

Lecture 10: Evolutionary timescales

MATH 242
11/4/2025

Recall **Moran process** with N individuals.

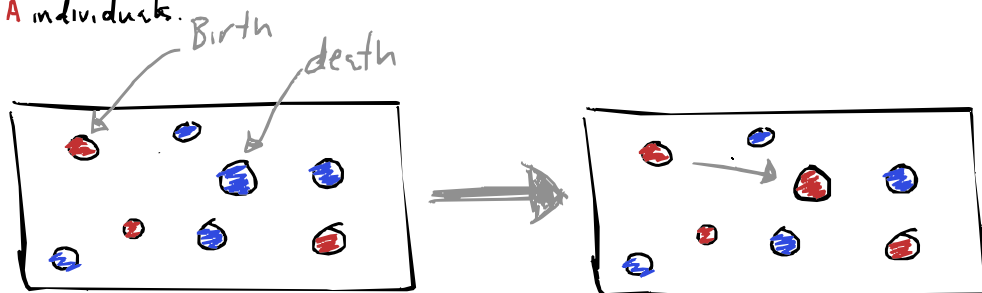
Two types of individuals: type **A** & type **B**

Type **A** has relative fitness $r > 1$, whereas type **B** has relative fitness 1.

At each time step, an individual is selected for birth with probability proportional to fitness.

An individual is selected for death uniformly at random. The individual selected for death is removed from the population.

The individual selected for birth places creates a duplicate of its type. The state space is $i = 0, 1, \dots, N$, representing the number of type **A** individuals.



(Example step in Moran process)

Previously, we showed that the fixation probability, of a population consisting of a single type **A** individual in a background population of type **B** individuals is given by:

$$p_1^A = \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^N}}$$

for $r \neq 1$, $\frac{1}{N}$ if $r = 1$

But what about the time taken? With ^{out} mutations,

the above process is pure selection-replication.

The assumption is that the time until a new mutation occurs takes much longer than the time until a mutant fixates/goes extinct.

Is this assumption valid?

The violation of this assumption can lead to clonal interference, a phenomenon the above model fails to explain.

It also may be the case that the environment changes after a long period of time, causing the mutant advantage to lessen, thus affecting the fixation probability.

Absorption times in Moran process

Let AT_i be the random variable representing the number of time steps—starting w/ i type **A** individuals—until the population ^(size N) becomes all type **A** (fixation) or all type **B** (extinction).

We can write:

$$AT_0 = AT_N = 0$$

$$AT_i = 1 + \epsilon_i^{\uparrow} \cdot AT_{i+1} + \epsilon_i^{\downarrow} \cdot AT_{i-1} + \epsilon_i^{\rightarrow} \cdot AT_i$$

↗
one time step

↑
indicator v.v.
for event #

indicator v.v.
for event
type **A**
stays the same
in one time step

for $i = 1, \dots, N-1$

type A increases
in one time step

Let $x_i = \mathbb{E} A_i$. Then by linearity of expectation,

$$x_0 = x_N = 0 \quad (= \mathbb{P}\{E\})$$

$$(*) \quad x_i = 1 + p_i^{\uparrow} \cdot x_{i+1} + p_i^{\downarrow} x_{i-1} + p_i^{\circ} x_i \quad \text{for } i=1, \dots, N-1.$$

Does this system look familiar?? It is "almost" the same system we solved for the fixation probability (what was different?)

Let matrix A be $N \times N$ s.t. $A_{i,i+1} = -p_i^{\uparrow}$, $A_{i,i-1} = -p_i^{\downarrow}$, and $A_{i,i} = 1 - p_i^{\circ} = 1 - (1 - p_i^{\uparrow} - p_i^{\downarrow}) = p_i^{\uparrow} + p_i^{\downarrow}$

Then for column vector $b \in \mathbb{R}^N$ w/ $b_0 = b_N = 0$, $b_i = 1$ otherwise, $(*)$ is soln to $Ax = b$.

