УДК 519.21
DOI https://doi.org/10.24144/2616-7700.2021.39(2).38-46

А. Б. Ільєнко[1], В. В. Стаматієва[2]

[1] Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”,
доцент кафедри математичного аналізу та теорії ймовірностей,
кандидат фізико-математичних наук, доцент
ilienko@matan.kpi.ua
ORCID: https://orcid.org/0000-0003-4828-0788

[2] Національний технічний університет України “Київський політехнічний інститут імені Ігоря Сікорського”,
аспірант
stamatieva56@gmail.com
ORCID: https://orcid.org/0000-0003-2721-8985

ГРАНИЧНА ТЕОРЕМА ДЛЯ ТОЧКОВИХ ПРОЦЕСІВ, ПОВ’ЯЗАНИХ З УЗАГАЛЬНЕНОЮ ЗАДАЧЕЮ ПРО ДНІ НАРОДЖЕННЯ

У роботі доведено граничну теорему для послідовності точкових процесів, які описують моменти $(r+1)$-х надходжень різних типів з загальної кількості в n типів в узагальненій задачі про дні народження. Класична задача про дні народження, відома з популярної літератури, відповідає параметрам $r = 1$ (достатньо одного збігу) та $n = 365$ (кількість днів у невисокосному році). Доведення базується на застосуванні техніки пуассонізації/депуассонізації. Цей результат далі використовується для простого доведення деяких класичних граничних теорем у задачі про дні народження, які фактично описують асимптотичну поведінку різних змістовних функціоналів від побудованих процесів.

Ключові слова: задача про дні народження, точковий процес, гранична теорема, груба збіжність, процес Пуассона, пуассонізація, депуассонізація, теорема про неперервне відображення.

1. Вступ. Задача про дні народження, яка є однією з класичних та найпопулярніших задач комбінаторної теорії ймовірностей, була вперше розглянута Р. фон Мізесом у 1939 р. Наведемо одне з її численних формулювань у вигляді, зручному для подальших узагальнень.

Розглянемо достатньо велику множину людей, які в послідовні цілі моменти часу один за одним заходять до кімнати. Позначимо через $T_1^{(365)}$ випадковий момент часу, коли до кімнати вперше зайшла людина, день народження якої збігається з днем народження когось з уже присутніх. Тут верхній індекс показує загальну кількість різних можливих днів народження (всі роки вважаються невисокосними), а нижній означає, що достатньо збігу з днем народження лише *одного* з присутніх.

У природний спосіб узагальнюючи це формулювання, розглянемо таку задачу. Нехай у послідовні цілі моменти часу до спостерігача один за одним надходять об’єкти, кожен з яких з імовірністю $\frac{1}{n}$ незалежно від інших належить до одного з n типів. Будемо позначати через $T_r^{(n)}$ перший момент часу, коли деякий з типів з’явився в $(r+1)$-й раз. Зокрема, $T_0^{(n)} = 1$, а $T_1^{(365)}$ відповідає інтерпретації, наведеній у попередньому абзаці.

У цій роботі ми пропонуємо дещо нетрадиційний підхід, що дозволяє легко отримувати граничні теореми для випадкових величин $T_r^{(n)}$ та їх узагальнень. Він базується на основному результаті роботи — граничній теоремі для послідовності точкових процесів $\big(\xi_r^{(n)}, n \ge 1\big)$, що описують моменти $(r+1)$-х надходжень усіх n типів. Під збіжністю точкових процесів ми традиційно розуміємо їх збіжність за розподілом у грубій топології, заданій на просторі локально скінченних точкових мір. Відомо (див., наприклад, теорему 4.20 у монографії [1]), що ця збіжність еквівалентна збіжності відповідних лічильних процесів у J_1-топології Скорохода.

У нещодавніх роботах [2] та [3] розглядалося питання про збіжність точкових процесів у іншій класичній задачі — задачі збирача купонів. В цій статті ми будемо застосовувати ідеї пуассонізації та депуассонізації схожі до тих, що були використані в [3].

