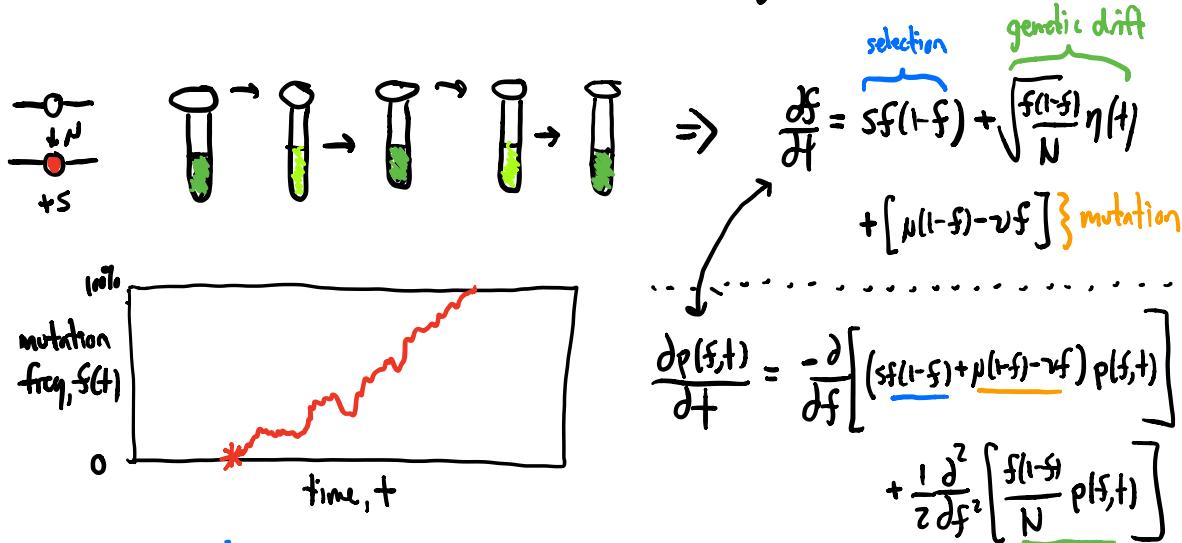


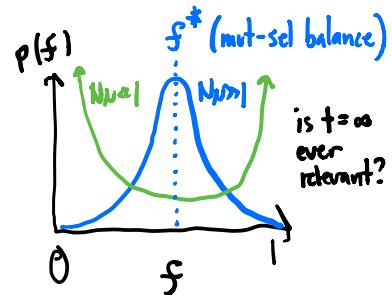
Announcements: PSET 2 updated to fix typo (thx Olivia + Xiran)

Last time: Quick review - how did we get here?

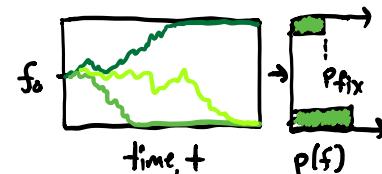


① "Moment hell": $\frac{\partial \langle f \rangle}{\partial t} = s \langle f \rangle - \langle f^2 \rangle$, etc...

② Stationary distribution $p(f,t) \xrightarrow{t \rightarrow \infty} p(f) \propto f^{2N\mu-1} (1-f)^{2N\nu-1} e^{-2Ns f}$
"mutation-selection-drift balance"

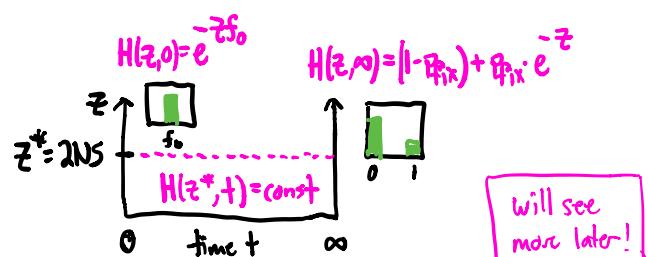


③ Fixation probability: $p_{fix}(N, s, f_0) = \frac{1 - e^{-2Ns f_0}}{1 - e^{-2Ns}}$
(no mutation, $\mu = \nu = 0$) "Kimura formula"



