

Seminar 1: Ветвящиеся процессы Гальтона-Ватсона

Carlos Buitrago

Table of contents

1	Средний размер поколения	3
2	Производящие функции	4
2.1	Основные свойства	4
2.2	Производящая функция размера поколения	5
3	Вероятность вырождения	5
4	Теорема о вероятности вырождения	7
5	Фазовый переход: теория против симуляции	8
6	Достаточно ли знать только среднее число потомков?	9
7	Общее число частиц	11
8	Время до вырождения	13
9	Интерпретации и приложения	15
9.1	Эпидемии	15
9.2	Распространение информации в сетях	15
9.3	Финансовые каскады	15
9.4	Machine Learning и Data Science	15

Пусть $\xi \in \mathbb{Z}_+$ — случайная величина, описывающая число потомков одной частицы.

Пусть $\{\xi_k^{(n)} : k, n \in \mathbb{N}\}$ — независимые случайные величины с тем же распределением, что и ξ .

Определим $X_0 = 1$, и для $n \geq 1$

$$X_n = \sum_{k=1}^{X_{n-1}} \xi_k^{(n)}.$$

Определение

Процесс $(X_n)_{n \geq 0}$ называется **ветвящимся процессом Гальтона–Ватсона** с законом размножения ξ .
Здесь

- X_n — число частиц в n -м поколении;
- $\xi_k^{(n)}$ — число потомков k -й частицы поколения $n - 1$.

Главный вопрос этой главы:

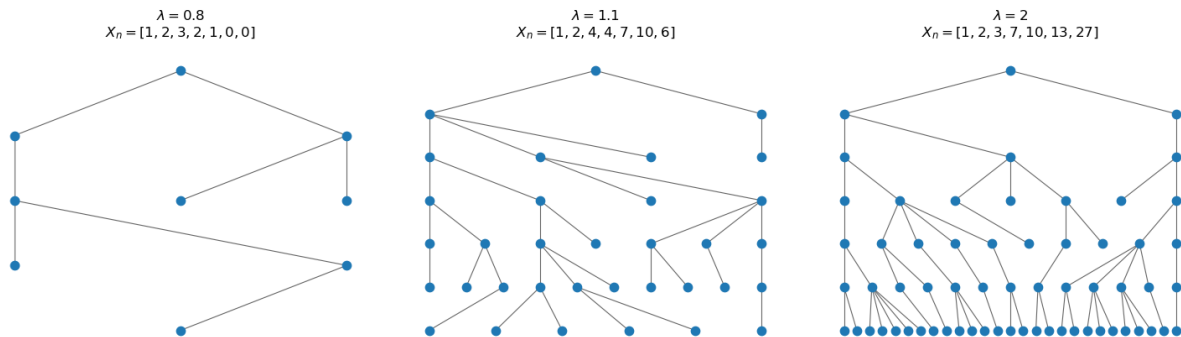
Какова вероятность того, что процесс когда-нибудь выродится?

💡 Как выглядит одна реализация процесса?

Для первого эксперимента возьмём $\xi \sim \text{Pois}(\lambda)$, поэтому $\mu = \mathbb{E}\xi = \lambda$.
Сравним три режима:

$$\lambda = 0.8, \quad \lambda = 1.1, \quad \lambda = 2.$$

Три реализации процесса Гальтона-Ватсона



💡 Три режима ветвящегося процесса

Главной характеристикой закона размножения является

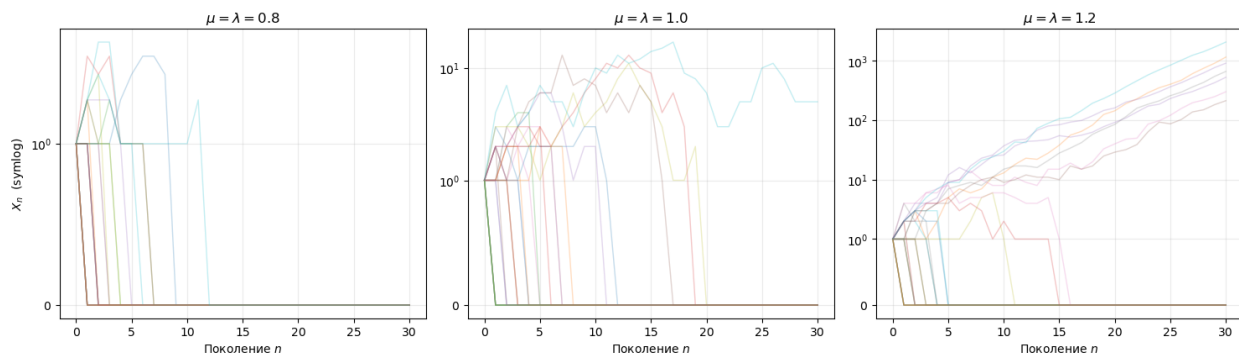
$$\mu = \mathbb{E}\xi.$$

Уже на уровне симуляций видно три качественно различных режима:

- $\mu < 1$ — **докритический**;
- $\mu = 1$ — **критический**;
- $\mu > 1$ — **надкритический**.

Особенно интересен критический случай: процесс может совершать большие случайные выбросы, хотя в дальнейшем всё равно вырождается с вероятностью 1.