2. Огляд деяких попередніх результатів. Явним формулам для числових характеристик випадкових величин $T_r^{(n)}$ та їх подальших узагальнень, а також граничним теоремам для таких величин присвячено безліч робіт. Наприклад, з майже очевидного зображення

$$\begin{aligned}\mathbb{E}T_1^{(n)} &= \sum_{k=1}^{n+1} \mathbb{P}\big\{T_1^{(n)} \ge k\big\} = \\ &= 1 + \sum_{k=1}^{n} \frac{n(n-1)\cdots(n-k+1)}{n^k} = \int_0^\infty \mathrm{e}^{-t}\Big(1+\frac{t}{n}\Big)^n \,\mathrm{d}t\end{aligned} \tag{1}$$

та результатів роботи [4] про асимптотичну поведінку Q-функції Рамануджана (яка якраз і задається другою сумою в (1)) випливають різноманітні досить точні оцінки величини $\mathbb{E}T_1^{(n)}$. Ми наведемо тут для повноти викладення лише початок класичного асимптотичного розкладу Рамануджана-Ватсона-Кнута, який в наших позначеннях може бути записаний у вигляді

$$\mathbb{E}T_1^{(n)} = \sqrt{\frac{\pi n}{2}} + \frac{2}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} - \frac{4}{135n} + \dots$$

Відповідні посилання можна знайти в бібліографії до процитованої роботи [4]. Зауважимо також, що при $n = 365$, тобто для класичної задачі про дні народження, $\mathbb{E}T_1^{(365)} \approx 24.62$. Це трохи відрізняється від медіанного значення 23, відомого з популярної літератури.

Аналог формули (1) для довільного $r \ge 1$ був отриманий в роботі [5]:

$$\mathbb{E}T_r^{(n)} = \int_0^\infty \mathrm{e}^{-t}\Big[S_r\Big(\frac{t}{n}\Big)\Big]^n \,\mathrm{d}t,$$

де позначено $S_r(t) = \sum_{j=0}^r \frac{x^j}{j!}$. На його основі в цій роботі також було одержано асимптотичну формулу

$$\mathbb{E}T_r^{(n)} \sim \sqrt[r+1]{(r+1)!}\,\Gamma\Big(\frac{r+2}{r+1}\Big) n^{\frac{r}{r+1}}, \qquad n \to \infty.$$

З результатів статті [6] (див. також [7]), зокрема, випливає гранична теорема для $T_r^{(n)}$:

$$\lim_{n\to\infty} \mathbb{P}\{n^{-\frac{r}{r+1}} T_r^{(n)} \le x\} = 1 - \mathrm{e}^{-\frac{x^{r+1}}{(r+1)!}}, \qquad x \ge 0. \tag{2}$$

Отже, граничним є розподіл Вейбула зі щільністю $\frac{x^r}{r!}\mathrm{e}^{-\frac{x^{r+1}}{(r+1)!}} \cdot \mathbb{1}\{x \geq 0\}$. Багато інших граничних теорем для споріднених випадкових величин та процесів можна знайти у класичній монографії [8], а також у численних журнальних публікаціях. Відзначимо, зокрема, роботи [7] та [9]. В першій з них запропоновано потужний метод вкладення в процес Пуассона, який ми будемо використовувати в цій роботі, а у другій за допомогою цього ж методу, зокрема, розглянуто схожу задачу з випадковими ймовірностями належності об'єктів до різних типів. Зауважимо також, що в порівняно нещодавній статті [10] розглядалися граничні теореми для моментів досягнення сумарної кількості в r збігів серед усіх об'єктів, що надійшли до цього часу.

Як було відзначено у вступі, основним результатом цієї роботи є гранична теорема для послідовності точкових процесів $\left(\xi_r^{(n)}, n \geq 1\right)$, що описують моменти $(r+1)$-х надходжень різних типів. З цього факту за допомогою теореми про неперервне відображення можна легко отримувати граничні результати для різноманітних функціоналів від процесів $\xi_r^{(n)}$, наприклад, такі, як формула (2) та її узагальнення.

3. Основний результат. Для фіксованих $r, n \geq 1$ розглянемо точковий процес $\xi_r^{(n)}$, атоми якого визначаються (у спеціальний спосіб нормованими) моментами $(r+1)$-х надходжень усіх n типів.

