

8/28/13

(0)

1-focus selection: a deterministic model

Q: How do we know the fate of a one-time mutation?

A: Consider clonal pop. of size $2N_0$ w/
fit growth rate W per time step

• single mutant arises with g.r. $W_m = W(1+s)$ ($s > -1$)

• after t time steps:

$$\# \text{ mutants} = 1 \cdot W_m^t = 1 \cdot W^t (1+s)^t$$

$$\# \text{ non-muts} = (2N_0 - 1) W^t$$

$$\begin{aligned} \text{pop size} = 2N_t &= \# \text{ muts} + \# \text{ non-muts} \\ &= 1 \cdot W^t (1+s)^t + (2N_0 - 1) W^t \\ &= W^t [1 \cdot (1+s)^t + (2N_0 - 1)] \end{aligned}$$

$$\begin{aligned} \bullet \quad q_t = \text{mut. freq at } t &= \frac{\# \text{ muts}}{2N_t} = \frac{W^t (1+s)^t}{W^t [(1+s)^t + (2N_0 - 1)]} \\ &= \frac{1 \cdot (1+s)^t}{1 \cdot (1+s)^t + (2N_0 - 1)} \end{aligned}$$

$$\Rightarrow \text{freq. non-muts} = 1 - q_t = \frac{2N_0 - 1}{1 \cdot (1+s)^t + (2N_0 - 1)}$$

$$\begin{aligned} \bullet \text{ Consider } \frac{q_t}{1 - q_t} &= \frac{1 \cdot (1+s)^t / [1 \cdot (1+s)^t + (2N_0 - 1)]}{(2N_0 - 1) / [1 \cdot (1+s)^t + (2N_0 - 1)]} \\ &= \frac{1}{2N_0 - 1} \cdot (1+s)^t = \frac{\frac{1}{2} 2N_0}{(2N_0 - \frac{1}{2} 2N_0)} (1+s)^t \\ &= \frac{q_0}{1 - q_0} (1+s)^t \end{aligned}$$

• If $s > 0$, $\frac{q_t}{1 - q_t} \rightarrow \infty$ as $t \rightarrow \infty$ "selective sweep"

• $-1 < s < 0$ $\frac{q_t}{1 - q_t} \rightarrow 0$ as $t \rightarrow \infty$ "selectively purged"

• Outcome independent of q_0 and W

• Conclude If we know s , can predict muta.'s fate with certainty

1/2/03

(1)

1-locus selection: Stochastic models

- Last time: Predict future allele freqs. with certainly given initial freqs & relative fitnesses
→ "Deterministic"

• 2 Problems w/ Deterministic Model

① Population & allele counts are integers → not all freqs. possible (e.g. $p = .75$ not possible with $N = 3$ or $2N = 6$)

- Possible soln: Round or "Integerize" recursions

- More natural soln solves 2nd problem:

② Survival, Reproduction, & Inheritance are random

e.g. Mendel's Rules $Aa \times Aa$ $\left\{ \begin{array}{l} \text{one offspring} \\ AA \text{ w.p. } 1/4 \\ Aa \text{ w.p. } 1/2 \\ \text{or } aa \text{ w.p. } 1/4 \end{array} \right\}$ Outcome certain

• Pop Gen Solution: Wright-Fisher Model (w/ selection)

Standard Version: for clones fitness $\frac{A}{1+s} \mid 1$

- Population size fixed & finite: $2N$

- includes A & a individuals

- Initial ~~Parent~~ genotypes: j A 's & $2N-j$ a 's

- After selection, expect $\frac{j(1+s)}{j(1+s) + (2N-j)}$ $\frac{j(1+s)}{j(1+s) + (2N-j)} = \frac{j(1+s)}{j(1+s) + 2N - j}$

- Form next generation by binomial sampling of $2N$ genotypes
 $X_t = \# A$ genes in $2N$

$$\text{Pr}(X_{t+1} = k \mid X_t = j) = \binom{2N}{k} \left(\frac{j(1+s)}{j(1+s) + 2N - j}\right)^k (1 - \frac{j(1+s)}{j(1+s) + 2N - j})^{2N-k} = g_{jk}$$

9/2/03

(2)

Comments

“Future of system depends only on current state”

1. Generates Markov Chain in the $X_t = \#$ of A alleles in gen. t
- since pop. size is fixed @ $2N$, can recover allele freq. at any time by dividing by $2N$.

2. “State Space” integers $0, 1, \dots, 2N$ ($= 2N+1$ values)
states 0 & $2N$ are absorbing in that $g_{00} = 0, k \neq 0; g_{2N} = 1$
 $g_{2Nk} = 0, k \neq 2N; g_{2N2N} = 1$

3. Outcomes uncertain in a given pop., but can compute the probabilities that a pop. is in a given state a gen. t by iteration:

$$Pr(X_t = k) = \sum_{j=0}^{2N} Pr(X_{t-1} = j) Pr(X_t = k | X_{t-1} = j) = \sum_{j=0}^{2N} Pr(X_{t-1} = j) g_{jk}$$

Math Note: Mathematicians will recognize this as matrix eqn.

