$$\boxed{\S \, 2 \quad \text{Notes}}$$

2.1.) Allele = a piece of DNA (one of $2N$ alleles)

$$Pr(\text{shows up}) = 1 - Pr(\text{does not show up})$$
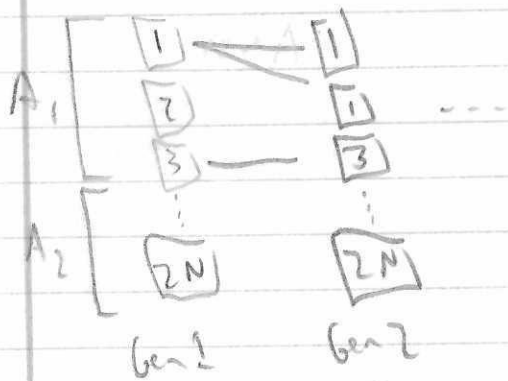
$$= 1 - \left(\frac{2N-1}{2N}\right)^{2N}$$

$$= 1 - \left(1 - \frac{1}{2N}\right)^{2N} \qquad \text{recall: } \left(1 + \frac{1}{N}\right)^{N} \xrightarrow[N \to \infty]{} e$$

$$\xrightarrow[N \to \infty]{} = 1 - e^{-1} \qquad\qquad \left(1 - \frac{1}{2N}\right)^{2N} \to e^{-1}$$

- Wright - Fisher Model:

A_1 ... A_2

Gen 1   Gen 2

sample w/ replacement
from previous generation

- Measure allele frequencies over the.

- Genetic Drift — change of allele frequencies ver. time
  ① Random
  ② Removes alleles
  ③ Direction of change is neutral

- Ex.) Pop. w/ 1 $A_1 A_2$ individual

Gen 1

$Pr(A_1 A_1) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$

$Pr(A_1 A_2) = Pr(A_2 A_1 \text{ or } A_1 A_2) = 2\left(\frac{1}{2}\right)^2 = \frac{1}{2}$

$Pr(A_2 A_2) = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$

Gen 2

$Pr(A_1 A_1) = \underbrace{\frac{1}{4} \cdot 1}_{\substack{A_1 A_1 \\ \text{in gen1}}} + \underbrace{\frac{1}{2} \cdot \frac{1}{4}}_{\substack{A_1 A_2 \\ \text{in gen1}}} + \underbrace{\frac{1}{4} \cdot 0}_{\substack{A_2 A_2 \\ \text{in gen1}}} = \frac{3}{8}$

$Pr(A_1 A_2) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot 0 = \frac{1}{4}$

$Pr(A_2 A_2) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot \frac{1}{4} + \frac{1}{4} \cdot 1 = \frac{3}{8}$

Gen 3

$Pr(A_1 A_2) = \underbrace{\frac{1}{4} \cdot \frac{1}{2}}_{\substack{A_1 A_2 \\ \text{in gen2}}} = \frac{1}{8}$

Gen 4

$Pr(A_1 A_2) = \frac{1}{8} \cdot \frac{1}{2} = \frac{1}{16}$

$P_r(\text{Hetero}) = \left(\frac{1}{2}\right)^t$ $t = $ generation #

- Let $H_t = $ Probability that two alleles drawn randomly from population differ by state

Here: $H_t = \left(\frac{1}{2}\right)^t$

- In general $H_t = \underline{H_0}\left(1 - \frac{1}{2N}\right)^t$ (let...)

$\uparrow$

initial probability of Hetero

2.3.) Let $X = $ the to homozygosity

$$E[X] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + \cdots + t\left(\frac{1}{2}\right)^t + \cdots$$

$$= \sum_{t=1}^{\infty} t\left(\frac{1}{2}\right)^t$$

$$= 2$$

$\uparrow$

mathematica or Wolfram α

- <u>Rate of Decay of Heterozygosity:</u>

- N diploid hermaphrodites

- $\mathcal{G} = Pr$ (two alleles identical by state | differ by origin)
  $\mathcal{L}$ almost same as homozygosity $= Pr$ (two alleles identical by state | randomly sampled <u>with</u> replacement)