Нехай дискретна випадкова величина $Y_{i,r}^{(n)}$, $1 \leq i \leq n$, задає момент надходження $(r+1)$-го об'єкта типу i. Її розподіл, очевидно, визначається формулою

$$\mathbb{P}\{Y_{i,r}^{(n)} = k\} = C_{k-1}^{r}\Big(\frac{1}{n}\Big)^{r+1}\Big(\frac{n-1}{n}\Big)^{k-r-1}, \qquad k \geq r+1, \tag{3}$$

і є одним з варіантів від'ємного біноміального розподілу. В цих термінах введена у вступі величина $T_r^{(n)}$ допускає зображення

$$T_r^{(n)} = \min\{Y_{1,r}^{(n)}, \ldots, Y_{n,r}^{(n)}\}.$$

Оскільки зі зростанням n значення $T_r^{(n)}$ також збільшуватимуться, то точковий процес, побудований за $Y_{i,r}^{(n)}$, не матиме жодної границі при $n \to \infty$. Тому для забезпечення збіжності до деякого граничного процесу випадкові величини $Y_{i,r}^{(n)}$ треба в деякий спосіб нормувати. Отже, розглянемо точковий процес $\xi_r^{(n)}$ вигляду

$$\xi_r^{(n)} = \sum_{i=1}^{n} \delta_{n^{-\frac{r}{r+1}} Y_{i,r}^{(n)}}, \tag{4}$$

де δ_a означає одиничну точкову міру, зосереджену в a. Основним результатом цієї роботи є теорема про збіжність процесів $\xi_r^{(n)}$ при $n \to \infty$ до деякого неоднорідного точкового процесу Пуассона.

Перед формулюванням цього результату коротко нагадаємо означення збіжності точкових процесів за розподілом у грубій топології (детальне викладення можна знайти, наприклад, у главі 4 в [1] або у розділах 3.4, 3.5 в [11]). Позначимо через $M_p(\mathbb{R})$ простір усіх локально скінченних точкових мір на $\mathbb{R}$. Для $\mu, \mu_1, \mu_2, \ldots \in M_p(\mathbb{R})$ кажуть, що μ_n збігаються до μ у грубій (vague) топології, якщо $\int_{\mathbb{R}} f\,\mathrm{d}\mu_n \to \int_{\mathbb{R}} f\,\mathrm{d}\mu$ для будь-якої фінітної неперервної функції $f\colon \mathbb{R} \to [0, +\infty)$. Факт такої збіжності позначають $\mu_n \xrightarrow{v} \mu$. Простір $M_p(\mathbb{R})$ з топологією цієї збіжності допускає повну сепарабельну метризацію, що у стандартний спосіб дозволяє розглядати збіжність точкових процесів (тобто випадкових точкових мір) $\xi, \xi_1, \xi_2 \ldots$ за розподілом у грубій топології. Її ми будемо позначати $\xi_n \xrightarrow{vd} \xi$.

Тепер ми можемо сформулювати основний результат роботи. Ми виключимо з розгляду випадок $r = 1$, оскільки в цій ситуації наш метод доведення не спрацьовує. Питання про справедливість теореми в цьому випадку залишається наразі відкритим.

Теорема 1. *Нехай $r \geq 2$ та ξ_r є неоднорідним точковим процесом Пуассона на $\mathbb{R}$ з мірою інтенсивності*

$$\lambda_r(\mathrm{d}x) = \frac{x^r}{r!} \mathbb{1}\{x \geq 0\}\, \mathrm{d}x. \tag{5}$$

Тоді $\xi_r^{(n)} \xrightarrow{vd} \xi_r$ при $n \to \infty$.

Зауважимо, що граничний процес ξ_r може бути явно побудований у досить простий спосіб. Розглянемо послідовність $(E_j, j \geq 1)$ незалежних однаково розподілених випадкових величин $E_j \sim \mathsf{Exp}(1)$. Покажемо, що атоми міри ξ_r знаходяться в точках $\sqrt[r+1]{(r+1)! \sum_{j=1}^k E_j}$, $k \geq 1$.

Дійсно, внаслідок теореми про інтервали для однорідного процесу Пуассона (див., наприклад, теорему 7.2 в [12]) атоми випадкової пуассонівської міри одиничної інтенсивності розташовані в точках $S_k = \sum_{j=1}^k E_j$. Те, що дія на неї функцією $h(x) = \sqrt[r+1]{(r+1)!\, x}$, $x \geq 0$, приводить до міри ξ_r з теореми 1, пояснюється теоремою про перетворення для процесів Пуассона (теорема 5.1 в [12]), адже для $[a, b] \subset [0, \infty)$ виконується рівність

$$h(\mathsf{Leb})[a, b] = \big(\mathsf{Leb} \circ h^{-1}\big)[a, b] = \mathsf{Leb}\Big[\frac{a^{r+1}}{(r+1)!}, \frac{b^{r+1}}{(r+1)!}\Big] = \int_a^b \frac{x^r}{r!}\, \mathrm{d}x = \lambda_r[a, b].$$

Тут Leb позначає міру Лебега і є мірою інтенсивності однорідного процесу Пуассона з параметром 1.

