On a model of evolution of subspecies

Hideki Tanemura (Keio University)

Boston-Keio-Tsinghua Workshop 2023
Probability and Statistics
(June 26th, 2023)

This talk is based on a joint work with Rahul Roy (ISI)

Introduction 1 (Preferential attachment model)

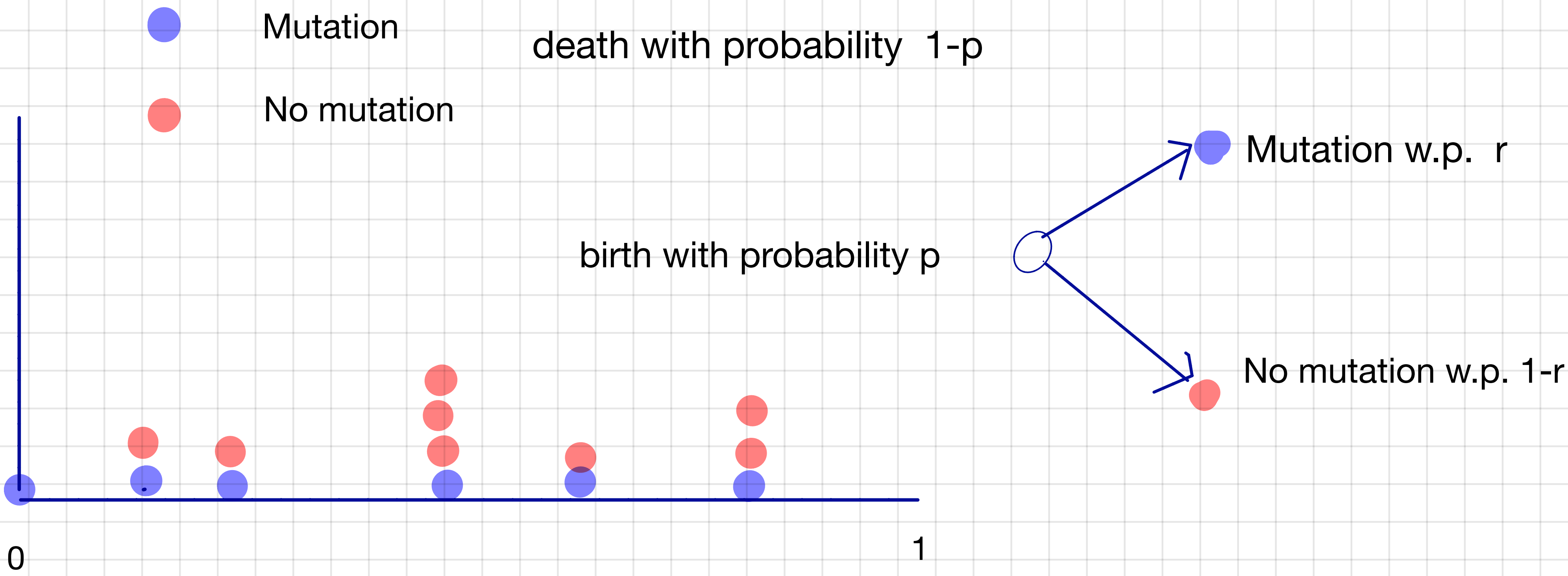
We consider the following model with preferential attachment (PAM);

(1) At time 0 there is one individual of fitness level 0.