  $\mathcal{G} = 1 \Rightarrow$ one state
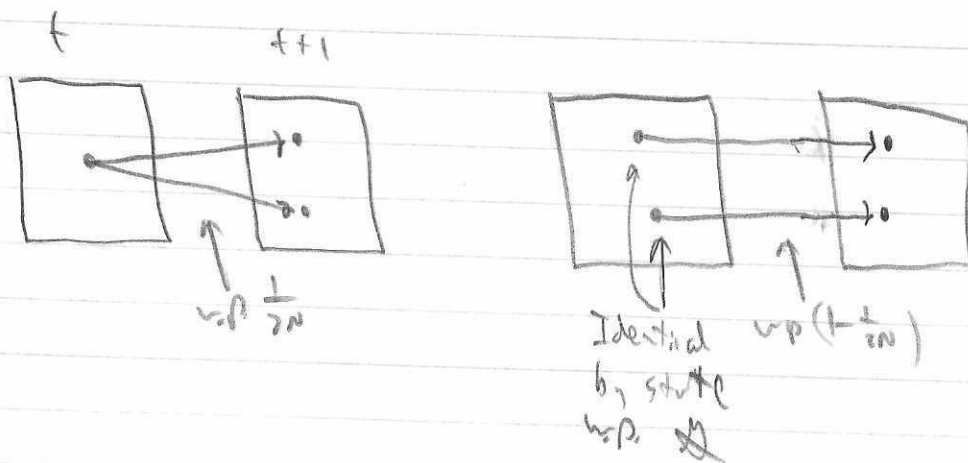 $\mathcal{G} = 0 \Rightarrow$ all differ by state

- $\mathcal{G}' = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\mathcal{G}$

  Next Generation

  previous generation

  $\mathcal{G}' = Pr(\text{come from same allele}) + Pr(\text{Not same allele}) \, Pr(\text{Previous gen identical})$

  $= \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\mathcal{G}'$

- Fix one allele in next generation

  t $\qquad$ $t+1$

  w.p. $\frac{1}{2N}$

  Identical by state w.p. $\mathcal{G}$ $\qquad$ w.p. $\left(1 - \frac{1}{2N}\right)$

- Let $H = 1 - \vartheta$

$$= P_r(\text{differ by state} \mid \text{differ by origin})$$

$$H' = 1 - \left(\frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\vartheta\right)$$

$$= \left(1 - \frac{1}{2N}\right)\left(1 - \vartheta\right)$$

$$= \left(1 - \frac{1}{2N}\right) H$$

$$= H - \frac{1}{2N} H$$

$$\Rightarrow \Delta_N H = H' - H = -\frac{1}{2N} H$$

    ↑ change in H

- $H \approx$ heterozygosity $= \boxed{1 - \sum p_n^2}$

                assumes ∧∧∧∧∧
                pop size

- Let H_t be heterozygosity after t generations

$$H_t = H_0 \left(1 - \frac{1}{2N}\right)^t$$

↑ heterozygosity goes to 0

↑ ie, one allele will eventually get fixed.

- let $\frac{1}{2} H_0 = H_0 \left(1 - \frac{1}{2N}\right)^t$ Time to halve heterozygosity

Solve for t $\Rightarrow$ $t = -\dfrac{\log(2)}{\log\left(1 - \frac{1}{2N}\right)}$

Note: $\log(1+x) \approx x$ for **small** x

$\Rightarrow t \approx \dfrac{-\log(2)}{-\frac{1}{2N}} = 2N \log(2)$

Taylor series

$\log(1+x) = \log(1) + \frac{1}{1-x}\Big|_{x=0} x +$ remainder

$\hspace{2cm} = x$

↳ large pop $\Rightarrow$ larger time to decay by half

- **Fixation Probability:**

Fixation = one allele becomes only allele

$\hspace{2cm}$ = Allele frequency

↳ Exactly one allele must be fixed, all have equal probability

- Homozygosity = Pr(Identical by state | two drawn w/ replacement)

$$= G$$

Let p_i be allele frequency of allele i

$$\sum_{v=1}^{\hat{n}} \underbrace{p_j \cdot p_i}_{\substack{\text{allele 1} \\ \text{and allele 2} \\ \text{sue = independent}}}$$

↗ disjoint events

Claim: $G = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right) \mathcal{G}$

proof: $\underbrace{Pr(\text{sme by origin})}_{\frac{1}{2N}} + \underbrace{Pr(\text{differ by origin})}_{\left(1 - \frac{1}{2N}\right)} \underbrace{Pr(\text{sme state | differ by orig})}_{\mathcal{G}}$ //

For moderate N, $G \approx \mathcal{G}$ and $H \approx \mathcal{H}$

- HWE still valid even w/ drift

    ⌐ HWE on the order of 1 generation
    ⌐ Drift on order of $2N$ generations
    ⌐ ie. deviation from HWE is $O\left(\frac{1}{2N}\right)$