4. Доведення. У роботі [7] Л. Хольст запропонував ідею пуассонізації, яка дозволяє істотно спростити доведення граничних теорем у задачах розміщення частинок по комірках і, зокрема, в задачі про дні народження. Розглянемо її застосування в нашому випадку.

Випадкові величини $Y_{i,r}^{(n)}$, $1 \leq i \leq n$, введені на початку попереднього розділу, є однаково розподіленими, але залежними. Факт цієї залежності пояснюється хоча б тим, що $Y_{i,r}^{(n)} \neq Y_{j,r}^{(n)}$ при $i \neq j$ — в один момент часу надходить лише один об'єкт. Для того, щоб позбавитися цієї залежності, розглянемо пуассонізовану постановку і вважатимемо, що об'єкти надходять не в послідовні цілі моменти часу, а в моменти $X_k^{(n)}$, $k \geq 1$, однорідного точкового процесу Пуассона з параметром 1. Отже, інтервали між такими надходженнями утворюватимуть послідовність незалежних $\mathsf{Exp}(1)$-розподілених випадкових величин $(E_j, j \geq 1)$. Для розрізнення типів різних об'єктів будемо вважати цей процес маркованим — кожному атому $X_k^{(n)}$ незалежно від інших і від самої послідовності $(X_k^{(n)}, k \geq 1)$ присвоїмо випадкову мітку $M_k^{(n)}$, рівномірно розподілену на множині типів $\{1, \ldots, n\}$. Тоді внаслідок теореми про прорідження для процесів Пуассона (див., наприклад, теорему 5.8 у [12]) послідовності

$X^{(n)}(i) = \big(X_k^{(n)}(i), k \ge 1\big)$, $1 \le i \le n$, моментів надходження об'єктів різних типів утворюватимуть *незалежні* однорідні точкові процеси Пуассона з параметром $\frac{1}{n}$.

Випадкова величина $X_{r+1}^{(n)}(i)$ описує момент $(r+1)$-го надходження i-го типу в пуассонізованій постановці, а отже, є пуассонізованим аналогом величини $Y_{i,r}^{(n)}$. При різних $i \in \{1, \dots, n\}$ ці величини є незалежними й однаково розподіленими — внаслідок вже цитованої теореми про інтервали вони розподілені за законом $\Gamma\big(r+1, \frac{1}{n}\big)$. Більш того, за побудовою величини $X_{r+1}^{(n)}(i)$ та $Y_{i,r}^{(n)}$ допускають простий каплінг:

$$X_{r+1}^{(n)}(i) = \sum_{j=1}^{Y_{i,r}^{(n)}} E_j, \tag{6}$$

причому $(Y_{i,r}^{(n)}, 1 \le i \le n)$ та $(E_j, j \ge 1)$ є незалежними. Останнє твердження випливає з того, що перша послідовність визначається значеннями міток $M_k^{(n)}$, а друга задає інтервали між моментами надходжень $X_k^{(n)}$.

Аналогічно (4) введемо пуассонізований точковий процес

$$\eta_r^{(n)} = \sum_{i=1}^{n} \delta_{n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i)}.$$

На першому кроці доведення ми покажемо, що при $n \to \infty$ виконується збіжність $\eta_r^{(n)} \xrightarrow{vd} \xi_r$. Для цього спочатку запишемо щільність розподілу $\Gamma\big(r+1, \frac{1}{n}\big)$-розподіленої величини $X_{r+1}^{(n)}(i)$:

$$f_r^{(n)}(x) = \frac{1}{r!\, n^{r+1}} x^r \mathrm{e}^{-\frac{x}{n}}, \qquad x \ge 0.$$

Тоді щільність розподілу кожного атома $n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i)$ випадкової міри $\eta_r^{(n)}$ має вигляд

$$\tilde{f}_r^{(n)}(x) = n^{\frac{r}{r+1}} f_r^{(n)}\big(n^{\frac{r}{r+1}} x\big) = \frac{1}{r!\, n} x^r \mathrm{e}^{-x n^{-\frac{1}{r+1}}}, \qquad x \ge 0.$$

Отже,

$$\lim_{n\to\infty} n \tilde{f}_r^{(n)}(x) = \frac{x^r}{r!} \mathbb{1}\{x \ge 0\}.$$