Generating function: $H(z,t) = \langle e^{-z f(t)} \rangle$

$\frac{\partial H}{\partial t} = \left[sz - \frac{z^2}{2N} \right] \left[\frac{\partial H}{\partial z} + \frac{\partial^2 H}{\partial z^2} \right] \Rightarrow$



Today: what can we learn from $P_{\text{fix}}(N, s, f_0) = \frac{1 - e^{-2Ns f_0}}{1 - e^{-2Ns}}$?

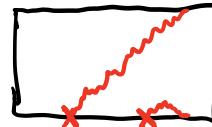
\Rightarrow battle between selection & genetic drift ($N \times s$)

a) If $N|s| \ll 1 \Rightarrow P_{\text{fix}} \approx f_0$ ("drift wins") ("weak selection" / "neutrality")

b) If $N|s| \gg 1$ ("strong selection")

$$P_{\text{fix}}(Ns, f_0) \approx \begin{cases} 1 & \text{if } s > 0; f_0 > \frac{1}{2Ns} \xrightarrow{\text{"selection wins"}} \\ 2Ns f_0 & \text{if } s > 0; f_0 < \frac{1}{2Ns} \xrightarrow{\text{"outcome uncertain...}} \\ e^{-2Ns|s|(1-f_0)} & \text{if } s < 0; \xrightarrow{\approx 0} \text{"selection wins"} \end{cases}$$

e.g. extrapolate to new mutation ($f_0 = \frac{1}{N}$)



$\Rightarrow P_{\text{fix}} \approx 2s$ (independent of N !) "Haldane's formula" (Haldane 1930s)

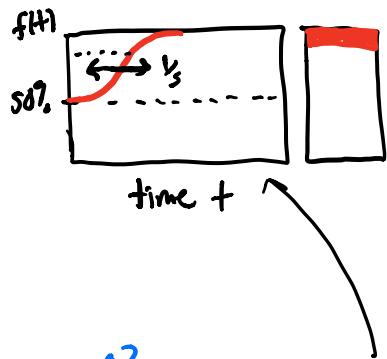
e.g. if $N = 10^7$ (per gen.) $s \approx 0.01$ \Rightarrow only 2% chance that mutation fixes!

(pretty beneficial
on lab timescales!)

\Rightarrow 98% of these mutations go extinct due to genetic drift!

Problem 4, Part 2

\Rightarrow but same mutant mixed @ 50-50
will rapidly + consistently take over!



what's going on here?

naively, as $N \rightarrow \infty$: $\frac{df}{dt} = sf(1-f) + \sqrt{\frac{f(1-f)}{N}} \eta(t)$ $\Rightarrow f(t) = \frac{f(0)e^{st}}{f(0)e^{st} + 1 - f(0)}$

(deterministic solution)

$$P_{\text{fix}} \approx \begin{cases} 1 & \text{if } s > 0 \\ 0 & \text{if } s \leq 0 \end{cases}$$

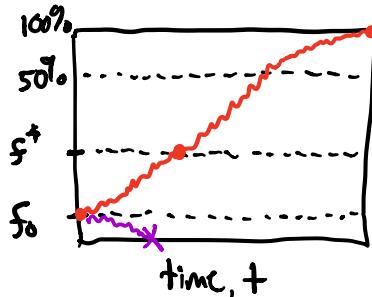
Somehow wrong...

How can we understand this?

