Ветвящиеся процессы

Владимир Алексеевич Ватутин

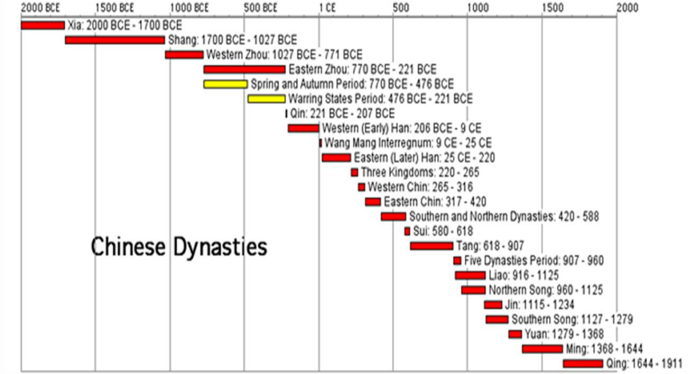
Математический институт им. В.А.Стеклова

В настоящее время официально принята следующая хронология Древнего Египта

Династии	Периоды	Годы
1-2	Раннее царство	3100-2686 гг. до н.э
3-6	Древнее царство	2686-2181 гг. до н.э.
7-10	1-й Переходный период	2181-2133 гг. до н.э.
11-12	Среднее Царство	2133-1786 гг. до н.э
13-17	2-й Переходный период	1786-1567 гг. до н.э.
18-20	Новое Царство	1567-1080 гг. до н.э
21-25	Позднее Новое Царство	1080-664 гг. до н.э.
26	Саисская династия	664-525 гг. до н.э
27-31	Позднее царство	525-332 гг. до н.э
	Птолемеи	332-30 гг. до н.э.
	Римляне	30 г. до н.э. — 642 г. н.э.
	Арабы	642 г. до н.э.— настоящее

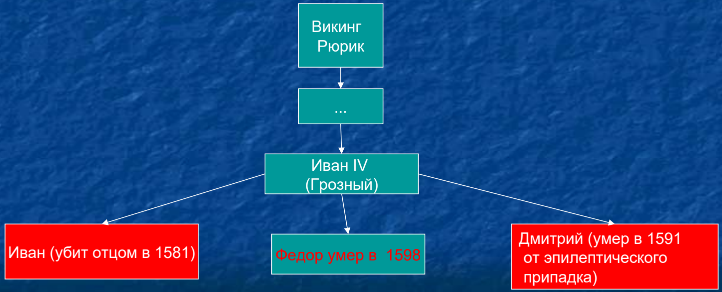
Считается что 390 египетских монархов принадлежали к тридцати одной четко определяемой династии.

CHINESE DYNASTIES



Иван Грозный убивает своего сына (Художник Илья Репин)

Династия Рюриков (8??-1598)



В конце 18 столетия Мальтус в книге

Malthus: Essay on the Principle of Population (1798)

писал о **геометрическом росте** населения Земли, но заметил, что при этом отдельные семейста вырождаются.

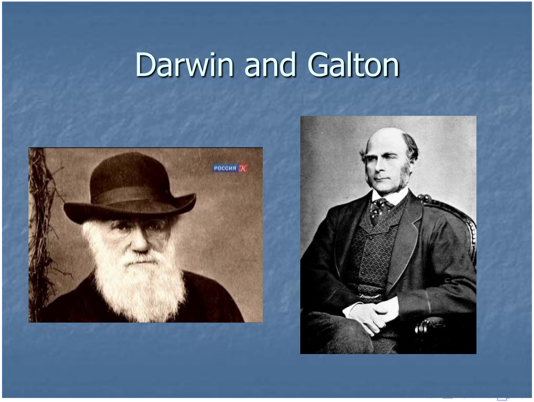
В качестве примера он привел город Берн (Швейцария), в котором в промежутке между **1583 и 1783 годами** 379 из 487 буржуазных семейств вымерли. Заметим, что

$$379/487 \approx 0,75(!)$$

Таким образом, примерно за 8 поколений исчезло 3/4 семейств!

Вопрос: вырождение отдельных семейств неизбежно или возможны другие сценарии?

Математические модели прпроцесса вырождения семейств были предложены независимо французским математиком **Бьенайме (Bienayme)** и английскими исследователями **Гальтоном (Galton) и Ватсоном (Watson)**.



Биенайме-Гальтон-Ватсон предложили следующую модель развития популяций по поколениям.

Имеется 1 частица с единичной продолжительностью жизни. Погибая, частица производит некоторое число потомков (может и ни одного). Новорожденные частицы имеют также единичную продолжительность жизни и, погибая, производят независимо друг от друга некоторое число потомков (может и ни одного) и так далее.

Какие вопросы интересны при исследовании эволюции популяций?

- Вероятность вырождения
- Если популяция не выродилась к моменту n, то сколько в ней индивидуумов.
- Каково расстояние до ближайшего общего предка всех частиц, существующих в популяции в далекий момент времени n.

Обозначим число возможных потомко частицы через ξ.

Будем предполагать, что с вероятностью p_0 частица в конце жизни не производит потомков, а с вероятностями $p_1, p_2, ..., p_n, ...$ производит $1, 2, ..., n, ...$ потомков, то есть,

$$p_k = \mathbf{P}\left(\xi = k\right), k = 0, 1, 2, ..., ...,$$

причем

$$p_0 + p_1 + p_2 + ..., +p_n + ... = 1.$$

Математическим ожиданием $\mathbf{E}\xi$ (или средним значением) величины ξ называется сумма

$$\mathbf{E}\xi = \sum_{k=0}^{\infty} k\mathbf{P}(\xi = k) = \sum_{k=0}^{\infty} kp_k.$$

Пример 0 Если частица в конце жизни либо **не отставляет потомства** с вероятностью $p_0 = \mathbf{P}(\xi = 0) = 1/4$, либо производит **одного потомка** с вероятностью $p_1 = \mathbf{P}(\xi = 1) = 1/4$, либо производит **двух потомков** с вероятностью $p_2 = \mathbf{P}(\xi = 2) = 1/2$, то

$$\begin{aligned} \mathbf{E}\xi &= 0 \times \mathbf{P}(\xi = 0) + 1 \times \mathbf{P}(\xi = 1) + 2 \times \mathbf{P}(\xi = 2) \\ &= 0 \times \frac{1}{4} + 1 \times \frac{1}{4} + 2 \times \frac{1}{2} = 1\frac{1}{4}. \end{aligned}$$

