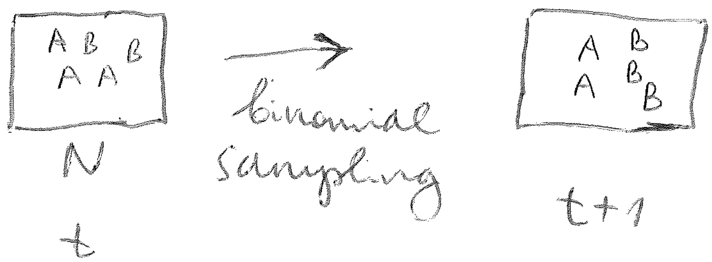
Kimura, 1962 {Lecture 8}

Fixation of mutant genes in a population



Consider a population of N genes in which the freq. of the allele A is p $(0 \le p \le 1)$.
The freq. of allele B is then $q = 1-p$

Define $u(p,t)$ = prob. that A is fixed (i.e. its freq. $= 1$) in $\le t$ generations (continuous time approx'n)

$$u(p, t+\delta t) = \int d(\delta p)\, \underbrace{f(p, p+\delta p\,;\, \delta t)}_{\substack{\text{prob. density} \\ p \to p+\delta p \text{ in } \delta t}} \times u(p+\delta p, t)$$

Introduce

$$\begin{cases} \lim\limits_{\delta t \to 0} \dfrac{1}{\delta t} \displaystyle\int (\delta p)\, f(p, p+\delta p\,;\, \delta t)\, d(\delta p) \equiv M, \\ \lim\limits_{\delta t \to 0} \dfrac{1}{\delta t} \displaystyle\int (\delta p)^2\, f(p, p+\delta p\,;\, \delta t)\, d(\delta p) \equiv V \end{cases}$$

Just as before, we obtain:

$$\frac{\partial u(p,t)}{\partial t} = \frac{V}{2}\frac{\partial^2 u}{\partial p^2} + M\frac{\partial u}{\partial p}$$

BCs: $u(0,t)=0, \quad u(1,t)=1$

Measure time in generations $\to$

$\to$ M, V are the mean & variance of the change of p ~~per~~ generation.

Consider $u(p) = \lim_{t\to+\infty} u(p,t)$:

$$\frac{\partial u}{\partial t} = 0 \Rightarrow \frac{V}{2}\frac{d^2u}{dp^2} + M\frac{du}{dp} = 0. \quad (*)$$

"steady state" $\quad u(0)=0, \quad u(1)=1$

(*) can be solved: $u(p) = \dfrac{\int_0^p G(x)\,dx}{\underbrace{\int_0^1 G(x)\,dx}_{=C_1}}$

Indeed, $\frac{du}{dp} = \frac{G(p)}{C_1}$; $\frac{d^2u}{dp^2} = G(p)\left[-\frac{2M(p)}{V(p)}\right]\frac{1}{C_1}$

$-G(p)M(p)\frac{1}{C_1} + \frac{M(p)G(p)}{C_1} = 0$, as expected!

$$G(x) = e^{-\int_0^x \frac{2M(y)}{V(y)}\,dy}$$

Note that $u(\frac{1}{N})$ is a chance of fixation of a "novel" mutant gene.

Now, for the binomial distr'n we have: $\begin{cases} \mu = np, & n = \text{\# trials},\ p = \text{prob. of success} \\ \sigma^2 = npq \end{cases}$

Prob. (i alleles of type A) =

$$= \frac{N!}{i!(N-i)!} \underset{\substack{\uparrow \\ \text{freq. of} \\ \text{allele A}}}{p^i}(1-p)^{N-i} .$$

No selection:

$$E(\Delta p | p) = \underbrace{\frac{E(i)}{N}}_{\substack{\text{new} \\ \text{freq,} \\ \text{on average}}} - \underbrace{p}_{\substack{\text{old} \\ \text{freq}}} = \frac{Np}{N} - p = 0$$

$$Var(\Delta p | p) = Var\left(\frac{i}{N} - p | p\right) = \frac{Var(i)}{N^2} = \frac{Np(1-p)}{N^2} =$$

$$= \frac{p(1-p)}{N} .$$

With selection,

	A	B	
	$1+s$	1	$\leftarrow$ prob. of survival
before selection $\Rightarrow$	p	q	

after selection $\Rightarrow$ $p' = \dfrac{p(1+s)}{p(1+s)+q}$ $\qquad \dfrac{q}{p(1+s)+q} = q'$