Тому з теореми Лебега про мажоровану збіжність випливає, що для будь-якої обмеженої борелівської множини $B \subset \mathbb{R}$ виконується

$$\lim_{n\to\infty} n \int_B \tilde{f}_r^{(n)}(x)\, \mathrm{d}x = \int_B \frac{x^r}{r!} \mathbb{1}\{x \ge 0\}\, \mathrm{d}x.$$

В термінах грубої збіжності це внаслідок твердження 3.12 з [11] еквівалентно тому, що

$$n \mathbb{P}\big\{n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i) \in \cdot\big\} \xrightarrow{v} \lambda_r(\cdot)$$

при $n \to \infty$. Тому факт збіжності $\eta_r^{(n)} \xrightarrow{vd} \xi_r$ випливає з незалежності $X_{r+1}^{(n)}(i)$ при різних $1 \le i \le n$ та відомого твердження про збіжність біноміальних точкових процесів до пуассонівського (див., наприклад, першу частину доведення твердження 3.21 у [11]).

Метою наступного кроку доведення буде депуассонізація, тобто перетворення $\eta_r^{(n)} \xrightarrow{vd} \xi_r$ у $\xi_r^{(n)} \xrightarrow{vd} \xi_r$. Для цього ми оцінимо ступінь близькості дограничних процесів $\eta_r^{(n)}$ та $\xi_r^{(n)}$.

Для довільних $\varepsilon > 0$ та $i \in \{1, \ldots, n\}$ з рівності (6) випливає, що

$$\mathbb{P}\Big\{\big|n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i) - n^{-\frac{r}{r+1}} Y_{i,r}^{(n)}\big| > \varepsilon\Big\} = \\ = \mathbb{P}\left\{\left|\sum_{j=1}^{Y_{i,r}^{(n)}} (E_j - 1)\right| > n^{\frac{r}{r+1}}\varepsilon\right\} = \mathbb{E}\,\mathbb{P}_{Y_{i,r}^{(n)}}\left\{\left|\sum_{j=1}^{Y_{i,r}^{(n)}} (E_j - 1)\right| > n^{\frac{r}{r+1}}\varepsilon\right\}, \tag{7}$$

де через $\mathbb{P}_{Y_{i,r}^{(n)}}$ позначено умовну ймовірність відносно випадкової величини $Y_{i,r}^{(n)}$. Зважаючи на незалежність $Y_{i,r}^{(n)}$ та $(E_j, j \geq 1)$, цю ймовірність можна легко оцінити за допомогою нерівностей Чебишова та Розенталя (останню див., наприклад, на с. 152 в [13]):

$$\mathbb{P}_{Y_{i,r}^{(n)}}\left\{\left|\sum_{j=1}^{Y_{i,r}^{(n)}} (E_j - 1)\right| > n^{\frac{r}{r+1}}\varepsilon\right\} \leq \\ \leq n^{-\frac{8r}{r+1}}\varepsilon^{-8}\,\mathbb{E}_{Y_{i,r}^{(n)}}\left(\sum_{j=1}^{Y_{i,r}^{(n)}} (E_j - 1)\right)^8 \leq C n^{-\frac{8r}{r+1}}\varepsilon^{-8}\Big(Y_{i,r}^{(n)}\Big)^4.$$

Тут $\mathbb{E}_{Y_{i,r}^{(n)}}$ позначає умовне математичне сподівання, C — деяку фіксовану константу, а саму нерівність треба розуміти в сенсі майже напевно.

Простими, хоча й дещо громіздкими обчисленнями на основі (3) легко переконатися, що при $n \to \infty$ виконується асимптотична рівність $\mathbb{E}\Big(Y_{i,r}^{(n)}\Big)^4 \sim C_r n^4$ з деякою фіксованою константою C_r. Тому, повертаючися до нерівності (7), отримуємо

$$\mathbb{P}\Big\{\big|n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i) - n^{-\frac{r}{r+1}} Y_{i,r}^{(n)}\big| > \varepsilon\Big\} \leq \\ \leq C n^{-\frac{8r}{r+1}}\varepsilon^{-8}\,\mathbb{E}\Big(Y_{i,r}^{(n)}\Big)^4 = C C_r n^{4-\frac{8r}{r+1}}\varepsilon^{-8} = C' n^{\frac{4-4r}{r+1}}\varepsilon^{-8}.$$