Свойства математического ожидания

$$\begin{aligned} \mathbf{E}(\xi_1 + \xi_2) &= \mathbf{E}\xi_1 + \mathbf{E}\xi_2, \\ \mathbf{E}(C\xi) &= C\mathbf{E}\xi. \end{aligned}$$

Целочисленные неотрицательные случайные величины ξ_1 и ξ_2 называются независимыми, если для любых неотрицательных целых чисел k_1 и k_2

$$\mathbf{P}\left(\xi_1 = k_1, \xi_2 = k_2\right) = \mathbf{P}\left(\xi_1 = k_1\right)\mathbf{P}\left(\xi_2 = k_2\right).$$

Если случайные величины ξ_1 и ξ_2 независимы, то

$$\mathbf{E}\left(\xi_1\xi_2\right) = \mathbf{E}\xi_1 \times \mathbf{E}\xi_2.$$

Ветвящиеся процессы Гальтона-Ватсона можно описать в терминах эволюции популяции частиц. В этой популяции первоначально имеется одна частица: $Z(0) = 1$. Эта частица имеет единичную продолжительность жизни. В конце жизни частица производит случайное число потомков ξ в соответствии с распределением

$$\mathbf{P}(\xi = k) = p_k,\ k = 0, 1, ...$$

Каждая из новорожденных частиц также имеет единичную продолжительность жизни и в конце ее производит (независимо от остальных частиц) случайное число потомков в соответствии с вероятностным распределением $\{p_k,\ k = 0, 1, ...\}$.

Таким образом, при $n \geq 0$

$$Z(n+1) = \xi_1^{(n)} + ... + \xi_{Z(n)}^{(n)},$$

где $\xi_i^{(n)}$ – число потомков i-й частицы n-го поколения $(i = 1, 2, ..., Z(n))$, причем $\xi_i^{(n)}$ при всех $i = 1, 2, \dots$ и $n = 0, 1, 2, \dots$ распределены как и ξ и независимы.

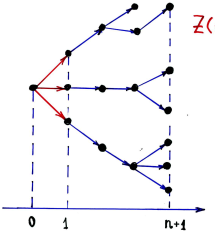
Справедливо и другое представление

$$Z(n+1) = Z_1(n) + ... + Z_{Z(1)}(n),$$

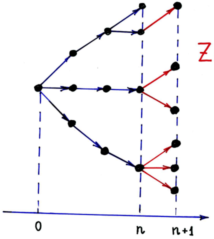
где

$$Z_i(n), i = 1, 2, ..., Z(1)$$

-число частиц в поколении $n+1$, которые являются потомками i-й частицы первого поколения.



$$Z(n+1)=Z_1(n)+\dots+Z_{Z(1)}(n)$$



$$Z(n+1)=\xi_1^{(n)}+\dots+\xi_{Z(n)}^{(n)}$$

$A := \mathbf{E}\xi$ – математическое ожидание числа потомков одной частицы.

Ветвящийся процесс Гальтона-Ватсона называется

- надкритическим, если $A > 1$,
- критическим, если $A = 1$,
- докритическим, если $A < 1$.

Математическое ожидание числа частиц в n-ом поколении Мы знаем, что

$$Z(n+1) = \xi_1^{(n)} + \ldots + \xi_{Z(n)}^{(n)},$$

Пусть $Z(0) = 1$. Тогда

$$\begin{aligned}
\mathbf{E}\left[Z(n+1)|Z(0)=1\right] &= \mathbf{E}\left[\xi_1^{(n)} + \ldots + \xi_{Z(n)}^{(n}|Z(0)=1\right] \\
&= \sum_{k=0}^{\infty} \mathbf{E}\left[\xi_1^{(n)} + \ldots + \xi_{Z(n)}^{(n}|Z(n)=k\right] \mathbf{P}\left(Z(n)=k|Z(0)=1\right) \\
&= \sum_{k=0}^{\infty} k\mathbf{E}\left[\xi_1^{(n)}|Z(n)=1\right] \mathbf{P}\left(Z(n)=k|Z(0)=1\right) \\
&= A\sum_{k=0}^{\infty} k\mathbf{P}\left(Z(n)=k|Z(0)=1\right) = A\mathbf{E}\left[Z(n)|Z(0)=1\right] \\
&= \ldots = A^n\mathbf{E}\left[Z(1)|Z(0)=1\right] = A^{n+1}.
\end{aligned}$$

Таким образом,

$$\mathbf{E}\left[Z(n)|Z(0)=1\right] = A^n.$$

Функция

$$f(s) := \mathbf{E}s^{\xi} = \sum_{k=0}^{\infty} \mathbf{P}(\xi = k)s^{k}.$$

называется **вероятностной производящей функцией** неотрицательной целочисленной случайной величины ξ.
Пример 1 Пусть

$$p_0 = \mathbf{P}\left(\xi = 0\right) = 1/4, p_1 = \mathbf{P}\left(\xi = 1\right) = 1/4, p_2 = \mathbf{P}\left(\xi = 2\right) = 1/2.$$

Тогда

$$\begin{aligned} f(s) := \mathbf{E}[s^{\xi}] &= \mathbf{P}\left(\xi = 0\right) s^{0} + \mathbf{P}\left(\xi = 1\right) s^{1} + \mathbf{P}\left(\xi = 2\right) s^{2} \\ &= \frac{1}{4} + \frac{1}{4}s + \frac{1}{2}s^{2}. \end{aligned}$$

Пример 2 Если

$$\mathbf{P}\,(\xi = k) = qp^k, k = 0, 1, \ldots \textbf{ где } p + q = 1, p, q > 0,$$

то

$$f(s) := \sum_{k=0}^{\infty} \mathbf{P}(\xi = k) s^k = \sum_{k=0}^{\infty} qp^k s^k = \frac{q}{1 - ps}.$$

- в дальнейшем мы будем использовать для такого распределения числа потомков термин **геометрическое распределение**.