Let $y_i \triangleq x_i - x_{i-1}$ and $\gamma_i \triangleq \frac{p_i^{\downarrow}}{p_i^{\uparrow}}$. Then $(*)$ simplifies to:

$$\begin{aligned} y_{i+1} &= -\frac{1}{p_i^{\uparrow}} + \gamma_i y_i = -\left(\sum_{j=2}^i \frac{1}{p_j^{\uparrow}} \cdot \prod_{k=j+1}^i \gamma_k \right) + \gamma_i \cdot \prod_{k=1}^i \gamma_k \\ &= -\left(\sum_{j=2}^i \frac{N^2}{j(N-j)} \cdot r^{j-i} \right) + x_1 \cdot r^{-i} \\ &= -\left(N^2 r^{-i} \sum_{j=2}^i \frac{r^j}{j(N-j)} \right) + x_1 r^{-i} \end{aligned}$$

Notice that $\sum_{i=0}^{N-1} y_{i+1} = \sum_{i=0}^{N-1} (x_{i+1} - x_i) = (x_1 - x_0) + (x_2 - x_1) + \dots + (x_N - x_{N-1})$

$$= x_N - x_0 = 0. \quad (1)$$

But also $\sum_{i=0}^{N-1} y_{i+1} = \sum_{i=0}^{N-1} \left[-\left(N^2 r^{-i} \sum_{j=2}^i \frac{r^j}{j(N-j)} \right) + x_1 r^{-i} \right]$

$$= x_1 \left(\sum_{i=0}^{N-1} r^{-i} \right) - N^2 \sum_{i=0}^{N-1} r^{-i} \sum_{j=2}^i \frac{r^j}{j(N-j)} \quad (2)$$

Combining (1) & (2) gives

$$X_1 = N^2 \cdot$$

$$\frac{\sum_{i=0}^{N-1} r^{-i} \sum_{j=2}^i \frac{r^j}{j(N-j)}}{\sum_{i=0}^{N-1} r^{-i}} \approx (1 - \frac{1}{r})^{-1}$$

$$\approx \sum_{i=0}^{N-1} \sum_{j=2}^i \frac{r^{j-i}}{j(N-j)}$$

$$\approx \frac{r}{1-r}$$

When $r=1$, we have $X_1^{(r=1)} = N \cdot \sum_{i=0}^{N-1} \sum_{j=2}^i \frac{1}{j(N-j)}$

$$= N \sum_{i=0}^{N-1} \sum_{j=2}^i \frac{1}{N} \cdot \left[\frac{1}{j} + \frac{1}{N-j} \right]$$

$$= \sum_{i=0}^{N-1} \left(\sum_{j=2}^i \left[\frac{1}{j} \right] + \sum_{j=2}^i \left[\frac{1}{N-j} \right] \right)$$

$$= \sum_{i=0}^{N-1} \left[-1 + H_i + \underbrace{\left(\frac{1}{N-i} + \dots + \frac{1}{N-2} \right)}_{H_{N-2} - H_{N-i-1}} \right]$$

$$H_i = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{i}$$

$$= \left(\sum_{i=0}^{N-1} H_{N-2} \right) + \left(\sum_{i=0}^{N-1} H_i - H_{N-i-1} \right) - N$$

$$\sim cN \ln N \quad \text{for } 1 < c < 2$$

$$= \Theta(N \log N).$$

When $r > 1$, (i) $\sum_{i=0}^{N-1} r^{-i} = \left(\frac{1 - \frac{1}{r} N}{1 - \frac{1}{r}} \right) \sim \frac{1}{1 - \frac{1}{r}} = \frac{r}{r-1}$ (looks familiar...)

(ii) $\sum_{i=0}^{N-1} r^{-i} \sum_{j=2}^i \frac{r^j}{j(N-j)} = \sum_{j=2}^{N-1} \frac{1}{j(N-j)} \underbrace{\sum_{i=j}^{N-1} r^{j-i}}_{\approx (1 - \frac{1}{r})^{-1}}$

$$\leq c \frac{H_N}{N} \cdot \frac{r}{r-1}$$

So $X_1^{(r>1)} \leq \frac{N^2 \cdot \frac{cH_N}{N} \cdot \frac{r}{r-1}}{\frac{r}{r-1}} = cN H_N$

$$\leq cN \ln N$$

(What happens for $r < 1$?)

$$= O(N \log N).$$

The above is for absorption times

The absorption time is sometimes also called the unconditional fixation time.

But in a lot of cases, we are interested in the (conditional) fixation time ^(denoted FT_i): the absorption time given fixation of type A .

Q: How do we solve for this quantity?

A: use Bayes' rule. ^{"conditioned on"}

Note that $FT_i = AT_i \mid \text{fixation}$
(fix = next timestep)

so by Bayes',

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P\{\varepsilon_i^* = 1 \mid \text{fixation}\} = \frac{P\{\text{fixation} \mid \varepsilon_i^* = 1\} \cdot P\{\varepsilon_i^* = 1\}}{P\{\text{fixation}\}}$$

What does this mean, formally?