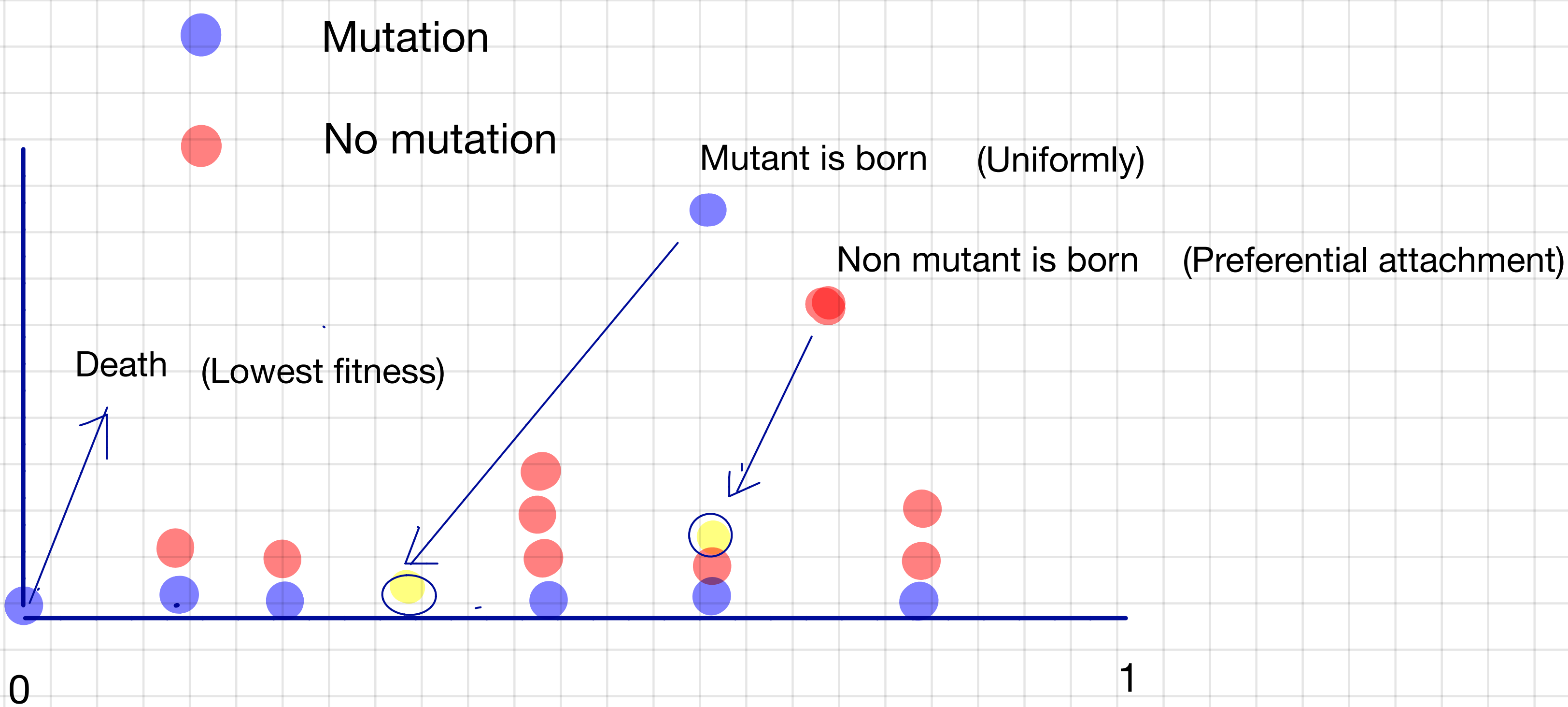
(2) With probability $p \in [0, 1]$ there is a new birth.
There are two possibilities.
(2-a) with probability $r \in [0, 1]$ a mutation is born and has a fitness f uniformly at random in $[0, 1]$, or
(2-b) with probability $1 - r$, the individual born has a fitness f with a probability proportional to the number of individuas with fitness f among the entire population present at that time.
Here we have caveat that, if there is no indvidual present at the time of birth, then the fitness of the individual is sampled uniformly in $[0, 1]$.

(3) With probability $1 - p$ there is a death event: an individual from the population with the smallest fitness is eliminated.

Preferential attachment model



Preferential attachment model



Suppose at time n the total population is of size N_n. Clearly $N_n \leq n$. The demographic distribution of this population is given by

$$X_n := \{(k_i, x_i) : k_i \in \mathbb{N}, x_i \in [0,1], i = 1, 2, \dots, \ell\},$$

where the total population at time n is divided according to their fitness levels $x_1, x_2, \dots, x_\ell$, with the size of the population with fitness x_i being exactly k_i.
Thus at time n, the population N_n is divided into exactly ℓ classes, each class being identified by its fitness level x_i.
In case there is no individual present at time n we take $X_n = \emptyset$.
The process X_n is a Markov process on the state space

$$\mathbb{X} = \{\{\emptyset\} \cup \{(k, x)\}_{x \in \Lambda} : (k, x) \in \mathbb{N} \times [0,1],\ \Lambda \subset [0,1],\ \sharp\Lambda < \infty\}.$$

(k, x) is the state that k individuals exist on the fitness x.

For a given $f \in (0,1)$, let L_n^f (respectively, R_n^f) denote the size of the population at time n whose fitness levels are in $[0,f]$ (respectively, in $(f,1]$), i.e.

$$L_n^f := \sum_{(k_i,x_i)\in X_n,\ x_i\le f} k_i,$$

$$R_n^f := \sum_{(k_i,x_i)\in X_n,\ x_i> f} k_i,$$

$$S_n := \sharp\{s\in[0,1] : (k,s)\in X_n \text{ for some } k\ge 1\}.$$

Clearly, $N_n = L_n^f + R_n^f$.

Remark.
S_n is not a Markov chain since its transition depends on the number of individuals at the lowest fitness.

Results: Macroscopic behaviour

The model exhibits a phase transition at a critical fitness level $f_c = \frac{1-p}{pr}$.

Theorem 1

(i) In case $p \leq 1-p$, the population dies out infinitely often almost surely in the sense that $P(N_n = 0 \text{ for infintely many } n) = 1$.

