Зміст

Розділ 1

Перестановки Юенса

1.1 Граничний розподіл нерухомих точок

Розглянемо ймовірнісний розподіл на групі перестановок S_n, заданий наступним чином:

$$\mathbb{P}(\{\pi\}) = \frac{\theta^{\mathrm{c}(\pi)}}{\theta(\theta+1)\dots(\theta+n-1)}, \pi \in S_n \tag{1.1}$$

де $\theta > 0$, а $\mathrm{c}(\pi)$ позначає кількість циклів у π. Цей розподіл також відомий як *міра Юенса.* Тут і далі, відповідні випадкові перестановки називатимемо *перестановками Юенса.*

Зауваження. *Якщо* $\theta = 1$, (1.1) *задає рівномірний розподіл, тобто* $\mathbb{P}(\{\pi\}) = \frac{1}{n!}$ *для всіх* $\pi \in S_n$.

Перед тим, як вводити подальші поняття, розглянемо і доведемо наступну лему:

Лема 1.1.1. *Нехай* σ *— випадкова перестановка на множині* $\{1, \dots, n\}$, *що задана розподілом* (1.1) *(тобто,* σ *є перестановкою Юенса з* S_n*). Для* $a \in \mathbb{R}$ *позначимо* $\lceil a \rceil = \min\{k \in \mathbb{Z} : k \geq a\}$. *Нехай* $\gamma \in [0, 1]$, *а* $X_n = \operatorname{card}\{i \in \{1, \dots, \lceil \gamma n \rceil\} : \sigma(i) = i\}$ *— кількість нерухомих точок* σ *серед перших* $\lceil \gamma n \rceil$ *натуральних чисел. Тоді* X_n *за розподілом збігається до* $\operatorname{Poiss}(\gamma\theta)$, *тобто*

$$\lim_{n\to\infty} \mathbb{P}\{X_n = k\} = \frac{(\gamma\theta)^k}{k!} e^{-\gamma\theta}, k \in \mathbb{N}_0 \tag{1.2}$$

Доведення. Отримаємо явну формулу для $\mathbb{P}\{X_n = k\}$, починаючи з випадку $k = 0$. Нехай F_i позначає множину перестановок, для яких i є нерухомою точкою. Тоді

$$\mathbb{P}\{X_n = 0\} = \mathbb{P}\left(F_1^C \cap F_2^C \cap \dots \cap F_{\lceil \gamma n \rceil}^C\right) = 1 - \mathbb{P}\left(F_1 \cup F_2 \cup \dots \cup F_{\lceil \gamma n \rceil}\right) =$$
$$= 1 - \sum_i \mathbb{P}(F_i) + \sum_{i<j} \mathbb{P}(F_i \cap F_j) - \dots + (-1)^{\lceil \gamma n \rceil} \mathbb{P}\left(F_1 \cap F_2 \cap \dots \cap F_{\lceil \gamma n \rceil}\right)$$

У цьому виразі $\lceil \gamma n \rceil$ однакових доданків виду $\mathbb{P}(F_i)$, $C^2_{\lceil \gamma n \rceil}$ однакових доданків виду $\mathbb{P}(F_i \cap F_j)$ і так далі. Це означає, що достатньо знайти вирази для цих ймовірностей лише для конкретних наборів індексів. Якщо 1 є нерухомою точкою перестановки π, то вона має містити «тотожній» цикл (1), тобто $\pi = (1) \circ \tilde{\pi}$, де $\tilde{\pi}$ є перестановкою множини $\{2, \dots, n\}$. Аналогічно, якщо 1 і 2 є нерухомими точками π, то $\pi = (1)(2) \circ \tilde{\pi}$. Отже,

$$\mathbb{P}\{1, 2, \dots, i \text{ are the fixed points of } \sigma\} = \sum_{\pi=(1)(2)\dots(i)\circ\tilde{\pi}\in S_n} \mathbb{P}(\{\pi\}) =$$
$$= \sum_{\pi=(1)(2)\dots(i)\circ\tilde{\pi}\in S_n} \frac{\theta^{\mathrm{c}(\pi)}}{\theta(\theta+1)\dots(\theta+n-1)} = [\mathrm{c}(\pi) > i] =$$
$$= \frac{\theta^i}{\theta(\theta+1)\dots(\theta+n-1)} \sum_{\pi=(1)(2)\dots(i)\circ\tilde{\pi}\in S_n} \theta^{\mathrm{c}(\pi)-i}$$

Остання сума є сумою ймовірностей розподілу Юенса (1.1) на S_{n-i}, але без константи нормалізації, тому дорівнює $\theta(\theta+1)\dots(\theta+n-i-1)$, отже

$$\mathbb{P}\left\{1,2,\dots,i \text{ are the fixed points of } \sigma\right\} = \frac{\theta^i}{(\theta+n-1)\dots(\theta+n-i)}$$

З цього отримуємо

$$\mathbb{P}\left\{X_n=0\right\} = \sum_{i=0}^{\lceil\gamma n\rceil}(-1)^i C^i_{\lceil\gamma n\rceil}\frac{\theta^i}{(\theta+n-1)\dots(\theta+n-i)}$$

$\mathbb{P}\left\{X_n=k\right\}$ для $k>0$ можна отримати аналогічно: існує $C^k_{\lceil\gamma n\rceil}$ способів вибрати k натуральних чисел, які будуть нерухомими точками, а для інших $\lceil\gamma n\rceil-k$ застосувати формулу, аналогічну до $\mathbb{P}\left\{X_n=0\right\}$:

$$\mathbb{P}\left\{X_n=k\right\} = C^k_{\lceil\gamma n\rceil}\sum_{i=0}^{\lceil\gamma n\rceil-k}(-1)^i C^i_{\lceil\gamma n\rceil-k}\frac{\theta^{i+k}}{(\theta+n-1)\dots(\theta+n-i-k)}$$

Тепер доведемо $\lim_{n\to\infty}\mathbb{P}\left\{X_n=k\right\} = \frac{(\gamma\theta)^k}{k!}e^{-\gamma\theta}$.