Moran process is a Markov chain w/ state space $\Omega = \{0, 1, \dots, N\}$, and at time t , the random variable $Z_t \in \Omega$ denotes # type A individuals.

$$\text{So } AT_i = AT \mid \{Z_0 = i\}.$$

$$\text{Also } FT_i = AT_i \mid \text{fixation}$$

$$= AT_i \mid \{\exists \tau \text{ s.t. } Z_\tau = N\}$$

$$= AT \mid \{Z_0 = i\} \cap \{\exists \tau \text{ s.t. } Z_\tau = N\}$$

$$\text{and } \varepsilon_i^* = 1 \mid \{Z_{t+1} = i+1\} \mid \{Z_t = i\}.$$

$$\text{So } P\{\varepsilon_i^\uparrow = 1 | \text{fixation}\} = P\{Z_{t+1} = i+1 | \{Z_t = i\} \wedge \{\exists t \text{ s.t. } Z_t = N\}\}$$

$$\begin{aligned} \text{Thus } P\{\text{fixation} | \varepsilon_i^\uparrow = 1\} &= P\{\exists t \text{ s.t. } Z_t = N | Z_t = i \text{ and } Z_{t+1} = i+1\} \\ &= P\{\exists t \text{ s.t. } Z_t = N | Z_{t+1} = i+1\} \\ &= p_A^{(i+1)}, \end{aligned}$$

$$\begin{aligned} P\{\text{fixation}\} &= P\{\exists t \text{ s.t. } Z_t = N | Z_t = i\} \\ &= p_A^{(i)}, \text{ and} \end{aligned}$$

$$\begin{aligned} P\{\varepsilon_i^\uparrow = 1\} &= P\{Z_{t+1} = i+1 | Z_t = i\} \\ &= p_i^\uparrow. \end{aligned}$$

$$\begin{aligned} \text{So } FT_i &= 1 + (\varepsilon_i^\uparrow | \text{fixation}) \cdot FT_{i+1} \\ &\quad + (\varepsilon_i^\downarrow | \text{fixation}) \cdot FT_{i-1} \\ &\quad + (\varepsilon_i^0 | \text{fixation}) \cdot FT_i. \end{aligned}$$

Setting $X'_i = \mathbb{E} FT_i$, we get

$$X'_0 = \text{undefined}$$

$$X'_N = 0$$

$$X'_i = 1 + \frac{p_A^{(i+1)}}{p_A^{(i)}} \cdot p_i^\uparrow X'_{i+1} + \frac{p_A^{(i-1)}}{p_A^{(i)}} p_i^\downarrow X'_{i-1} + p_i^0 X'_i$$

for $i=1, \dots, N-1$

$$\text{We know } p_A^{(i)} = \frac{1-r^{-i}}{1-r^{-N}} \text{ for } r \neq 1 \text{ and } p_A^{(i)} = i/N \text{ for } r=1$$

$$\text{So } \frac{p_A^{(i+1)}}{p_A^{(i)}} = \frac{1-r^{-(i+1)}}{1-r^{-N}} \quad \frac{p_A^{(i-1)}}{p_A^{(i)}} = \frac{1-r^{-(i-1)}}{1-r^{-N}}$$

for $r \neq 1$ $\frac{p_A^{(i)}}{p_A^{(i)}} = \frac{1-r^{-i}}{1-r^{-i}}$ and $\frac{p_A^{(i)}}{p_A^{(i)}} = \frac{1-r^{-i}}{1-r^{-i}}$

and for $r=1$ $\frac{p_A^{(i+1)}}{p_A^{(i)}} = \frac{i+1}{i} = 1 + \frac{1}{i}$ and $\frac{p_A^{(i-1)}}{p_A^{(i)}} = \frac{i-1}{i} = 1 - \frac{1}{i}$ for $i=1, \dots, N-1$