Тепер легко оцінити ймовірність того, що принаймні один атом випадкової міри $\eta_r^{(n)}$ відхилиться від відповідного атома міри $\xi_r^{(n)}$ більше, ніж на ε: за напівадитивністю

$$\mathbb{P}\Big\{\big|n^{-\frac{r}{r+1}} X_{r+1}^{(n)}(i) - n^{-\frac{r}{r+1}} Y_{i,r}^{(n)}\big| > \varepsilon \text{ для деякого } 1 \leq i \leq n\Big\} \leq \\ \leq n \cdot C' n^{\frac{4-4r}{r+1}}\varepsilon^{-8} = C' n^{\frac{5-3r}{r+1}}\varepsilon^{-8}. \tag{8}$$

Зафіксуємо довільний відрізок $[a, b] \subset [0, \infty]$. Тоді з нерівності (8) випливає оцінка

$$\mathbb{P}\big\{\xi_r^{(n)}[a, b] \neq \eta_r^{(n)}[a, b]\big\} \leq C' n^{\frac{5-3r}{r+1}}\varepsilon^{-8} + \\ + \mathbb{P}\big\{\eta_r^{(n)}[a - \varepsilon, a + \varepsilon] \geq 1\big\} + \mathbb{P}\big\{\eta_r^{(n)}[b - \varepsilon, b + \varepsilon] \geq 1\big\}. \tag{9}$$

Сенс цієї нерівності є досить прозорим: рівність $\xi_r^{(n)}[a,b] = \eta_r^{(n)}[a,b]$ може не виконуватися внаслідок лише двох причин — або деякі атоми $\xi_r^{(n)}$ значно відхиляються від відповідних атомів $\eta_r^{(n)}$, або деякі атоми $\eta_r^{(n)}$ розташовані досить близько до кінців відрізка $[a,b]$. Тому, покладаючи в (9) $\varepsilon = n^{-\frac{1}{25}}$ і позначаючи $\alpha_r = \frac{5-3r}{r+1} + \frac{8}{25}$, дістаємо

$$\mathbb{P}\big\{\xi_r^{(n)}[a,b] \neq \eta_r^{(n)}[a,b]\big\} \leq C' n^{\alpha_r} + \\ + \mathbb{P}\big\{\eta_r^{(n)}[a - n^{-\frac{1}{25}}, a + n^{-\frac{1}{25}}] \geq 1\big\} + \mathbb{P}\big\{\eta_r^{(n)}[b - n^{-\frac{1}{25}}, b + n^{-\frac{1}{25}}] \geq 1\big\}.$$

Оскільки, як неважко пересвідчитися, $\alpha_r < 0$ при $r \geq 2$, перший доданок у правій частині прямує до нуля при $n \to \infty$. Решта два доданки поводять себе так само — цей інтуїтивно очевидний факт є наслідком леми 3.5 з [3]. Отже, $\mathbb{P}\big\{\xi_r^{(n)}[a,b] \neq \eta_r^{(n)}[a,b]\big\} \to 0$ при $n \to \infty$. Більш того, з напівадитивності ймовірності випливає, що відрізок $[a,b]$ тут можна замінити на скінченне об'єднання таких відрізків $U = \cup_{i=1}^{l}[a_i, b_i]$. Тому за нерівністю каплінга для будь-якого $k \geq 0$ виконується

$$\big|\mathbb{P}\{\xi_r^{(n)}(U) = k\} - \mathbb{P}\{\eta_r^{(n)}(U) = k\}\big| \leq \mathbb{P}\big\{\xi_r^{(n)}(U) \neq \eta_r^{(n)}(U)\big\} \to 0, \qquad n \to \infty.$$

Разом зі збіжністю

$$\mathbb{P}\big\{\eta_r^{(n)}(U) = k\big\} \to \mathbb{P}\big\{\xi_r(U) = k\big\}, \qquad n \to \infty,$$

яка випливає з доведеного вище факту про $\eta_r^{(n)} \xrightarrow{vd} \xi_r$, це дає

$$\mathbb{P}\big\{\xi_r^{(n)}(U) = k\big\} \to \mathbb{P}\big\{\xi_r(U) = k\big\}, \qquad n \to \infty.$$

Отже, за теоремою 4.15 з [1] $\xi_r^{(n)} \xrightarrow{vd} \xi_r$, що й треба було довести.