Ясно, что

$$f'(s) = \sum_{k=0}^{\infty} k\mathbf{P}(\xi = k)s^{k-1}.$$

Следовательно,

$$f'(1) = \sum_{k=0}^{\infty} k\mathbf{P}(\xi = k) = \mathbf{E}\xi.$$

Элементарные свойства вероятностных производящих функций

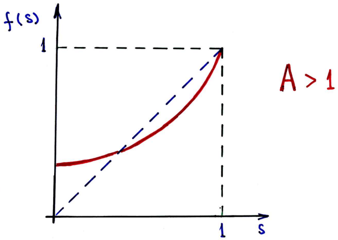
Пусть

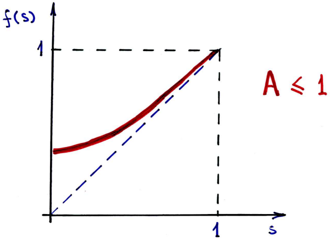
$$f(s) = \sum_{k=0}^{\infty} \mathbf{P}(\xi = k)\, s^k = \sum_{k=0}^{\infty} p_k s^k$$

– вероятностная производящая функция. В дальнейшем, чтобы исключить тривиальные случаи, будем предполагать, что $p_0 + p_1 < 1$. При этом условии

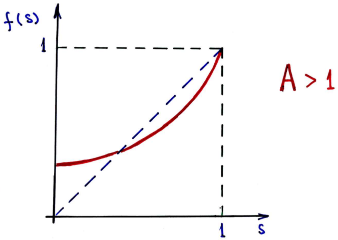
- $f(s)$ строго выпуклая функция на $[0, 1]: \quad f'(s) > 0,\ f''(s) > 0;$
- $f(0) = p_0 = \mathbf{P}(Z(1) = 0 | Z(0) = 1);$

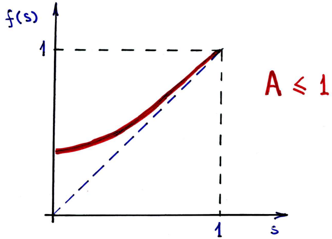
- Если $A \le 1$, то
$$f(s) > s, s \in [0,1).$$
Действительно, $f(1) = 1$ и
$$f'(s) - 1 = \sum_{k=0}^{\infty} k\mathbf{P}(\xi = k)s^{k-1} - 1 < 0$$
при $s \in [0,1)$. Следовательно, разность
$$f(s) - s, s \in [0,1],$$
достигает минимального значения в точке $s = 1$.


f(s)
1
1
s
A > 1


f(s)
1
1
s
A ≤ 1

- Если $A > 1$, то уравнение $f(s) = s$ имеет единственный корень r на полуинтервале $[0, 1)$, причем $f(s) > s$ при $s < r$ и $f(s) < s$ при $s > r$.
  Действительно, положим $\phi(s) = f(s) - s$. Тогда $\phi(1) = 0$ и
  $$\phi'(1) = f'(1) - 1 = A - 1 > 0.$$
  Поэтому $\phi(s) < 0$ в точках $s < 1$ достаточно близких к 1. С другой стороны, $\phi(0) \geq 0$. Значит есть точка s_0, в которой $\phi(s_0) = 0$. Двух таких точек быть не может, так как $\phi(s) = f(s) - s$ - выпуклая функция и не может пересечь отрезок [0,1] в трех точках.


f(s)
1
1
s
A > 1


f(s)
1
1
s
A ≤ 1

Итерациями производящей функции $f(s)$ называется последовательность

$$f_0(s) := s, \quad f_{n+1}(s) := f(f_n(s)),\, n \geq 0.$$

В биологических приложениях весьма часто используется производящая функция

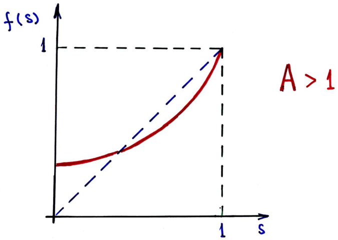
$$f(s) = q + ps^2, \quad q > 0, p > 0, q + p = 1,$$

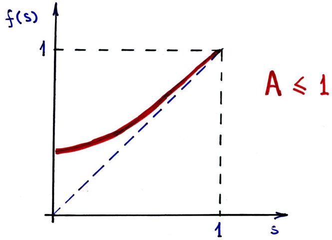
при помощи которой можно исследовать процесс деления клеток.

$$f_2(s) = f(f_1(s)) = q + pf_1^2(s) = q + p\left(q + ps^2\right)^2 = q + pq^2 + 2qp^2s^2 + p^3s^4$$

$$f_3(s) = f(f_2(s)) = q + pf_2^2(s) = q + p\left(q + pq^2 + 2qp^2s^2 + p^3s^4\right)^2$$

и так далее.


f(s)
1
1
s
A > 1


f(s)
1
1
s
A ≤ 1

Свойство: Если η_1 и η_2 независимые неотрицательные целочисленные случайные величины, то

$$\mathbf{E}[s^{\eta_1+\eta_2}] = \mathbf{E}[s^{\eta_1}]\mathbf{E}[s^{\eta_2}].$$

В терминах популяции производящую функцию можно записать в виде

$$f(s) = \mathbf{E}\left[s^{Z(1)}|Z(0) = 1\right].$$

Если же $Z(0) = k \geq 2$, то

$$\mathbf{E}\left[s^{Z(1)}|Z(0) = k\right] = \mathbf{E}\left[s^{Z_1(1)+\ldots+Z_k(1)}|Z(0) = k\right]$$

$$= \prod_{i=1}^{k} \mathbf{E}\left[s^{Z_i(1)}|Z_i(0) = 1\right] = f^k(s).$$

Полезность итераций производящих функций в теории ветвящихся процессов во многом определяется тем, что если $Z(0) = 1$ и

$$F(n,s) = \mathbf{E}s^{Z(n)} = \sum_{k=0}^{\infty} \mathbf{P}\left(Z(n) = k\right) s^k,$$

то

$$\begin{aligned}
F(n+1,s) := \quad & \mathbf{E}s^{Z(n+1)} = \sum_{k=0}^{\infty} \mathbf{E}\left[s^{Z(n+1)} \middle| Z(n) = k\right] \mathbf{P}\left(Z(n) = k\right) \\
= \quad & \sum_{k=0}^{\infty} \mathbf{E}\left[s^{Z(1)} \middle| Z(0) = k\right] \mathbf{P}\left(Z(n) = k\right) \\
= \quad & \sum_{k=0}^{\infty} \mathbf{P}\left(Z(n) = k\right) f^k(s) = F(n, f(s)) \\
= \quad & \mathbf{E}(f(s))^{Z(n)} = F(n, f(s)) = \ldots = f_{n+1}(s).
\end{aligned}$$