(ii) In case $1-p < pr$ (i.e. $f_c < 1$) the size of the population goes to infinity as $n \to \infty$ a.s. and the fitness levels of most of the population lies in the interval $[f_c, 1]$, in the sense that

$$P(\lim_{n\to\infty} \frac{R_n^f}{N_n} = 1) = 1 \text{ and } P(\liminf_{n\to\infty} \frac{R_n^{f_c} - R_n^f}{N_n} > 0) = 1 \text{ for any } f > f_c.$$

(iii) In case $pr \leq 1-p < p$ (i.e. $f_c > 1$ and $1-p < p$), the size of the population goes to infinity as $n \to \infty$ a.s., and the fitness levels of most of the population lies near 1, in the sense that

$$P(\lim_{n\to\infty} N_n = \infty) = 1, \text{ and } P(\lim_{n\to\infty} \frac{R_n^f}{N_n} = 1) = 1 \ \forall f < f_c \wedge 1 = 1.$$

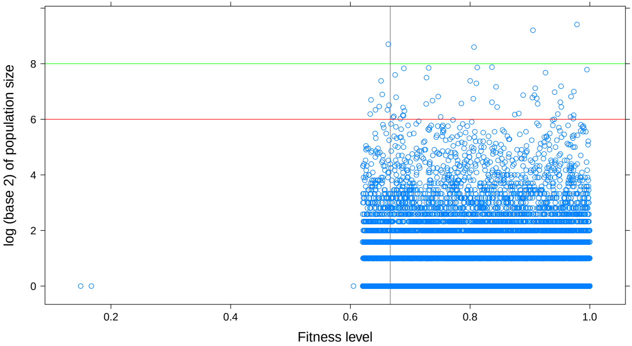


Figure: Population (in $\log_2$ scale) at various fitness levels.

Results: Microscopic behaviour of case (ii) $(1 - p < rp)$

$U_n^k(f)$: the number of fitness levels $x \in [f, 1]$, each of which has exactly k individuals at time n.

Then $S_n = \sum_k U_n^k(0)$, the total number of fitness levels at time n.
We define the empirical distribution of size and fitness on $\mathbb{N} \times [0,1]$:

$$H_n(A) = \begin{cases} \frac{\sum_{(k,f)\in A} U_n^k(f) - U_n^k(f+)}{S_n} & S_n > 0 \\ \delta_{(0,0)}(A). & S_n = 0 \end{cases}$$

Theorem 2 Let $1 - p < pr$ (i.e. $f_c < 1$). H_n converges weakly to $p_k \cdot \frac{1}{1 - f_c} \mathbf{1}_{[f_c,1]}(x)dx$, where

$$p_k = \frac{2p-1}{p(1-r)} B\left(1 + \frac{2p-1}{p(1-r)}, k\right), \quad k \in \mathbb{N}, \tag{1}$$

where $B(a, b)$ is the Beta function with parameter $a, b > 0$.

Remark for the case $1 - p < pr$ (i.e. $f_c < 1$)

(i) Let $F_n(f)$ denote the empirical distribution of fitnesses at time n. As a corollary of Theorems 1 and 2, the following Glivenko-Cantelli-type result holds:

$$F_n(f) \to \frac{f - f_c}{1 - f_c} \quad \text{as } n \to \infty, \quad \text{a.s.}$$

(ii) The distribution characterized by (1) is the Yule-Simon distribution with parameter $\frac{2p-1}{p(1-r)}$. Since

$$B(s, k) = \mathcal{O}(k^{-s}), \quad k \to \infty,$$

the probability density p_k, $k \in \mathbb{N}$ has m-th moment if and only if

$$m < \frac{2p - 1}{p(1 - r)}.$$

For $1 - p < rp$ (i.e. $f_c < 1$), $1 < \frac{2p-1}{p(1-r)}$.

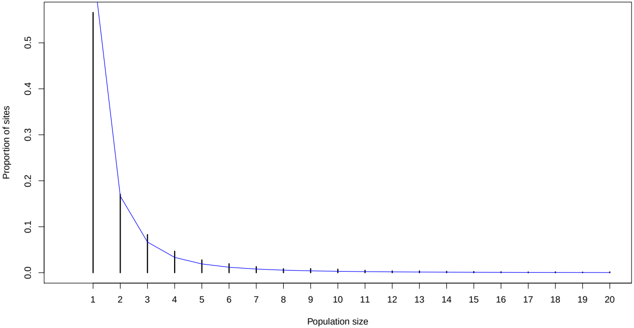


Figure: Theoretical and observed proportion of fitnesses with respect to population size.

Results: Case (iii) $pr \leq 1 - p < p$ (i.e. $f_c \geq 1$)

We introduce the increasing sequence $\{\xi_n\}$ defined by

$$\xi_n := \max\{x : \text{an individual of fitness } x \text{ was eliminated by time } n\}.$$

Put $\gamma := \dfrac{1 - p - pr}{p(1 - r)} \in [0, 1)$.

Theorem 3 Let $pr \leq 1 - p < p$ (i.e. $f_c \geq 1$). As $n \to \infty$,

(i) in case $pr = 1 - p$ (i.e. $f_c = 1$ and $\gamma = 0$), the expected number of individuals at a fitness level $x \in (\xi_n, 1)$ at time n is $\mathcal{O}(\log n)$.
In addition $1 - \xi_n = \mathcal{O}(1/\log n)$.