Note: $P_{\text{fix}} = \frac{1 - e^{-2Ns f_0}}{1 - e^{-2Ns}} \approx 1$ when $f_0 > \frac{1}{2Ns}$, even when $f_0 \ll 1$

\Rightarrow i.e. outcome only uncertain when $f_0 \lesssim \frac{1}{2Ns}$ ($\ll 1$ when $Ns \gg 1$)

\Rightarrow break into two parts: $f_0 \rightarrow f^*$ + $f^* \rightarrow 1$ ($w/ f^* \ll 1$)

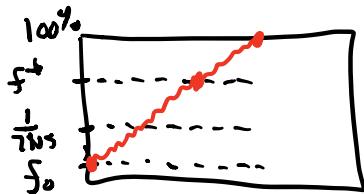


\Rightarrow must have

$$P_{\text{fix}}(f_0) = \Pr(f_0 \rightarrow f^*) \times P_{\text{fix}}(f^*)$$

$$\Rightarrow \Pr(f_0 \rightarrow f^*) = \frac{\Pr_{\text{fix}}(f_0)}{\Pr_{\text{fix}}(f^*)} \approx \frac{2Ns f_0}{\Pr_{\text{fix}}(f^*)} \quad \text{if } f_0 \ll \frac{1}{2Ns}$$

① if $f^* \gg \frac{1}{2Ns}$ $\Rightarrow \Pr(f_0 \rightarrow f^*) \approx 2Ns f_0 \approx \Pr_{\text{fix}}(f_0)$



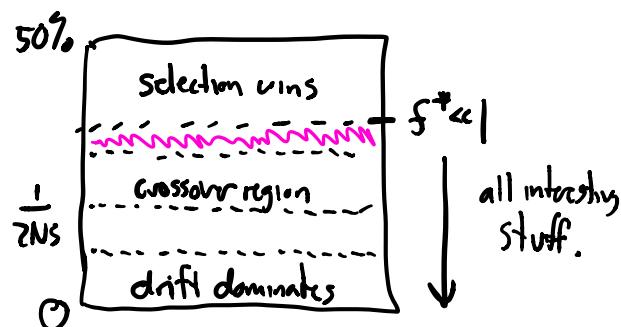
[all uncertainty in mut'n's fate
takes place between $0 \leq f \leq f^*(\ll 1)$]

\Rightarrow i.e. "selection wins" if $f(t) \gg \frac{1}{2Ns}$

② if $f^* \ll \frac{1}{2Ns}$ $\Rightarrow \Pr(f_0 \rightarrow f^*) = \frac{2Ns f_0}{2Ns f^*} = \frac{f_0}{f^*} \quad (\text{independent of } s!)$

\Rightarrow i.e., similar to neutral mutation for $f(t) \ll \frac{1}{2Ns} \ll 1$

\Rightarrow interesting partitioning of frequency space:

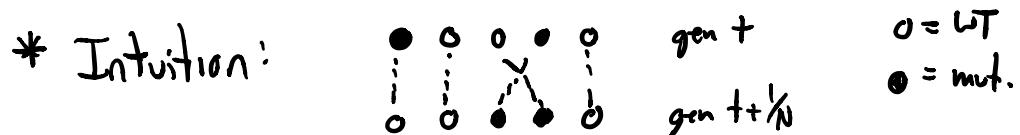


\Rightarrow important for evolution because new mutations: $f_0 = \frac{1}{N} \ll \frac{1}{Ns}$

when $f \ll 1$: evolutionary model reduces to

$$\boxed{\frac{df}{dt} = sf + \sqrt{\frac{f}{N}}\eta(t) + [n - nf]}$$

known as
"linear branching
process"



\Rightarrow Let's first consider no mutations, $n=2=0$

\Rightarrow SDE is linear (by construction)

e.g. for $\langle f(t) \rangle \Rightarrow d_t \langle f \rangle = s \langle f \rangle$ (no moment hell!)

$$\Rightarrow \langle f(t) \rangle = f(0) e^{st}$$

(same as deterministic model)