Много независимых траекторий X_n



1 Средний размер поколения

Из определения

$$X_{n+1} = \sum_{k=1}^{X_n} \xi_k^{(n+1)}$$

получаем

$$\mathbb{E}(X_{n+1} \mid X_n) = \mu X_n, \quad \mu = \mathbb{E}\xi.$$

По формуле полного математического ожидания,

$$\mathbb{E}X_{n+1} = \mu \mathbb{E}X_n.$$

Так как $X_0 = 1$, то по индукции

$$\boxed{\mathbb{E}X_n = \mu^n.}$$

💡 Проверим формулу методом Монте-Карло

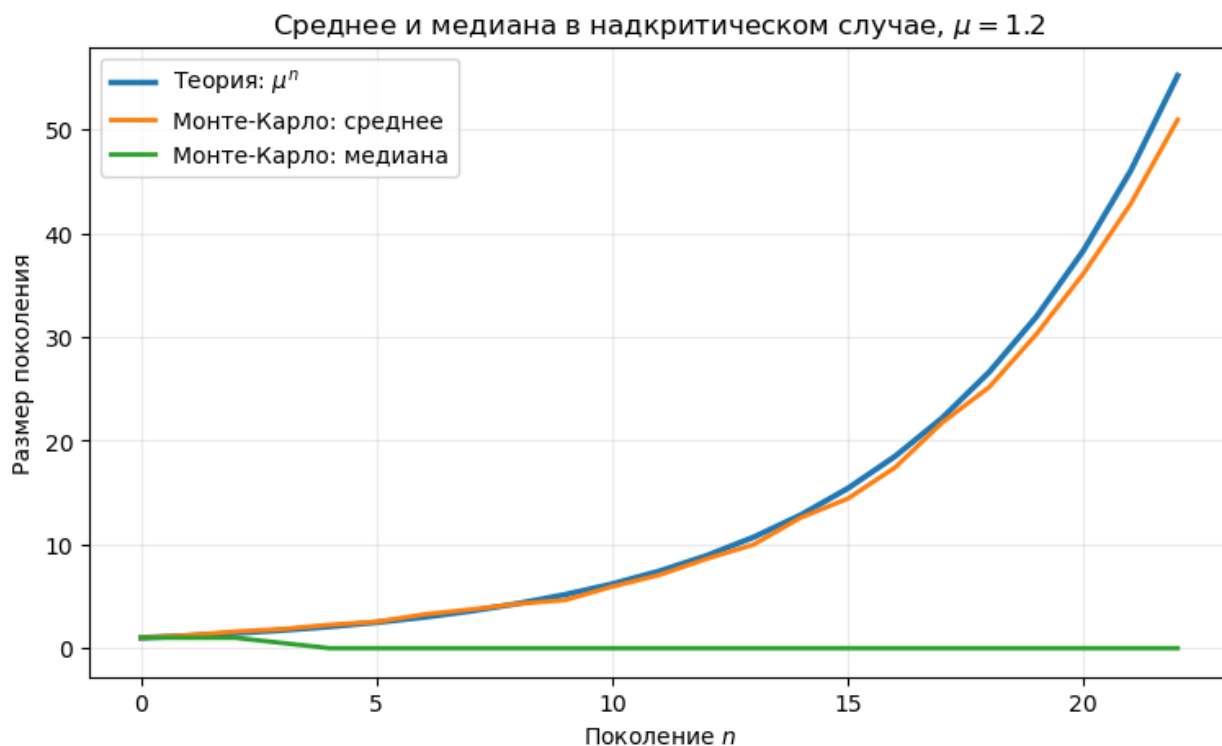
Сгенерируем N независимых реализаций процесса и вычислим

$$\widehat{\mathbb{E}X_n} = \frac{1}{N} \sum_{j=1}^N X_n^{(j)}.$$

Сравним это с

$$\mathbb{E}X_n = \mu^n.$$

Одновременно посмотрим на медиану X_n .



2 Производящие функции

Определение

Пусть ξ — случайная величина. Её **производящей функцией** называется

$$G_{\xi}(z) = \mathbb{E}z^{\xi}.$$

Для неотрицательной целочисленной случайной величины

$$G_{\xi}(z) = \sum_{k=0}^{\infty} z^k P(\xi = k), \quad |z| < 1.$$

2.1 Основные свойства

1.

$$G_{\xi}(1) = 1.$$

2. Если $\mathbb{E}\xi < \infty$, то

$$G'_{\xi}(1) = \mathbb{E}\xi.$$

3. Если ξ и η независимы, то

$$G_{\xi+\eta}(z) = G_{\xi}(z)G_{\eta}(z).$$

4. Для $\xi \in \mathbb{Z}_+$:

$$G_{\xi}(0) = P(\xi = 0).$$

5. Внутри единичного круга производящая функция бесконечно дифференцируема.
6. Вероятности можно восстановить по производной:

$$P(\xi = k) = \frac{1}{k!} G_\xi^{(k)}(0).$$

2.2 Производящая функция размера поколения

Пусть $(X_n)_{n \geq 0}$ — процесс Гальтона–Ватсона с законом размножения ξ .