Математическое ожидание числа индивидуумов n-го поколения можно вычислить, опираясь на формулы

$$
\begin{aligned}
\mathbf{E}Z(n) &= \left(\mathbf{E}s^{Z(n)}\right)' |_{s=1} = (f_n(s))' |_{s=1} \\
&= f'(f_{n-1}(s)) \, (f_{n-1}(s))' |_{s=1} \\
&= \prod_{k=0}^{n-1} f'(f_k(s))|_{s=1} = (f'(1))^n = A^n.
\end{aligned}
$$

Вероятность вырождения

Как мы уже знаем,

$$f_n(s) = \mathbf{E} s^{Z(n)} = \sum_{k=0}^{\infty} \mathbf{P}\left(Z(n) = k\right) s^k.$$

Поэтому вероятность вырождения процесса к моменту n равна

$$\mathbf{P}(Z(n) = 0) = f_n(0).$$

Так как справедлива импликация $\{Z(n) = 0\} \Rightarrow \{Z(n+1) = 0\}$, то

$$f_n(0) = \mathbf{P}(Z(n) = 0) \leq \mathbf{P}(Z(n+1) = 0) = f_{n+1}(0).$$

Отсюда следует, что последовательность

$$P(n) = \mathbf{P}(Z(n) = 0) = f_n(0), \quad n = 1, 2 \ldots,$$

монотонно возрастая, стремится к вероятности вырождения процесса, которую мы обозначим буквой P:

$$\lim_{n \to \infty} P(n) = P.$$

Тогда

$$P(n) = f_n(0) = f(f_{n-1}(0)) = f(P(n-1)).$$

Переходя к пределу и пользуясь непрерывностью функции $f(s)$, получаем

$$P = f(P).$$

Отсюда $P = 1$, если $A \le 1$.

Если же $A > 1$, то $P = r$, где r - минимальный неотрицательный корень уравнения $s = f(s)$. Действительно, если

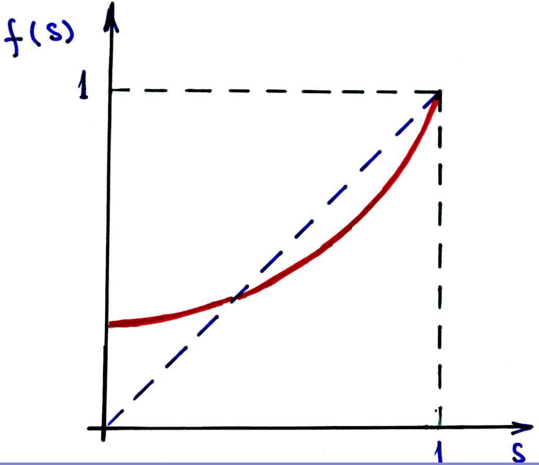
$$f(0) = \mathbf{P}\,(\xi = 0) = 0,$$

то $P = r = 0$. Если же $f(0) > 0$, то $\mathbf{P}(Z(n) = 0) \le r$. Действительно, если бы $\mathbf{P}(Z(n) = 0) > r$, то тогда

$$\mathbf{P}(Z(n+1) = 0) = f(\mathbf{P}(Z(n) = 0)) > f(r) = r$$

и, следовательно, было бы выполнено неравенство

$$\mathbf{P}(Z(n+1) = 0) = f(\mathbf{P}(Z(n) = 0)) < \mathbf{P}(Z(n) = 0).$$



A > 1

Таким образом, **докритические и критические процессы** вырождаются с вероятностью 1, в то время как вероятность вырождения **надкритического** процесса $P < 1$ и является наименьшим неотрицательным корнем уравнения

$$f(s) = s, s \in [0, 1).$$

Пример 3 Для бинарного расщепления частиц, т.е. в случае, когда в конце жизни частица либо умирает с вероятностью q, либо производит двух потомков с вероятностью p, для нахождения вероятности вырождения необходимо решить уравнение $q + px^2 = x$. В результате получаем $P = q/p$, если $p > 1/2$, и $P = 1$ в ином случае.

Пример 4 Для чисто геометрического закона распределения числа потомков решение уравнения

$$q/(1 - xp) = x$$

имеет вид: $P = q/p$ при $p > 1/2$, и $P = 1$ при $p \leq 1/2$. В этом случае $A = p/q$, и, следовательно, $P = 1/A$.

В частности, в процессе Гальтона–Ватсона с чисто геометрическим законом распределения и математическим ожиданием числа потомков $A = 1.2$ удвоение математического ожидания числа частиц происходит за четыре поколения: если $Z(0) = N$, то

$$\mathbf{E}[Z(4)] = N(1.2)^4 = N \times 2.07,$$

в то время как вероятность вырождения **отдельного семейства** больше 80 %

$$P = 1/A = 1/1.2 = 0.83,$$

если $Z(0) = 1$.

Вычисление итераций для чисто геометрического закона распределения числа потомков

Рассмотрим производящую функцию

$$f(s) = \sum_{k=0}^{\infty} qp^k s^k = \frac{1}{1 + A(1-s)},$$

где $A = p/q$. Ясно, что

$$f'(s) = \frac{A}{(1 + A(1-s))^2}, \text{ следовательно}, f'(1) = A = p/q.$$

Далее, несложно видеть, что

$$1 - f(s) = 1 - \frac{1}{1 + A(1-s)} = \frac{A(1-s)}{1 + A(1-s)}$$

и

$$\frac{1}{1 - f(s)} - \frac{1}{A(1-s)} = \frac{A(1-s) + 1}{A(1-s)} - \frac{1}{A(1-s)} = 1.$$

Таким образом, для итераций $f_n(s)$, $n = 1, 2, \ldots,$ справедлива цепочка равенств

$$\frac{1}{1 - f_n(s)} - \frac{1}{A(1 - f_{n-1}(s))} = \frac{1}{1 - f(f_{n-1}(s))} - \frac{1}{A(1 - f_{n-1}(s))} = 1,$$

то есть

$$\frac{1}{1 - f_n(s)} = 1 + \frac{1}{A(1 - f_{n-1}(s))} = 1 + \frac{1}{A} + \frac{1}{A^2(1 - f_{n-2}(s))} = \ldots.$$

Отсюда следует, что

$$\begin{aligned} \frac{1}{1 - f_n(s)} &= 1 + \frac{1}{A} + \frac{1}{A^2} + \ldots + \frac{1}{A^{n-1}} + \frac{1}{A^n(1 - s)} \\ &= \begin{cases} \frac{A^n - 1}{A^{n-1}(A-1)} + \frac{1}{A^n(1-s)}, & \text{если} \quad A \neq 1, \\ n + \frac{1}{1-s}, & \text{если} \quad A = 1. \end{cases} \end{aligned}$$

