

**Л.Лукашевська, О.Б.Рудик**  
**Дослідження умов збереження ієрархії результатів**  
**учасників учнівської олімпіади з інформатики**

### **Передмова**

У роботі викладено результати дослідження відповідності оцінок учнів міста Києва, отриманих ними під час проведення III і IV етапів Всеукраїнської учнівської олімпіади з інформатики та відбірково-тренувальних зборів команди міста Києва до IV етапу у 2002–2003 роках.

**Первинні дані** (бали, отримані учнями на III і IV етапах Всеукраїнської учнівської олімпіади з інформатики та відбірково-тренувальних зборів команди міста Києва до IV етапу) запозичено з таких сайтів:

- <http://www.uoi.kiev.ua> (Українські учнівські олімпіади з інформатики);
- <http://kievoi.narod.ru> (Київські учнівські олімпіади з інформатики).

Ці самі сайти містять порядок проведення відповідних заходів.

#### **Мета дослідження:**

- довести або спростувати слушність технології підготовки команди міста Києва;
- виявити причин зростання ймовірності випадкових порушень ієрархії результатів;
- виробити рекомендації щодо удосконалення проведення відбірково-тренувальних зборів до IV етапу олімпіади з інформатики.

**Актуальність дослідження** визначається важливістю роботи з обдарованими дітьми як з представниками майбутньої інтелектуальної еліти взагалі і створення умов добросовісної та досконалої конкуренції між ними зокрема.

## **1. Коефіцієнт рангової кореляції Спірмена**

Порядковий характер шкали вимірювання результатів та невеликі довжини вибірки суттєво звужують коло статистичних методів, які можна використати у роботі. Результати отримано шляхом обчислення коефіцієнта рангової кореляції  $r_s$  Спірмена й подальшим порівнянням знайденої величини з критичним значенням. Подамо опис критерію, обраного для виконання дослідження, запозичений з монографії [1].

**Призначення:** визначення сили й напрями кореляційного зв'язку між двома ознаками.

#### **Обмеження:**

- нижня межа кількості вимірювань — 5;
- верхня межа кількості вимірювань — до 40 — визначається наявними таблицями критичних значень.

#### **Гіпотези:**

$H_0$  — кореляція між змінними (ієрархіями) не відрізняється від 0;

$H_1$  — кореляція між змінними (ієрархіями) достовірно відмінна від 0.

#### **Алгоритм**

1. Даним вимірюванням кожної змінної приписуємо ранги — номери відповідних елементів впорядкованого за зростанням масиву. Якщо кілька вимірювань мають один і той же результат, то всім їм приписуємо ранг, що є середнім арифметичним номерів відповідних елементів впорядкованого масиву. Таким чином, сума рангів усіх результатів вимірювання для вибірки довжини  $n$  дорівнює  $n(n + 1)/2$ .

2. Знаходимо  $\sum_{j=1}^n \Delta_j^2$  — суму квадратів різниць рангів.

3. Знаходимо поправки на наявність однакових рангів:

$$T_A = \sum_a (a^3 - a)/12; \quad T_B = \sum_b (b^3 - b)/12$$

де  $a, b$  — об'єми кожної групи однакових рангів відповідно для результатів вимірювання двох змінних. Ці поправки дорівнюють 0, якщо результати вимірювань різні.

4. Обчислюємо коефіцієнт рангової кореляції:

$$r_s = 1 - 6 \cdot \frac{\sum_{j=1}^n \Delta_j^2 + T_A + T_B}{n \cdot (n^2 - 1)}.$$

5. За таблицею 1, запозиченою з публікацій [1–2] визначаємо  $r_c$  — критичне значення коефіцієнту рангової кореляції для даного  $n$ .

6. Якщо  $r_c \leq r_s$ , то приймаємо гіпотезу  $H_0$ , інакше —  $H_1$ .

**Таблиця 1**

**Критичні значення коефіцієнту рангової кореляції  $r_c$**

$n$	$\rho$		$n$	$\rho$		$n$	$P$		$n$	$\rho$	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
5	0,94	–	14	0,54	0,68	23	0,42	0,53	32	0,36	0,45
6	0,85	–	15	0,52	0,66	24	0,41	0,52	33	0,34	0,45
7	0,78	0,94	16	0,50	0,64	25	0,40	0,51	34	0,34	0,44
8	0,72	0,88	17	0,48	0,62	26	0,39	0,50	35	0,33	0,43
9	0,68	0,83	18	0,47	0,60	27	0,38	0,49	36	0,33	0,43
10	0,64	0,79	19	0,46	0,58	28	0,38	0,48	37	0,33	0,43
11	0,61	0,76	20	0,45	0,57	29	0,37	0,48	38	0,32	0,41
12	0,58	0,73	21	0,44	0,56	30	0,36	0,47	39	0,32	0,41
13	0,56	0,70	22	0,43	0,54	31	0,36	0,46	40	0,31	0,40

У таблиці 1 *рівень значимості критерію  $\rho$*  — це ймовірність помилки I роду, тобто помилки відхилення нульової гіпотези  $H_0$  за умови, що вона справджується. Склалися традиційні уявлення про *низький, достатній і високий рівні* статистичної значимості як такі, що не перевищують відповідно 0,05 (5%), 0,01 (1%) і 0,001 (0,1%).

