Новые результаты о распределении Миттаг-Леффлера

Чернышенко Екатерина Глебовна

Кафедра теории вероятностей, механико-математический факультет
МГУ имени М.В. Ломоносова

18 ноября, 2025

Определение процесса Юла.

Определение

Процесс Юла (Феллер (1952) [10])
Рассм. совокупность элементов, которые могут порождать новые элементы, но не могут исчезать. Пусть за короткий промежуток времени h каждый элемент с вероятностью $\lambda h + o(h)$ производит новый элемент; постоянная λ определяет скорость разрастания совокупности. Если не имеется никакого взаимодействия между элементами и в момент t объем совокупности равен n, то вероятность увеличения объема за время $(t, t+h)$ равна $n\lambda h + o(h)$. Вероятность $P_n(t)$ того, что объем совокупности равен ровно n, удовлетворяет уравнению

$$P_n'(t) = -n\lambda P_n(t) + (n-1)\lambda P_{n-1}(t) \quad (n > 1)$$

Если i - объем совокупности в момент $t = 0$, то

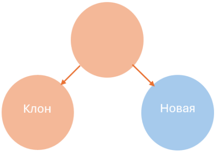
$$P_n(t) = \binom{n-1}{n-i} e^{-i\lambda t} \left(1 - e^{-it}\right)^{n-t}$$

при $n \geqslant i$, в то время как, конечно, $P_n(t) = 0$ при $n < i$.

Описание модели.

Модель

Рассмотрим популяцию частиц двух типов: новых и клонов. Начиная с одной новой частицы в момент времени $t = 0$, полагаем, что каждая частица производит потомство на протяжении своей бесконечной жизни. Клоны - это частицы, идентичные родителю. Новые - это частицы, отличные от всех предыдущих частиц в популяции.



Применение в генетике

Процесс Юла моделирует появление и фиксацию нейтральных мутаций. Он является основным инструментом для анализа геномных данных и реконструкции демографической истории.

Описание модели.

Параметр модели

В данной модели фиксируется один параметр $0 \leq \alpha \leq 1$ - интенсивность пуассоновского процесса.

Эволюции частиц

За малый промежуток времени h возможны следующие эволюции частицы, являющиеся **пуассоновскими процессами**:

• Если частица является **новой**, то она может произвести **новую** частицу с вероятностью $\alpha h + o(h)$.

• Если частица является **новой**, то она может произвести своего **клона** с вероятностью $(1 - \alpha)h + o(h)$.

• Если частица является **клоном**, то она может произвести **только клона** с вероятностью $h + o(h)$.

• Наконец за малое h частица может никого не произвести.

Объекты исследования

Процессы Юла

Объектами исследования являются два процесса Юла:
• $N_t :=$ число всех особей в момент времени t.
• $N_t^* :=$ число новых особей (особей различных цветов) в момент времени t.
Отметим, что$N_0^* = N_0 = 1$, и $1 \leq N_t^* \leq N_t$ для всех $t \geq 0$.

Параметры процессов Юла

• Процесс $(N_t^*, t \geq 0)$ - это процесс Юла с параметром α. Если на момент времени t $N_t^* = i$, то параметр перехода к состоянию $N_t^* = i+1$ равен $i\alpha$, так как новые порождаются только новыми с параметром α.
• Процесс $(N_t, t \geq 0)$ - это процесс Юла с параметром 1. Если на момент времени t $N_t = i$, а $N_t^* = j$, то число клонов равно $i - j$. Тогда параметр перехода из $N_t = i$ в $N_t = i + 1$ равен i, так как $\alpha j + (1 - \alpha)j + 1(i - j) = \alpha j + j - \alpha j + i - j = j + i - j = i$.

Предельная теорема для числа различно окрашенных особей в популяции

Теорема

Пусть $0 < \alpha < 1$.
Тогда

$$\frac{N_t^*}{N_t^\alpha} \to \zeta \quad \textit{п.н.},$$

где ζ *- положительная случайная величина с распределением Миттаг-Леффлера с параметром* α.

Доказательство

Шаг 1. Связь процессов N_t^* и N_t с числом занятых столиков K_n из процесса китайского ресторана

Предположим, что каждая новая частица окрашивается в цвет, отличный от всех имеющихся цветов в популяции, а каждый клон окрашивается в тот же цвет, что и его родитель. Тогда число новых особей N_t^* к моменту времени t является количеством различных цветов к моменту t.

Тогда скажем, что:

- Число N_t^* цветов $=$ число m занятых столиков;
- Число всех особей N_t $=$ число всех вошедших посетителей n;
- Число q_i особей одного i-го цвета $=$ число q_i посетителей за i-м столиком.

Доказательство

Тогда на каждом следующем малом промежутке времени h рождается $N_t + 1$-я частица:
• **Новая** (т.е. нового цвета, тогда цветов станет $N_t^* + 1$, или, что то же самое, - $n+1$-й посетитель сел за незанятый столик).
• или **клон** (т.е. цвета родителя, тогда число цветов не изменилось, или, что то же самое, - $n+1$-й посетитель сел за уже занятый столик).