Лемма

Для всех $n \geq 0$

$$G_{X_{n+1}}(z) = G_{X_n}(G_\xi(z)).$$

Доказательство

Зафиксируем $m \geq 0$. На событии $\{X_n = m\}$ имеем

$$X_{n+1} = \sum_{k=1}^m \xi_k^{(n+1)}.$$

Поэтому из независимости потомств различных частиц

$$\mathbb{E}(z^{X_{n+1}} \mid X_n = m) = \prod_{k=1}^m \mathbb{E} z^{\xi_k^{(n+1)}} = G_\xi(z)^m.$$

Тогда

$$\begin{aligned} G_{X_{n+1}}(z) &= \mathbb{E} z^{X_{n+1}} \\ &= \mathbb{E} [\mathbb{E}(z^{X_{n+1}} \mid X_n)] \\ &= \mathbb{E} [G_\xi(z)^{X_n}] \\ &= G_{X_n}(G_\xi(z)). \end{aligned}$$

Поскольку $X_0 = 1$, имеем $G_{X_0}(z) = z$.

Последовательно применяя предыдущую лемму,

$$G_{X_n}(z) = \underbrace{G_\xi \circ G_\xi \circ \dots \circ G_\xi}_{n \text{ раз}}(z).$$

В частности,

$$G_{X_{n+1}}(z) = G_\xi(G_{X_n}(z)).$$

3 Вероятность вырождения

Обозначим $q_n = P(X_n = 0)$ — вероятность того, что процесс выродился к моменту n .

Также обозначим $q = P(\exists n : X_n = 0)$ — вероятность того, что процесс когда-либо вырождается.

Поскольку из $X_n = 0$ следует $X_{n+1} = 0$,

$$\{X_n = 0\} \subseteq \{X_{n+1} = 0\},$$

и поэтому

$$q_n \leq q_{n+1}.$$

По непрерывности вероятностной меры,

$$q = \lim_{n \rightarrow \infty} q_n.$$

Лемма

Вероятность вырождения q является решением уравнения

$$s = G_\xi(s).$$

Доказательство

$$q \leftarrow q_n = \mathbb{P}(X_n = 0) = G_{X_n}(0) = G_\xi(G_{X_{n-1}}(0)) = G_\xi(q_{n-1}) \rightarrow G_\xi(q).$$

💡 Геометрия уравнения вырождения

Вероятность вырождения удовлетворяет

$$q = G_\xi(q).$$

Поэтому нас интересуют точки пересечения

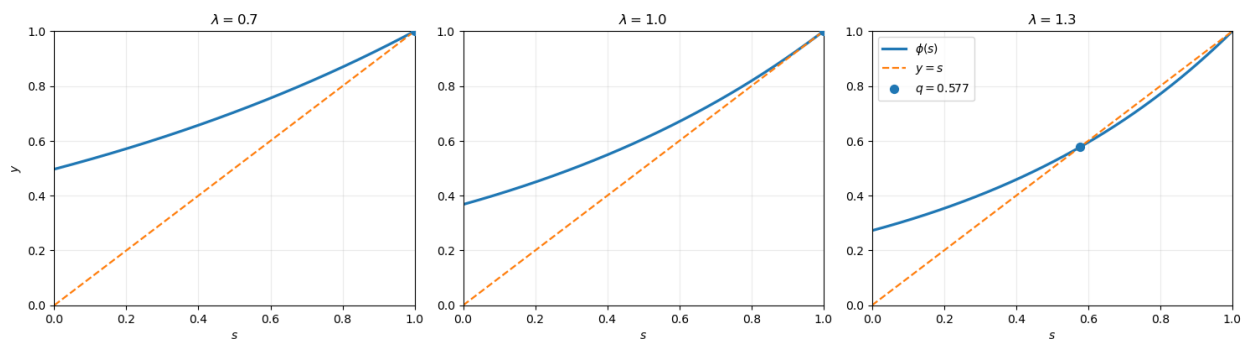
$$y = G_\xi(s) \quad \text{и} \quad y = s.$$

Для пуассоновского закона $G_\xi(s) = e^{\lambda(s-1)}$.

Сравним

$$\lambda = 0.7, \quad \lambda = 1, \quad \lambda = 1.3.$$

Геометрия уравнения $s = \phi(s)$



Вопрос: что делать, если на $[0, 1]$ решений несколько?

Т.к. $\varphi_\xi(1) = 1$, всегда есть решение $s = 1$.

4 Теорема о вероятности вырождения

Пусть $\mu = \mathbb{E}\xi$ и предположим, что

$$P(\xi = 1) \neq 1.$$

Теорема

1. Если $\mu \leq 1$, то уравнение $s = G_\xi(s)$ имеет на $[0, 1]$ единственное решение $s = 1$, и

$$q = 1.$$

2. Если $\mu > 1$, то существует единственный корень $s_0 \in [0, 1)$, и

$$q = s_0 < 1.$$

Иными словами,

$$q = \min\{s \in [0, 1] : G_\xi(s) = s\}.$$

💡 Как последовательность q_n находит вероятность вырождения?

Мы знаем, что

$$q_{n+1} = G_\xi(q_n), \quad q_0 = 0.$$

Это обычная итерация функции.