Таким образом, если $A \neq 1$, то

$$1 - f_n(s) = \frac{A^n(A-1)(1-s)}{A(A^n-1)(1-s) + A - 1}.$$

Если же $A = 1$, то

$$1 - f_n(s) = \frac{1}{n + (1-s)^{-1}}.$$

При помощи полученных равенств нетрудно вычислить вероятность выживания процесса по крайней мере в течение n поколений: если $A = p/q \neq 1$, то

$$\mathbf{P}\left(Z(n) > 0\right) = 1 - f_n(0) = \frac{A^n(A-1)}{A(A^n-1) + A - 1}.$$

Если же $A = 1$, то

$$\mathbf{P}\left(Z(n) > 0\right) = \frac{1}{n+1}.$$

В частности, если $A > 1$, то

$$\lim_{n\to\infty} \mathbf{P}\left(Z(n) > 0\right) = \lim_{n\to\infty} \frac{A^{n+1}(1-1/A)}{A^{n+1}-1} = 1 - \frac{1}{A},$$

а если $A < 1$, то при $n \to \infty$

$$\mathbf{P}\left(Z(n) > 0\right) \sim (1-A)A^n.$$

Теорема сходимости преобразований Лапласа

Преобразование Лапласа для распределения неотрицательной целочисленной случайной величины ζ

$$\varphi_\zeta(\lambda) = \sum_{k=0}^{\infty} \mathbf{P}(\zeta = k) e^{-\lambda k} = f_\zeta(e^{-\lambda}).$$

Если же распределение неотрицательной случайной величины ζ имеет плотность, то есть

$$\mathbf{P}(\zeta \leq x) = \int_0^x g_\zeta(y) dy,$$

где $g_\zeta(y) \geq 0,$ то по определению

$$\varphi_\zeta(\lambda) = \int_0^\infty e^{-\lambda y} g_\zeta(y) dy.$$

Theorem

Если последовательность неотрицательных случайных величин

$$\zeta_1, \zeta_2, \dots$$

такова, что для любого $\lambda \geq 0$

$$\lim_{n \to \infty} \varphi_{\zeta_n}(\lambda) = \varphi(\lambda),$$

и $\varphi(0) = 1$, *то*

$$\varphi(\lambda) = \varphi_{\zeta}(\lambda)$$

для некоторой случайной величины ζ, *причем*

$$\lim_{n \to \infty} \mathbf{P}(\zeta_n \leq x) = \mathbf{P}\left(\zeta \leq x\right)$$

во всех точках $x \geq 0$, *в которых функция*

$$G(x) = \mathbf{P}(\zeta \leq x)$$

непрерывна.

Предельная теорема для распределения числа частиц в надкритических процессах Гальтона–Ватсона

Theorem

Если $A > 1$ и $f''(1) < \infty$, то существует невырожденная случайная величина ζ такая, что

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z(n)}{A^n} \le x\right) = \mathbf{P}\left(\zeta \le x\right),$$

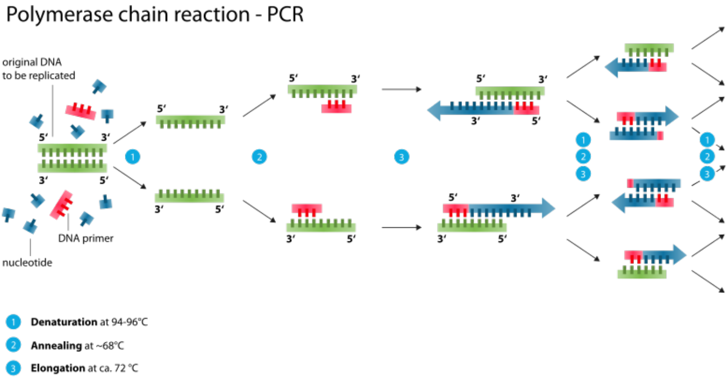
во всех точках $x \ge 0$, в которых функция

$$G(x) = \mathbf{P}(\zeta \le x)$$

непрерывна, причем

$$\mathbf{P}(\zeta = 0) = P = \mathbf{P}(Z(n) = 0 \text{ при некотором } n).$$

Важно для приложений: Уже столкнулись в своей жизни - Тест ПЦР.


Polymerase chain reaction - PCR
original DNA
to be replicated
5′
3′
3′
5′
DNA primer
nucleotide
1
2
3
1 Denaturation at 94-96°C
2 Annealing at ~68°C
3 Elongation at ca. 72 °C

ПЦР тест

Кери Муллис и Майкл Симит - Нобелевская премия по химии за 1993 год

Производящая функция

$$f(s) = p_0 + p_1 s + p_2 s^2.$$

ТОЛЬКО ГЕОМЕТРИЧЕСКОЕ РАСПРЕДЕЛЕНИЕ . Если

$$f(s) = \frac{q}{1 - ps} = \frac{1}{1 + A(1 - s)},$$

где $A = p/q > 1,$ то

$$f_n(s) = 1 - \frac{A^n(A-1)(1-s)}{A(A^n - 1)(1-s) + A - 1}$$

и при $n \to \infty$

$$f_n(0) = 1 - \frac{A^n(A-1)}{A(A^n - 1) + A - 1} \sim 1 - \frac{A-1}{A} = \frac{1}{A}.$$

$$f_n(s) = 1 - \frac{A^n(A-1)(1-s)}{A(A^n-1)(1-s)+A-1}.$$

Для

$$\zeta_n := \frac{Z(n)}{A^n}$$

имеем при $n \to \infty$

$$\begin{aligned}
\mathbf{E}e^{-\lambda\zeta_n} &= 1 - \frac{A^n(A-1)(1-e^{-\lambda A^{-n}})}{A(A^n-1)(1-e^{-\lambda A^{-n}})+A-1} \\
&\to 1 - \frac{\lambda(A-1)}{\lambda A + A - 1} \\
&= \frac{1}{A} + \left(1 - \frac{1}{A}\right)\frac{(1-\frac{1}{A})}{\lambda + 1 - \frac{1}{A}} = \mathbf{E}e^{-\lambda\zeta}.
\end{aligned}$$

Обращая преобразование Лапласа предельной величины, находим

$$\mathbf{P}\left(\zeta \le x\right) = \frac{1}{A} + (1 - \frac{1}{A})(1 - e^{-(1-\frac{1}{A})x}), \quad x \ge 0.$$