Для любого момента времени t, такого что $N_t = n$, число новых особей N_t^* совпадает по распределению с числом K_n:

$$N_t^* \stackrel{d}{=} K_n \quad \text{при условии } N_t = n.$$

Доказательство

Шаг 2. Введение случайной последовательности моментов времени

Определим моменты времени τ_n как моменты, когда общее число особей достигает n:

$$\tau_n = \inf\{t > 0 : N_t = n\}.$$

Тогда:

$$N_{\tau_n} = n, \quad N^*_{\tau_n} = K_n.$$

Поэтому для τ_n процесс N^*_t и K_n - это одно и то же. Поэтому для τ_n для N^*_t выполняется то же, что и для K_n.

Доказательство

Шаг 3. Предельная теорема о сходимости почти наверное для K_n

Теорема Питмана (2006, [5])

Пусть $0 < \alpha < 1$ и $\theta > -\alpha$.
Предположим, что выполнены следующие условия на моменты:
•$[\mathrm{K}_n] < \infty$ для всех $n \geq 1$
•$[\mathrm{K}_n^2] < \infty$ для всех $n \geq 1$
• Существуют константы $C_1, C_2 > 0$ такие, что для всех $n \geq 1$:
$\mathrm{E}[\mathrm{K}_n] \leq C_1 n^\alpha$
$E[K_n^2] \leq C_2 n^{2\alpha}$
Тогда

$$\lim_{n\to\infty} \frac{K_n}{n^\alpha} = \zeta \quad \text{п.н.},$$

где ζ - это строго положительная случайная величина с непрерывной плотностью $\frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} x^{\theta/\alpha} g_\alpha(x)$, $x > 0$ и $g_\alpha(x) = \frac{1}{\pi\alpha} \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k!} \Gamma(\alpha k + 1) sin(\pi\alpha k) x^{k-1}$, $x > 0$ - это плотность Миттаг-Леффлера.

Доказательство

В нашей модели $\theta = 0$, но идею доказательства для теоремы для K_n мы приведем для общего двухпараметрического случая.

Док-во теоремы Питмана (B. Bercu, S. Favaro (2024) [11]).

Везде далее мы считаем, что $0 < \alpha < 1$, $\theta > -\alpha$.

Шаг 3.1: Построение мартингала

Из модели китайского ресторана имеем:

$$K_{n+1} = K_n + \xi_{n+1}$$

где $\xi_{n+1} \mid F_n \sim$ Бернулли(p_n) с параметром

$$p_n = \frac{\alpha K_n + \theta}{n + \theta}, \quad F_n = \sigma(K_1, \ldots, K_n)$$

Тогда

$$E[K_{n+1} \mid F_n] = K_n + p_n = \left(1 + \frac{\alpha}{n+\theta}\right) K_n + \frac{\theta}{n+\theta}$$

Обозначим $\beta_n = 1 + \frac{\alpha}{n+\theta}$, тогда

$$E\left[K_{n+1} + \frac{\theta}{\alpha} \mid F_n\right] = \beta_n \left(K_n + \frac{\theta}{\alpha}\right)$$

Доказательство

Введём последовательность:

$$b_n = \prod_{k=1}^{n-1} \beta_k^{-1} = \prod_{k=1}^{n-1} \left(\frac{k+\theta}{k+\alpha+\theta} \right)$$

и определим мартингал:

$$M_n = b_n \left(K_n + \frac{\theta}{\alpha} \right)$$

Проверим мартингальное свойство:

$$\begin{aligned} E[M_{n+1} \mid F_n] &= b_{n+1} E\left[K_{n+1} + \frac{\theta}{\alpha} \mid F_n \right] \\ &= b_{n+1}\beta_n \left(K_n + \frac{\theta}{\alpha} \right) \\ &= b_n \left(K_n + \frac{\theta}{\alpha} \right) = M_n \end{aligned}$$

Доказательство

Шаг 3.2: Вычисление квадратичной вариации

Имеем:

$$M_{n+1} - M_n = b_{n+1}(\xi_{n+1} - E[\xi_{n+1} \mid F_n])$$

Следовательно:

$$E[(M_{n+1} - M_n)^2 \mid F_n] = b_{n+1}^2 \cdot p_n(1 - p_n),$$

где

$$p_n(1 - p_n) = \frac{(\alpha K_n + \theta)(n - \alpha K_n)}{(n + \theta)^2}$$

Квадратичная вариация:

$$\langle M \rangle_n = \sum_{k=0}^{n-1} [(M_{k+1} - M_k)^2 \mid F_k] =$$

$$= \left(\frac{\alpha + \theta}{\alpha}\right)^2 + \sum_{k=1}^{n-1} b_{k+1}^2 \cdot \frac{(\alpha K_k + \theta)(k - \alpha K_k)}{(k + \theta)^2}$$