$$\mathbb{P}\left\{X_n=k\right\} = \frac{(\lceil\gamma n\rceil)!}{k!(\lceil\gamma n\rceil-k)!}\sum_{i=0}^{\lceil\gamma n\rceil-k}(-1)^i\frac{(\lceil\gamma n\rceil-k)!}{i!(\lceil\gamma n\rceil-k-i)!}\frac{\theta^{i+k}}{(\theta+n-1)\dots(\theta+n-i-k)} =$$

$$= \frac{\theta^k}{k!}\sum_{i=0}^{\lceil\gamma n\rceil-k}(-1)^i\frac{\theta^i}{i!}\frac{\lceil\gamma n\rceil(\lceil\gamma n\rceil-1)(\lceil\gamma n\rceil-2)\dots(\lceil\gamma n\rceil-k-i+1)}{(\theta+n-1)\dots(\theta+n-i-k)}$$

Нехай N достатньо велике і $N<\lceil\gamma n\rceil-k$, тоді $\mathbb{P}\left\{X_n=k\right\}$ можна розбити на дві суми — S_1 від 0 до $N-1$ та S_2 від N до $\lceil\gamma n\rceil-k$.

$$\frac{k!}{\theta^k}\cdot|S_2| \le \sum_{i=N}^{\lceil\gamma n\rceil-k}\frac{\theta^i}{i!}\frac{\lceil\gamma n\rceil(\lceil\gamma n\rceil-1)(\lceil\gamma n\rceil-2)\dots(\lceil\gamma n\rceil-k-i+1)}{(\theta+n-1)\dots(\theta+n-i-k)} \le$$

$$\le \sum_{i=N}^{\lceil\gamma n\rceil-k}\frac{\theta^i}{i!}\left(\frac{\lceil\gamma n\rceil}{\theta+n-\lceil\gamma n\rceil}\right)^{i+k} \le \sum_{i=N}^{\lceil\gamma n\rceil-k}\frac{\theta^i}{i!}\left(\frac{\lceil\gamma n\rceil}{n-\lceil\gamma n\rceil}\right)^{i+k}$$

Оскільки $\lim_{n\to\infty}\frac{\lceil\gamma n\rceil}{n-\lceil\gamma n\rceil} = \frac{\gamma}{1-\gamma}$ для $\gamma\in[0,1)$, $\frac{\lceil\gamma n\rceil}{n-\lceil\gamma n\rceil}\le C=C(\gamma)$, то

$$\frac{k!}{\theta^k}\cdot|S_2| \le C^k\sum_{i=N}^{\lceil\gamma n\rceil-k}\frac{\theta^i}{i!}C^i \le C^k\sum_{i=N}^{\infty}\frac{\theta^i}{i!}C^i \to 0, N\to\infty$$

Якщо $\gamma=1$,

$$\frac{\lceil\gamma n\rceil(\lceil\gamma n\rceil-1)(\lceil\gamma n\rceil-2)\dots(\lceil\gamma n\rceil-k-i+1)}{(\theta+n-1)\dots(\theta+n-i-k)} = \frac{n(n-1)(n-2)\dots(n-k-i+1)}{(\theta+n-1)\dots(\theta+n-i-k)} \le$$

$$\le \frac{n(n-1)(n-2)\dots(n-k-i+1)}{(n-1)\dots(n-i-k)} = \frac{n}{n-i-k} \to 1, n\to\infty$$

тому цей дріб теж обмежений і $\lim_{N\to\infty}S_2=0$ також справджується. Що стосується S_1, то для фіксованого N

$$\lim_{n\to\infty}S_1 = \frac{\theta^k}{k!}\sum_{i=0}^{N-1}(-1)^i\frac{\theta^i}{i!}\lim_{n\to\infty}\frac{\lceil\gamma n\rceil(\lceil\gamma n\rceil-1)(\lceil\gamma n\rceil-2)\dots(\lceil\gamma n\rceil-k-i+1)}{(\theta+n-1)\dots(\theta+n-i-k)} =$$

$$= \frac{\theta^k}{k!}\sum_{i=0}^{N-1}(-1)^i\frac{\theta^i}{i!}\gamma^{i+k} = \frac{(\gamma\theta)^k}{k!}\sum_{i=0}^{N-1}(-1)^i\frac{(\gamma\theta)^i}{i!} \to \frac{(\gamma\theta)^k}{k!}e^{-\gamma\theta}, N\to\infty$$

□

Користуючись позначеннями з леми 1.1.1, X_n можна інтерпретувати як $P_n([0,\gamma])$, де P_n є деякою випадковою точковою мірою, тому результат леми можна записати як

$$\lim_{n\to\infty} \mathbb{P}\{P_n([0,\gamma]) = k\} = \mathbb{P}\{N([0,\gamma]) = k\}, k \in \mathbb{N}_0$$

де N є точковим процесом Пуассона з мірою інтенсивності $\theta \cdot \mathrm{Leb}$ on $[0,1]$. Виявляється, що має місце узагальнення цієї збіжності:

Теорема 1.1.2. *Точковий процес P_n грубо збігається за розподілом до точкового процесу Пуассона N з мірою інтенсивності $\Lambda = \theta \cdot \mathrm{Leb}$ на $[0,1]$ ($P_n \xrightarrow{vd} N, n \to \infty$).*

Як сказано в [1], *груба збіжність за розподілом* $\xi_n \xrightarrow{vd} \xi$ послідовності випадкових мір ξ_n на деякому просторі S до деякої випадкової міри ξ означає, зо $\mathbb{E}\varphi(\xi_n) \to \mathbb{E}\varphi(\xi)$ для кожної обмеженої функції φ, неперервної відносно *грубої топології* на просторі (невипадкових) мір на S. У цій грубій топології, $\mu_n \xrightarrow{v} \mu$ означає $\int_S f d\mu_n \to \int_S f d\mu$ для всіх обмежених неперервних функцій f на S з обмеженим носієм. Оскільки $S = [0,1]$, усі функції на S мають обмежений носій, а їх власна обмеженість випливає з неперервності.