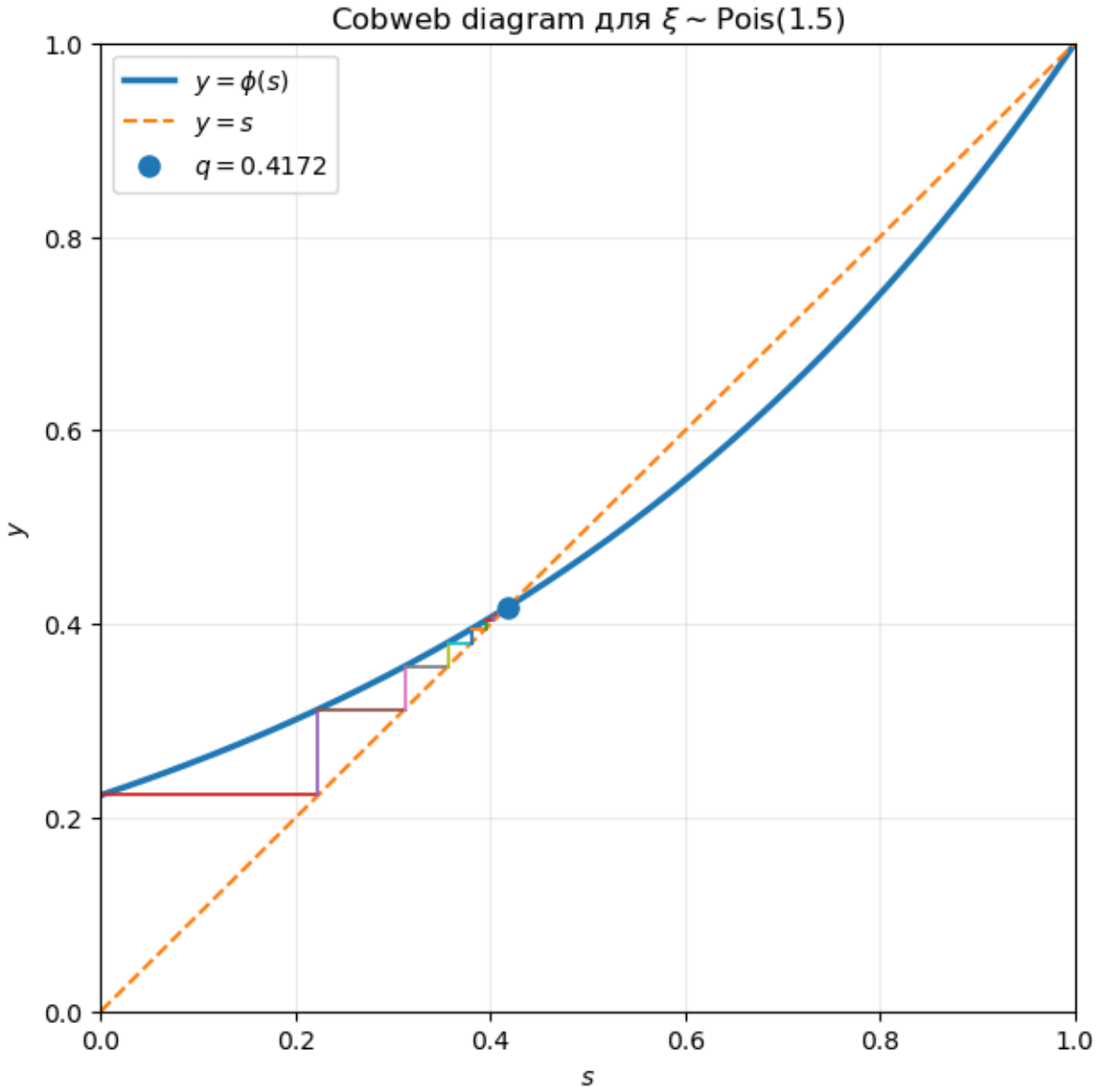
Её можно визуализировать с помощью **cobweb diagram**:

1. начинаем с $q_0 = 0$;
2. идём вертикально до $y = G_\xi(s)$;
3. идём горизонтально до $y = s$;
4. повторяем.

Последовательность

$$q_0, q_1, q_2, \dots$$

стремится к наименьшей неподвижной точке G_ξ .