Let's look at the $r=1$ case,

We have, for $i=1, \dots, N-1$,

$$x'_i = 1 + \left(1 + \frac{1}{i}\right) p_i^{\uparrow} x'_{i+1} + \left(1 - \frac{1}{i}\right) p_i^{\downarrow} x'_{i-1} + (1 - p_i^{\uparrow} - p_i^{\downarrow}) x'_i$$

$$\Leftrightarrow -1 = (p_i^{\uparrow} x'_{i+1} - p_i^{\uparrow} x'_i) - (p_i^{\downarrow} x'_i - p_i^{\downarrow} x'_{i-1})$$

$$+ \frac{p_i^{\uparrow}}{i} x'_{i+1} - \frac{p_i^{\downarrow}}{i} x'_{i-1}$$

Let $y'_i \triangleq x'_i - x'_{i-1}$

(Notice when $r=1$, $p_i^{\uparrow} = p_i^{\downarrow} = \frac{i(N-i)}{N^2} \triangleq p_i$)

$$= p_i y'_{i+1} - p_i y'_i + \frac{p_i}{i} (y'_{i+1} + y'_i)$$

$$\Leftrightarrow \frac{-1}{p_i} = \left(1 + \frac{1}{i}\right) y'_{i+1} - \left(1 - \frac{1}{i}\right) y'_i$$

$$\Leftrightarrow y'_{i+1} = \frac{-1}{p_i \left(1 + \frac{1}{i}\right)} + \frac{1 - \frac{1}{i}}{1 + \frac{1}{i}} y'_i$$

$$= \frac{-1}{q_i} + \eta_i y'_i \quad \text{for} \quad q_i \triangleq p_i \cdot \left(1 + \frac{1}{i}\right) = \frac{i(N-i)}{N^2} \cdot \left(1 + \frac{1}{i}\right)$$

$$\eta_i \triangleq \frac{1 - \frac{1}{i}}{1 + \frac{1}{i}}$$

$$= \left(- \sum_{j=2}^i \frac{1}{q_j} \prod_{k=j+1}^i \eta_k \right) + y'_2 \prod_{j=2}^i \eta_j$$

$\uparrow$
(note $y'_i = -\frac{1}{2} p_i$, $p_1 = \frac{N-1}{N^2}$)

$$\sum_{i=1}^{N-1} y'_{i+1} = (x'_2 - x'_1) + (x'_3 - x'_2) + \dots + (x'_N - x'_{N-1})$$

$$= x'_N - x'_1$$

$$\Omega(N^2)$$

$$\text{So } x_1^{(r+1)} = \Theta(N^2).$$

When $r=1$, (expected) absorption time is $\Theta(N \log N)$, but (expected) fixation time is $\Theta(N^2)$

Suppose $u > 0$ is the probability that when an individual gives birth, it gives birth to an entirely new (never before seen) type.

Let M be the number of timesteps until a brand new mutant appears.

For the Moran process (pure selection process) to be applicable, we need to show that for typical mutation rates, it is likely the case that an old mutant absorbs before a new mutant appears.

$\hookrightarrow \approx \text{Geom}(u)$

$$\begin{aligned} \mathbb{P}\{M < AT_1\} &= \mathbb{P}\{AT_1 > M\} \\ &= \sum_{m \geq 1} \underbrace{\mathbb{P}\{AT_1 > m \mid M = m\}}_{\substack{\text{marker} \\ \text{ineq}} \rightarrow \leq \frac{\mathbb{E}[AT_1]}{m} \approx \frac{CN \ln N}{m}} \cdot \mathbb{P}\{M = m\} \\ &\leq (uN \ln N) \sum_{m \geq 1} \frac{(1-u)^{m-1}}{m} \\ &= \frac{uN \ln N}{1-u} \sum_{m \geq 1} \frac{(1-u)^m}{m} \\ &= (uN \ln N) \int_0^u \frac{1}{1-x} dx \end{aligned}$$

$$\begin{aligned}
&= \frac{C u N \ln N}{1-u} \sum_{m \geq 1} \int_u^1 (1-x)^{m-1} dx \\
&= \frac{C u N \ln N}{1-u} \sum_{k \geq 0} \int_u^1 (1-x)^k dx \\
&= \frac{C u N \ln N}{1-u} \int_u^1 \underbrace{\left(\sum_{k \geq 0} (1-x)^k \right)}_{= \frac{1}{1-(1-x)} = \frac{1}{x}} dx \\
&= \frac{C u N \ln N}{1-u} \int_u^1 \frac{1}{x} dx \\
&= \frac{C u N \ln N}{1-u} \left[\ln |x| \right]_u^1 = 0 - \ln u = \ln \frac{1}{u} \\
&= \frac{C u \ln \frac{1}{u}}{1-u} \cdot N \ln N
\end{aligned}$$