Наведемо визначення точкового процесу Пуассона (з [2]). Нехай μ — міра Радона (радонова міра?) на σ-алгебрі борелевих підмножин множини S — $\mathcal{B}(S)$. Точковий процес N називається *процесом Пуассона* або *випадковою мірою Пуассона* з мірою інтенсивності μ, якщо N задовольняє наступні умови:

1. Для будь-якої $F \in \mathcal{B}(S)$ та будь-якого невід'ємного цілого числа k

$$\mathbb{P}\{N(F) = k\} = \begin{cases} \frac{(\mu(F))^k}{k!} \exp\{-\mu(F)\}, & \mu(F) < \infty \\ 0, & \mu(F) = \infty \end{cases}$$

2. Для будь-якого натурального k, якщо $F_1, \dots, F_k$ з $\mathcal{B}(S)$ попарно не перетинаються, то $(N(F_i), 1 \le i \le k)$ є незалежними в сукупності випадковими величинами.

Теорема 4.15 з [1] формулює критерій грубої збіжності точкових процесів:

Теорема (збіжність точкових процесів, Калленберг)**.** *Нехай $\xi, \xi_1, \xi_2, \dots$ — точкові процеси на S, де ξ простий, $\mathcal{U} \subset \hat{\mathcal{S}}_\xi$ — фіксоване (dissecting ring), а $\mathcal{I} \subset \mathcal{U}$ — напів-кільце. Тоді $\xi_n \xrightarrow{vd} \xi$ тоді і тільки тоді, коли*

1. $\lim_{n\to\infty} \mathbb{P}\{\xi_n(U) = 0\} = \mathbb{P}\{\xi(U) = 0\}, U \in \mathcal{U}$,
2. $\limsup_{n\to\infty} \mathbb{P}\{\xi_n(I) > 1\} \le \mathbb{P}\{\xi(I) > 1\}, I \in \mathcal{I}$.

Формулювання цієї теореми потребує додаткових пояснень. Для невипадкової міри μ $\hat{\mathcal{S}}_\mu$ означає клас борелевих множин $B \subset S$ з $\mu(\partial B) = 0$, а для випадкової міри ξ $\hat{\mathcal{S}}_\xi$ позначає $\hat{\mathcal{S}}_{\mathbb{E}\xi}$. Клас $\mathcal{C}$ обмежених борелевих підмножин S називається dissecting, якщо кожна відкрита множина $G \subset S$ представляється у вигляді зліченного об'єднання множин з $\mathcal{C}$, а кона обмежена борелева множина $B \subset S$ покривається скінченною кількістю множин з $\mathcal{C}$.

Розглянемо клас $\mathcal{X}$ of скінченних диз'юнктних об'єднань інтервалів $\langle a,b\rangle \subset [0,1]$, де $\langle a,b\rangle$ позначає одне з $[a,b]$, (a,b), $[a,b)$ чи $(a,b]$. Для точкового процесу Пуассона N з мірою інтенсивності $\Lambda = \theta \cdot \mathrm{Leb}$ на $[0,1]$ (який є простим), $\mathbb{E}N(\partial B) = \Lambda(\partial B)$, тому для всіх $B \subset \mathcal{X}$ $\mathbb{E}N(\partial B) = 0$, бо ∂B складається зі скінченного об'єднання окремих точок. Це означає, що $\hat{\mathcal{S}}_N = \mathcal{X}$. Також, $\mathcal{X}$ є кільцем і dissecting класом, оскільки всі необхідні умови очевидно виконуються. Отже, для доведення теореми 1.1.2, можна використати теорему Калленберга про збіжність для $\xi_n = P_n$, $\xi = N$ та $\hat{\mathcal{S}}_N = \mathcal{U} = \mathcal{I} = \mathcal{X}$.

Доведення теореми 1.1.2. Нехай $\langle\gamma_1,\delta_1\rangle, ..., \langle\gamma_m,\delta_m\rangle$ $(\gamma_1 < \delta_1 < \gamma_2 < ... < \gamma_m < \delta_m)$ — набір інтервалів в $[0,1]$, що попарно не перетинаються, $\lfloor a\rfloor = \max\{k\in\mathbb{Z} : k\le a\}$, $I_j = \{\lfloor\gamma_j n\rfloor+1, \lceil\delta_j n\rceil\}$ і $I = \bigvee_{j=1}^m I_j \in \mathcal{X}$. Позначимо $Y_n = \operatorname{card}\{i\in I : Z_n(i) = i\}$, тоді $P_n(I) = Y_n$. Як у і лемі 1.1.1, позначимо F_i множину перестановок, для яких i є нерухомою точкою, тоді

$$\mathbb{P}\{Y_n = 0\} = 1 - \mathbb{P}\left(\bigcup_{i\in I} F_i\right)$$

Нехай $\mathcal{M} = \operatorname{card} I = \sum_{j=1}^m (\lceil\delta_j n\rceil - \lfloor\gamma_j n\rfloor)$ і, аналогічно лемі 1.1.1,

$$\mathbb{P}\{Y_n = k\} = C_{\mathcal{M}}^k \sum_{i=0}^{\mathcal{M}-k} (-1)^i C_{\mathcal{M}-k}^i \frac{\theta^{i+k}}{(\theta+n-1)\dots(\theta+n-i-k)}$$

Оскільки $\lim_{n\to\infty}\frac{\mathcal{M}}{n} = \sum_{j=1}^m(\delta_j-\gamma_j)$, повторенням доведення збіжності у лемі 1.1.1, отримуємо

$$\lim_{n\to\infty}\mathbb{P}\{Y_n = k\} = \frac{1}{k!}\left(\theta\sum_{j=1}^m(\delta_j-\gamma_j)\right)^k \exp\left\{-\theta\sum_{j=1}^m(\delta_j-\gamma_j)\right\}, k\in\mathbb{N}_0$$

Оскільки $\Lambda(I) = \theta\cdot\operatorname{Leb}(I) = \theta\sum_{j=1}^m(\delta_j-\gamma_j)$, то

$$\lim_{n\to\infty}\mathbb{P}\{P_n(I) = 0\} = \mathbb{P}\{N(I) = 0\}, I\in\mathcal{X}$$