(ii) in case $pr < 1 - p < p$ (i.e. $f_c > 1$ and $\gamma \in (0, 1)$), the expected number of individuals at a fitness level $x \in (\xi_n, 1)$ at time n is $\mathcal{O}(n^\gamma)$.
In addition $1 - \xi_n = \mathcal{O}(n^{-\gamma})$.

Proof of Theorem 1

Lemma 4 (i) Let $f_c = \frac{1-p}{pr} < 1$.
(a) For $f < f_c$ we have

$$\lim_{n\to\infty} \frac{L_n^f}{N_n} = 0 \quad \text{a.s.} \tag{2}$$

$$P(L_n^f = 0 \text{ infinitely often}) = 1. \tag{3}$$

(b) For $f > f_c$ we have

$$P(L_n^f = 0 \text{ infinitely often}) = 0. \tag{4}$$

(ii) Let $1 - p < p$ and $f_c \geq 1$.
(a) For $f < 1$ we have (2) and (3).
(b) For $f = 1$ we have (4).

Proof of Theorem 1

(i) is obtained by the random walk comparison.
(ii) is derived from (i) of Lemma 4.

Since $N_n = L_n^f + R_n^f$, from (2) we have that $P(\lim_{n\to\infty} \frac{R_n^{f_c}}{N_n} = 1) = 1$.
Moreover, considering the birth rate pr of mutants, as $n \to \infty$

$$\frac{\sharp \text{ of them with a fitness between } (a,b) \subset [f_c, 1]}{n} \to pr(b-a), \text{ a.s.}$$

by an application of the strong law of large numbers. Thus we have

$$\liminf_{n\to\infty} \frac{R_n^b - R_n^a}{N_n} \geq \frac{p(b-a)}{2p-1}, \quad \text{a.s.}$$

This completes the proof of the statement of part (ii) of Theorem 1.
(iii) is derived from (ii) of Lemma 4.

Proof of Lemma 4

For a fixed $f \in (0,1)$, the pair (L_n^f, R_n^f) is a spatially inhomogeneous Markov chain on $\mathbb{Z}_+ \times \mathbb{Z}_+$, $(\mathbb{Z}_+ = \mathbb{N} \cup \{0\})$ with transition probabilities given by

(LR) If $(L_n^f, R_n^f) \in \mathbb{N} \times \mathbb{N}$

$$(L_{n+1}^f, R_{n+1}^f) = \begin{cases} (L_n^f + 1, R_n^f) & \text{w. p. } fpr + p(1-r)\dfrac{L_n^f}{N_n} \\ (L_n^f, R_n^f + 1) & \text{w. p. } (1-f)pr + p(1-r)\dfrac{R_n^f}{N_n} \\ (L_n^f - 1, R_n^f) & \text{w. p. } 1-p. \end{cases}$$

with the boundary condition (bc-1)-(bc-3) (bc-1) If $(L_n^f, R_n^f) = (0,0)$

$$(L_{n+1}^f, R_{n+1}^f) = \begin{cases} (1,0) & \text{w. p. } fp \\ (0,1) & \text{w. p. } (1-f)p \\ (0,0) & \text{w. p. } 1-p \end{cases}$$

(bc-2) If $(L_n^f, R_n^f) \in \{0\} \times \mathbb{N}$

$$(L_{n+1}^f, R_{n+1}^f) = \begin{cases} (1, R_n^f) & \text{w. p. } fpr \\ (0, R_n^f + 1) & \text{w. p. } (1-f)pr + p(1-r) \\ (0, R_n^f - 1) & \text{w. p. } 1-p \end{cases}$$

(bc-3) If $(L_n^f, R_n^f) \in \mathbb{N} \times \{0\}$

$$(L_{n+1}^f, R_{n+1}^f) = \begin{cases} (L_n^f + 1, 0) & \text{w. p. } fpr + p(1-r) \\ (L_n^f, 1) & \text{w. p. } (1-f)pr \\ (L_n^f - 1, 0) & \text{w. p. } 1-p \end{cases}$$

Modified Markov chain (Ep)

The idea of the proof is that, since for $f < f_c \wedge 1$, R_n^f will be much larger than L_n^f. We stochastically bound the spatially inhomogeneous Markov chain by a spatially homogeneous Markov chain, and study the modified Markov chain.

Define the spatially homogeneous Markov chains $(L_n^f(\varepsilon), R_n^f(\varepsilon))$, $\varepsilon \in (0,1)$ on $\mathbb{Z}_+ \times \mathbb{Z}_+$ by

(Ep) If $(L_n^f(\varepsilon), R_n^f(\varepsilon)) \in \mathbb{N} \times \mathbb{N}$

$$(L_{n+1}^f(\varepsilon), R_{n+1}^f(\varepsilon)) = \begin{cases} (L_n^f(\varepsilon)+1, R_n^f(\varepsilon)) & \text{w. p. } fpr + p(1-r)\varepsilon \\ (L_n^f(\varepsilon), R_n^f(\varepsilon)+1) & \text{w. p. } (1-f)pr + p(1-r)(1-\varepsilon) \\ (L_n^f(\varepsilon)-1, R_n^f(\varepsilon)) & \text{w. p. } 1-p. \end{cases} \tag{5}$$

with the boundary condition (bc-1)-(bc-3).