Условная предельная теорема для критических процессов

Theorem

Если $A = f'(1) = 1,\ f''(1) = 2B \in (0, \infty)$, *то*

$$Q(n) = \mathbf{P}\left(Z(n) > 0\right) \sim \frac{1}{Bn},\ n \to \infty,$$

и для любого $\lambda \in [0, \infty)$

$$\lim_{n\to\infty} \mathbf{E}\left[\exp\left\{-\lambda \frac{Z(n)}{Bn}\right\} \Big| Z(n) > 0\right] = \frac{1}{1+\lambda}.$$

Замечание Поскольку

$$\frac{1}{1+\lambda} = \int_0^\infty e^{-\lambda x} e^{-x} dx,$$

то

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z(n)}{Bn} \le y \Big| Z(n) > 0\right) = \int_0^y e^{-x} dx = 1 - e^{-y}.$$

Теорему об асимптотике вероятности невырождения опубликовал **А.Н.Колмогоров в 1938 году**, в Трудах Томского университета **на немецком языке!**


СССР в 1945–1991 гг.
СЕВЕРНЫЙ ЛЕДОВИТЫЙ ОКЕАН
о. Врангеля
Земля Франца-Иосифа
Северная Земля
Новосибирские о-ва
БАРЕНЦЕВО МОРЕ
Новая Земля
КАРСКОЕ МОРЕ
МОРЕ ЛАПТЕВЫХ
ВОСТОЧНО-СИБИРСКОЕ МОРЕ
БЕРИНГОВО МОРЕ
Анадырь
Командорские о-ва (СССР)
Петропавловск-Камчатский
Магадан
Охотск
ОХОТСКОЕ МОРЕ
о. Сахалин
Курильские о-ва (СССР)
Хабаровск
Владивосток
ЯПОНСКОЕ МОРЕ
ТОКИО
КОРЕЯ
Якутск
Диксон
Дудинка
Нарьян-Мар
Салехард
Архангельск
Петрозаводск
Ленинград
Сыктывкар
МОСКВА
Горький
Йошкар-Ола
Казань
Ижевск
Саратов
Куйбышев
Уфа
Краснодар
Астрахань
Омск
Томск
Новосибирск
Красноярск
Иркутск
Чита
Улан-Удэ
оз. Байкал
РОССИЙСКАЯ СОВЕТСКАЯ ФЕДЕРАТИВНАЯ СОЦИАЛИСТИЧЕСКАЯ РЕСПУБЛИКА
ФРГ
ГДР
ДАНИЯ
ШВЕЦИЯ
НОРВЕГИЯ
ФИНЛЯНДИЯ
БЕРЛИН
КОПЕНГАГЕН
СТОКГОЛЬМ
ХЕЛЬСИНКИ
ПРАГА
ПОЛЬША
ВАРШАВА
Калининград
ТАЛЛИН
РИГА
ВИЛЬНЮС
МИНСК
БЕЛОРУССКАЯ ССР
УКРАИНСКАЯ ССР
КИЕВ
КИШИНЕВ
Симферополь
ЧЕРНОЕ МОРЕ
ТУРЦИЯ
ТБИЛИСИ
ЕРЕВАН
БАКУ
КАСПИЙСКОЕ МОРЕ
ТЕГЕРАН
ИРАН
ТУРКМЕНСКАЯ ССР
АШХАБАД
УЗБЕКСКАЯ ССР
АРАЛЬСКОЕ МОРЕ
КАЗАХСКАЯ ССР
Семипалатинск
оз. Балхаш
ТАШКЕНТ
ДУШАНБЕ
ТАДЖИКСКАЯ ССР
ФРУНЗЕ
КИРГИЗСКАЯ ССР
АЛМА-АТА
АФГАНИСТАН
КИТАЙ
МОНГОЛИЯ
УЛАН-БАТОР
Цифрами на карте обозначены:
1 Эстонская ССР
2 Латвийская ССР
3 Литовская ССР
4 Молдавская ССР
5 Грузинская ССР
6 Армянская ССР
7 Азербайджанская ССР
YOURSITE.RU

Только геометрическое распределение

$$f(s) = \frac{q}{1-ps} = \frac{1}{2-s}, \quad p = q = \frac{1}{2}.$$

Тогда

$$f'(s) = \frac{1}{(2-s)^2}, \quad f''(s) = \frac{2}{(2-s)^3}.$$

Поэтому

$$f'(1) = 1, \quad f''(1) = 2 = 2B, \quad \text{значит} \quad B = 1.$$

Далее,

$$1 - f_n(s) = \left(n + \frac{1}{1-s}\right)^{-1} = \frac{1-s}{n(1-s)+1}$$

и

$$\mathbf{P}\left(Z(n) > 0\right) = 1 - f_n(0) = \frac{1}{n+1} \sim \frac{1}{Bn}.$$

Отсюда

$$\mathbf{E}\left[s^{Z(n)}|Z(n)>0\right] = \frac{\mathbf{E}\left[s^{Z(n)};Z(n)>0\right]}{\mathbf{P}\left(Z(n)>0\right)}$$
$$= \frac{\mathbf{E}\left[s^{Z(n)}\right]-\mathbf{E}\left[s^{Z(n)};Z(n)=0\right]}{\mathbf{P}\left(Z(n)>0\right)}$$
$$= \frac{f_n(s)-f_n(0)}{1-f_n(0)}$$
$$= 1-\frac{1-f_n(s)}{1-f_n(0)} = 1-\frac{(1-s)\,(n+1)}{n(1-s)+1}.$$

и, следовательно,

$$\lim_{n\to\infty}\mathbf{E}\left[e^{-\lambda Z(n)/n}|Z(n)>0\right] = 1-\lim_{n\to\infty}\frac{\left(1-e^{-\lambda/n}\right)(n+1))}{n(1-e^{-\lambda/n})+1}$$
$$= 1-\frac{\lambda}{\lambda+1} = \frac{1}{\lambda+1}.$$

Вероятность невырождения докритических процессов Мы знаем, что вероятность вырождения докритических процессов равна 1. А как быстро стремится к нулю вероятность невырождения процесса $\mathbf{P}(Z(n) > 0)$ при $n \to \infty$?
При $A < 1$ имеем

$$\begin{aligned} Q(n) \quad : &= \mathbf{P}(Z(n) \geq 1 | Z(0) = 1) \\ &= \sum_{k=1}^{\infty} \mathbf{P}(Z(n) = k | Z(0) = 1) \\ &\leq \sum_{k=1}^{\infty} k\mathbf{P}(Z(n) = k | Z(0) = 1) = \mathbf{E}Z(n) = A^n. \end{aligned}$$

Таким образом, вероятность невырождения в докритических процессах Гальтона-Ватсона убывает по крайней мере с экспоненциальной скоростью. Но сколь точна эта оценка (хотя бы по порядку)?

