Теор. ймовірност. та матем. статист.
Вип. 53, 1995, стор. 69–75

Teor. Ĭmovirnost. ta Matem. Statyst.
No. 53, 1995, pp. 69–75

ВЕРХНЯ МЕЖА ДЛЯ ЧИСЛА ПОВНИХ ВІДОБРАЖЕНЬ

УДК 519.21

К. КУПЕР ТА І. М. КОВАЛЕНКО

Резюме. Повним відображенням на множині G з бінарною операцією "$\circ$" називається така бієкція $\Theta: G \to G$, для якої відображення $\eta(x) = x \circ \Theta(x)$ також є бієкцією. В статті одержано асимптотичні оцінки для частки повних відображень серед перестановок порядку n, якщо "$\circ$" є операцією додавання за модулем n.

Наслідуючи поняття Denes, Keedwell [2], назвімо повним відображенням на множині G з бінарною операцією $\circ$ таку бієкцію $\Theta: G \to G$, для якої відображення $\eta(x) = x \circ \Theta(x)$ також є бієкцією.

Зупинимося на випадкові, коли $(G, \circ)$ є групою цілих чисел $\{0, 1, \ldots, n-1\}$ з додаванням за модулем n. Наприклад, у групі $\{\mathbb{Z}_5, +\}$ тотожне відображення є повним, оскільки

$$\Theta = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix}, \qquad \eta = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 2 & 4 & 1 & 3 \end{pmatrix}.$$

Множина бієкцій $\mathbb{Z}_n \to \mathbb{Z}_n$ — то є множина S_n перестановок на $\{0, 1, \ldots, n-1\}$.

Цілком природно поставити таке питання. Якщо $G_n = \{\sigma \in S_n : \sigma$ є повне відображення на $(\mathbb{Z}, +)\}$, то що можна сказати про $|G_n|/|S_n|$? Будь-яка група G непарного порядку має повне відображення (див. теорему 1, 4, 3 [2]); це наслідок тієї властивості, що кожен елемент G має єдиний квадратний корінь. У випадку групи $(\mathbb{Z}, +)$ непарного порядку можна тривіальним способом побудувати повні відображення, виходячи з тотожньої перестановки, другий рядок якої піддано циклічному зсуву.

Коли ж n парне, то жодного повного відображення не існує, в чому можна переконатися на основі теореми Paige [4], яку ми цитуємо за джерелом [2], де її наведено як теорему 1.4.5.

Теорема 1. *Якщо $(G, \circ)$ є скінченною групою порядку n, що має повне відображення, то існує таке впорядкування її елементів $a_1, a_2, \ldots, a_n$, що $a_1 a_2 \ldots a_n = e$, де e — одиниця G. У випадку $(\mathbb{Z}_n, +)$ маємо $\frac{1}{2}n(n-1) \equiv 0 \pmod{n}$ лише тоді, коли n ділить $\frac{1}{2}n(n-1)$, тобто n має бути непарним.*

Нашим першим результатом є така верхня оцінка.

1991 *AMS Mathematics Subject Classification.* Primary 60C50.

Роботу І. М. Коваленка виконано під час наукового візиту до групи STORM Університету Північного Лондону. Роботу К. Купера підтримала дослідницька група STORM цього університету.

Теорема 2. *Існує така константа* $c \geq 0.08854$, *що для достатньо великого* n

$$\frac{|G_n|}{|S_n|} \leq \exp\{-cn\}.$$

Ми пропонуємо читачеві два доведення цієї теореми. Перше доведення базується на мартингальній нерівності; воно дає трохи слабішу оцінку 0.06766 для константи c. Прямий комбінаторний метод дозволяє одержати оцінку 0.08854.

На даний момент автори не спроможні якось посилити оцінку. Проте, якщо розглянути не всі можливі перестановки, а лише ті, що задовольняють певним умовам, притаманним "випадковій" перестановці, то тут маємо асимптотичну межу $\exp\{-(1-\varepsilon)n\}$, де ε — довільно мале додатне число.