Modified Markov chain (Epn)

We also prepare an auxiliary Markov process. Let $\mathbf{a} = \{a_n\}_{n\in\mathbb{N}}$ is a sequence in $[0,1]$. We introduce the temporally inhomogeneous Markov chains $(L_n^f(\mathbf{a}), R_n^f(\mathbf{a}))$ on $\mathbb{Z}_+ \times \mathbb{Z}_+$, such that

(Epn) If $(L_n^f(\mathbf{a}), R_n^f(\mathbf{a})) \in \mathbb{N} \times \mathbb{N}$

$$(L_{n+1}^f(\mathbf{a}), R_{n+1}^f(\mathbf{a}) = \begin{cases} (L_n^f(\mathbf{a})+1, R_n^f(\mathbf{a})) & \text{w. p. } fpr + p(1-r)a_{n-1} \\ (L_n^f(\mathbf{a}), R_n^f(\mathbf{a})+1) & \text{w. p. } (1-f)pr + p(1-r)(1-a_{n-1}) \\ (L_n^f(\mathbf{a})-1, R_n^f(\mathbf{a})) & \text{w. p. } 1-p. \end{cases}$$

with the boundary condition (bc-1)-(bc-3).

Note that $N_n(\mathbf{a}) := L_n^f(\mathbf{a}) + R_n^f(\mathbf{a}) = N_n$.
We use this process when $\mathbf{a} = \{a_n\}$ is an addapted process. In this case the process is regarded as a Markov process on a random media.

We couple the processes $\{(L_n^f(\mathbf{a}), R_n^f(\mathbf{a})) : n \geq 1\}$ s.t. if $a_n \leq a'_n$ for $n \in \mathbb{N}$ then

$$L_n^f(\mathbf{a}) \leq L_n^f(\mathbf{a}'), \qquad R_n^f(\mathbf{a}) \geq R_n^f(\mathbf{a}') \quad \text{for all } n \geq 1. \tag{6}$$

We have, for $\rho_n^f := \frac{L_n^f}{N_n}$,

$$\begin{aligned} & L_n^f = L_n^f\left(\{\rho_n^f\}\right), \quad R_n^f = R_n^f\left(\{\rho_n^f\}\right), \\ & L_n^f(0) \leq L_n^f \leq L_n^f(1), \quad R_n^f(1) \leq R_n^f \leq R_n^f(0). \end{aligned} \tag{7}$$

By the law of large numbers we have

$$\lim_{n\to\infty} \frac{L_n^f(\varepsilon)}{n} = [fpr + p(1-r)\varepsilon - 1 + p]_+ \text{ and } \lim_{n\to\infty} \frac{N_n}{n} = 2p-1, \text{ a.s.,}$$

and so, for $\rho_n^f(\varepsilon) := \frac{L_n^f(\varepsilon)}{N_n}$, we have

$$\lim_{n\to\infty} \rho_n^f(\varepsilon) = \left[\frac{fpr - 1 + p}{2p-1} + \frac{p(1-r)\varepsilon}{2p-1}\right]_+ \qquad \text{a.s.} \tag{8}$$

We introduce the linear function defined by

$$h(x) = h^f(x) := \frac{fpr - 1 + p}{2p - 1} + \frac{p(1-r)}{2p-1}x.$$

Note that $\frac{p(1-r)}{2p-1} > 0$.

Suppose that $f < f_c \wedge 1$. A simple calculation shows that,

$$h(0) = \frac{fpr - 1 + p}{2p - 1} < 0, \; h(1) \le 1 - \frac{2p}{(1-r)pr} < 1.$$

Let ε_0 be the point such that $h(\varepsilon_0) = 0$, i.e.

$$\varepsilon_0 = \frac{(1-p) - fpr}{p(1-r)} > 0$$

Note that ε_0 could be ≥ 1. Put $h^{[1]}(\cdot) = h(\cdot)$ and $h^{[k]+1}(\cdot) = h^{[k]}(h(\cdot))$ for any $k \in \mathbb{N}$. The we see that there exists $k_0 \in \mathbb{N}$ such that

$$[h^{[k]}(1)]_+ = 0, \quad k \ge k_0. \tag{9}$$

From (7)

$$\rho_n^f(0) \le \rho_n^f \le \rho_n^f(1)$$

and from (8)

$$0 = [h(0)]_+ \le \limsup_{n\to\infty} \rho_n^f \le [h(1)]_+ \quad \text{a.s.}$$

Noting that $\rho_n^f = \rho_n^f(\{\rho_n^f\})$, from (6)

$$\rho_n^f(\{\rho_n^f(0)\}) \le \rho_n^f \le \rho_n^f(\{\rho_n^f(1)\}) \quad \text{a.s.}$$

and from (8) and the continuity of the function h

$$0 \le \limsup_{n\to\infty} \rho_n^f \le [h([h(1)]_+)]_+ = [h^{[2]}(1)]_+ \quad \text{a.s.}$$

Repeating this procedure, we have for any $k \in \mathbb{N}$.

$$0 \le \limsup_{n\to\infty} \rho_n^f \le [h^{[k]}(1)]_+ \quad \text{a.s.}$$

From (9) we obtain (2) if $f < f_c \wedge 1$.