Доказательство

Шаг 3.3: Ограниченность квадратичной вариации
Возьмём математическое ожидание:

$$E[\langle M\rangle_n] = \left(\frac{\alpha+\theta}{\alpha}\right)^2 + \sum_{k=1}^{n-1} b_{k+1}^2 \cdot E\left[\frac{(\alpha K_k+\theta)(k-\alpha K_k)}{(k+\theta)^2}\right]$$

$$\leq \left(\frac{\alpha+\theta}{\alpha}\right)^2 + \sum_{k=1}^{n-1} b_{k+1}^2 \cdot \frac{\alpha k E[K_k]+\theta k}{(k+\theta)^2}$$

Используя условие $E[K_k] \leq C_1 k^\alpha$ и асимптотику $b_k \sim C \cdot k^{-\alpha}$, получаем:

$$b_k^2 \cdot \frac{E[K_k]}{k} \sim C^2 k^{-2\alpha} \cdot C_1 k^{\alpha-1} = O(k^{-1-\alpha})$$

Поскольку $\alpha > 0$, ряд $\sum k^{-1-\alpha}$ сходится, следовательно:

$$\sup_{n\geq 1} E[\langle M\rangle_n] < \infty$$

Доказательство

Шаг 3.4: Сходимость мартингала
Из ограниченности $E[\langle M\rangle_n]$ следует ограниченность $E[M_n^2]$, поскольку для мартингала:

$$E[M_n^2] = E[\langle M\rangle_n] + E[M_0^2]$$

По теореме сходимости мартингалов Дуба, (M_n) сходится почти наверное и в L^2 к некоторой случайной величине $M_{\alpha,\theta}$:

$$M_n \to M_{\alpha,\theta} \quad \text{п.н.}$$

Доказательство

Шаг 3.5: Переход к сходимости K_n/n^α

Из асимптотики b_n:

$$\lim_{n\to\infty} n^\alpha b_n = \frac{\Gamma(\alpha+\theta+1)}{\Gamma(\theta+1)} =: C_b > 0$$

Поскольку $M_n = b_n(K_n + \frac{\theta}{\alpha})$, имеем: $K_n = \frac{M_n}{b_n} - \frac{\theta}{\alpha}$, $\quad \frac{K_n}{n^\alpha} = \frac{M_n}{b_n n^\alpha} - \frac{\theta}{\alpha n^\alpha}$

Правая часть сходится почти наверное:

$$\frac{M_n}{b_n n^\alpha} \to \frac{M_{\alpha,\theta}}{C_b}, \quad \frac{\theta}{\alpha n^\alpha} \to 0$$

Определяя $\zeta = \frac{M_{\alpha,\theta}}{C_b}$, получаем:

$$\frac{K_n}{n^\alpha} \to \zeta \quad \text{п.н.}$$

Доказательство

Шаг 4. Возврат от τ_n к t

Из предыдущего шага:

$$\frac{K_n}{n^\alpha} \xrightarrow[\text{п.н.}]{n\to\infty} \zeta.$$

Или, что то же самое $\frac{N^*_{\tau_n}}{N^\alpha_{\tau_n}} \xrightarrow[\text{п.н.}]{n\to\infty} \zeta$. Но при $n \to \infty$ $\tau_{n+1} - \tau_n \sim \frac{1}{n} \to 0$. Поэтому предел $\frac{N^*_t}{N^\alpha_t}$ сходится к случайной величине ζ, имеющей распределение Миттаг-Леффлера с параметром α.

- G. Undy Yule *A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F.R.S.*, Phil. Trans. R. Soc. Lond. B 1925 213, 21-87
- W. Ewens *The sampling theory of selectively neutral alleles.*, Theor. Popul. Biol., 3:87 – 112, 1972.
- *D.M. Blei, T.L. Griffiths, M.I. Jordan*, The Nested Chinese Restaurant Process and Bayesian Nonparametric Inference of Topic Hierarchies, 2010.
- *H. Yamato, M. Sibuya*, Moments of some statistics of Pitman Sampling Formula, Bulletin of informatics and cybernetics. 32 (1), pp.1-10, 2000-06. Research Association of Statistical Sciences, 2000
- *J.Pitman*, *Combinatorial Stochastic Processes.*, Lecture Notes in Math. 1875. Springer, Berlin, 2006
- *J.Pitman*, *Exchangeable and partially exchangeable random partitions .*, Probab. Theory Retat. Fields 102, 145 - 158 (1995).

- *J.Pitman, Brownian motion, Bridge, Excursion and Meander characterizedby sampling at independent uniform times*, Electronic Journal of Probability, Vol. 4 (1999) Paper no. 11, pages 1-33.
- *D.G. Kendall, Branching processes since 1873*, 1966.
- *А. А. Апарин, Г. А. Попов, Е. Б. Яровая, О распределении времени пребывания случайного блуждания в точке многомерной решетки.*, Теория вероятн. и ее примен., 66:4 (2021), 657–675; Theory Probab. Appl., 66:4 (2022), 522–536
- *В. Феллер, Введение в теорию вероятностей и ее приложения, Том 1*, 1952.
- *B. Bercu, S. Favaro, A martingale approach to Gaussian fluctuations and laws of iterated logarithm for Ewens-Pitman model*, 2024.