Comp Note: Straightforward to iterate on a computer — provided N not too large.

Analysis Note Many (most?) properties of evol. interest are difficult to generalize from these recursions.

4. Diploid pop'n: same treatment (multinomial samplg of genotypes leads to binomial samplg of alleles)

5. W.F model is perhaps plausible biologically, but not realistic

• Amazing fact: This doesn't matter since just about any ^{real} population can be represented in this framework by a suitably-chosen “effective pop. size”

6. For selection, expected allele freqs. \neq deterministic allele freqs.

9/2/03

(3)

7. W-F model is "easy" to simulate:

Algorithm: a. If $X_t = j$, ~~set~~ compute $p_{jt}^* = \frac{jC(t+s)}{jS + 2N}$

b. $X_{t+1} =$ binomial r.v. with parameters $2N, p_{jt}^*$

$T = 2N$ trials, each of which has prob. of "success" ($\in A$) = p_{jt}^* (HW)

Prob. Fixation of a new mutation ~~HW~~

- Deterministic theory says $p \rightarrow 1$ if A favored ($s > 0$) and $p_0 \neq 0$ (even for a new mutation $p_0 = \frac{1}{2N}$)
- Stochastic theory: fixation not certain, especially for a new mutation

• Let u_i = ultimate prob. fixation when $p_{ij}(t) = \frac{i}{2N}$
 BCs $u_0 = 0$ and $u_{2N} = 1$
 For $i = 1, 2, \dots, 2N-1$

$$\begin{aligned} u_i &= g_{i0}u_0 + g_{i1}u_1 + \dots + g_{i2N}u_{2N} \\ &= g_{i1}u_1 + \dots + g_{i2N-1}u_{2N-1} + g_{i2N} \quad \text{since } u_0 = 0 \text{ \& } u_{2N} = 1 \\ &= \sum_{j=1}^{2N-1} g_{ij}u_j + g_{i2N} \end{aligned}$$

Set $\underline{u} = \{u_i\}_{i=1}^{2N-1}$, $\underline{Q} = \{g_{ij}\}_{i,j=1}^{2N-1}$, $\underline{r} = \{g_{i2N}\}_{i=1}^{2N-1}$

eqns $\Leftrightarrow \underline{u} = \underline{Q}\underline{u} + \underline{r} \Rightarrow \underline{u} = (\underline{I} - \underline{Q})^{-1}\underline{r}$

i.e. can compute fixation probs — in principle

- practical only for fairly small pops. (HW)
- difficult to assess impact of various factors (initial freq., strength of sel.)

• Better approach for larger pops? Branching Process or diffusion approximations.
 e.g. Prob fixation of new adv. allele $\approx 2s \ll 1$

9/4/03

(4)

Prob. Fixation of a new mutation: Branching Process Approximation

- Consider mutation A arising as single copy in large pop'n.
- A produces $1+s$ times more offspring than a (= others)
- Let $P = \text{prob. A will ultimately fix}$
- $\Rightarrow 1-P = \text{prob A leave no descendants}$

• If A leaves j offspring in the next generation, then all of these leave no descendants w.p. $(1-P)^j$

• Assume $\text{Pr}(j \text{ offspring}) = \text{Poisson distn. w/ mean } 1+s$

$\Rightarrow 1-P = \sum_{j=0}^{\infty} e^{-(1+s)} \frac{(1+s)^j}{j!} (1-P)^j = e^{- (1+s)P}$

recall: $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$

$\therefore P = \sqrt[1+s]{1-P}$

• From this can show $P \approx 2s$ e.g. If $s = .1$
 $P \approx .2$
 (any!)

"pt" $s = \frac{-\log(1-P) - P}{P} = \frac{P}{2} + \frac{P^2}{3} + \frac{P^3}{4} + \dots$

\Rightarrow If s small, P small. Ignoring terms of P^2 + higher
 $\Rightarrow P \approx 2s$.

$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$

Better (HW) Write $\ln(1-P) = -P(1+s) = -Pe^x$ where $x = sP(1+s)$

Expand both sides in TS. around $P + x$, resp.

$-P - \frac{P^2}{2} - \frac{P^3}{3} - \dots = -P(1+x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots)$

$P \neq 0 \Rightarrow 1 + \frac{P^2}{2} + \frac{P^3}{3} + \frac{P^4}{4} - \dots = 1+x + \frac{x^2}{2!} + \dots$

Now Assume $P = Ax + Bx^2 + Cx^3$; plug into lhs, match coeffs; solve for A, B, C