5. Приклад застосування. У розділі 4 роботи [3] наведені деякі застосування граничної теореми для точкових процесів, пов'язаних з задачею збирача купонів. Не повторюючи тут аналогічні міркування для нашої задачі, ми обмежимося лише найпростішим прикладом застосування теореми 1 і покажемо, як з неї можна отримати граничну формулу (2). Перш за все зробимо одне зауваження, необхідне для доведення.

Зауваження 1. *Результат теореми 1 залишається справедливим, якщо розглядати випадкові міри* $\xi_r^{(n)}$ *та* ξ_r *на компактифікованому просторі* $[0,\infty]$. *Таку компактифікацію можна забезпечити, наприклад, уведенням повної метрики* $d(x,y) = |\operatorname{arctg} x - \operatorname{arctg} y|$, $x, y \in [0,\infty]$. *Доведення повністю аналогічне доведенню теореми 1.*

Перейдемо до доведення формули (2). Оскільки $\xi_r^{(n)} \xrightarrow{vd} \xi_r$, то перший атом міри $\xi_r^{(n)}$ збігається за розподілом до першого атома міри ξ_r. Цей факт випливає з теореми про неперервне відображення та того, що функціонал "перший атом точкової міри на компакті $[0,\infty]$" (див. зауваження 1) внаслідок твердження 3.13 з [11] є неперервним у грубій топології. Перший атом міри $\xi_r^{(n)}$ можна записати як $n^{-\frac{r}{r+1}} \min_{1\leq i\leq n} Y_{i,r}^{(n)} = n^{-\frac{r}{r+1}} T_r^{(n)}$. Розподіл першого атома A_r граничної пуассонівської міри ξ_r можна легко знайти: для $x \geq 0$ з формули (5) маємо

$$\mathbb{P}\{A_r \leq x\} = 1 - \mathbb{P}\big\{\xi_r[0,x] = 0\big\} = 1 - \mathrm{e}^{-\lambda_r[0,x]} = 1 - \mathrm{e}^{-\frac{x^{r+1}}{(r+1)!}}.$$

Звідси безпосередньо випливає граничний результат (2) для $r \geq 2$.

Формулу (2) можна істотно узагальнювати на зразок того, як це було зроблено в теоремі 4.1 у [3]. З міркувань, наведених після формулювання теореми 1, випливає нескінченновимірний аналог формули (2). Ми залишимо його без детального доведення, оскільки воно проводиться повністю аналогічно доведенню згаданої теореми 4.1.

Теорема 2. *Для фіксованого $r \geq 2$ нехай $T_{r,k}^{(n)}$, $1 \leq k \leq n$, означає перший момент часу, коли деякі k з n типів надійшли принаймні по $r+1$ разів кожний, і покладемо $T_{r,k}^{(n)} = 0$ при $k > n$. Розглянемо такі випадкові елементи простору числових послідовностей $\mathbb{R}^\infty$:*

$$V_r^{(n)} = \left(n^{-\frac{r}{r+1}} T_{r,k}^{(n)}, k \geq 1\right), \qquad V_r = \left(\left((r+1)! \sum_{j=1}^{k} E_j\right)^{\frac{1}{r+1}}, k \geq 1\right).$$

Тут, як і раніше, E_j, $j \geq 1$, є незалежними $\mathsf{Exp}(1)$-розподіленими випадковими величинами.

Тоді $V_r^{(n)} \xrightarrow{d} V_r$ при $n \to \infty$ у просторі $\mathbb{R}^\infty$.

6. Висновки та перспективи подальших досліджень. У роботі показано, що при $r \geq 2$ послідовність точкових процесів $\left(\xi_r^{(n)}, n \geq 1\right)$, які описують моменти $(r+1)$-х надходжень усіх n типів в узагальненій задачі про дні народження, збігається за розподілом у грубій топології до деякого неоднорідного точкового процесу Пуассона. Цей результат дає змогу досить просто отримувати граничні теореми для різноманітних змістовних функціоналів від таких процесів.

Надалі було б цікаво розповсюдити основний результат роботи на випадок $r = 1$, в якому техніка пуассонізації/депуассонізації не спрацьовує. Іншим можливим напрямком досліджень є вивчення збіжності точкового процесу, який об'єднує розглянуті процеси при різних r.