Исчерпывающий ответ на этот вопрос дан в следующей теореме. Пусть $\ln^+ x = \ln x$ при $x > 1$ и $\ln^+ x = 0$ при $x \le 1$.

Theorem

Если $A < 1$, *то*

$$Q(n) \sim KA^n(1 + o(1)), \quad K > 0,$$

тогда и только тогда, когда

$$\mathbf{E}\xi \ln^+ \xi = \sum_{k=1}^{\infty} p_k k \ln k < \infty.$$

Условная предельная теорема для распределения числа частиц в докритических процессах

Поскольку докритические ветвящиеся процессы вырождаются с вероятностью 1, изучение свойств таких процессов в момент n естественно проводить при условии $Z(n) > 0$, т.е. при условии невырождения процесса к моменту n.

Theorem

Если $A < 1$, *то для любого* $k = 1, 2, \ldots$

$$\lim_{n\to\infty} \mathbf{P}(Z(n) = k | Z(n) > 0) = P_k^*, \quad \sum_{k=1}^{\infty} P_k^* = 1.$$

Для геометрического распределения

Для докритических ветвящихся процессов со средним $A = p/q$ и геометрическим распределением числа частиц

$$f(s) = \frac{q}{1-ps} = \frac{1}{1+A(1-s)}, \quad p < q,$$

имеем

$$\begin{aligned} 1 - f_n(s) &= \frac{A^n(A-1)(1-s)}{A(A^n-1)(1-s)+A-1} \\ &\sim \frac{A^n(A-1)(1-s)}{-A(1-s)+A-1} = \frac{A^n(A-1)(1-s)}{-1+As} \end{aligned}$$

и

$$\mathbf{P}\left(Z(n) > 0\right) = 1 - f_n(0) \sim \frac{A^n(A-1)}{-1}.$$

Отсюда

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{E}\left[s^{Z(n)}|Z(n)>0\right] &= \lim_{n\to\infty}\left[1-\frac{1-f_n(s)}{1-f_n(0)}\right] \\
&= 1-\frac{1-s}{1-As} = \frac{s(1-A)}{1-As} = \frac{s(q-p)}{q-ps} \\
&= \frac{s(q-p)}{q-p+p(1-s)} = \frac{s}{1+A^*(1-s)} = f^*(s),
\end{aligned}$$

где $A^* = p/(q-p)$. Таким образом, в пределе у нас снова будет геометрическое распределение, но не на множестве $\{0,1,2,...\}$, а на множестве $\{1,2,3,...\}$.

Северный гладкий кит

Вырождение северных гладких китов

- Среднее число особей самок = A
- Нынешний размер популяции самок - примерно 150
- (Данные из статьи Caswell H., Fujimagara, M., Brault S.,1999)

	A<1 (!!!)	0.988	0.976
Вероятность выживания >0.99 не менее n лет	n>	357	177
Вероятность вырождения >0.99 в течение n лет	n<	796	395

Редуцированные процессы

Редуцированные процессы Пусть

$$\{Z(k),\ k \geq 0\}$$

– ветвящийся процесс Гальтона-Ватсона, а $Z(m,n)$– число частиц в этом процессе в момент $m \leq n$, которые имеют **непустое потомство** в момент n. Процесс

$$\{Z(m,n), 0 \leq m \leq n\}$$

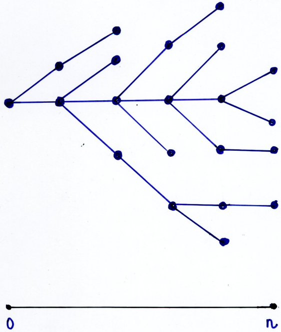
называется редуцированным процессом, построенным по набору $Z(k), k = 0, 1, ..., n$.

Для редуцированного ветвящегося процесса момент

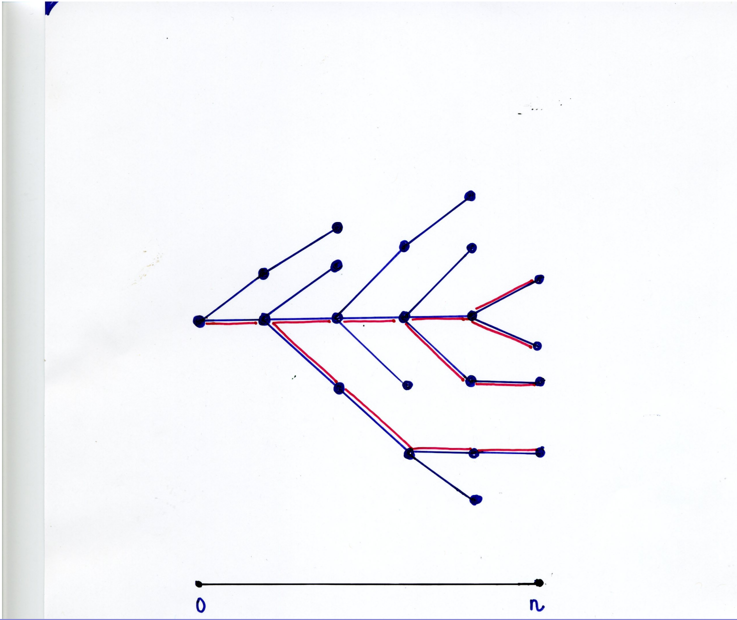
$$\tau = \max\{0 \leq k < n : Z(k,n) = 1\}$$

называется моментом рождения ближайшего общего предка всех частиц, существующих в момент n. Очевидно, что

$$\{\tau \geq m\} = \{Z(m,n) = 1\}.$$



Full! Tree


0
n

Для редуцированного ветвящегося процесса момент

$$\tau = \max\{0 \leq k < n : Z(k,n) = 1\}$$

называется моментом рождения ближайшего общего предка всех частиц, существующих в момент n. Очевидно, что

$$\{\tau \geq m\} = \{Z(m,n) = 1\}.$$

Редуцированные надкритические процессы

Поскольку надкритический процесс с положительной вероятностью существует бесконечно долго, то естественно ожидать, что в редуцированном надкритическом процессе момент рождения ближайшего общего предка всех частиц, существующих в момент n (при условии невырождения процесса к моменту n) находится в начале эволюции процесса.