Нехай

$$\begin{pmatrix} 1 & \dots & n \\ \alpha(1) & \dots & \alpha(n) \end{pmatrix}$$

є довільна перестановка. Позначимо $\delta_{ij} = 1$, якщо $\alpha(i) = j$, $\delta_{ij} = 0$ у противному разі,

$$N_k(a) = \sum_{i=1}^{k} \sum_{j=a}^{a+m-1 \pmod n} \delta_{ij}.$$

Очевидно,

$$\frac{1}{n} \sum_{a} N_k(a) = \frac{km}{n}.$$

1. Розглянемо відхилення

$$\Delta_k(a) = N_k(a) - \frac{km}{n}$$

та максимальне відхилення по клітинках спеціального виду

$$\Delta_m^+(a) = \max\{\Delta_k(a), a \in \mathbb{Z}_n, k = m, 2m, \dots, [(n-m)/m]m\}.$$

Позначимо через $G_n(z)$ множину таких перестановок з класу G_n, для яких

$$\Delta_m^+ \leq z.$$

Теорема 3. *Нехай* $n \to \infty$, *а* m *та* z *змінюються так, що*

$$m = o(n), \qquad z = o(m).$$

Тоді для довільного $\varepsilon > 0$ *при достатньо великому* n *маємо*

$$\frac{|G_n(z)|}{|S_n|} \leq e^{(1-\varepsilon)n}.$$

Для скорочення словесного тексту ми будемо називати перестановки з класу G_n "повними перестановками".

2. *Доведення, що* $c \geq 0.06766$.

Наступне доведення базується на мартингальній нерівності, що наведено, наприклад, в оглядовій статті McDiarmid [3]. Для повноти наведемо необхідні поняття та формулювання цієї нерівності.

Нехай (V, d) — скінченний метричний простір. Послідовність розбиттів $((P_k, c_k): k = 0, \dots, n)$ визначається так: $P_0, P_1, \dots, P_n$ — це зростаюча послідовність розбиттів V, що починається з тривіального розбиття P_0 з єдиною підмножиною V і закінчується дискретним розбиттям P_n множини V на одноелементні підмножини. Вимагається, щоб із співвідношення $A \in P_k$ для довільного $k \geq 1$ випливало

співвідношення $A \subset B \in P_{k-1}$. Послідовність $c_0, c_1, \ldots, c_k, \ldots, c_n$ додатних чисел має таку властивість. Для довільного $k \geq 1$ з того, що $A, B \in P_k$ та $A, B \subset C \in P_{k-1}$, випливає, що існує бієкція $\varphi\colon A \to B$, для якої

$$d(x, \varphi(x)) \leq c_k$$

для будь-якого $x \in A$.

Для розглядуваного випадку $V = S_n$. Якщо $\alpha = (\alpha(i))$ та $\beta = (\beta(i))$ — дві перестановки, позначимо

$$d(\alpha, \beta) = \sum_{j:\alpha(j)\neq\beta(j)} 1.$$

Ця функція визначатиме метрику. Послідовність розбиттів буде визначено наступним чином. Дві перестановки α та β віднесемо до тієї самої підмножини A розбиття P_k множини S_n, якщо в них перші k компонент співпадають, тобто $\alpha(1) = \beta(1), \ldots, \alpha(k) = \beta(k)$. Очевидно, така послідовність розбиттів є зростаючою, оскільки при $A, B \in P_k$, $A \neq B$; $A, B \in P_{k-1}$ компоненти α і β вперше не співпадають на k-му числі. Очевидно, будь-яке $\alpha \in A$ має певне $\alpha(k)$; те ж саме й для $\beta \in B$. Для будь-якого $\alpha \in A$ маємо $\alpha(i) = \beta(k)$ при деякому $i > k$. Означимо $\varphi(\alpha)$ як таке $\beta \in B$, для якого елементи $\alpha(k)$, $\alpha(i)$ перестановлено, а решта елементів незмінні. У цьому випадку $d(\alpha, \varphi(\alpha)) \leq 2$, звідси випливає, що $c_k = 2$.

Теорема 4. *Припустимо, що скінченний метричний простір (V, d) має послідовність розбиттів $((P_k, c_k), k = 0, \ldots, n)$. Нехай функція f на V задовольняє умові*

$$|f(x) - f(y)| \leq d(x, y)$$

для довільних $x, y \in V$. Нехай також X є випадкова величина, рівномірно розподілена на V. Тоді для довільного $t > 0$

$$\mathsf{P}\big(f(X) - \mathsf{E}(f(X)) \geq t\big) \leq \exp\Big\{-2t^2 / \sum c_k^2\Big\}.$$