## 2. Результати дослідження

Досліджено рангову кореляцію окремо для кожної пари рангів. Ранги виставлено у порядку зростання кількості отриманих балів, протилежному до порядку присудження місць.

У таблицях 2–7 подано такі величини:

$R_3$  — ранги згідно з результатами III етапу;

$R_S$  — ранги, згідно з результатами відбірково-тренувальних зборів;

$R_4$  — ранги згідно з результатами VI етапу;

$D_{3S} = R_3 - R_S$ ;  $D_{34} = R_3 - R_4$ ;  $D_{S4} = R_S - R_4$  — різниці рангів.

Таблиця 2

Результати 2002 року ( $n = 9$ )

Прізвище, ім`я	$R_3$	$R_S$	$R_4$	$D_{3S}$	$D_{34}$	$D_{S4}$	$D^2_{3S}$	$D^2_{34}$	$D^2_{S4}$
Корюкалов Олександр	2	1	1	1	1	0	1	1	0
Павличко Віктор	8	2	2	6	6	0	36	36	0
Знов'як Юрій	1	3	4	-2	-3	-1	4	9	1
Гігіняк Володимир	9	4	7	5	2	-3	25	4	9
Михайлова Марія	5	5	5	0	0	0	0	0	0
Ліссов Павло	3,5	6	8	-2,5	-4,5	-2	6,25	20,25	4
Нестерук Володимир	3,5	7	6	-3,5	-2,5	1	12,25	6,25	1
Веденський Кирило	7	8	3	-1	4	5	1	16	25
Ткачук Володимир	6	9	9	-3	-3	0	9	9	0
<b>Суми квадратів різниць рангів</b>							94,5	101,5	40
$r_s$							0,20833	0,15	0,(6)
<b>Гіпотеза</b>							$H_1$	$H_1$	$H_1$

Лише для відповідності результатів відбірково-тренувальних зборів і IV етапу  $r_s \approx r_c$  з точністю до 1% для рівня значимості 5%. У всіх інших випадках коефіцієнт  $r_s$  суттєво менший за  $r_c$ . Цей факт має доволі прозоре тлумачення. У 2001–2002 навчальному році і II (районний), і III (міський) етап проходили під загальним гаслом „Примітивна графіка, алгебра многочленів над полем раціональних чисел, розв’язування систем лінійних рівнянь на олімпіадах з інформатики”. Саме під такою назвою пізніше були опубліковано опис завдань. Такі завдання не є самими ймовірними, щонайменше для IV етапу. А завдання відбірково-тренувальних зборів вибиралася збалансованими щодо ймовірності випадання певної теми на олімпіаді з інформатики.

Чому ж тоді порушено кореляцію результатів відбірково-тренувальних зборів і IV етапу? Є доволі просте і прозоре пояснення. Найбільший вклад у суму квадратів різниць рейтингів вніс учень 11 класу, прізвище якого вилілено курсивом. З усіх членів команди лише він того року брав участь у IV етапі олімпіади з математики, де посів II місце з правом участі у відбірково-тренувальних зборах команди України до Міжнародної олімпіади з математики. Зауважимо, що:

При цьому:

- III етап олімпіади з математики відбувся за чотири тижні до I туру III етапу олімпіади з інформатики у тому самому місті;
- IV етап олімпіади з математики закінчився за тиждень до початку IV етапу олімпіади.

Неможливість зняти сильне психічне навантаження у такий короткий строк — очевидна причина порушення спаду результативності. Якщо вилучити результати цього учня з розгляду, отримаємо наступну таблицю 3.

Таблиця 3

Результати 2002 року без врахування результатів учасника  
IV етапу олімпіади з математики ( $n = 8$ )

Прізвище, ім`я	$R_S$	$R_4$	$D_{S4}$	$D^2_{S4}$
Корюкалов Олександр	1	1	0	0
Павличко Віктор	2	2	0	0
Знов'як Юрій	3	3	0	0
Гігняк Володимир	4	6	-2	4
Михайлова Марія	5	4	1	1
Лісов Павло	6	7	-1	1
Нестерук Володимир	7	5	2	4
Ткачук Володимир	8	8	0	0
<b>Суми квадратів різниць рангів</b>				10
				$r_s$ 0,88
<b>Гіпотеза</b>				$H_0$

**Висновок.** При розгляді результатів лише тих членів команди на IV етап 2002 року, хто не мав стресового навантаження у вигляді олімпіади з іншого предмету з перспективою участі у Міжнародній олімпіаді, спостерігаємо відповідність між результатами відбірково-тренувальних зборів і IV етап на достатньому рівні статистичної значимості.