\Rightarrow e.g. for 1-way mutation ($n>0, v=0$)

$$\Rightarrow d_t \langle f \rangle = s \langle f \rangle + n \Rightarrow \langle f(t) \rangle = f(0) e^{st} + \frac{n}{s} (e^{st} - 1)$$

$$\Rightarrow \text{e.g. } s < 0 \Rightarrow \langle f(t) \rangle \approx \frac{n}{|s|} \left(\begin{array}{l} \text{just like} \\ \text{deterministic mut-} \\ \text{sel balance} \end{array} \right)$$

\Rightarrow can extend to higher moments:

e.g. $N=0 \Rightarrow$ can show that $\frac{d\langle f^2 \rangle}{dt} = 2s\langle f^2 \rangle + \frac{\langle f \rangle}{N}$

↑
from deterministic part.
↑
from noise term.

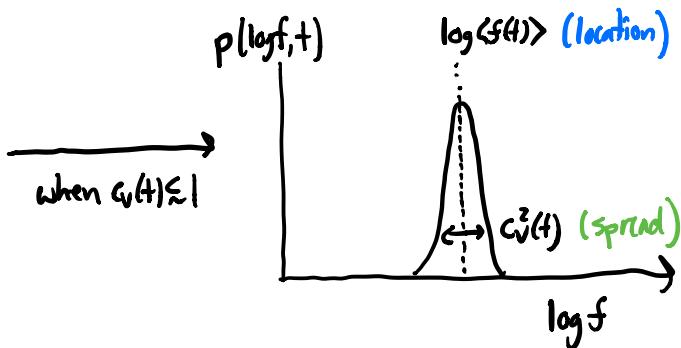
↑
know this
from before.

$$\Rightarrow \langle f(t)^2 \rangle = f_0^2 e^{2st} + \frac{f_0 e^{st}(e^{-st})}{Ns}$$

$$\Rightarrow \text{Coefficient of Variation} \quad C_v^2(t) \equiv \frac{\text{Var}(f(t))}{\langle f(t) \rangle^2} = \frac{\langle f^2 \rangle - \langle f \rangle^2}{\langle f \rangle^2} = \frac{1 - e^{-st}}{Ns f_0}$$

useful for visualizing spread of dist'n in log space

(i.e. how uncertain @ order of magnitude level?)

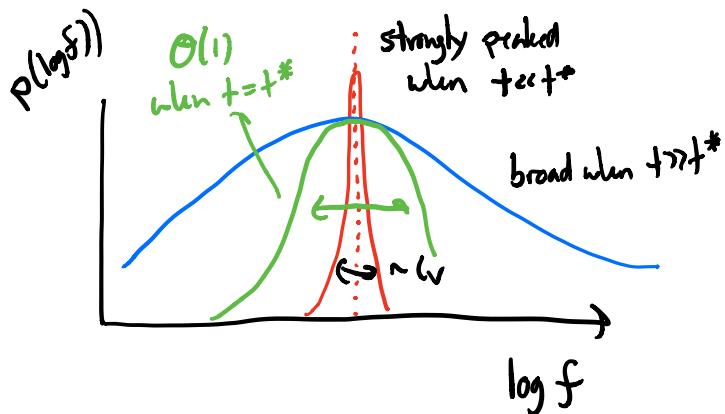


$$\textcircled{1} \quad \text{when } s > 0 \text{ + } f_0 \gg \frac{1}{Ns} \Rightarrow C_v(t) \leq \frac{1}{Ns f_0}$$

≈ 1 for all time!

② In contrast, when $f_0 \ll \frac{1}{Ns}$, (or if $s \leq 0$)

$$C_V(t) = \frac{1 - e^{-st}}{Ns f_0} \approx \begin{cases} \sim 1 & \text{if } t \ll t^* \\ \gg 1 & \text{if } t \gg t^* \end{cases} \quad \text{for some } t^*$$



can solve for t^* by setting $C_V^2(t^*) = \frac{1 - e^{-st^*}}{Ns f_0} \sim 1$

can solve for t^* :

$$t^* \sim \begin{cases} \infty & \text{if } s > 0, f_0 \gg \frac{1}{Ns} \\ Nf & \text{if } f_0 \ll \frac{1}{Ns} \\ \frac{1}{s} \log(Ns/f_0) & \text{if } s < 0; f_0 \gg \frac{1}{Ns} \end{cases}$$

will discuss
the intuition
behind these
expressions later!