У розглянутому випадку

$$I_\alpha = \big\{j \in \{0, 1, \ldots, n-1\}\colon i + \alpha(i) = j\big\}.$$

Означимо $f(\alpha) = |I_\alpha|$. Зазначимо, що S_n має рівномірну міру. Нам потрібно довести, що наведене означення f задовольняє нерівності

$$|f(\alpha) - f(\beta)| \leq d(\alpha, \beta) \tag{1}$$

для довільних α та β. Позначимо $I_\alpha \cap I_\beta$ через $I_{\alpha\beta}$, $I_\alpha \cap \overline{I}_\beta$ через $I_{\alpha\overline{\beta}}$. Нехай $j \in I_{\alpha\overline{\beta}}$. Тоді для деякого i маємо $i + \alpha(i) = j$ і в той же час $i + \beta(i) \neq j$. Значить, $\alpha(i) \neq \beta(i)$, індекси i є різні для різних $j \in I_{\alpha\overline{\beta}}$, отже

$$|I_{\alpha\overline{\beta}}| \leq d(\alpha, \beta).$$

З іншого боку

$$f(\alpha) - f(\beta) = |I_{\alpha\overline{\beta}}| - |I_{\overline{\alpha}\beta}| \leq |I_{\alpha\overline{\beta}}|.$$

Таким чином

$$f(\alpha) - f(\beta) \leq d(\alpha, \beta).$$

Завдяки симетрії отримуємо (1).

Твердження, що $\mathsf{E}(|I|)$ асимптотично еквівалентне $n(1 - e^{-1})$, є наслідком того, що число нулів серед $i + \alpha(i) \pmod n$ асимптотично пуассонівське з параметром 1.

У цьому можна переконатися, розглянувши факторіальні моменти або ж безпосередньо на основі теореми 4А з [1], що оцінює похибку даної апроксимації числом $2/n$. Таким чином,

$$\mathsf{P}(\alpha \in G_n) = \mathsf{P}(|J_a| = n) \le \mathsf{P}\big([|J_a| - \mathsf{E}(|J|)] \ge ne^{-1}(1 + o(1))\big)$$
$$\le \exp\left\{-\frac{e^{-2}}{2}(1 + o(1))n\right\},$$

що й треба було довести.

3. *Доведення того, що* $c \ge 0.08854$.

Спочатку опишемо метод генерування перестановок за допомогою вибірки без повернення. Вибираємо (x_1, y_1) випадково таким чином, щоб

$$\mathsf{P}(x_1 = i, y_1 = j) = \frac{1}{n^2}, \qquad 0 \le i,\ j \le n-1,$$

і означаємо $X_1 = (x_1)$, $Y_1 = (y_1)$. При довільному k вибираємо x_k з множини $\mathbb{Z}_n \setminus \{x_1, \dots, x_{k-1}\}$, y_k з множини $\mathbb{Z}_n \setminus \{y_1, \dots, y_{k-1}\}$. Позначивши $X_k = (x_1, \dots, x_k)$, $Y_k = (y_1, \dots, y_k)$, маємо

$$\mathsf{P}\big((x_k, y_k) = (i, j) \colon X_{k-1}, Y_{k-1}\big) = \frac{1}{(n-k+1)^2}$$

для $i \notin \{x_1, \dots, x_{k-1}\}$, $j \notin \{y_1, \dots, y_{k-1}\}$.

Нехай $z_k = x_k + y_k \pmod n$. Ми скажемо, що в момент k випадає невдача, якщо $z_k = \in \{z_1, \dots, z_{k-1}\}$. Якщо за час від 1 до n не буде жодної невдачі, тоді перестановка

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{pmatrix}$$

є повною. Очевидно,

$$|G| = n!\, \mathsf{E}[p_2(X_1, Y_1) \cdots p_n(X_{n-1}, Y_{n-1})],$$

де $p_k(X_{k-1}, Y_{k-1})$ є умовна ймовірність успіху (тобто відсутності невдачі) в момент k при данних X_{k-1}, Y_{k-1}). Множина $I = \{(i, j) \colon 0 \le i, j \le n-1\}$ складається з дозволених та недозволених позицій для нової точки (x_k, y_k). Існує $(n-k+1)^2$ дозволених позицій (i, j). Спробуймо оцінити, скільки з них задовольняє умові $i + j \pmod n \notin \{z_1, \dots, z_{k-1}\}$. Для довільної точки $(i, j) \in I$ може статися будь-яка з трьох подій:

$$A = \big\{i \in \{x_1, \dots, x_{k-1}\}\big\},$$
$$B = \big\{j \in \{y_1, \dots, y_{k-1}\}\big\},$$
$$C = \big\{i + j \pmod n \in \{z_1, \dots, z_{k-1}\}\big\}.$$

Маємо:

$$p_k(X_{k-1}, Y_{k-1}) = \frac{n^2 - N(A \cup B \cup C)}{(n-k+1)^2}.$$

Згідно принципу включення та виключення маємо

$$N(A \cup B \cup C) = N(A) + N(B) + N(C) - N(AB) - N(BC) - N(AC) + N(ABC).$$

За умови, що до моменту $k-1$ включно не було жодної невдачі,

$$N(A) = N(B) = N(C) = n(k-1),$$
$$N(AB) = N(BC) = N(AC) = n(k-1)^2.$$

Що ж до $N(ABC)$, маємо очевидну подвійну нерівність $k-1 \le N(ABC) \le (k-1)^2$.

