Noble Names, Branching Processes, and Fixation Probabilities

Joachim Hermisson

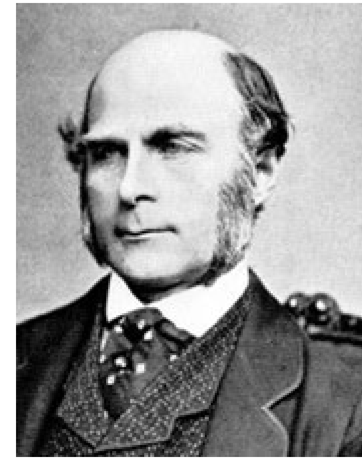
Mathematics & Max Perutz Labs,

University of Vienna

The fate of aristocratic family names

A problem of inheritance inspires new mathematics

"The decay of the families of men who occupied conspicuous positions in the past times has been a subject of frequent remark and has given rise to various conjectures …" [Galton and Watson 1874]



Sir Francis Galton (1822-1911)

Conjecture:

- Aristocrats (and "other men of genius") have reduced fertility → trade-off?
- Population only maintained by proletarians

Degradation risk!

Henry William Watson (1827 - 1903)

Galton:

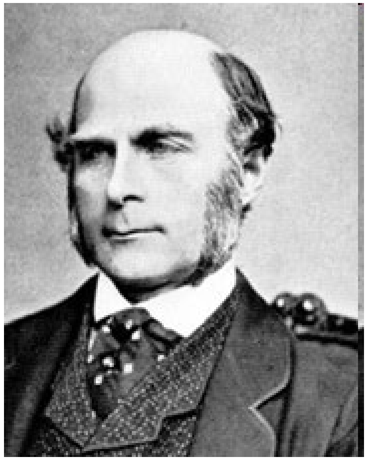
- It may also be just chance: *Need a model!*

Noble names, branching processes, and fixation probabilities

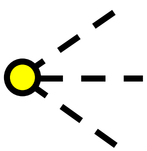
Galton's branching model

- Z_0 founders of noble families in generation $n = 0$
- Each founder j can have $k_j = 0,1,2,3,\ \ldots$ sons

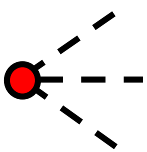
$\longrightarrow$ independently and with identical probability p_k



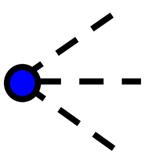
Sir Francis Galton
(1822-1911)



$k_1 = ?$



$k_2 = ?$



$k_3 = ?$

$Z_0 \longrightarrow Z_1$

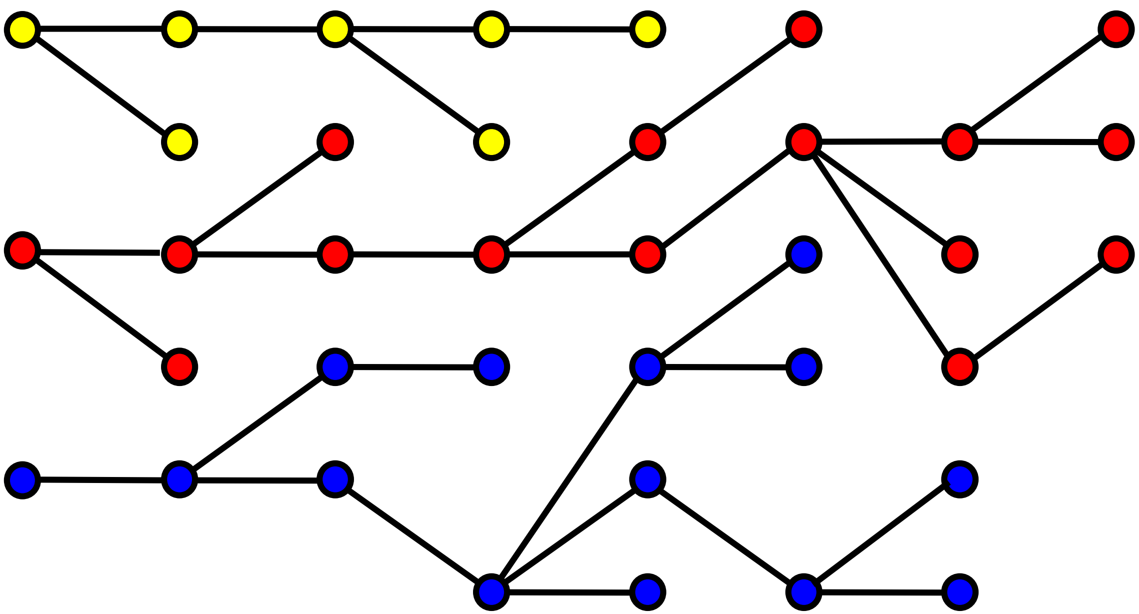
Henry William Watson
(1827 - 1903)