If $f < f_c \wedge 1$, then $fpr + p(1-r)\varepsilon < 1-p$ for sufficiently small $\varepsilon > 0$, and $L_n^f(\varepsilon)$ is recurrent. From (2) we see that L_n^f hits the origin infinitely often. This proves (3).

If $f_c < f \leq 1$, then $1-p < fpr$, and $L_n^f(0)$ is transient. From (7) we have (4).

If $f = 1$ and $1-p < p$, then $L_n^f = N_n$, and we have (4).

This completes the proof of Lemma 4.

Proof of Theorem 2

$A_k(t_1, n)$, $k, t_1, n \in \mathbb{N}$: the event that a mutant born at time t_1 gets $k-1$ attachments until time n,

$$q_k(t_1, n) := P(A_k(t_1, n)).$$

Lemma 5 For the preferential attachment model with $p = 1$, i.e., no deaths, we have

$$E\left[\left\{\frac{1}{n}\sum_{t_1=1}^{n}(\mathbf{1}_{A_k(t_1,n)} - q_k(t_1, n))\right\}^2\right] \to 0 \text{ as } n \to \infty.$$

Lemma 6 Let $p = 1$. For each $k \in \mathbb{N}$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t_1=1}^{n} q_k(t_1, n) = \frac{r}{1-r}B\left(\frac{2-r}{1-r}, k\right) = rp_k.$$

Proof of Lemma 6

Proof. For $k = 1$, we have

$$q_1(t_1, n) = r \prod_{j=t_1+1}^{n} \left(1 - \frac{1-r}{j}\right),$$

since the number of individuals at time $j-1$ is j and the probability that the mutant who arrived at time t_1 gets an attachment at time j is $\frac{1-r}{j}$.
For $k = 2$

$$q_2(t_1, n) = r \sum_{t_2=t_1+1}^{n} \Big\{ \prod_{j=t_1+1}^{t_2-1} \big(1 - \frac{1-r}{j}\big)\Big\} \frac{1-r}{t_2} \Big\{ \prod_{j=t_2+1}^{n} \big(1 - \frac{2(1-r)}{j}\big)\Big\},$$

where t_2 is the time of the first attachment. Similarly for each $k \in \mathbb{N}$

$$q_k(t_1, n) = r \sum_{t_1<t_2<\cdots<t_k\le n} \prod_{\ell=1}^{k} \prod_{j=t_\ell+1}^{t_{\ell+1}} \big(1 - \frac{\ell(1-r)}{j}\big) \prod_{\ell=1}^{k-1} \frac{\ell(1-r)}{t_{\ell+1} - \ell(1-r)},$$

where $t_{k+1} = n$, and we used the equation $\frac{\ell(1-r)}{t_{\ell+1}} \frac{1}{1-\frac{\ell(1-r)}{t_{\ell+1}}} = \frac{\ell(1-r)}{t_{\ell+1}-\ell(1-r)}$.

By using Stirling's formula we see that

$$\prod_{j=t_\ell+1}^{t_{\ell+1}} (1-\frac{\ell(1-r)}{j}) = \prod_{j=t_\ell+1}^{t_{\ell+1}} \frac{j-\ell(1-r)}{j} \sim (\frac{t_\ell}{t_{\ell+1}})^{\ell(1-r)}, \quad t_\ell, t_{\ell+1} \to \infty.$$

Now letting $n \to \infty$ and taking $t_\ell = nx_\ell$ we have

$$\frac{1}{n}\sum_{t_1=1}^{n} q_k(t_1, n) \sim r\int_{0<x_1<\cdots<x_k<1} dx_1 \cdots dx_k \prod_{\ell=1}^{k} (\frac{x_\ell}{x_{\ell+1}})^{\ell(1-r)} \prod_{\ell=1}^{k-1} \frac{\ell(1-r)}{x_{\ell+1}}$$

$$= r(1-r)^{k-1}\int_0^1 dx_1 x_1^{1-r} \prod_{\ell=2}^{k}\int_{x_1}^1 dx_\ell\, x_\ell^{-r} = r\int_0^1 dx_1 x_1^{1-r}(1-x_1^{1-r})^{k-1}$$

$$= \frac{r}{1-r}\int_0^1 dy\, y^{\frac{1}{1-r}}(1-y)^{k-1} = \frac{r}{1-r}B\left(\frac{2-r}{1-r}, k\right).$$

Then we have Lemma 6.