Так як $\mathbb{P}\{P_n(I) > 1\} = 1 - (\mathbb{P}\{P_n(I) = 0\} + \mathbb{P}\{P_n(I) = 1\})$ і $\mathbb{P}\{P_n(I) = 1\} \to \mathbb{P}\{N(I) = 1\}$ для $I\in\mathcal{X}$, отримуємо навіть більше, ніж треба:

$$\lim_{n\to\infty}\mathbb{P}\{P_n(I) > 1\} = \mathbb{P}\{N(I) > 1\}, I\in\mathcal{X}$$

Значення цих двох границь доводять $P_n \xrightarrow{vd} N, n\to\infty$. □

Для наступних досліджень будуть важливі перестановки з принаймні однією нерухомою точкою. З 1.1.2, для $\gamma = 1$ $\mathbb{P}\{X_n = 0\} = \sum_{i=0}^n(-1)^i C_n^i\frac{\theta^i}{(\theta+n-1)\dots(\theta+n-i)} \to e^{-\theta}, n\to\infty$. Введемо ще один точковий процес P_n', що визначений для борелевих множин $F\in\mathcal{B}([0,1])$ як

$$\mathbb{P}\{P_n'(F) = k\} = \mathbb{P}\{P_n(F) = k \mid P_n(F) > 0\} = \begin{cases}\frac{\mathbb{P}\{P_n(F)=k\}}{1-\mathbb{P}\{P_n(F)=0\}}, & k>0\\ 0, & k=0\end{cases} \tag{1.3}$$

Повторенням доведення 1.1.2 можна отримати наступний результат:

Теорема 1.1.3. *Точковий процес P_n' грубо збігається за розподілом до «обумовленого» точкового процесу Пуассона N' з мірою інтенсивності $\Lambda = \theta\cdot\operatorname{Leb}$ на $[0,1]$, для якого*

$$\mathbb{P}\{N'(F) = k\} = \frac{(\Lambda(F))^k}{k!}\frac{\exp\{-\Lambda(F)\}}{1-\exp\{-\Lambda(F)\}}$$

для всіх $F\in\mathcal{B}([0,1])$ ьта $k\in\mathbb{N}$.

Точковий процес N' можна назвати обумовленим процесом Пуассона, бо

$$\mathbb{P}\{N(F) = k \mid N(F) > 0\} = \frac{\mathbb{P}\{N(F) = k, N(F) > 0\}}{\mathbb{P}\{N(F) > 0\}} = \begin{cases}\frac{\mathbb{P}\{N(F)=k\}}{1-\exp\{-\Lambda(F)\}}, & k>0\\ 0, & k=0\end{cases}$$

тому $\mathbb{P}\{N(F) = k \mid N(F) > 0\} = \mathbb{P}\{N'(F) = k\}$ для $k\in\mathbb{N}$.

1.2 Статистичні властивості нерухомих точок

Результати теорем 1.1.2 та 1.1.3 можуть бути застосовані з теоремою про неперервне відображення з [3]. В термінах грубої збіжності точкових процесів, її можна сформулювати наступним чином:

Теорема 1.2.1 (теорема про неперервне відображення)**.** *Нехай φ є неперервним відображенням з простору $\mathcal{M}^p_{[0,1]}$ точкових мір на $[0,1]$ з грубою топологією в $\mathbb{R}$ зі стандартною топологією. Якщо ξ_n — послідовність точкових процесів, що грубо збігається за розподілом до ξ, $\xi_n \xrightarrow{vd} \xi$, тоді $\varphi(\xi_n) \xrightarrow{d} \varphi(\xi)$, тобто послідовність випадкових величин $\varphi(\xi_n)$ збігається за розподілом до $\varphi(\xi)$.*

Зауваження. *За означенням збіжності за розподілом, якщо φ є обмеженою, то також має місце $\mathbb{E}\varphi(\xi_n) \to \mathbb{E}\varphi(\xi)$.*

Перед дослідженням граничних розподілів для деяких відображень, варто навести твердження 3.13 з [2], яке можна сформулювати наступним чином:

Лема 1.2.2. *Нехай μ_n — грубо збіжна послідовність точкових мір в $\mathcal{M}^p_{[0,1]}$, тобто $\mu_n \xrightarrow{v} \mu$. Тоді*

$$\mu_n = \sum_{i=1}^{k} \delta_{x_i^{(n)}},\ \mu = \sum_{i=1}^{k} \delta_{x_i}$$

де δ_x є мірою Дірака, зосередженою в x, а $x_i^{(n)} \to x_i, n \to \infty$ для $i = 1, \ldots, k$.

Це означає, що будь-яка неперервна функція з $\mathbb{R}^k$ (або принаймні $[0,1]^k$) в $\mathbb{R}$ задає неперервне відносно грубої топології відображення.

1.2.1 Найменша та найбільша нерухомі точки

Для точкової міри μ відображення можна визначити два відображення $\min(\mu) = \sup\{x \in [0,1] : \mu($ та $\max(\mu) = \inf\{x \in [0,1] : \mu([x,1]) = 0\}$, що ставлять у відповідність цій міри її найменший та найбільший атом, де для порожньої множини за домовленістю $\sup\varnothing = 0$ та $\inf\varnothing = 1$. Нехай $\mu_n \xrightarrow{v} \mu$. Оскільки $\min\{x_1, \ldots, x_k\}$ та $\max\{x_1, \ldots, x_k\}$ є неперервними функціями з $\mathbb{R}^k$ в $\mathbb{R}$, з леми 1.2.2 випливає, що $\min(\mu)$ та $\max(\mu)$ є неперервними відносно грубої топології.