5 Фазовый переход: теория против симуляции

Для $\xi \sim \text{Pois}(\lambda)$ имеем

$$G(s) = e^{\lambda(s-1)}.$$

Поэтому вероятность вырождения является наименьшим решением

$$q = e^{\lambda(q-1)}.$$

Теорема утверждает

$$q = 1, \quad \lambda \leq 1,$$

а при $\lambda > 1$

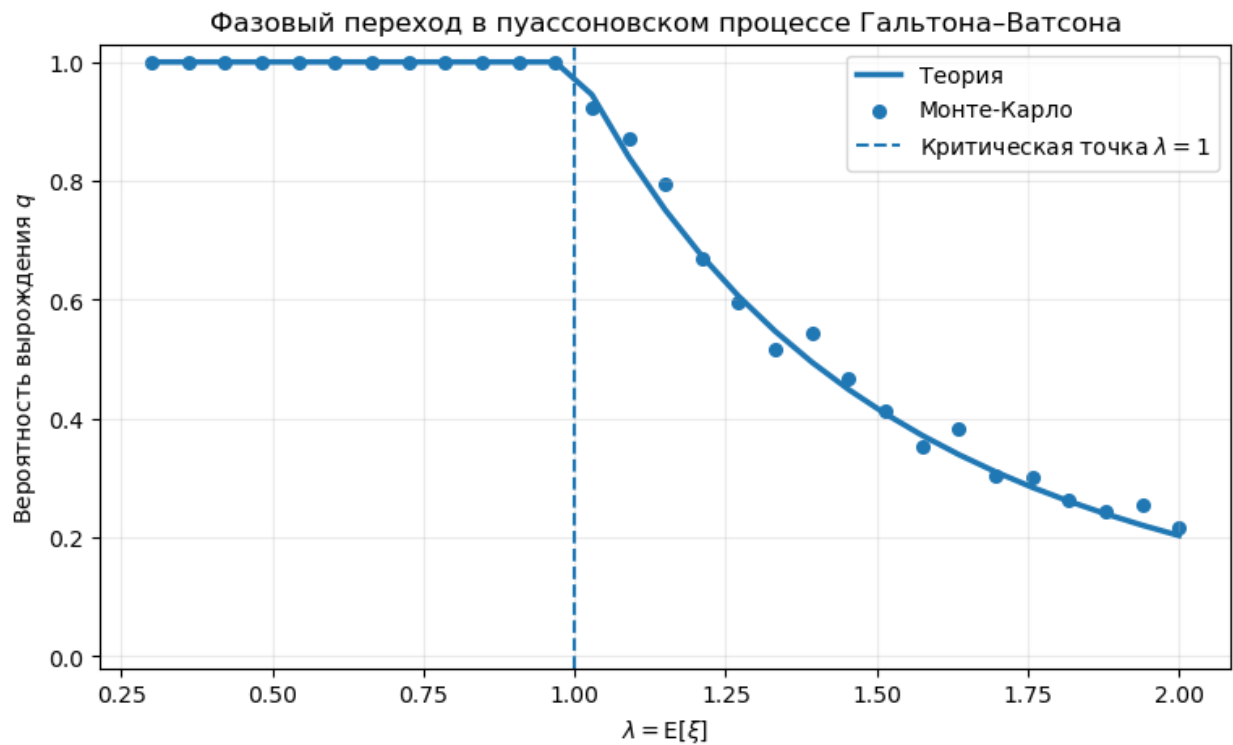
$$q < 1.$$

💡 Численный эксперимент

Для каждого λ :

1. найдём теоретическое значение q ;
2. сгенерируем много независимых процессов;
3. оценим вероятность вырождения методом Монте-Карло.

В надкритическом режиме в численной симуляции процесс будем считать практически выжившим, если его популяция достигла большого порога.



6 Достаточно ли знать только среднее число потомков?

Теорема показывает, что

$$\mu = E\xi$$

определяет, происходит ли вырождение с вероятностью 1:

$$\mu \leq 1 \implies q = 1.$$

Но при $\mu > 1$ само значение μ **не определяет** q .

Причина в том, что

$$q = \min\{s \in [0, 1] : G_{\xi}(s) = s\},$$

а производящая функция зависит от всего закона распределения ξ .

💡 Один и тот же $\mathbb{E}\xi$, но разные вероятности вырождения

Возьмём три закона с одинаковым средним

$$\mathbb{E}\xi = 1.2.$$

Первый:

$$\xi_1 \sim \text{Pois}(1.2).$$

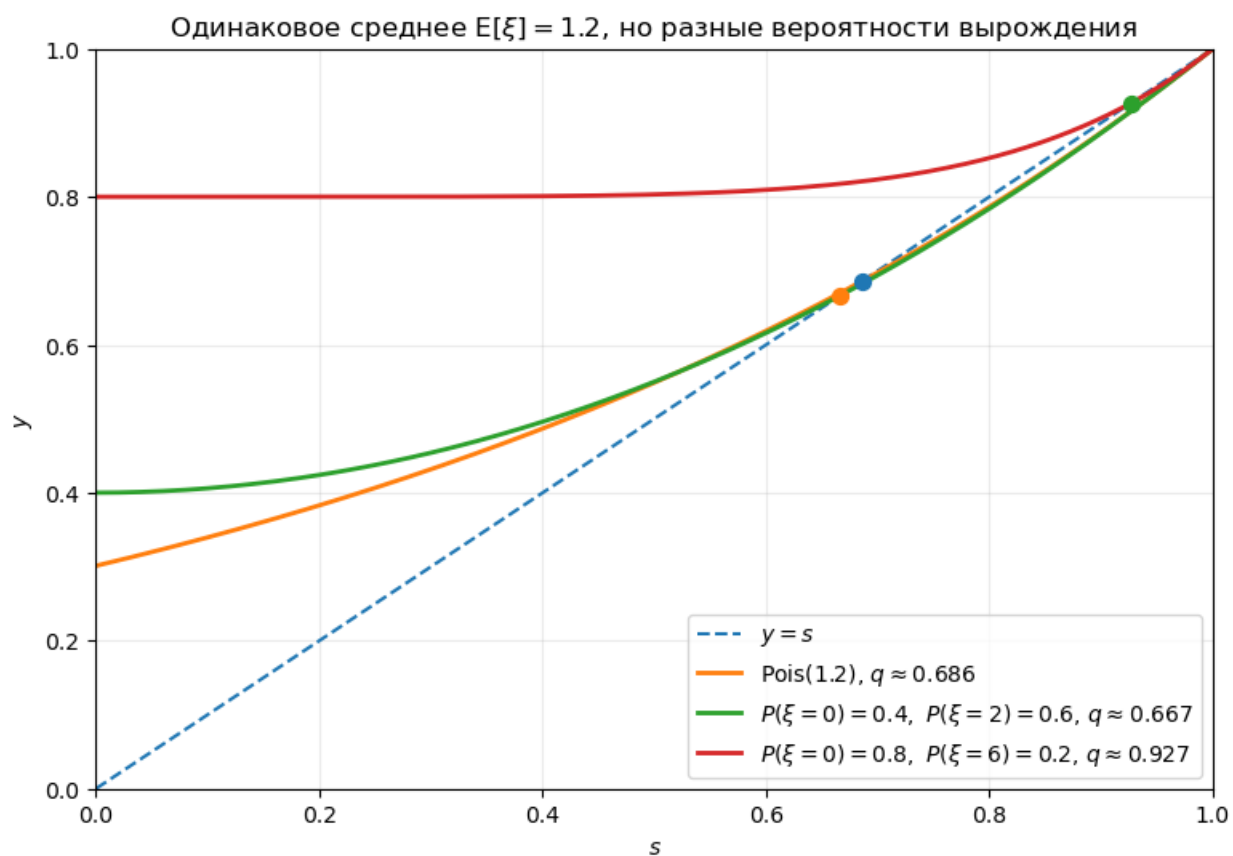
Второй:

$$\xi_2 = \begin{cases} 0, & P = 0.4, \\ 2, & P = 0.6. \end{cases}$$

Третий:

$$\xi_3 = \begin{cases} 0, & P = 0.8, \\ 6, & P = 0.2. \end{cases}$$

У них одинаковое среднее, но различные дисперсии и различные производящие функции.



Сравнение распределений:

$$\begin{aligned} & \$\mathrm{Pois}(1.2) \$ \quad E[xi] = 1.20, \quad \mathrm{Var}(xi) = 1.20, \quad q = 0.6863 \\ & \$P(\backslash xi=0)=0.4, \backslash P(\backslash xi=2)=0.6 \$ \quad E[xi] = 1.20, \quad \mathrm{Var}(xi) = 0.96, \quad q = 0.6667 \\ & \$P(\backslash xi=0)=0.8, \backslash P(\backslash xi=6)=0.2 \$ \quad E[xi] = 1.20, \quad \mathrm{Var}(xi) = 5.76, \quad q = 0.9265 \end{aligned}$$

7 Общее число частиц

Обозначим

$$Y_n = 1 + X_1 + \dots + X_n$$

— общее число частиц до момента n включительно.

Утверждение

Производящие функции удовлетворяют рекурсии

$$G_{Y_{n+1}}(z) = zG_\xi(G_{Y_n}(z)).$$

Доказательство

По определению производящей функции,

$$G_{Y_n}(z) = \mathbb{E}[z^{Y_n}] = \sum_{k=0}^{\infty} \mathbb{E}[z^{Y_n} \mathbf{1}_{\{X_1=k\}}].$$

Зафиксируем событие $\{X_1 = k\}$. Обозначим через A_i общее число потомков i -й частицы первого поколения до момента n . Тогда

$$Y_n = 1 + A_1 + \dots + A_k \quad \text{при условии } X_1 = k.$$

При этом

$$A_i \stackrel{d}{=} Y_{n-1}, \quad i = 1, \dots, k,$$

а случайные величины $A_1, \dots, A_k$ независимы и не зависят от X_1 .

Следовательно,

$$\begin{aligned} G_{Y_n}(z) &= \sum_{k=0}^{\infty} \mathbb{E}[z^{1+A_1+\dots+A_k} \mathbf{1}_{\{X_1=k\}}] \\ &= z \sum_{k=0}^{\infty} \mathbb{E}[z^{A_1}] \dots \mathbb{E}[z^{A_k}] \mathbb{P}(X_1 = k) \\ &= z \sum_{k=0}^{\infty} [G_{Y_{n-1}}(z)]^k \mathbb{P}(X_1 = k) \\ &= z G_\xi(G_{Y_{n-1}}(z)). \end{aligned}$$

Таким образом,

$$\boxed{G_{Y_n}(z) = z G_\xi(G_{Y_{n-1}}(z))}.$$

Утверждение

Пусть $\mathbb{P}(\xi = 0) \neq 1$. Тогда для всех $z \in (0, 1)$ выполнено

$$G_{Y_n}(z) < G_{Y_{n-1}}(z),$$

и, значит, существует предел

$$g(z) = \lim_{n \rightarrow \infty} G_{Y_n}(z).$$

Доказательство

Для $n = 1$ имеем

$$G_{Y_1}(z) = \mathbb{E}z^{1+X_1} = zG_\xi(z) < z = G_{Y_0}(z).$$

Шаг индукции. Предположим, что

$$G_{Y_{n-1}}(z) < G_{Y_{n-2}}(z).$$

Тогда, поскольку производящая функция G_ξ монотонно возрастает,

$$\begin{aligned} G_{Y_n}(z) &= zG_\xi(G_{Y_{n-1}}(z)) \\ &< zG_\xi(G_{Y_{n-2}}(z)) \\ &= G_{Y_{n-1}}(z). \end{aligned}$$