this converges since $|1-x| < 1$ $\forall x \in [u, 1]$

$$\text{So } \underbrace{\mathbb{P}\{M < AT_1\}}_{\text{"bad"}} \leq \frac{C u \ln \frac{1}{u}}{1-u} \cdot N \ln N$$

$$\begin{aligned}
\text{So if } u < \frac{1}{Nd}, \text{ then } \mathbb{P}\{\text{bad}\} &\leq \frac{\frac{C}{Nd} \ln Nd}{1 - \frac{1}{Nd}} N \ln N \\
&\stackrel{\downarrow}{=} \frac{1-u \geq 1 - \frac{1}{Nd}}{=} \frac{d C N (\ln N)^2}{\frac{Nd-1}{Nd}} \\
&= \frac{d C N (\ln N)^2}{Nd-1} \\
&= o(1) \text{ when } d > 1
\end{aligned}$$

So if $u < \frac{1}{N}$, then we typically only have at

most two types in the population at a time.

Kimura rate of neutral evolution (revisited)

- $u > 0$ is mutation probability (to brand new mutant)
- N is population size
- p_r is fixation probability of a single mutant w/ fitness r
- K is the # of mutant fixations per generation (N time steps)

Kimura's argument:

- Nu is (expected) # new mutants per generation
- When $r=1$, mutants have probability $p_{r=1} = 1/N$ chance of fixing

$$\Rightarrow K = Nu \cdot p_{r=1} = Nu \cdot (1/N) = u$$

$$\boxed{K = u}$$

So rate of mutant fixations equals mutation rate in neutral evolution

But this argument uses the fact $p_{r=1} = 1/N$ which is a result w/o mutation.

Using infinite alleles model for mutation, mutant substitution becomes more difficult.

Also exasperated by fixation/absorption time:
 "time" of a mutant was "instantly"

if the fate of a mutation is realized, no concerns.

Imagine case where $u \approx 1$. Then every birth likely gives rise to a new mutant. So fixation is extremely difficult because

$$p_i^* = (1-u) \frac{i}{N} \cdot \frac{N-i}{N} \approx 0,$$

Note that if $r > 1$, then Kimura's calculation

$$k = Nu \cdot p_{fix} = Nu \cdot \frac{1 - \frac{1}{r}}{1 - \frac{1}{r^2}} \quad \left(\begin{array}{l} \text{depends on} \\ \text{pop. size,} \\ \text{selective advantage,} \\ \text{\& mutation rate!} \end{array} \right)$$

But in data, we see "constancy" in rate of evolution. So Kimura argues this means most fixated mutants are neutral.



Time in the Moran process

```
In [ ]: import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
import random
import math
import enum
from dataclasses import dataclass
```

```
In [ ]: class Type(enum.Enum):
    A = 0
    B = 1

@dataclass(frozen=True)
class SimulationResult:
    N: int
    fixation: bool
    time: int
```

```
In [ ]: def simulate(N):
    i = 1 # number of type A individuals
    t = 0 # time
    while 0 < i < N:
        birther_type, = random.choices([Type.A, Type.B], weights=(i, N-i))
        dier_type, = random.choices([Type.A, Type.B], weights=(i, N-i))
        if birther_type == Type.A and dier_type == Type.B:
            i += 1
        if birther_type == Type.B and dier_type == Type.A:
            i -= 1
        t += 1
    return SimulationResult(N=N, fixation=i==N, time=t)
```

```
In [ ]: N = 100
TRIALS = 1000
data = [
    simulate(n)
    for n in range(1, N+1)
    for _ in range(TRIALS)
]
df = pd.DataFrame(data)
```

```
In [ ]: df.sample(n=10)
```

