IBM 360/370 у вітчизняній серії машин ЄС ще раз підтверджує перспективність саме такого шляху українізації комп"ютерів.

Стаття надійшла до редколегії 16.07.90

УДК 519.21

І.Д.Квіт

ПОРІВНЯННЯ ДВОХ ВИБІРОК

Нехай

$$x_{(1)} < x_{(2)} < \ldots < x_{(\bar{i})} < \ldots < x_{(k)} \qquad /1/$$

та

$$y_{(1)} < y_{(2)} < \ldots < y_{(\bar{j})} < \ldots < y_{(l)} \qquad /2/$$

варіаційні ряди як завгодно зрізаних вибірок обсягів відповідно m і n незалежних спостережень над абсолютно неперервними випадковими змінними ξ та η, де $\bar{i}, \bar{j}$ – середні ранги. Кожний з рядів /1/ та /2/ може бути побудований на основі, наприклад, повної вибірки, підмножини повної вибірки, справа зрі[illegible]ної вибірки, з двох боків зрізаної вибірки і т.п. Слід перевірити гіпотезу про те, що вибірки однорідні, тобто, що ξ і η стохастично еквівалентні та, отже, мають однакову функцію розподілу ймовірностей

$$P\{\xi \leqslant x\} \equiv P\{\eta \leqslant x\}. \qquad /3/$$

Якщо гіпотеза /3/ істинна, то перетворення рядів /1/ і /2/ за допомогою відповідно підібраного того самого розподілу ймовірностей $F(x)$ утворює два однорідні ряди з інтервалу /0,1/. На основі методики, описаної в праці [1], кожному з рядів $F(x_{(\bar{i})})$, $(i=1,\ldots,k)$ та $F(y_{(\bar{j})}), (j=1,\ldots,l)$ ставиться у відповідність варіаційний ряд повної вибірки

$$r^*_{x(1)} < r^*_{x(2)} < \ldots < r^*_{x(i)} < \ldots < r^*_{x(k)} \qquad /4/$$

та

$$r^*_{y(1)} < r^*_{y(2)} < \ldots < r^*_{y(j)} < \ldots < r^*_{y(l)} \qquad /5/$$

обсягів відповідно k та l.

Таким чином, перевірка гіпотези /3/ про однорідність вибірок з варіаційними рядами /I/ і /2/ зводиться до перевірки гіпотези про однорідність повних варіаційних рядів /4/ і /5/. Останню гіпотезу можемо перевірити, наприклад, за допомогою критерію інверсій Вілкоксона /порівн. [2] /. Нагадаємо, що при $k \geq 4$, $l \geq 4$ та $k + l \geq 20$ область прийому гіпотези на рівні значущості $\alpha = 0{,}05$ міститься між $a - 1{,}96\sigma$ і $a + 1{,}96\sigma$, де $a = \frac{kl}{2}$, $\sigma^2 = \frac{kl}{2}(k+l+1)$.

Приклад. Дано повний варіаційний ряд вибірки обсягу $m = 10$ / x /: 264 422 548 663 775 891 1017 1161 1346 1643 та справа зрізаної вибірки обсягу $n = 15$
/ y /: 216 342 439 524 603 680 756 833 912 997.
Перевірити гіпотезу про однорідність вибірок.

Перетворення рядів / x / та / y / виконаємо, наприклад, за допомогою розподілу

$$F(x) = 1 - e^{-\frac{x}{1000}}, \quad x > 0.$$

Ряд / x / відразу переходить у варіаційний ряд /4/

$r^*_{x(i)}$: 0,2320264 0,3442659 0,4218950 0,4846968 0,5392962 0,5897546 0,6383216 0,6868271 0,7397206 0,8066011

Ряд / y / переводиться у варіаційний ряд справа зрізаної вибірки на інтервалі $(0,1)$

$r_{y(i)}$: 0,1942647 0,2896517 0,3553192 0,4078528 0,4528323 0,4933829 0,5304592 0,5652569 0,5982800 0,6310151

Ряд $r_{y(i)}$ за допомогою формули

$$r'_i = 1 - \left(\frac{1 - r_{(i)}}{1 - r_{(i-1)}}\right)^{16-i}, \quad (i = 1, \dots, 10), \quad r_{(0)} \equiv 0$$

переходить у повну випадкову вибірку обсягу 10 зі значеннями на інтервалі $(0,1)$

r'_{yi} : 0,9608376 0,8286444 0,7166336 0,6394097 0,5806343 0,5369846 0,4954031 0,4598973 0,4247828 0,3995084.

За допомогою формули

$$r^*_{(j)} = 1 - \prod_{i=1}^{j} (1 - r'_i)^{\frac{1}{11-i}}$$

з ряду r'_{yi} одержуємо варіаційний ряд /5/

$r^*_{y(j)}$: 0,2767523 0,4054826 0,4921822 0,5610407 0,6202278
0,6744317 0,7256037 0,7765382 0,8305196 0,8982285.

Спільний варіаційний ряд для $r^*_{x(i)}$ та $r^*_{y(j)}$ набуває вигляду

$$x\ y\ x\ y\ x\ x\ y\ x\ y\ x\ y\ x\ y\ x\ y\ x\ y\ x\ y\ y.$$

Кількість інверсій y відповідно x дорівнює I+2+2+3+4+5+6+ +7+8=38, а x відносно y : I+2+4+5+6+7+8+9+I0+I0=62, причому 38+62=I00= kl . Тут a = I0·I0:2=50, σ^2 = I0·I0/I0+I0+I/: : I2=I75, σ = I3,22875 . Область прийому гіпотези на рівні значущості α = 0,05 задається інтервалом /24,07; 75,93/. Оскільки емпіричні значення статистики Вілкоксона потрапляють в область прийому, то гіпотеза про однорідність вибірок заданих рядами / x / та / y / не суперечить експериментальним даним. Методом, описаним у праці [1], можна перевірити, що кожна вибірка / x / та / y / взята з популяції Релея при параметрі масштабу σ = I000 і параметрі форми ν = 2.

I. К в і т І.Д., М о с к в я к Є.В. Гіпотеза про розподіл // У цьому ж Віснику. С.32-40. 2. К в і т І.Д. Статистична змінна. Львів, I974. Ч.I.

Стаття надійшла до редколегії 15.03.90