Galton's branching model

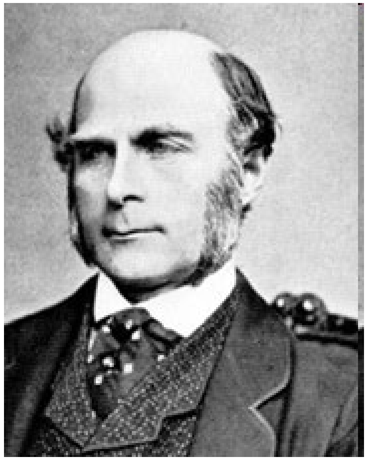
- Z_0 founders of noble families in generation $n = 0$
- Each founder j can have $k_j = 0,1,2,3, \ldots$ sons

$\rightarrow$ independently and with identical probability p_k

- Iterate with offspring generation



$Z_0 \rightarrow Z_1 \rightarrow Z_2 \rightarrow \ldots Z_n$



Sir Francis Galton (1822-1911)

Henry William Watson (1827 - 1903)

Galton's branching model

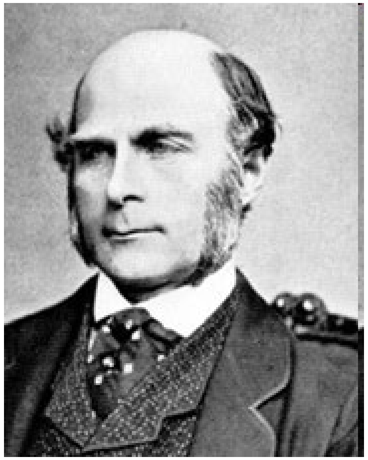
A Galton-Watson process $\{Z_n\}_{n\geq 0}$ with offspring distribution $\{p_k\}_{k\geq 0}$ is a Markov chain with values in $\mathbb{Z}_+$ and transition probabilities

$$P[Z_{i+1} = k | Z_i = m] = p_k^{*m}$$

where $\{p_k^{*m}\}$ is the m-fold convolution of $\{p_k\}$

(i.e., the distribution of the sum of m i.i.d. random variables, each with distribution $\{p_k\}$)

Due to independence, we can use $Z_0 = 1$ as default initial state ("fate of one family")



Sir Francis Galton (1822-1911)

Henry William Watson (1827 - 1903)

Ref. Watson's insights

π_n: probability for extinction by generation n

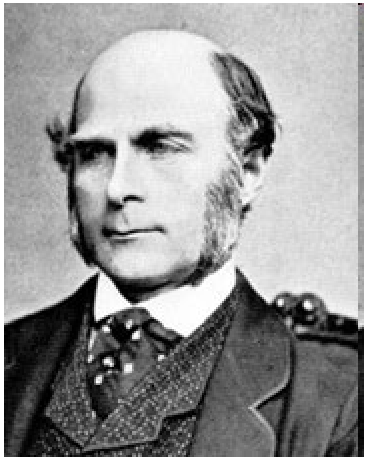
use generating function of offspring distribution

$$\varphi(t) = \sum_{k=0}^{\infty} p_k t^k$$

Recursion: $\pi_{n+1} = \varphi(\pi_n)$

$\pi_n \to \pi_\infty$ (monotonic and bounded)