Следовательно, последовательность $\{G_{Y_n}(z)\}_{n \geq 0}$ убывает и ограничена снизу нулём, поэтому существует предел

$$g(z) = \lim_{n \rightarrow \infty} G_{Y_n}(z).$$

Утверждение

Функция $g(z)$ удовлетворяет уравнению

$$g(z) = zG_\xi(g(z)).$$

Доказательство

По предыдущему утверждению

$$G_{Y_n}(z) = zG_\xi(G_{Y_{n-1}}(z)).$$

Переходя к пределу при $n \rightarrow \infty$ и используя непрерывность производящей функции G_ξ , получаем

$$\lim_{n \rightarrow \infty} G_{Y_n}(z) = zG_\xi\left(\lim_{n \rightarrow \infty} G_{Y_{n-1}}(z)\right).$$

Таким образом,

$$g(z) = zG_\xi(g(z)).$$

💡 Насколько большим может оказаться вымерший процесс?

Даже если

$$q = 1,$$

общее число частиц

$$Y = \sum_{n=0}^{\infty} X_n$$

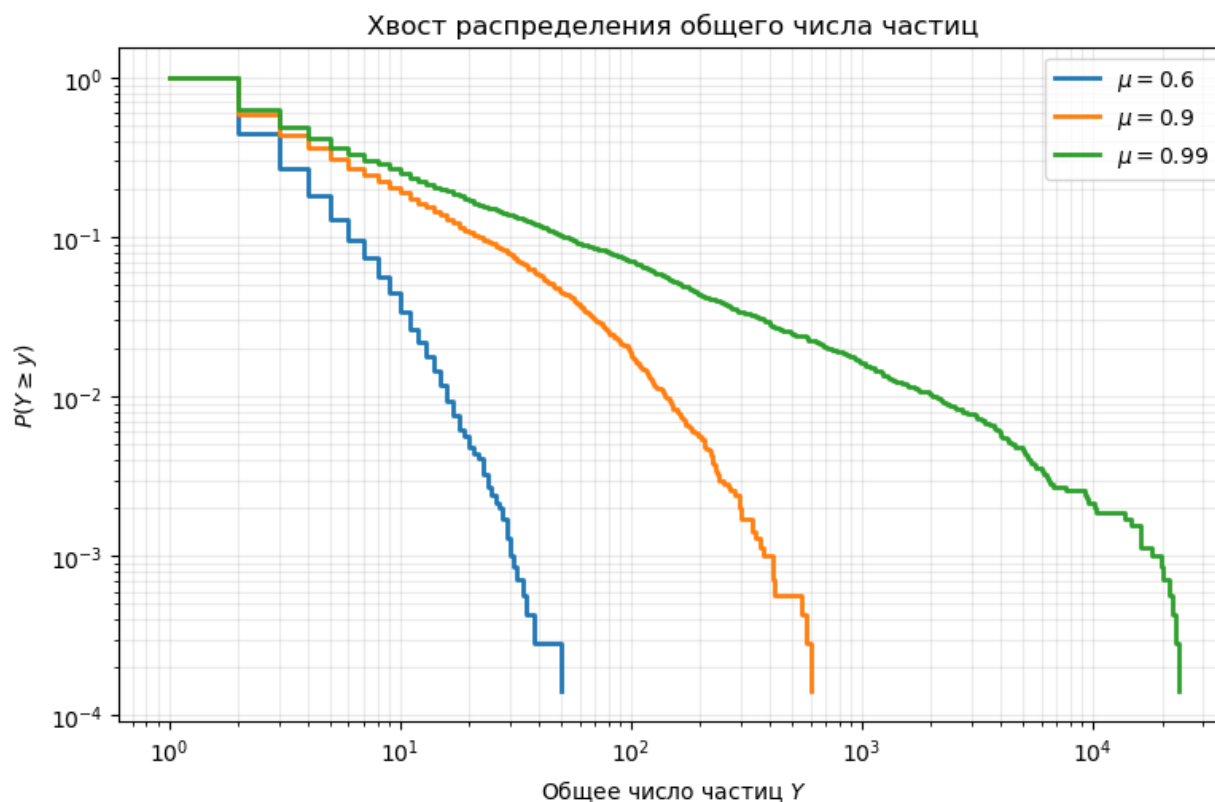
может оказаться очень большим.

Особенно интересно приближение к критическому режиму

$$\mu \uparrow 1.$$

Сравним

$$\mu = 0.6, \quad \mu = 0.9, \quad \mu = 0.99.$$



8 Время до вырождения

Определим

$$T = \inf\{n \geq 0 : X_n = 0\}.$$

Для докритического и критического процессов

$$P(T < \infty) = 1,$$

но распределение T зависит от близости к критической точке.