Незважаючи на результат теореми 1.1.3, простіше отримати розподіл $\min(N)$ та $\max(N)$, ніж $\min(N')$ та $\max(N')$, оскільки умовний розподіл $\mathbb{P}\{N(F) = k \mid N([0,1]) = m\}$ є відомим (твердження 3.8, [4]) — це біноміальний процес з розміром m та розподілом $\mathrm{Unif}\,(0,1)$. Іншим корисним фактом є такий: нехай $U_1, U_2, \ldots, U_m$ є тоді розподіли $U_{(1)}$ та $U_{(m)}$ задаються

$$\mathbb{P}\left\{U_{(1)} \le x\right\} = \begin{cases} 0, & x < 0 \\ 1-(1-x)^m, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases},\ \mathbb{P}\left\{U_{(m)} \le x\right\} = \begin{cases} 0, & x < 0 \\ x^m, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

Отже, розподіли $\min(N)$ та $\max(N)$ задаються

$$\mathbb{P}\{\min(N) \le x\} = \sum_{m=0}^{\infty} \mathbb{P}\{\min(N) \le x \mid N([0,1]) = m\}\, \mathbb{P}\{N([0,1]) = m\} = \tag{1.4}$$

$$= \mathbb{1}\{x \ge 1\} \cdot e^{-\theta} + \sum_{m=1}^{\infty} \mathbb{P}\{\min(N) \le x \mid N([0,1]) = m\} \frac{\theta^m}{m!} e^{-\theta} =$$

$$= \begin{cases} 0, & x < 0 \\ \sum_{m=1}^{\infty} \left(1-(1-x)^m\right) \frac{\theta^m}{m!} e^{-\theta} = 1 - e^{-\theta x}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

$$\mathbb{P}\left\{\max(N) \le x\right\} = \sum_{m=0}^{\infty} \mathbb{P}\left\{\max(N) \le x \mid N([0,1]) = m\right\} \mathbb{P}\left\{N([0,1]) = m\right\} = \tag{1.5}$$
$$= \mathbb{1}\left\{x \ge 0\right\} \cdot e^{-\theta} + \sum_{m=1}^{\infty} \mathbb{P}\left\{\max(N) \le x \mid N([0,1]) = m\right\} \frac{\theta^m}{m!} e^{-\theta} =$$
$$= \begin{cases} 0, & x < 0 \\ e^{-\theta} + \sum_{m=1}^{\infty} x^m \frac{\theta^m}{m!} e^{-\theta} = e^{\theta(x-1)}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases}$$

Ці розподіли є змішаними, бо $\mathbb{P}\left\{N([0,1]) = 0\right\} = e^{-\theta}$ і тому $\mathbb{P}\left\{\min(N) = 1\right\} = \mathbb{P}\left\{\max(N) = 0\right\} = e^{-\theta}$. Відповідні умовні розподіли є абсолютно неперервними:

$$\mathbb{P}\left\{\min(N) \le x \mid \min(N) < 1\right\} = \begin{cases} 0, & x < 0 \\ \frac{1-e^{-\theta x}}{1-e^{-\theta}}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases} \tag{1.6}$$

$$\mathbb{P}\left\{\max(N) \le x \mid \max(N) > 0\right\} = \begin{cases} 0, & x < 0 \\ \frac{e^{\theta x}-1}{e^{\theta}-1}, & 0 \le x < 1 \\ 1, & x \ge 1 \end{cases} \tag{1.7}$$

Умови $\{\min(N) < 1\}$ та $\{\max(N) > 0\}$ еквівалентні $\{N([0,1]) > 0\}$, тому умовні розподіли (1.6) та (1.7) задають безумовні розподіли $\min(N')$ та $\max(N')$.

Для P'_n $\min(P'_n) = \sup\left\{x \in [0,1] : \mathrm{card}\left\{i \in \{1, \dots, \lceil xn \rceil\} : Z(i) = i\right\} = 0\right\}$ означає супремум значень x, для яких усі $i \in \{1, \dots, \lceil xn \rceil\}$ під дією перестановки опиняються не на своїх місцях, аналогічно $\max(P'_n)$ означає інфімум x, для $i \in \{\lfloor xn \rfloor + 1, \dots, n\}$ опиняються не на своїх місцях. Легко побачити, що такі x можуть набувати лише значень, що є пропорційними $\frac{1}{n}$.

```
--- CDF plot ---
```

Обчислення $\mathbb{P}\left\{\frac{k-1}{n} < \min(N') \le \frac{k}{n}\right\}$ та $\mathbb{P}\left\{\frac{k-1}{n} < \max(N') \le \frac{k}{n}\right\}$ дає приблизну частку перестановок, для яких k є відповідно найменшою та найбільшою нерухомою точкою.

```
--- comparison table ---
```

Також, $\mathbb{E}\min(N') = \int_0^1 \left(1 - \frac{1-e^{-\theta x}}{1-e^{-\theta}}\right) dx = \frac{1}{\theta} + \frac{1}{1-e^{\theta}}$ та $\mathbb{E}\max(N') = \int_0^1 \left(1 - \frac{e^{\theta x}-1}{e^{\theta}-1}\right) dx = 1 - \frac{1}{\theta} + \frac{1}{e^{\theta}-1}$, тому $\lim_{n\to\infty} \mathbb{E}\min(P'_n) = \frac{1}{\theta} + \frac{1}{1-e^{\theta}}$ та $\lim_{n\to\infty} \mathbb{E}\max(P'_n) = 1 - \frac{1}{\theta} + \frac{1}{e^{\theta}-1}$.

1.2.2 Сума нерухомих точок

Граничний розподіл суми нерухомих точок можна отримати, користуючись функціоналом Лапласа точкового процесу Пуассона. З [3], для процесу Пуассона з мірою інтенсивності $\theta \cdot \mathrm{Leb}$ на $[0,1]$, цей функціонал задається

$$\psi_N(f) = \mathbb{E}\exp\left\{-\int_0^1 f(x)dN\right\} = \exp\left\{-\theta \int_0^1 \left(1 - e^{-f(x)}\right) dx\right\} \tag{1.8}$$

для вимірних, невід'ємних, обмежених функцій f на $[0,1]$.