Таким чином,

$$1 - \frac{k-1}{n-k+1} \le p_k(X_{k-1}, Y_{k-1}) \le \frac{(k-1)(n-2k+3)}{(n-k+1)^2}.$$

З верхньої оцінки маємо

$$\begin{aligned}\mathsf{E}[p_2(X_1,Y_1)\cdots p_n(X_{n-1},Y_{n-1})] &\le \prod_{1\le k\le (n+3)/2}\left(1-\frac{(k-1)(n-2k+3)}{(n-k+1)^2}\right)\\ &= \exp\left\{\sum_{1\le k\le (n+3)/2}\ln\left(1-\frac{(k-1)(n-2k+3)}{(n-k+1)^2}\right)\right\}\\ &\le \exp\left\{(1+o(1))n\int_0^{1/2}\ln\left(1-\frac{x(1-2x)}{(1-x)}\right)dx\right\}.\end{aligned}$$

Чисельним інтегруванням для інтегралу можна одержати наближене значення (-0.0885474).

4. *Оцінка числа повних перестановок при обмеженні на частоти.*

Вибираємо перестановки $\alpha = (\alpha(i))$ лише з класу $G_n(z)$. Нехай $\alpha(1), \dots, \alpha(mr)$ вже вибрано. Підрахуймо число можливостей для вибору $(\alpha(mr+1), \dots, \alpha(mr+m))$. Якщо до моменту $mr-r$ включно не було жодної невдачі, то це число дорівнює $mn - 2m^2r + L$, де L — число точок (i,j) цілочислового прямокутника $\{rm+1, \dots, rm+m\} \times \{1, \dots, n\}$, для кожної з яких

$$j \in \{\alpha(1), \dots, \alpha(mr)\}$$

та

$$i + j \in \{\nu + \alpha(\nu) \pmod n, 1 \le \nu \le rm\}$$

Таким чином, L є число точок перетину rm "вертикальних" відрізків та rm "похилих" відрізків. Уздовж "похилого" відрізку з нижньою точкою $(rm+m, a)$ трапляється $N_{rm}(a)$ перетинів з "вертикальними" відрізками. Таким чином,

$$L = N_{rm}(a_1) + \cdots + N_{rm}(a_{rm}) \le \frac{r^2m^3}{n} + rm\Delta_m^+.$$

Оскільки за умов $x_1 + \cdots + x_m = c$, $x_1 \ge 0, \dots, x_m \ge 0$ максимум функції $x_1 \cdots x_m$ досягається при $x_1 = \cdots = x_m = 1/m$, то число можливостей для вибору $(\alpha(mr+1), \dots, \alpha(mr+m))$ не перевищує

$$\begin{aligned}&\left(n - 2mr + \frac{r^2m^2}{n} + rm\Delta_m^+\right)^m\\ &\quad = n^m\left(\left(1-\frac{rm}{n}\right)^2 + \frac{r}{n}\Delta_m^+\right)^m \le n^m\left(\left(1-\frac{rm}{n}\right)^2 + \frac{r}{z}\right)^m.\end{aligned}$$

Таким чином, якщо $sm \le n$, то

$$|G_n(z)| \le n^{sm}\prod_{r=0}^{s-1}\left(\left(1-\frac{rm}{n}\right)^2 + \frac{rz}{n}\right)^m (n-sm)!$$

і отже

$$|G_n(z)| \le n^n\prod_{r=0}^{s-1}\left(\left(1-\frac{rm}{n}\right)^2 + \frac{rz}{n}\right)^m. \qquad (2)$$

Фіксуємо деяке δ, $0 < \delta < 1$ і виберемо s із умови $sm \le (1-\delta)n < (s+1)m$. Для логарифму добутку в правій частині (2) маємо оцінку

$$m\sum_{r=0}^{s-1}\ln\left(\left(1-\frac{rm}{n}\right)^2+\frac{rz}{n}\right)=n(I_0+I_1),$$

де

$$I_0=\frac{2m}{n}\sum_{r=0}^{s-1}\ln\left(1-\frac{rm}{n}\right),$$

$$I_1=\frac{m}{n}\sum_{r=0}^{s-1}\ln\left(1-\frac{rz}{n}\left(1-\frac{rm}{n}\right)^{-2}\right).$$