Proof of Theorem 2

When $p = 1$ From Lemmas 5 and 6 we have

$$\frac{1}{n}\sum_{t_1=1}^{n} \mathbf{1}_{A_k(t_1,n)} \to \frac{r}{1-r} B\left(\frac{2-r}{1-r}, k\right) \text{ as } n \to \infty, \quad \text{in probability.}$$

Noting that $\lim_{n\to\infty} \frac{S_n}{n} = r$, a.s. we have

$$\lim_{n\to\infty} \frac{\sum_{f\in(0,1)} U_n^k(f) - U_n^k(f+)}{S_n} = \frac{1}{1-r} B\left(\frac{2-r}{1-r}, k\right) = p_k. \qquad (10)$$

Noting that the fitness levels are uniformly distributed on $[0, 1]$ independently, and preferential attachment does not depend on the position of fitnesses, we obtain Theorem 2 from (10) for $p = 1$.

Next we consider the case where $p \in (0,1)$. We introduce another Markov process $\hat{X}_n$, $n \in \mathbb{Z}_+$, which is a pure birth process, as follows:

1. At time 0 there exists one individual at a fitness uniformly distributed on $(f_c, 1)$.
2. with probability $p(1-rf_c)$ there is a new birth. There are two possibilities;
   - with probability $\hat{r} := \dfrac{pr(1-f_c)}{p(1-rf_c)} = \dfrac{p(1+r)-1}{2p-1}$ a mutant is born with a fitness uniformly distributed in $[f_c, 1]$,
   - With probability $1-\hat{r} := \dfrac{p(1-r)}{p(1-rf_c)} = \dfrac{p(1-r)}{2p-1}$ the individual born has a fitness f with a probability proportional to the number of individuals of fitness f and we increase the corresponding population of fitness f by 1.
3. With probability $1-p(1-rf_c)$ nothing happens, i.e. neither a birth nor a death occurs.

X_n, $n \in \mathbb{Z}_+ \Rightarrow q_k$, S_n and U_n, $\quad \hat{X}_n$, $n \in \mathbb{Z}_+$, $\Rightarrow \hat{q}_k$, $\hat{S}_n$ and $\hat{U}_n$.

Then by the same argument as the case $p = 1$ we see that

$$\frac{1}{n}\sum_{t_1=1}^{n} \tilde{q}_k(t_1, n) \sim p(1 - rf_c)\frac{\hat{r}}{1-\hat{r}} B\left(\frac{2-\hat{r}}{1-\hat{r}}, k\right) \text{ and } \lim_{n\to\infty}\frac{\hat{S}_n}{n} = pr(1 - f_c).$$

Hence

$$\lim_{n\to\infty} \frac{\sum_{f\in(0,1)} \hat{U}_n^k(f) - \hat{U}_n^k(f+)}{\hat{S}_n} = \frac{1}{1-\hat{r}} B\left(\frac{2-\hat{r}}{1-\hat{r}}, k\right) = \hat{p}_k.$$

We know that for any ε, deletions of individuals in $[f_c + \varepsilon, 1)$ occur finitely often almost surely as $n \to \infty$.

$$\lim_{n\to\infty} \frac{\sum_{f\in(0,1)} U_n^k(f) - U_n^k(f+)}{S_n} = \lim_{n\to\infty} \frac{\sum_{f\in(0,1)} \hat{U}_n^k(f) - \hat{U}_n^k(f+)}{\hat{S}_n} \quad \text{a.s.}$$

and so (10) for $p \in (0, 1]$.

Noting that the fitnesses are uniformly distributed on $[0, 1]$ independently, and preferential attachment does not depend on the position of fitnesses, we obtain Theorem 2 from (10).

Proof of Theorem 3

Lemma 7 Let Y_k, $k \in \mathbb{Z}_+$ be a reflecting random walk with negative drift on $\mathbb{Z}_+$. Then there exists $c > 0$ such that

$$\sum_{n=1}^{\infty} P(\max_{1 \le k \le n} Y_k \ge c \log n) < \infty.$$

Lemma 8 Let $pr < 1 - p < p$ (i.e $f_c > 1$). Let $M_k = \sharp$ of individuals with fitness level $< \xi_k$ at time k. There exists $c' > 0$ such that

$$\sum_{n=1}^{\infty} P(\max_{1 \le k \le n} M_k \ge c' \log n) < \infty.$$

We first prove (ii). From **Lemma 8** we see that as $n \to \infty$,

$$\sharp \text{ of individuals with fitness level } < \xi_n \text{ present at time } n \text{ is } \mathcal{O}(\log n), \text{ a.s.}$$

From the strong law of large numbers and the central limit theorem we see that as $n \to \infty$

$$N_n = (2p-1)n + \mathcal{O}(\sqrt{n}), \text{ a.s.}$$

and for any $0 < a < b \leq 1$

$$\sharp \text{ of fitness levels born in } (a,b) = pr(b-a)n + \mathcal{O}(\sqrt{(b-a)n}), \text{ a.s.}$$