Позначатимемо $\mathrm{sum}(N)$ суму атомів точкового процесу Пуассона N. Для будь-якої точкової міри μ, $\mathrm{sum}(\mu) = \int_0^1 x d\mu$. Перетворення Лапласа невід'ємної випадкової величини X задається $\mathcal{L}\left\{X\right\}(p) = \mathbb{E}e^{-pX}$. Якщо порівняти це означення з (1.8), можна побачити, що перетворення Лапласа $\mathrm{sum}(N)$ дорівнює значенню $\psi_N(f)$ для $f(x) = px$. Пряме обчислення дає наступний результат:

$$\mathcal{L}\left\{\mathrm{sum}(N)\right\}(p) = \exp\left\{-\theta\left(1 + \frac{1}{p}(e^{-p} - 1)\right)\right\} \tag{1.9}$$

Оскільки розподіл $\mathrm{sum}(N)$ є сумішшю абсолютно неперервного розподілу та дискретного з атомом в 0, можна знайти перетворення Лапласа лише абсолютно неперервної частини, що також буде перетворення для $\mathrm{sum}(N')$.

$$\mathcal{L}\left\{\mathrm{sum}(N)\right\}(p) = \mathbb{E}e^{-p\cdot\mathrm{sum}(N)} = 1\cdot\mathbb{P}\left\{\mathrm{sum}(N) = 0\right\} + \mathbb{E}e^{-p\cdot\mathrm{sum}(N')}\cdot\mathbb{P}\left\{\mathrm{sum}(N) > 0\right\} =$$
$$= e^{-\theta} + \mathcal{L}\left\{\mathrm{sum}(N')\right\}(p)\cdot(1 - e^{-\theta}) \Rightarrow \mathcal{L}\left\{\mathrm{sum}(N')\right\}(p) = \frac{1}{1 - e^{-\theta}}\left(\mathcal{L}\left\{\mathrm{sum}(N)\right\}(p) - e^{-\theta}\right) =$$
$$= \frac{e^{-\theta}}{1 - e^{-\theta}}\exp\left\{-\frac{\theta}{p}(e^{-p} - 1)\right\}$$

Оскільки $\mathrm{sum}(N')$ є абсолютно неперервною випадковою величиною, $\mathcal{L}\left\{\mathrm{sum}(N')\right\}(p)$ є перетворенням Лапласа для щільності, тому перетворення Лапласа для функції розподілу $\mathrm{sum}(N')$ задається

$$\mathcal{L}\left\{F_{\mathrm{sum}(N')}(x)\right\}(p) = \frac{e^{-\theta}}{1 - e^{-\theta}}\frac{1}{p}\exp\left\{-\frac{\theta}{p}(e^{-p} - 1)\right\} \tag{1.10}$$

Знаходження оберненого перетворення для (1.10) є доволі складним.

Хоча, існує інший підхід до знаходження $F_{\mathrm{sum}(N)}(x)$:

$$\mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = \sum_{m=0}^{\infty}\mathbb{P}\left\{\mathrm{sum}(N) \le x \mid N([0,1]) = m\right\}\mathbb{P}\left\{N([0,1]) = m\right\} =$$
$$= \mathbb{1}\left\{x \ge 0\right\}\cdot e^{-\theta} + \sum_{m=1}^{\infty}\mathbb{P}\left\{\mathrm{sum}(N) \le x \mid N([0,1]) = m\right\}\frac{\theta^m}{m!}e^{-\theta}$$

Умовні розподіли $\mathbb{P}\left\{\mathrm{sum}(N) \le x \mid N([0,1]) = m\right\}$ є розподілами Ірвіна-Голла — розподілами суми m незалежних випадкових величин з розподілом $\mathrm{Unif}\,(0,1)$. Їх функція розподілу наступна:

$$\begin{cases} 0, & x < 0 \\ \frac{1}{m!}\sum_{k=0}^{\lfloor x\rfloor}(-1)^k C_m^k (x-k)^m, & 0 \le x < m \\ 1, & x \ge m \end{cases}$$

Для кожного інтервалу $[n, n+1)$, $n \in \mathbb{N}\cup\{0\}$, $\mathbb{P}\left\{\mathrm{sum}(N) \le x\right\}$ може бути виражена через $I_\nu(z)$, $\nu \in \mathbb{R}$ — модифіковані функції Бесселя першого роду:

$$I_\nu(z) = \left(\frac{1}{2}z\right)^\nu \sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}z\right)^k}{k!\Gamma(\nu + k + 1)}$$

Наприклад:

$$x \in [0,1),\ \mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = e^{-\theta}\left(1 + \sum_{m=1}^{\infty} \frac{\theta^m}{(m!)^2} x^m\right) = e^{-\theta} \cdot I_0(2\sqrt{\theta x})$$

$$x \in [1,2),\ \mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = e^{-\theta}\left(1 + \frac{\theta}{1!} + \sum_{m=2}^{\infty} \frac{\theta^m}{(m!)^2} x^m - \sum_{m=2}^{\infty} \frac{\theta^m}{(m-1)!m!}(x-1)^m\right) =$$
$$= e^{-\theta}\left(I_0(2\sqrt{\theta x}) - \sqrt{\theta(x-1)} \cdot I_1(2\sqrt{\theta(x-1)})\right)$$

$$x \in [2,3),\ \mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = e^{-\theta}\left(1 + \frac{\theta}{1!} + \frac{\theta^2}{2!} + \sum_{m=3}^{\infty} \frac{\theta^m}{(m!)^2} x^m - \sum_{m=3}^{\infty} \frac{\theta^m}{(m-1)!m!}(x-1)^m + \right.$$
$$\left. + \sum_{m=3}^{\infty} \frac{\theta^m}{(m-2)!m!}(x-2)^m\right) = e^{-\theta}\left(I_0(2\sqrt{\theta x}) - \sqrt{\theta(x-1)} \cdot I_1(2\sqrt{\theta(x-1)}) + \right.$$
$$\left. + \frac{1}{2!}\theta(x-2) I_2(2\sqrt{\theta(x-2)})\right)$$