💡 Что происходит около критической точки?

Рассмотрим

$$\xi \sim \text{Pois}(\lambda)$$

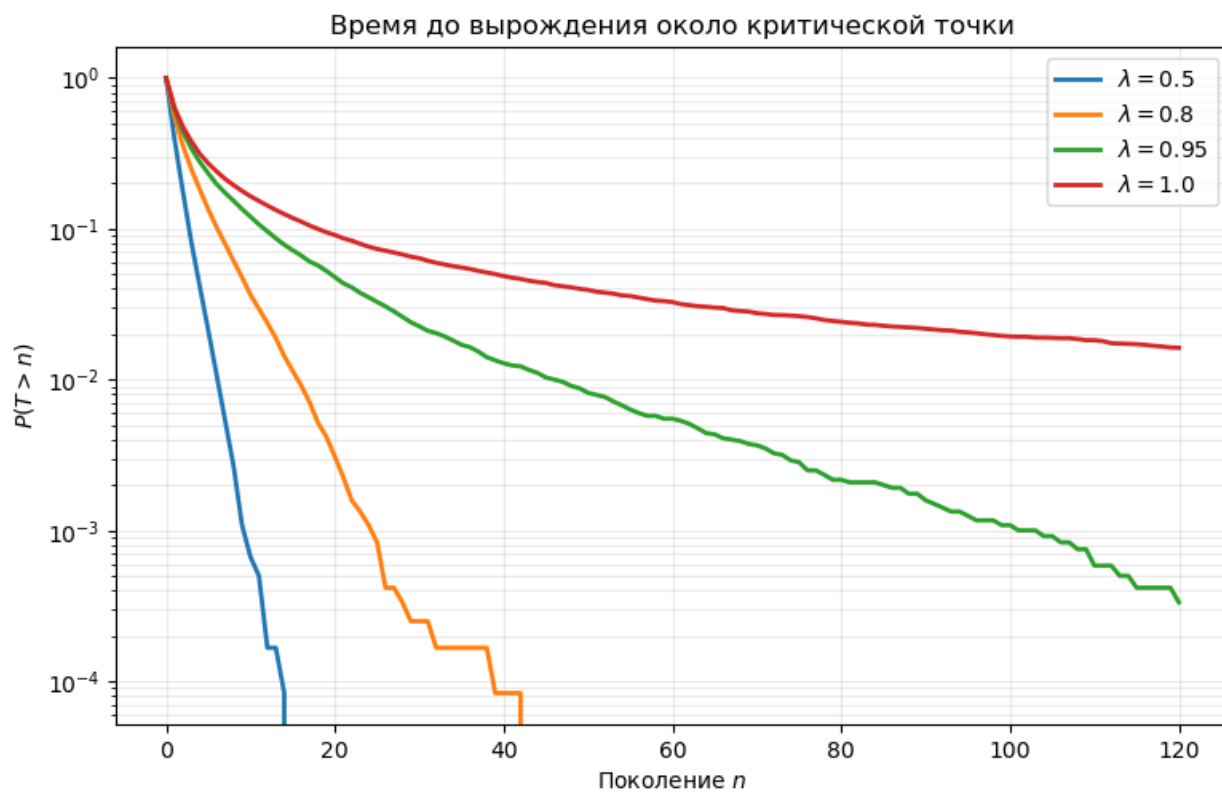
при

$$\lambda = 0.5, \quad 0.8, \quad 0.95, \quad 1.$$

Оценим

$$P(T > n).$$

Несмотря на почти верное вырождение во всех четырёх случаях, около $\lambda = 1$ процесс может существовать очень долго.



9 Интерпретации и приложения

i Где возникают ветвящиеся процессы?

Несмотря на простоту модели, процессы Гальтона–Ватсона возникают как естественное приближение во многих задачах.

9.1 Эпидемии

На ранней стадии эпидемии один заражённый человек заражает случайное число новых людей. Если обозначить это число через ξ , то

$$\mu = \mathbb{E}\xi$$

играет роль эффективного коэффициента воспроизводства.

При $\mu < 1$ цепочка заражений быстро прекращается, а при $\mu > 1$ существует положительная вероятность большой вспышки.

9.2 Распространение информации в сетях

Каждый пользователь может передать сообщение случайному числу новых пользователей.

На ранней стадии, пока повторные контакты и циклы редки, такая динамика естественно приближается ветвящимся процессом.

9.3 Финансовые каскады

В моделях финансового заражения дефолт одного участника может вызвать дефолты нескольких других участников.

Если сеть разрежена и каскад находится на ранней стадии, число новых дефолтов, вызванных одним дефолтом, можно приближённо описывать законом размножения ξ .

Условие

$$\mathbb{E}\xi > 1$$

указывает на возможность самоподдерживающегося каскада.

9.4 Machine Learning и Data Science

Ветвящиеся идеи возникают при изучении:

- каскадов распространения информации;
- случайных деревьев;
- локального исследования разреженных графов;
- случайных сетей;
- алгоритмов поиска, порождающих случайное число новых состояний.

Во многих таких моделях процесс Гальтона–Ватсона служит первым приближением локального поведения.