Theorem

Если $A > 1$, *то для при любом* $m = 0, 1, \dots$

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(m,n)} | Z(n) > 0\right] = \frac{f_m(P + (1-P)s) - P}{1-P},$$

где величина $P < 1$ *есть вероятность вырождения рассматриваемого процесса.*

Доказательство. Поскольку при $m \leq n$

$$\mathbf{E}\left[s^{Z(m,n)}; Z(n)=0\right] = \mathbf{P}\left(Z(n)=0\right) = f_n(0)$$

и

$$\begin{aligned} \mathbf{E}\left[s^{Z(m,n)}\right] &= \mathbf{E}\left[\mathbf{E}\left[s^{Z(m,n)}|Z(m)\right]\right] \\ &= \mathbf{E}\left[(f_{n-m}(0)+(1-f_{n-m}(0))s)^{Z(m)}\right] \\ &= f_m\left(f_{n-m}(0)+(1-f_{n-m}(0))s\right), \end{aligned}$$

то

$$\begin{aligned} \mathbf{E}\left[s^{Z(m,n)}|Z(n)>0\right] &= \frac{\mathbf{E}\left[s^{Z(m,n)}; Z(n)>0\right]}{\mathbf{P}\left(Z(n)>0\right)} \\ &= \frac{\mathbf{E}\left[s^{Z(m,n)}\right]-\mathbf{E}\left[s^{Z(m,n)}; Z(n)=0\right]}{\mathbf{P}\left(Z(n)>0\right)} \\ &= \frac{f_m\left(f_{n-m}(0)+(1-f_{n-m}(0))s\right)-f_n(0)}{1-f_n(0)}. \end{aligned}$$

Отсюда, переходя к пределу при $n\to\infty$, используя непрерывность $f_m(y)$ по $y\in[0,1]$ и вспоминая, что $\lim_{n\to\infty} f_n(0)=P$, следует утвержение теоремы.

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(m,n)} | Z(n) > 0\right] = \frac{f_m(P + (1-P)s) - P}{1-P},$$

Момент рождения ближайшего общего предка

$$\mathbf{P}\left\{\tau \geq m | Z(n) > 0\right\} = \mathbf{P}\left\{Z(m,n) = 1 | Z(n) > 0\right\}.$$

Поэтому

$$\lim_{n\to\infty} \mathbf{P}\left\{\tau \geq m | Z(n) > 0\right\} = \lim_{n\to\infty} \mathbf{P}\left\{Z(m,n) = 1 | Z(n) > 0\right\}.$$

Имеем

$$\left(\frac{f_m(P + (1-P)s) - P}{1-P}\right)' = f_m'(P + (1-P)s).$$

И при $s = 0$ имеем

$$f_m'(P) = f'(f_{m-1}(P)) f_{m-1}'(P)) = f'(P) f_{m-1}'(P) = \left(f'(P)\right)^m.$$

Редуцированные критические процессы

Для $t \in [0,1]$ будем интерпретировать произведение nt как целое число $[nt]$.

Theorem

Если

$$A = f^{'}(1) = 1, \; f''(1) = 2B \in (0, \infty),$$

то

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(nt,n)} \,|\, Z(n) > 0\right] = \frac{s(1-t)}{1-st}$$

и, следовательно, для любого $k = 1, 2, \ldots$

$$\lim_{n\to\infty} \mathbf{P}\left(Z(nt,n) = k \,|\, Z(n) > 0\right) = (1-t)t^{k-1}.$$

Геометрический случай Мы знаем, что

$$1-f_n(s)=\left(n+\frac{1}{1-s}\right)^{-1}, \qquad 1-f_n(s)=\frac{1}{n+1}.$$

Далее,

$$\begin{aligned}
\mathbf{E}\left[s^{Z(m,n)}|Z(n)>0\right] &= \frac{\mathbf{E}\left[s^{Z(m,n)};Z(n)>0\right]}{\mathbf{P}\left(Z(n)>0\right)} \\
&= \frac{\mathbf{E}\left[s^{Z(m,n)}\right]-\mathbf{E}\left[s^{Z(m,n)};Z(n)=0\right]}{\mathbf{P}\left(Z(n)>0\right)} \\
&= \frac{f_m(f_{n-m}(0)+(1-f_{n-m}(0))s)-f_n(0)}{1-f_n(0)} \\
&= 1-\frac{1-f_m(f_{n-m}(0)+(1-f_{n-m}(0))s)}{1-f_n(0)}.
\end{aligned}$$

Имеем

$$\begin{aligned}
&1-f_m(f_{n-m}(0)+(1-f_{n-m}(0))s) \\
&= \left(m+\frac{1}{1-f_{n-m}(0)-(1-f_{n-m}(0))s}\right)^{-1}
\end{aligned}$$

Теперь при $m = nt$ имеем

$$1 - f_m(f_{n-m}(0) + (1 - f_{n-m}(0))s) \sim \frac{1-s}{n(1-ts)}.$$

Отсюда,

$$\frac{1 - f_m(f_{n-m}(0) + (1 - f_{n-m}(0))s)}{1 - f_n(0)} \sim \frac{1-s}{1-ts}$$

и

$$\mathbf{E}\left[s^{Z(m,n)} | Z(n) > 0\right] \sim 1 - \frac{1-s}{1-ts} = \frac{s(1-t)}{1-ts}.$$

Расстояние до ближайшего общего предка

$$\mathbf{E}\left[s^{Z(m,n)}|Z(n)>0\right] \sim 1-\frac{1-s}{1-ts}=\frac{s(1-t)}{1-ts}.$$

Отсюда

$$\lim_{n\to\infty}\mathbf{P}\left\{\tau\geq m|Z(n)>0\right\}=\lim_{n\to\infty}\mathbf{P}\left\{Z(m,n)=1|Z(n)>0\right\}=1-t.$$

Тогда для $D(n)=n-\tau$

$$\lim_{n\to\infty}\mathbf{P}\left\{D(n)\leq nt|Z(n)>0\right\}=\lim_{n\to\infty}\mathbf{P}\left\{\tau\geq n(1-t)|Z(n)>0\right\}=t.$$