В загальному випадку:

$$x \in [n, n+1),\ \mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = e^{-\theta} \sum_{k=0}^{n} (-1)^k \frac{1}{k!} \left(\theta(x-k)\right)^{\frac{k}{2}} I_k\left(2\sqrt{\theta(x-k)}\right)$$

Отже, функція розподілу $\mathrm{sum}(N)$ задається

$$F_{\mathrm{sum}(N)}(x) = \mathbb{P}\left\{\mathrm{sum}(N) \le x\right\} = \begin{cases} 0, & x < 0 \\ e^{-\theta} \sum\limits_{k=0}^{\lfloor x \rfloor} (-1)^k \frac{1}{k!} \left(\theta(x-k)\right)^{\frac{k}{2}} I_k\left(2\sqrt{\theta(x-k)}\right), & x \ge 0 \end{cases} \tag{1.11}$$

Таким чином, функція розподіу $\mathrm{sum}(N')$ може бути виражена через $F_{\mathrm{sum}(N)}(x)$ настуним чином:

$$\mathbb{P}\left\{\mathrm{sum}(N') \le x\right\} = \mathbb{P}\left\{\mathrm{sum}(N) \le x \mid \mathrm{sum}(N) > 0\right\} = \frac{1}{1 - e^{-\theta}}\left(F_{\mathrm{sum}(N)}(x) - e^{-\theta}\right) \tag{1.12}$$

`--- CDF plot ---`

При цьому, $\mathbb{E}\,\mathrm{sum}(N)$ значно простіше знайти за формулою повного математичного сподівання, оскільки для $m > 0$ $\mathbb{E}\left(\mathrm{sum}(N) \mid N([0,1]) = m\right) = \frac{m}{2}$ як математичне сподівання суми m незалежних випадкових величин з розподілом $\mathrm{Unif}\,(0,1)$:

$$\mathbb{E}\,\mathrm{sum}(N) = 0 \cdot \mathbb{P}\left\{N([0,1]) = 0\right\} + \sum_{m=1}^{\infty} \frac{m}{2} \frac{\theta^m}{m!} e^{-\theta} = \frac{e^{-\theta}}{2} \sum_{m=1}^{\infty} \frac{\theta^m}{(m-1)!} = \frac{\theta}{2}$$

1.2.3 Найменші і найбільші спейсинги

Визначимо граничні розподілу найменшого і найбільшого спейсингів — відстаней між нерухомими точками.

Зауваження. *Щоб застосувати тут теоретичні результати, що стосуються випадкового розбиття інтервалів, зручно вважати* $\min(N)$ *and* $1 - \max(N)$ *спейсингами. Для випадкової перестановки* $\{1, \dots, n\}$ *це означатиме вважати* 0 *та* $n+1$ *«штучними» нерухомими точками.*

Нехай $U_1, U_2, \ldots, U_n$ — незалежні випадкові величин з розподілом $\mathrm{Unif}\,(0,1)$, що розділяють відрізок $[0,1]$ на $n+1$ інтервалів з довжинами $S_1, S_2, \ldots, S_{n+1}$, або, у відсортованому вигляді, $S_{(1)}^{[n+1]} < S_{(2)}^{[n+1]} < \cdots < S_{(n+1)}^{[n+1]}$ (тут і далі, $S_{(i)}$ позначає i-ту порядкову статистику, а $S_{(i)}^{[n]}$ — те ж саме, але з вказанням n як кількості цих статистик). Розподіли $S_{(k)}^{[n+1]}$ отримано у багатьох роботах (наприклад, [5], [6]). Зокрема, для $x \in [0,1]$:

$$\mathbb{P}\left\{S_{(1)}^{[n+1]} > x\right\} = \left((1-(n+1)x)_+\right)^n \tag{1.13}$$

$$\mathbb{P}\left\{S_{(n+1)}^{[n+1]} > x\right\} = \sum_{j=1}^{n+1} (-1)^{j-1} C_{n+1}^j \left((1-jx)_+\right)^n \tag{1.14}$$

де $x_+ = \max(x, 0)$.

Отже, розподіли найменшого $\text{s-min}(N)$ та найбільшого $\text{s-max}(N)$ спейсингів між атомами N задаються (з домовленістю $S_{(1)}^1 = 1$)

$$\mathbb{P}\left\{\text{s-min}(N) > x\right\} = \sum_{n=0}^{\infty} \mathbb{P}\left\{S_{(1)}^{[n+1]} > x\right\} \mathbb{P}\left\{N([0,1]) = n\right\} \tag{1.15}$$

$$\mathbb{P}\left\{\text{s-max}(N) > x\right\} = \sum_{n=0}^{\infty} \mathbb{P}\left\{S_{(n+1)}^{[n+1]} > x\right\} \mathbb{P}\left\{N([0,1]) = n\right\} \tag{1.16}$$

Хоча явні вирази для (1.15) та (1.16), скоріш за все, доволі складні, цікаво звернути увагу на дві випадкові величини з такими ж розподілами.

Добре відомо (наприклад, [5]), що для незалежних величин $X_1, X_2, \ldots, X_n$ з розподілом $\mathrm{Exp}\,(1)$ мають місце наступні три рівності:

$$(S_1, S_2, \ldots, S_n)^T \stackrel{d}{=} \left(\frac{X_1}{\sum_{i=1}^n X_i}, \frac{X_2}{\sum_{i=1}^n X_i}, \ldots, \frac{X_n}{\sum_{i=1}^n X_i}\right)^T \tag{1.17}$$

$$\left(S_{(1)}, S_{(1)}, \ldots, S_{(n)}\right)^T \stackrel{d}{=} \left(\frac{X_{(1)}}{\sum_{i=1}^n X_i}, \frac{X_{(2)}}{\sum_{i=1}^n X_i}, \ldots, \frac{X_{(n)}}{\sum_{i=1}^n X_i}\right)^T \tag{1.18}$$

$$X_{(i)} \stackrel{d}{=} \frac{X_n}{n} + \frac{X_{n-1}}{n-1} + \cdots + \frac{X_{n-i+1}}{n-i+1} \tag{1.19}$$

Виявляється, (1.18) та (1.19) можна узагальнити в наступну неочікувану рівність:

Лема 1.2.3. *Для порядкових статистик спейсингів між незалежними величинами з розподілом* $\mathrm{Unif}\,(0,1)$ *та та незалежних величин* $X_1, X_2, \ldots, X_n$ *з розподілом* $\mathrm{Exp}\,(1)$ *має місце*

$$S_{(i)}^{[n]} \stackrel{d}{=} \frac{\frac{X_n}{n} + \frac{X_{n-1}}{n-1} + \cdots + \frac{X_{n-i+1}}{n-i+1}}{\sum_{i=1}^n X_i}, i = 1, \ldots, n \tag{1.20}$$

Доведення. Позначимо спейсинги між $X_1, X_2, \ldots, X_n$ через $\Delta_1 = X_{(1)}$, $\Delta_i = X_{(i)} - X_{(i-1)}, i = 2, \ldots, n$. З [7] відомо, що всі Δ_i незалежні та мають розподіли $\mathrm{Exp}\,(n-i+1)$. Отже, праву частину $S_{(i)} \stackrel{d}{=} \frac{X_{(i)}}{\sum_{j=1}^n X_j}$ можна переписати як

$$\frac{X_{(i)}}{\sum_{j=1}^n X_j} = \frac{X_{(i)}}{\sum_{j=1}^n X_{(j)}} = \frac{\Delta_1 + \cdots + \Delta_i}{\Delta_1 + (\Delta_1 + \Delta_2) + \cdots + (\Delta_1 + \cdots + \Delta_n)}$$

Введемо нові незалежні випадкові величини $Y_i = (n-i+1)\Delta_i$ з розподілом $\mathrm{Exp}\,(1)$. В термінах Y_i, верхню рівність можна переписати як

$$\frac{X_{(i)}}{\sum_{j=1}^n X_j} = \frac{\sum_{j=1}^i \frac{Y_j}{n-j+1}}{\sum_{j=1}^n Y_j}$$

Оскільки X_i та Y_i незалежні та мають однакові розподіли, то отримуємо (1.20). $\square$

Окремими випадками леми 1.2.3 є $S_{(1)}^{[n]} \stackrel{d}{=} \frac{X_1}{n\sum_{i=1}^n X_i}$ та $S_{(n)}^{[n]} \stackrel{d}{=} \frac{\sum_{i=1}^n \frac{X_i}{n-i+1}}{\sum_{i=1}^n X_i} \stackrel{d}{=} \frac{\sum_{i=1}^n \frac{X_i}{i}}{\sum_{i=1}^n X_i}$. Разом з (1.15) та (1.16) вони приводять до наступних рівностей за розподілом:

$$\text{s-min}(N) \stackrel{d}{=} \frac{X_1}{(\nu+1)\sum_{i=1}^{\nu+1} X_i}, \ \text{s-max}(N) \stackrel{d}{=} \frac{\sum_{i=1}^{\nu+1} \frac{X_i}{i}}{\sum_{i=1}^{\nu+1} X_i} \tag{1.21}$$

де ν має розподіл $\text{Poiss}\,(\theta)$, а $(X_i, i \in \mathbb{N})$ незалежні і мають розподіл $\text{Exp}\,(1)$.

$\mathbb{E}\,\text{s-min}(N)$ та $\mathbb{E}\,\text{s-max}(N)$ можна знайти з (1.21). Нехай $n \in \mathbb{N} \cup \{0\}$, тоді

$$\mathbb{E}\left(\frac{X_1}{(n+1)\sum_{i=1}^{n+1} X_i}\right) = \frac{1}{(n+1)^2} \cdot \mathbb{E}\left(\frac{X_1}{\sum_{i=1}^{n+1} X_i} + \dots + \frac{X_{n+1}}{\sum_{i=1}^{n+1} X_i}\right) = \frac{1}{(n+1)^2}$$

$$\mathbb{E}\left(\frac{\sum_{i=1}^{n+1} \frac{X_i}{i}}{\sum_{i=1}^{n+1} X_i}\right) = \sum_{i=1}^{n+1} \frac{1}{i} \cdot \mathbb{E}\left(\frac{X_i}{\sum_{i=1}^{n+1} X_i}\right) = \frac{1}{n+1} \cdot \sum_{i=1}^{n+1} \frac{1}{i}$$

Оскільки $\mathbb{P}\{\nu = n\} = \frac{\theta^n}{n!}e^{-\theta}$, то

$$\mathbb{E}\,\text{s-min}(N) = \frac{e^{-\theta}}{\theta} \sum_{n=1}^{\infty} \frac{\theta^n}{n \cdot n!} = \frac{e^{-\theta}}{\theta} \int_0^\theta \frac{e^t - 1}{t} dt$$

$$\mathbb{E}\,\text{s-max}(N) = \frac{e^{-\theta}}{\theta} \sum_{n=1}^{\infty} \frac{H_n}{n!} \theta^n$$

де $H_n = \sum_{k=1}^n \frac{1}{k}$ — n-те гармонічне число Зокрема, для $\theta = 1$ (випадок рівномірного розподілу) $\mathbb{E}\,\text{s-min}(N) \approx 0.48483$ and $\mathbb{E}\,\text{s-max}(N) \approx 0.7966$.

Бібліографія

[1] Olav Kallenberg. *Random Measures, Theory and Applications.* Springer International Publishing, 2017.

[2] Sidney I. Resnick. *Extreme Values, Regular Variation and Point Processes.* Springer New York, 1987.

[3] Sidney I. Resnick. Crash course II: Weak convergence; implications for heavy-tail analysis. In *Heavy-Tail Phenomena*, pages 39–69. Springer New York, 2007.

[4] Günter Last and Mathew Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[5] Lars Holst. On the lengths of the pieces of a stick broken at random. *Journal of Applied Probability*, 17(3):623–634, 1980.

[6] Iosif Pinelis. Order statistics on the spacings between order statistics for the uniform distribution, 2019.

[7] Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A First Course in Order Statistics (Classics in Applied Mathematics).* Society for Industrial and Applied Mathematics, USA, 2008.