Список використаної літератури

1. Kallenberg O. Random measures, theory and applications. Springer, 2017. 707 p.
2. Glavaš L., Mladenović P. New limit results related to the coupon collector's problem. *Stud. Sci. Math. Hung.* 2018. Vol. 55, No. 1. P. 115–140.
3. Ilienko A. Convergence of point processes associated with coupon collector's and Dixie cup problems. *Electron. Commun. Probab.* 2019. Vol. 24, No. 51. P. 1–9.
4. Flajolet Ph., Grabner P.J., Kirschenhofer P., Prodinger H. On Ramanujan's Q-function. *J. Comput. Appl. Math.* 1995. Vol. 58. P. 103–116.
5. Klamkin M.S., Newman D.J. Extensions of the birthday surprise. *J. Comb. Theory.* 1967. Vol. 3. P. 279–282.
6. Dwass M. More birthday surprises. *J. Comb. Theory.* 1969. Vol. 7. P. 258–261.
7. Holst L. On birthday, collectors', occupancy and other classical urn problems. *Int. Stat. Rev.* 1986. Vol. 54, No. 1. P. 15–27.
8. Колчин В.Ф., Севастьянов Б.А., Чистяков В.П. Случайные размещения. Москва: Наука, 1976. 224 с.
9. Holst L. Extreme value distributions for random coupon collector and birthday problems. *Extremes.* 2001. Vol. 4, No. 2. P. 129–145.
10. Arratia R., Garibaldi S., Kilian J. Asymptotic distribution for the birthday problem with multiple coincidences, via an embedding of the collision process. *Random Struct. Algorithms.* 2016. Vol. 48, No. 3. P. 480–502.
11. Resnick S.I. Extreme values, regular variation, and point processes. Springer, 1987. 332 p.
12. Last G., Penrose M. Lectures on the Poisson process. Cambridge University Press, 2017. 307 p.

13. Gut A. Probability: a graduate course. Second edition. Springer, 2013. 625 p.

Ilienko A. B., Stamatieva V. V. A limit theorem for point processes associated with the generalized birthday problem.

We prove a limit theorem for the sequence of point processes, which describe the moments of $(r+1)$-th arrivals of different types out of the total number of n types in the generalized birthday problem. The classic birthday problem, known from popular literature, corresponds to the parameters $r = 1$ (one match is enough) and $n = 365$ (the number of days in a common year). The proof is based on an application of the poissonization/depoissonization technique. This result is further used to simply prove some classic limit theorems in the birthday problem, which actually describe the asymptotic behavior of various meaningful functionals of the constructed processes.

Keywords: birthday problem, point process, limit theorem, vague convergence, Poisson process, poissonization, depoissonization, continuous mapping theorem.

References

1. Kallenberg, O. (2017). Random measures, theory and applications. *Springer.*
2. Glavaš, L., & Mladenović, P. (2018). New limit results related to the coupon collector's problem. *Stud. Sci. Math. Hung., 55*(1), 115–140.
3. Ilienko, A. (2019). Convergence of point processes associated with coupon collector's and Dixie cup problems. *Electron. Commun. Probab., 24*(51), 1–9.
4. Flajolet, Ph., Grabner, P. J., Kirschenhofer, P., & Prodinger, H. (1995). On Ramanujan's Q-function. *J. Comput. Appl. Math., 58*, 103–116.
5. Klamkin, M. S., & Newman, D. J. (1967). Extensions of the birthday surprise. *J. Comb. Theory, 3*, 279–282.
6. Dwass, M. (1969). More birthday surprises. *J. Comb. Theory, 7*, 258–261.
7. Holst, L. (1986). On birthday, collectors', occupancy and other classical urn problems. *Int. Stat. Rev., 54*(1), 15–27.
8. Kolchin, V. F., Sevastianov, B. A., & Chistyakov, V. P. (1976). Sluchaynyye razmeshcheniya [Random allocations]. *Moscow: Nauka* [in Russian].
9. Holst, L. (2001). Extreme value distributions for random coupon collector and birthday problems. *Extremes, 4*(2), 129–145.
10. Arratia, R., Garibaldi, S., & Kilian, J. (2016). Asymptotic distribution for the birthday problem with multiple coincidences, via an embedding of the collision process. *Random Struct. Algorithms, 48*(3), 480–502.
11. Resnick, S. I. (1987). Extreme values, regular variation, and point processes. *Springer.*
12. Last, G., & Penrose, M. (2017). Lectures on the Poisson process. *Cambridge University Press.*
13. Gut, A. (2013). Probability: a graduate course. Second edition. *Springer.*

Одержано 15.10.2021