Таблиця 4

Результати 2003 року ( $n = 10$ )

Прізвище, ім`я	$R_3$	$R_S$	$R_4$	$D_{3S}$	$D_{34}$	$D_{S4}$	$D^2_{3S}$	$D^2_{34}$	$D^2_{S4}$
Брюхацький Дмитро	2,5	1	1	1,5	1,5	0	2,25	2,25	0
Шурига Олексій	2,5	2	2,5	0,5	0	-0,5	0,25	0	0,25
Корюкалов Олександр	4	3	7	1	-3	-4	1	9	16
Корнієнко Антон	6	4	2,5	2	3,5	1,5	4	12,25	2,25
<u>Петушков</u> Тарас	1	5	9	-4	-8	-4	16	64	16
Гриненко Андрій	8	6	6	2	2	0	4	4	0
Кордубан Дмитро	7	7	4	0	3	3	0	9	9
<i>Челноков Володимир</i>	10	8	5	2	5	3	4	25	9
<i>Терлецький Юрій</i>	5	9	8	-4	-3	1	16	9	1
Знов'як Юрій	9	10	10	-1	-1	0	1	1	0

<b>Суми квадратів різниць рангів</b>	48,5	135,5	53,5
$r_s$	0,70303	0,17273	0,67273
<b>Гіпотеза</b>	$H_0$	$H_1$	$H_0$

Учні 10 класу, прізвища яких виділено курсивом, брали участь у IV етапі олімпіади з математики і *не* здобули права взяти участь у відбірково-тренувальних зборах до Міжнародної олімпіади з математики. Їхні результати не впливають найсуттєвіше на зменшення коефіцієнту рангової кореляції. Найсуттєвіший мають результати учня, чие прізвище виділено підкреслюванням. Особливості результатів цього учня — сталий приріст (на 4 позиції) при переході від III етапу до зборів і від зборів до IV етапу. Можливо, усвідомлення того, які вимоги висувають до учасника олімпіади з інформатики, в учня виникло лише під час тривалого процесу відбірково-тренувальних зборів і призвело до I місця на IV етапі. Принагідно зазначимо, що таке саме порушення ієрархії результатів у позитивному напрямку при переході від III етапу до відбірково-тренувальних зборів мав ще один учень. Відсутність такого ж приросту *відносних* результатів при переході від зборів до IV етапу у нього пояснюється не участю у IV етапі олімпіади з *математики*, а виходом на лідерські позиції.

Якщо вилучити результати учня, який *на відбірково-тренувальних зборах демонстрував найбільший приріст відносного результату, стартуючи з найгіршого результату на III етапі без виходу на лідерські позиції у відбірково-тренувальних зборах*, то отримаємо наступну таблицю 5.

**Таблиця 5**

**Результати 2003 року ( $n = 9$ )**

<b>Прізвище, ім`я</b>	$R_3$	$R_S$	$R_4$	$D_{3S}$	$D_{34}$	$D_{S4}$	$D^2_{3S}$	$D^2_{34}$	$D^2_{S4}$
Брюхацький Дмитро	1,5	1	1	0,5	0,5	0	0,25	0,25	0
Шурига Олексій	1,5	2	2,5	-0,5	-1	-0,5	0,25	1	0,25
Корюкалов Олександр	3	3	7	0	-4	-4	0	16	16
Корнієнко Антон	5	4	2,5	1	2,5	1,5	1	6,25	2,25
Гриненко Андрій	7	5	6	2	1	-1	4	1	1
Кордубан Дмитро	6	6	4	0	2	2	0	4	4
Челноков Володимир	9	7	5	2	4	2	4	16	4
Терлецький Юрій	4	8	8	-4	-4	0	16	16	0
Знов'як Юрій	8	9	9	-1	-1	0	1	1	0
<b>Суми квадратів різниць рангів</b>							26,5	61,5	27,5
$r_s$							0,775	0,47917	0,7(6)
<b>Гіпотеза</b>							$H_0$	$H_1$	$H_0$

**Висновок.** Після вилучення з розгляду результатів учня, який мав найвищий позитивний приріст результатів на відбіркових тренувальних зборах, спостерігаємо відповідність між результатами відбірково-тренувальних зборів і IV етап на достатньому рівні статистичної значимості.

### **Література**

- [1] Сидоренко Е.В. Математические методы обработки в психологии. – СПб., Речь, 2001. – 349 с.
- [2] Урбах В.Ю. Биометрические методы. Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине. – М.: Наука, 1964. – 415 с.