Зважаючи на те, що $m=o(n)$, маємо:

$$I_0\sim 2\int_0^{1-\delta}\ln(1-t)\,dt.$$

Оскільки

$$\int_0^1\ln(1-t)\,dt=-1,$$

то можна вибрати δ таким чином, що буде

$$I_0<-2+\frac{\varepsilon}{2} \tag{3}$$

при достатньо великому n. Суму I_1 оцінюємо так:

$$I_1\le\frac{m}{n}\sum_{r=0}^{s-1}\frac{rz}{n\delta^2}\sim\frac{zm}{2n^2\delta^2}s^2\sim\frac{zm}{2n^2\delta^2}\left(\frac{n(1-\delta)}{m}\right)^2<\frac{z}{m\delta^2}$$

при достатньо великому n. Оскільки, за умовою, $z=o(m)$, маємо оцінку

$$I_1<\frac{\varepsilon}{2}. \tag{4}$$

Підставивши (3) та (4) у (2) та замінивши $(n-sm)!$ на більше число n^{n-sm}, отримаємо оцінку

$$|G_n(z)|<n^n e^{-(2-\varepsilon)n},$$

що справджується при достатньо великому n. Таким чином,

$$\frac{|G_n(z)|}{|S_n|}<e^{-(1-\varepsilon)n},$$

при достатньо великому n, що й стверджує теорема 3.

Природно виникає таке питання. Нехай з множини S_n випадково вибирають підстановку α. З якою ймовірністю відбудеться подія $\{\Delta_m^+\le z\}$? Для ймовірності протилежної події маємо оцінку

$$\mathsf{P}\{\Delta_m^+>z\}\le n^s\max_{0\le k\le n}\mathsf{P}\left(N_k(0)>\frac{km}{n}+z\right). \tag{5}$$

Позначимо $p(i)=\mathsf{P}(N_k(0)=i)$. З комбінаторних міркувань маємо оцінку

$$p(i)=\binom{m}{i}k^{[i]}(n-k)^{[m-i]}/n^{[m]},$$

де $a^{[j]}$ означає $a(a-1)\cdots(a-j+1)$. Позначивши $c = n/(n-m)$, $k = n$, $\tau = 1-t$, дістанемо нерівність

$$p(i) = \binom{m}{i} t^i (\tau c)^{m-i},$$

Звідки при довільному дійсному λ

$$\sum_{i=0}^{m} e^{\lambda i} p(i) \le (te^{\lambda} + \tau c)^m.$$

Звідси при довільному $\lambda > 0$

$$\mathsf{P}(N_k(0) \ge u) \le e^{-\lambda u}(te^{\lambda} + \tau c)^m.$$

Підставивши $u = tm + z$, $e^{-\lambda} = (mt\tau - zt)/(mt\tau + zt)$, після деяких перетворень дістанемо оцінку

$$\ln \mathsf{P}(N_k(0) \ge tm + z) \lesssim m\left(\frac{\tau m}{n-m} - \frac{z^2}{2m^2 t\tau}\right),$$

що справджується при $z = o(m)$.

З урахуванням множника n^s в правій частині (5) маємо:

$$\ln \mathsf{P}(\Delta_m^+ \ge z) \lesssim \frac{n}{m}\ln n + \frac{m^2}{n-m} - \frac{2z^2}{m}. \tag{6}$$

Узявши, наприклад, $m \sim n\sigma$, $z \sim m\sigma$, де $\sigma = \sigma_n$ — повільно спадна функція n, дістанемо асимптотичну оцінку

$$\ln \mathsf{P}(\Delta_m^+ \ge z) \le \exp\{-\sigma^2 n(1 + o(1))\}. \tag{7}$$

Таким чином, при належному виборі змінної z випадкова перестановка потрапить до класу $G_n(z)$ з імовірністю, що прямує до 1 при $n \to \infty$.

Література

1. A. D. Barbour, L. Holst, and S. Janson, *Poisson approximation*, Clarendon Press, Oxford, 1992.
2. J. Dénes and A. D. Keedwell, *Latin squares and their applications*, English University Press, London, 1974.
3. C. McDiarmid, *On the method of bounded differences*, Surveys in combinatorics, LMS Lecture Note Series (J. Siemons, ред.), т. 141, Cambridge University Prees, Cambridge, 1989.
4. L. J. Paige, *Complete mappings of finite groups*, Pacific J. Math. 1 (1951), 111–116.

School of Mathematical Sciences, University of North London, 166-220 Holloway Road, London N7 8DB, United Kingdom

252207, Київ, пр. Академіка Глушкова, 40, Інститут кібернетики НАН України

Надійшла 16.09.94