przekraczającego wartości 16 jest wówczas bliska jedności. Jeśli częstość bezbłędnej oceny przekracza 0,75, to analogiczna własność dotyczy błędu estymacji nie przekraczającego 8. Mały błąd estymacji implikuje niewielką różnicę między oceną i rzeczywistą relacją.

5<sup>0</sup>. W przypadku, gdy nie wszystkie rozkłady porównań parami spełniają założenia Z1 i Z2, wyniki estymacji można uznać za nieakceptowalne, tj.

zero nie jest dominantą ani medianą rozkładu częstości błędów estymacji; sytuacja taka ma miejsce dla:  $N=1$  i  $\alpha \leq 0,4167$  oraz  $N=3$  i  $\alpha = 0,3334$ .

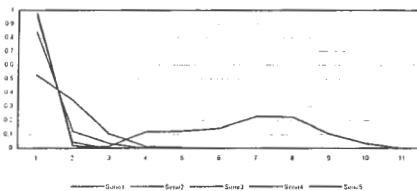
Powyższa analiza wyników symulacji pozwala określić: liczbę porównań  $N$  oraz wartość prawdopodobieństwa  $\alpha$  niezbędne do uzyskania wymaganej precyzji estymacji. Należy podkreślić, że rozkład błędu porównania każdej pary, zastosowany w badaniu symulacyjnym (quasi-równomierny), jest niekorzystny z punktu widzenia estymacji; wykorzystanie rozkładu ściśle monotonicznego, w którym większym, co do modułu, błędom odpowiadają mniejsze prawdopodobieństwa, zapewnia polepszenie precyzji ocen.

#### 4. Podsumowanie

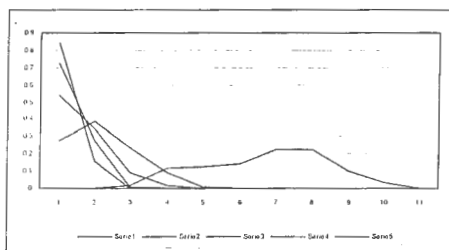
Wyniki badań symulacyjnych, omówionych w niniejszej pracy, są podstawą do następujących - syntetycznych wniosków: • estymatory relacji porządku oparte na różnicy rang, zaproponowane przez autora, charakteryzują się dobrą precyzją, przy słabych założeniach o błędach porównań; • najlepszą precyzję estymacji uzyskuje się w przypadku relacji mocnego porządku; • estymator oparty na sumie różnic między funkcją określającą postać relacji, a porównaniami, wykazuje znaczną przewagę precyzji nad estymatorem opartym na medianach z porównań; • precyzja estymacji polepsza się znacząco wraz z liczbą porównań  $N$ ; można wyznaczyć liczebność porównań zapewniającą ocenę bezbłędną lub o nieznacznym błędzie.

Można postulować, że powyższe wnioski znajdują zastosowanie w przypadku zbiorów  $X$  o większej liczbie elementów, niż w przeprowadzonym badaniu. Postulat ten potwierdzono dla liczebności nie większych niż 12, przy węższym zakresie symulacji – ze względu na

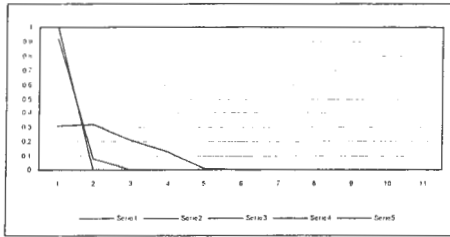
nakład obliczeń. Należy dodać, że wyznaczenie ocen relacji dla  $n > 12$  wymaga stosowania innych algorytmów niż przegląd pełny, tj. algorytmów optymalizacji dyskretnej lub algorytmów przybliżonych.