Thus: $\pi_\infty = \varphi(\pi_\infty)$ fixed point of $\varphi(t)$
(φ continuous)

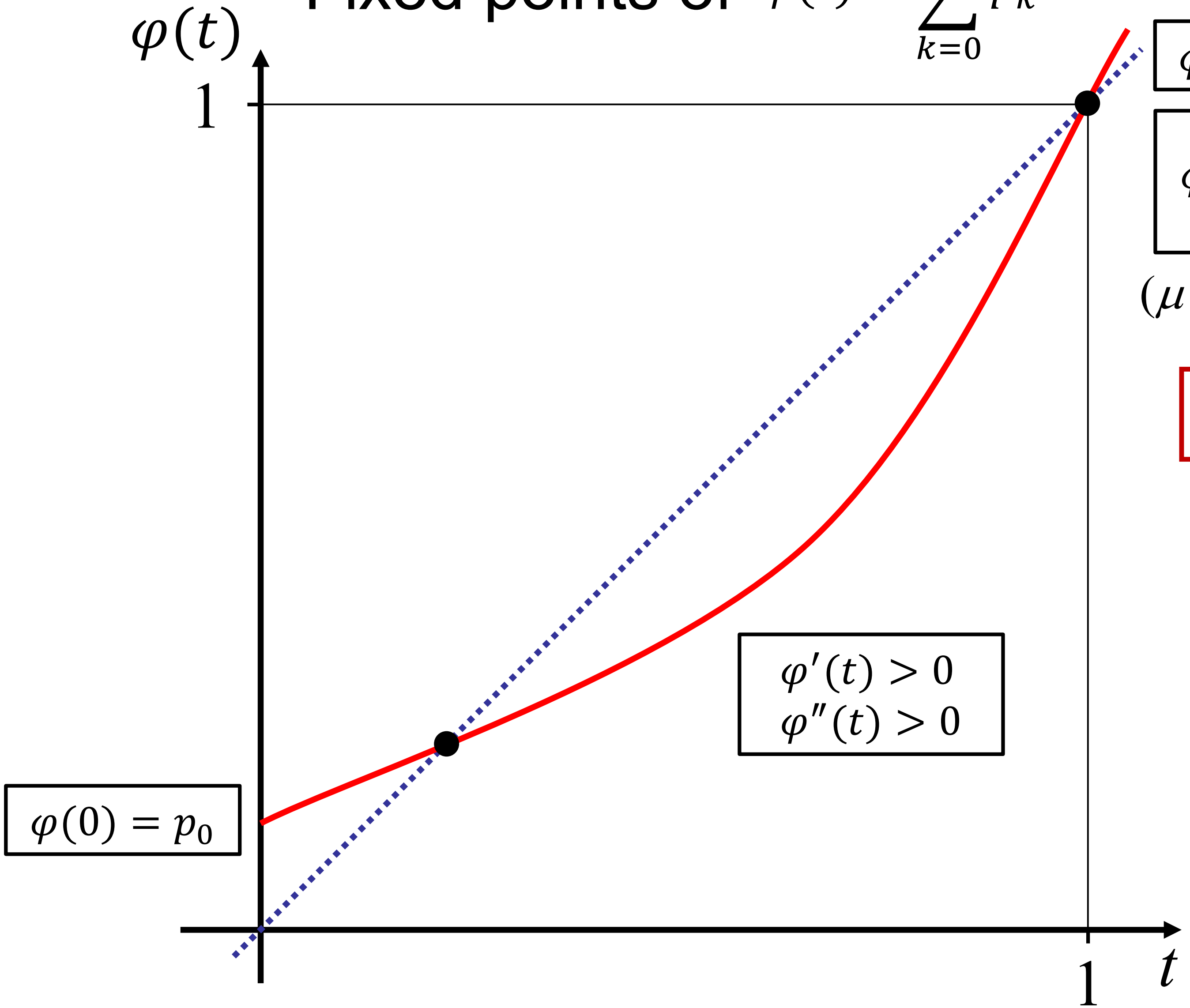


Sir Francis Galton
(1822-1911)

Henry William Watson
(1827 - 1903)