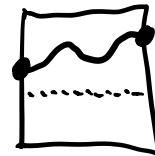
\Rightarrow ideally solve for full dist'n of $f(t)$:

e.g. Fokker-Planck eq: $\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial f} \left[s f \rho \right] + \frac{1}{2} \frac{\partial^2}{\partial f^2} \left[\frac{f}{N} \rho \right]$
 (hard!)

\Rightarrow easier to focus on MGF: $H(z,t) \equiv \langle e^{-zf(t)} \rangle = \int e^{-zf} \rho(f,t) df$

\hookrightarrow satisfies $\frac{\partial H}{\partial t} = \left[sz - \frac{z^2}{2N} \right] \frac{\partial H}{\partial z}$ w/ $H(z,0) = e^{-zf_0}$

solved via "method of characteristics"



(p. 4 of notes)

\Rightarrow can show that $H(z,t) = \exp \left[-\frac{f_0 z e^{st}}{1 + \frac{z}{2Ns} (e^{st} - 1)} \right]$

\hookrightarrow formally, can invert to obtain $\rho(f,t)$ (tricky)
 (hard to interpret)

\Rightarrow can learn a lot by focusing on $H(z,t)$ directly.

e.g. recall: Expand in powers of \bar{z} :

$$H(\bar{z}, t) \approx 1 - \bar{z} \langle f(t) \rangle + \frac{\bar{z}^2}{2} \langle f(t)^2 \rangle + \dots$$

$$\approx 1 - \bar{z} f_0 e^{st} \dots$$

can rewrite generating function in suggestive form:

$$H(\bar{z}, t) = \exp \left[- \frac{\bar{z} \langle f(t) \rangle}{1 + \frac{\bar{z}}{2} \langle f(t) \rangle \cdot c_V^2(t)} \right] \Leftarrow \begin{matrix} \text{not gaussian!} \\ H = e^{-\mu \bar{z} + \frac{1}{2} \sigma^2 \bar{z}^2} \end{matrix}$$

But, for $t \ll t^*$ [$c_V^2(t) \ll 1$]

$$H(\bar{z}, t) \approx \exp \left[- \bar{z} \langle f(t) \rangle + \frac{\bar{z}^2}{2} \underbrace{\langle f(t) \rangle^2}_{\text{Var}(f(t))} c_V^2(t) \right] \rightarrow \begin{matrix} \text{will be} \\ \text{a Gaussian} \\ (\text{in bulk}) \\ \omega / \text{mean } \langle f(t) \rangle \\ \propto c_V \ll 1 \end{matrix}$$

("case 1" dist'n)
mean + spread.

outside this regime: $H(z, t) \equiv \langle e^{-z \cdot f(t)} \rangle$

"time dependent
extinction
prob"

$$\text{when } z \rightarrow \infty \quad H(z, t) \approx 0 \times (1 - P_{\text{ext}}(t)) + 1 \times P_{\text{ext}}(t)$$

$$\approx P_{\text{ext}}(t)$$

$$\Rightarrow \text{In our case: } P_{\text{ext}}(t) = \exp \left[- \frac{2Ns f_0}{1 - e^{-st}} \right] = \exp \left[- \frac{z}{c_v^2(t)} \right]$$

can also define survival probability:

$$P_{\text{survive}}(t) \equiv 1 - P_{\text{ext}}(t) \xrightarrow{t \rightarrow \infty} P_{\text{fix}}(N, s, f_0)$$