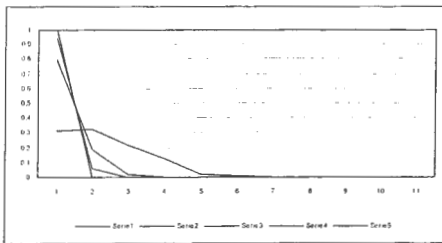
Rys. 1. Wykres częstości błędów relacji  $3 \times 9$ ,  $\alpha = 0,3334$ ;  
Estymator oparty na sumie różnic



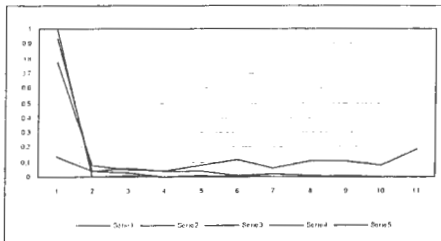
Rys. 2. Wykres częstości błędów relacji  $3 \times 9$ ,  $\alpha = 0,3334$ ;  
Estymator oparty na medianach z porównań



Rys. 3. Wykres częstości błędów relacji  $3 \times 9$ ,  $\alpha=0.5000$ :  
Estymator oparty na sumie różnic

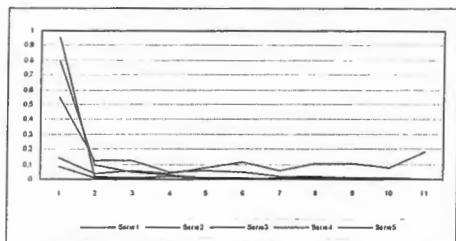


Rys. 4. Wykres częstości błędów relacji  $3 \times 9$ ,  $\alpha=0.5000$ :  
Estymator oparty na medianach

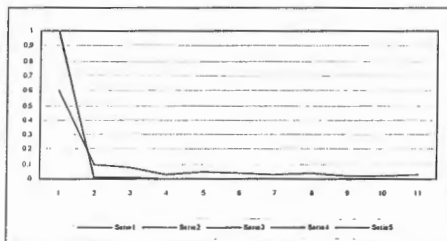


Rys. 5. Wykres częstości błędów relacji  $9 \times 9$ ,  $\alpha=0.3334$ :  
Estymator oparty na sumie różnic

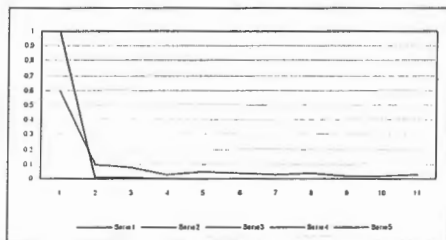




Rys. 6. Wykres częstości błędów relacji  $9 \times 9$ ,  $\alpha=0,3334$ ;  
Estymator oparty na medianach



Rys. 7. Wykres częstości błędów relacji  $9 \times 9$ ,  $\alpha=0,5000$ ;  
Estymator oparty na sumie różnic



Rys. 8. Wykres częstości błędów relacji  $9 \times 9$ ,  $\alpha=0,5000$ ;  
Estymator oparty na medianach

## Literatura

- David H. A. (1988) The Method of Paired Comparisons. 2<sup>nd</sup> ed. Ch. Griffin, London.
- Hoeffding W. (1963) Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc., vol. 58, pp. 13-30.
- Klukowski L. (1994) Some probabilistic properties of the nearest adjoining order method and its extensions. Annals of Operational Research, vol. 51, pp. 241 - 261.
- Klukowski L. (2000) The nearest adjoining order method for pairwise comparisons in the form of difference of ranks. Annals of Operational Research, vol. 97, pp. 357 -378.
- Klukowski (2007) Estimation of the Preference Relation on the Basis of Medians from Pairwise Comparisons in the Form of Difference of Ranks. W: Kurzynski (et al) red., Computer Recognition Systems 2, Springer, Berlin Heidelberg New York.
- Klukowski L. (2008a) Estimation of the preference relation on the basis of multiple pairwise comparisons in the form of difference of ranks. Control and Cybernetics 37, No. 3, 711 – 729.
- Klukowski (2008b) Estymacja relacji porządku na podstawie porównań parami z błędami losowymi. W: Trzaskalik T. (red.) MODELOWANIE PREFERENCJI A RYZYKO '09, 61 – 76.
- Slater P. (1961) Inconsistencies in a schedule of paired comparisons. Biometrika, vol. 48, pp. 303 - 312.