Noble names, branching processes, and fixation probabilities

Fixed points of $\varphi(t) = \sum_{k=0}^{\infty} p_k t^k$



$\varphi(1) = 1$

$\varphi'(1) = \sum_k k \cdot p_k = \mu$

(μ average offspring number)

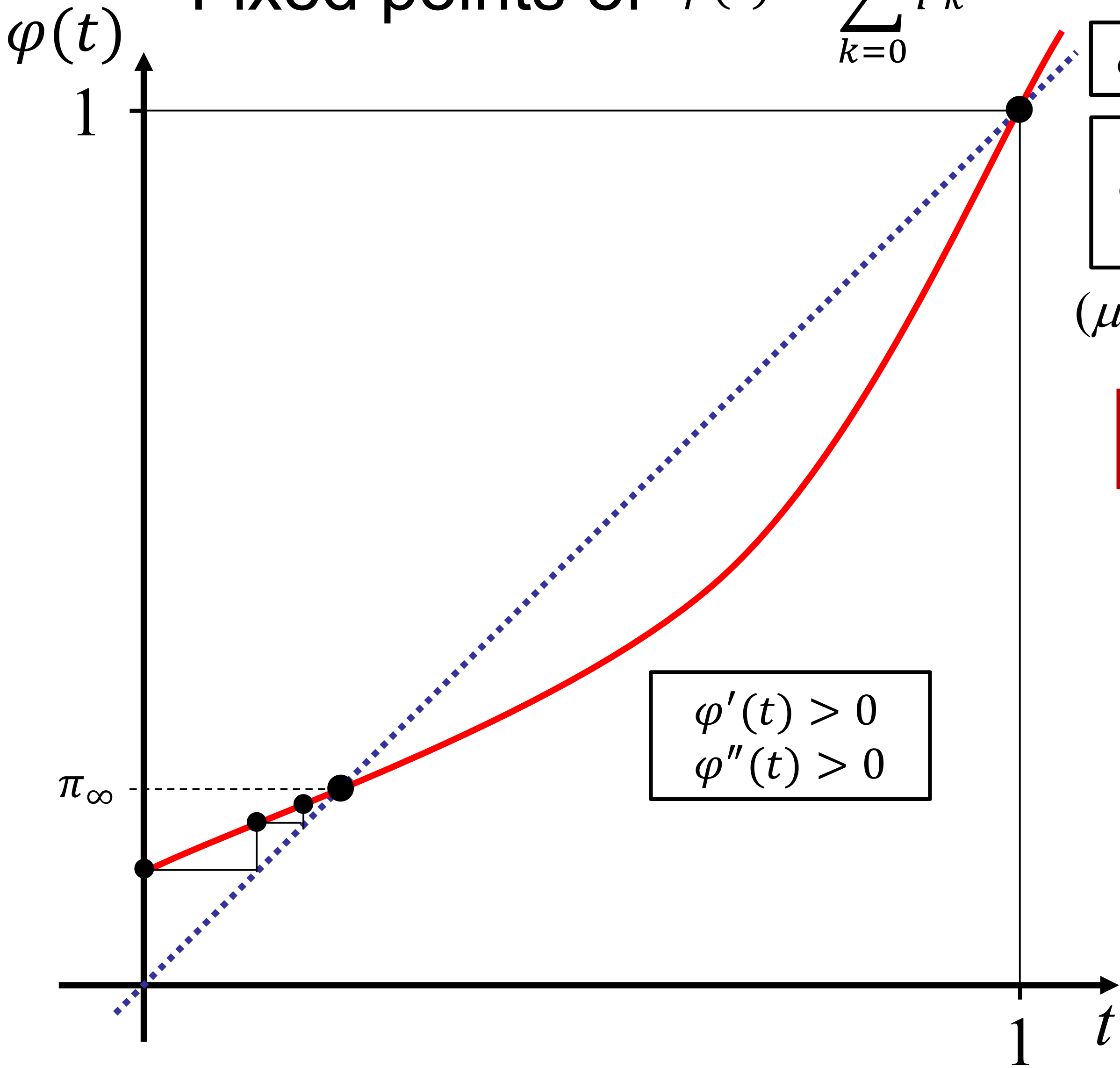
Case: $\mu > 1$

Assume:

- $p_0 > 0$
- $p_0 + p_1 < 1$

Noble names, branching processes, and fixation probabilities

Fixed points of $\varphi(t) = \sum_{k=0}^{\infty} p_k t^k$



$\varphi(1) = 1$

$\varphi'(1) = \sum_k k \cdot p_k = \mu$

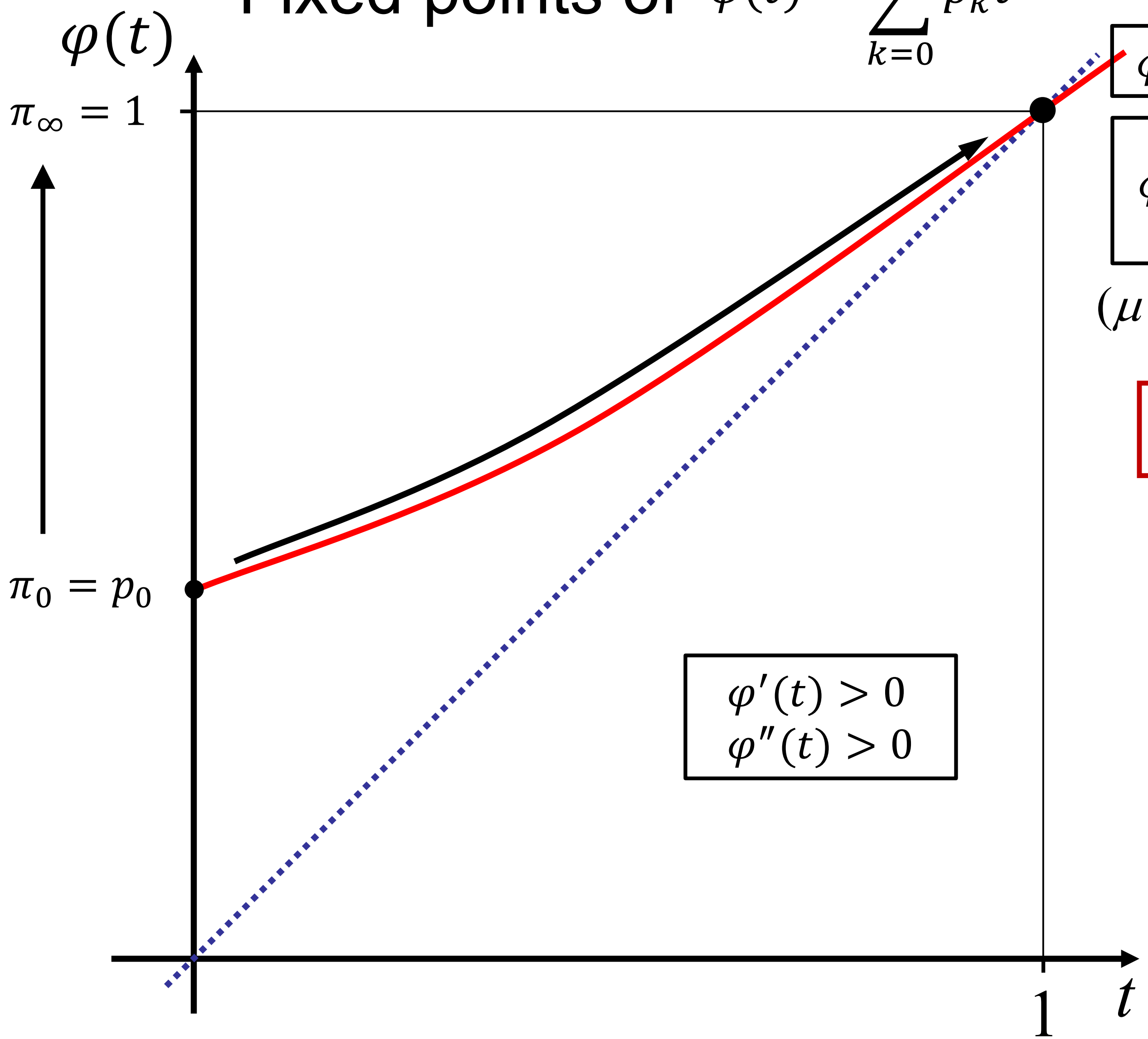
(μ average offspring number)