Thus, as $n \to \infty$

$$\sharp \text{ of individuals in } (\xi_n, 1] \text{ is } (2p-1)n + \mathcal{O}(\sqrt{n}), \text{ a.s.} \tag{11}$$

$$\sharp \text{ of fitness levels } \geq \xi_n \text{ at time } n \text{ is } pr(1-\xi_n)n + \mathcal{O}(\sqrt{(1-\xi_n)n}), \text{ a.s.} \tag{12}$$

Hence, from ths same argument as in the proof of Lemma 7 with (11) and (12), we have

$$
\begin{aligned}
&P(\text{a mutant is born at time } t \text{ of fitness } x \geq \xi_n \text{ and } N_n^x = k|\xi_n) \\
&\approx pr(1-\xi_n) \sum_{t=t_1<t_2<\cdots<t_k\leq n} \left[\prod_{\ell=1}^{k-1} \prod_{j=t_\ell+1}^{t_{\ell+1}-1} \left(1 - \frac{\ell p(1-r)}{(2p-1)j}\right) \right] \\
&\qquad \times \left[\prod_{j=t_k+1}^{n} \left(1 - \frac{kp(1-r)}{(2p-1)j}\right) \right] \left[\prod_{\ell=1}^{k-1} \frac{\ell p(1-r)}{(2p-1)t_{\ell+1}} \right].
\end{aligned}
$$

where $f_n \approx g_n$ implies $c \leq f_n/g_n \leq c'$ with some $c, c' > 0$.
Put $\beta := \frac{p(1-r)}{2p-1}$.
Suppose that $t_\ell = ns_\ell$ for some sequence $0 < s_1 < \cdots s_k < 1$.

$$
\prod_{j=t_\ell+1}^{t_{\ell+1}-1} \left(1 - \frac{\ell}{j}\beta\right) \sim \left(\frac{s_\ell}{s_{\ell+1}}\right)^{\ell\beta} \text{ and } \prod_{j=t_k+1}^{n} \left(1 - \frac{k}{j}\beta\right) \sim \left(\frac{t_k}{n}\right)^{k\beta} = s_k^{k\beta}.
$$

$$\frac{E(\#\{\text{ fitness levels} \geq \xi_n \text{ with } k \text{ individuals at time } n\}|\xi_n)}{\sharp \text{ of fitness levels} \geq \xi_n \text{ at time } n \text{ is}}$$

$$\sim \frac{\sum_{t=1}^{n} P(\text{a mutant born at time } t \text{ of fitness level } x \geq \xi_n \text{ and } N_n^x = k|\xi_n)}{pr(1-\xi_n)n}$$

$$\approx \frac{1}{n} \sum_{1 \leq t = t_1 < t_2 < \cdots < t_k < n} \left[\prod_{\ell=1}^{k-1} \left(\frac{s_\ell}{s_{\ell+1}} \right)^{\ell\beta} \right] s_k^{k\beta} \left[\prod_{\ell=1}^{k-1} \frac{1}{n} \frac{\ell\beta}{s_{\ell+1}} \right]$$

$$\sim \beta^{k-1} \int_0^1 \left[\int_0^{s_1} \cdots \int_0^{s_1} \prod_{\ell=2}^{k} s_\ell^{\beta-1} ds_\ell \right] s_1^\beta ds_1 = \int_0^1 ds_1 s_1^\beta (1-s_1^\beta)^{k-1}.$$

Making the substitution $y = s_1^\beta$ we have

$$\int_0^1 ds_1 s_1^\beta (1-s_1^\beta)^{k-1} = \frac{1}{\beta} \int_0^1 y^{1-\frac{\beta-1}{\beta}} (1-y)^{k-1} dy = \frac{1}{\beta} B(\frac{\beta+1}{\beta}, k).$$

Noting that $B(\frac{\beta+1}{\beta}, k) \approx k^{-\frac{\beta+1}{\beta}}$, we have

$$E(\#\text{individuals with a fitness level } x \geq \xi_n \text{ at time } n\ )$$
$$\approx \sum_{k=1}^{n} kB(\frac{\beta+1}{\beta}, k) \approx \sum_{k=1}^{n} k^{-1/\beta} \approx n^{1-\frac{1}{\beta}} \quad \text{for } \beta > 1.$$

Then Theorem 3 (ii) is now proved by taking $\gamma := 1 - \frac{1}{\beta} = \frac{1-p-pr}{(1-r)p}$.
We prove (i). From (ii) we see that ξ_n is asymptotically larger that $n^{-\varepsilon}$ for any ε. Then we see that the number of individuals with fitness levels $\leq \xi_n$ is $\mathcal{O}(n^{1/2})$, $n \to \infty$. Hence, by the same procedure we have

$$E(\#\text{individuals at time } n \text{ with a fitness level } x \geq \xi_n\ )$$
$$\approx \sum_{k=1}^{n} kB(\frac{\beta+1}{\beta}, k) \approx \sum_{k=1}^{n} k^{-1/\beta} \approx \log n \quad \text{for } \beta = 1.$$

Thank you for your attention