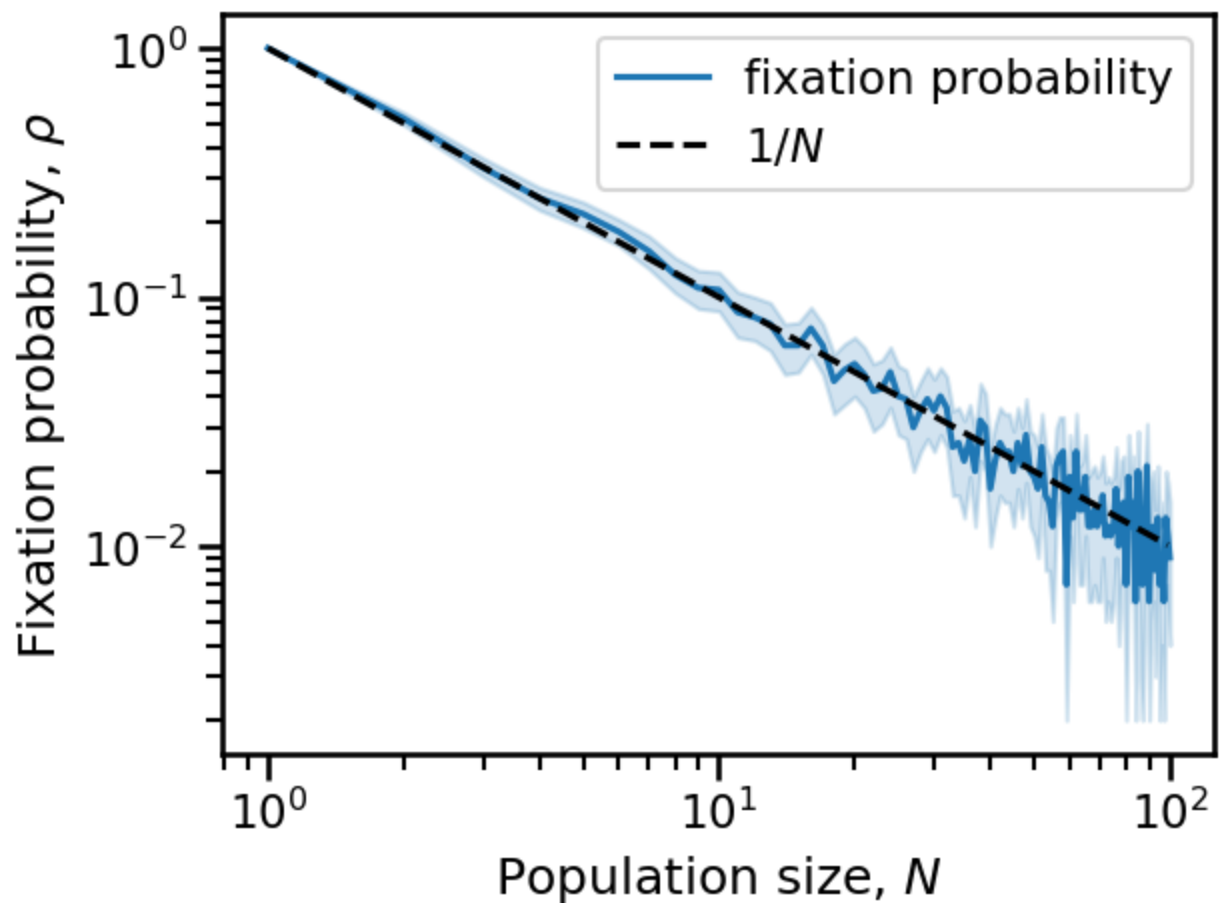
Out[]: **N** **fixation** **time**

74471	75	False	44
450	1	True	0
29843	30	False	76
13888	14	False	27
82113	83	False	321
23766	24	False	15
67306	68	False	9
54063	55	False	106
10483	11	False	30
62505	63	False	1448

```
In [ ]: with sns.plotting_context(context='talk'):
    g = sns.lineplot(
        data=df,
        x='N',
        y='fixation',
        label='fixation probability',
    )

    Ns = np.arange(1, N)
    sns.lineplot(
        x=Ns,
        y=1/Ns,
        label='$1/N$',
        ax=g,
        color='black',
        linestyle='--',
    )

    _ = g.set(
        xscale='log',
        yscale='log',
        xlabel='Population size, $N$',
        ylabel='Fixation probability, $\rho$',
    )
```



```
In [ ]: with sns.plotting_context(context='talk'):
    g = sns.lineplot(
        data=df,
        x='N',
        y='time',
        label='absorption time',
    )
    sns.lineplot(
        data=df[df['fixation']==True],
        x='N',
        y='time',
        label='fixation time',
        ax=g,
    )

    Ns = np.arange(1, N)
    sns.lineplot(
        x=Ns,
        y=Ns*np.log(Ns),
        label='$N \cdot \ln N$',
        ax=g,
        color='black',
        linestyle='--',
    )
    sns.lineplot(
```

```

x=Ns,
y=Ns**2,
label='$N^2$',
ax=g,
color='black',
linestyle='-.-',
)

= g.set(
xscale='log',
yscale='log',
xlabel='Population size, $N$',
ylabel='Time, $T$',
)

```