Case: $\mu > 1$

Assume:

- $p_0 > 0$
- $p_0 + p_1 < 1$

Fixed points of $\varphi(t) = \sum_{k=0}^{\infty} p_k t^k$



$\varphi(1) = 1$

$\varphi'(1) = \sum_k k \cdot p_k = \mu$

(μ average offspring number)

Case: $\mu \leq 1$

Assume:

- $p_0 > 0$
- $p_0 + p_1 < 1$

Extinction probability

Thus: $\pi_\infty = \varphi(\pi_\infty)$ smallest fixed point of $\varphi(t)$

For $\mu = \sum_k k \cdot p_k$ average offspring number:

$$\left.\begin{array}{ll} 1.\ \mu < 1 & \textit{subcritical} \\ 2.\ \mu = 1 & \textit{critical} \end{array}\right\} \quad \pi_\infty = 1$$

$$3.\ \mu > 1 \quad \textit{supercritical} \qquad \pi_\infty < 1$$

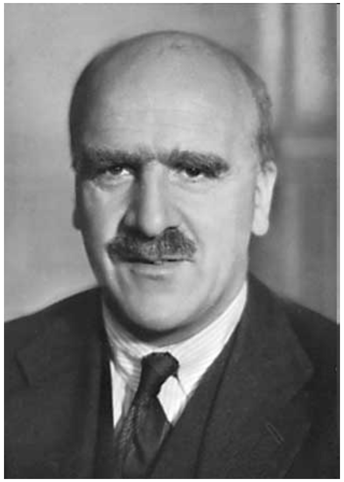
- Galton and Watson overlooked the smaller fixed point and concluded that all family names must die out because of chance alone
- Lotka (1931): $\pi_\infty \approx 0.82$ for US white males (1920 data)

Fixation probability

Ronald A. Fisher

The spread of a rare beneficial mutant through a population can be described as a supercritical branching process

[Fisher 1922, Haldane 1927]



J. B. S. Haldane

The fate of a beneficial mutant is decided while it is rare

- When frequent: loss very unlikely → eventual fixation (frequency 1)
- While rare: independent reproduction!

➢ Mutant population can be described by a branching process

➢ Fixation probability follow as:

$$p_{fix} = 1 - \pi_{\infty}$$

Fixation probability

Average offspring number

Wildtype: $\mu_{wt} = 1$ (constant population size)

Mutant: $\mu_m = 1 + s$ (typical s : $10^{-4} - 10^{-2}$ → "*slightly supercritical*")

Taylor expansion of the fixed point equation:

$$\pi_\infty = 1 - p_{fix} = \varphi\left(1 - p_{fix}\right) \approx \varphi(1) - p_{fix}\varphi'(1) - \frac{1}{2}p_{fix}^2\varphi''(1)$$

where:

$$\varphi(1) = 1 \quad ; \quad \varphi'(1) = \mu_m$$

$$\varphi''(1) = \sum_{k=2}^{\infty} k(k-1)p_k = \sigma_m^2 + \mu_m(\mu_m - 1)$$

(σ_m^2 variance of the offspring distribution)

Fixation probability

Solve for p_{fix}:

$$p_{fix} \approx \frac{2(\mu_m - 1)}{\sigma_m^2 + \mu_m(\mu_m - 1)} \quad = \quad \frac{2s}{\sigma_m^2} + O(s^2)$$

In particular, Wright-Fisher model (~ Poisson offspring distribution):

$$\sigma_m^2 = \mu_m = 1 + hs \quad \Rightarrow \quad \boxed{p_{fix} \approx 2hs}$$

(Haldane 1927)

(all mutants in heterozygotes)

Typical s : $10^{-4} - 10^{-2}$ $\Rightarrow$

almost all beneficial mutations in a population are lost because of random fluctuations (genetic drift)