$p' + q' = 1$

$$p' - p \equiv \Delta_s p = \frac{p(1+s) - p^2(1+s) - pq}{p(1+s)+q} =$$

$$= \frac{pqs}{\underbrace{p(1+s)+q}_{1+ps \approx 1 \text{ if } s \text{ is small}}} \approx pqs .$$

So, $E(\Delta p | p) = sp(1-p)$.

In our notation, $\frac{2M(y)}{V(y)} = 2NS$ &

$G(x) = e^{-2N sx}$

$$\overset{\shortparallel}{\frac{2\,[Spq]}{pq/N}}$$

Finally, $u(p) = \frac{1-e^{-2NSp}}{1-e^{-2NS}}$. Kimura's formula

$$u\left(\frac{1}{N}\right) = \frac{1-e^{-2S}}{1-e^{-2NS}} \approx \frac{2S}{1-e^{-2NS}}.$$

$$\lim_{S\to 0} u\left(\frac{1}{N}\right) = \frac{2S}{2NS} = \frac{1}{N}$$, as expected.

$u(p) = p$ is a well-known neutral result.

$\boxed{\text{Bollt, 2005}}$

Consider a complete graph (all nodes connected to one another):

(ex. $N=3$ nodes) [unweighted, undirected]

MFPT from node i to j:

$$T_{ij} = \underbrace{\frac{1}{N-1}}_{\text{go directly to } j} + \underbrace{\frac{N-2}{N-1}}_{\text{go to node } h \neq j}(1+T_{hj}) =$$

$$\frac{N-2}{N-1}N + \frac{1}{N-1} = \frac{N^2-2N+1}{N-1} = N-1, \text{ as expected}$$

$$= \frac{1}{N-1} + \frac{N-2}{N-1}(1+T_{ij}) \Rightarrow \boxed{T_{ij} = N-1}$$

$$\text{MRT:} \quad T_{ii} = \underbrace{1}_{\text{must step off}} + \overbrace{T_{ji}}^{N-1} = N$$

Furthermore, let $F_{ij}(t)$ be the prob. that FPT $i \to j$ is exactly t. Then

$$F_{ij}(t) = \frac{1}{N-1}\delta_{t,1} + \overbrace{\frac{N-2}{N-1}}^{\text{prob. not to reach } j} F_{ij}(t-1) \quad (**)$$

Introduce a generating f'n:

$$\hat{F}_{ij}(x) = \sum_{t=0}^{\infty} F_{ij}(t)\, x^t$$

Then we can act with

$\sum_{t=0}^{\infty} x^t$ on $(**)$:

$$\hat{F}_{ij}(x) = \frac{x}{N-1} + \sum_{t=0}^{\infty} x\, x^{t-1} \frac{N-2}{N-1} F_{ij}(t-1) =$$

$$= \frac{x}{N-1} + \frac{N-2}{N-1} x \underbrace{\sum_{t'=0}^{\infty} x^{t'} F_{ij}(t')}_{\hat{F}_{ij}(x)}$$

Then $\hat{F}_{ij}(x) = \dfrac{x}{(N-1) - (N-2)x}$

Taylor expand to read off $F_{ij}(t)$:

$$\hat{F}_{ij}(x) \simeq \frac{x}{N-1}\left(1 + \frac{N-2}{N-1} x + \dots\right) \Rightarrow \begin{cases} F_{ij}(1) = \dfrac{1}{N-1} \\ F_{ij}(2) = \dfrac{N-2}{(N-1)^2} \\ \dots \end{cases}$$

Note that $T_{ij} = \sum_{t=0}^{\infty} t\, F_{ij}(t) = N-1$, as expected