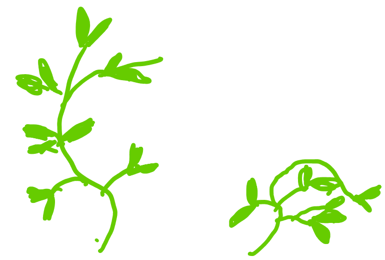
3.7 population genetics and Hardy-Weinberg

Saturday, February 20, 2021 12:53 PM

Mendelian genetics

- Genetics is the study of heredity and inherited variation
  - e.g. Hair color, eye color, pea plant height
- DNA (deoxyribonucleic acid) carries genetic information, in the form of genes
- Diploid organisms have two copies of every gene, sometimes slightly different, which are the variants of the gene called alleles
- With sexual reproduction, the offspring gets one copy from each parent, which causes mixing of the alleles.



Aa, AA aa

Punnett squares

	A	a
A	AA	Aa
A	AA	Aa

	A	a
A	AA	Aa
a	Aa	aa

Thm 3.2: Hardy-Weinberg Law

p_t = proportion of alleles in pop. at time t that are "A"

q_t = proportion at time t of "a"

Assume in a parent population, a particular gene has alleles A and a, with initial proportions p_0 and q_0. In addition:

1. Mating is random.
2. There is no variation in the number of progeny from parents of different genotypes
3. All genotypes have equal survival probability.
4. There is no immigration nor emigration.
5. There are no mutations.
6. Generations are nonoverlapping.

$$p = \frac{2NP_{AA} + NP_{Aa}}{2N} = P_{AA} + \frac{P_{Aa}}{2}$$

$$q = 1 - p = \frac{P_{Aa}}{2} + P_{aa}$$

Then in generation t, the allele frequencies do not change:

$$p_t = p_0 \text{ and } q_t = q_0.$$

Additionally, the genotypic frequencies do not change from the second generation onwards:

$$P_{AA} = p_0^2,\ P_{Aa} = 2p_0q_0, \text{ and } P_{aa} = q_0^2.$$

Proof of Hardy-Weinberg Law

Case 1: AA × AA

Frequency: $P_{AA} \cdot P_{AA} = P_{AA}^2$

Offspring:

	A	A
A	AA	AA
A	AA	AA

= all AA

Table 3.1

Mating	Frequency	Offspring Fraction AA	Aa	aa	Next generation AA	Aa	aa
AA x AA	P_{AA}^2	1	0	0	P_{AA}^2	0	0
AA x Aa	$2P_{AA}P_{Aa}$	1/2	1/2	0	$P_{AA}P_{Aa}$	$P_{AA}P_{Aa}$	0
AA x aa	$2P_{AA}P_{aa}$	0	1	0	0	$2P_{AA}P_{aa}$	0
Aa x Aa	P_{Aa}^2	1/4	1/2	1/4	$\frac{P_{Aa}^2}{4}$	$\frac{P_{Aa}^2}{2}$	$\frac{P_{Aa}^2}{4}$
Aa x aa	$2P_{Aa}P_{aa}$	0	1/2	1/2	0	$P_{Aa}P_{aa}$	$P_{Aa}P_{aa}$
aa x aa	P_{aa}^2	0	0	1	0	0	P_{aa}^2

$$P_{AA}' = P_{AA}^2 + P_{AA}P_{Aa} + \frac{P_{Aa}^2}{4} = \left(P_{AA} + \frac{P_{Aa}}{2}\right)^2 = p^2$$

$$P_{Aa}' = P_{AA}P_{Aa} + 2P_{AA}P_{aa} + \frac{P_{Aa}^2}{2} + P_{Aa}P_{aa}$$
$$= 2\left(P_{AA} + P_{Aa}/2\right)\left(P_{aa} + P_{Aa}/2\right) = 2pq$$

$$P_{aa}' = q^2$$

Case 2: AA × Aa

Frequency: $P_{AA}P_{Aa} + P_{Aa}P_{AA} = 2P_{AA}P_{Aa}$

Offspring:

	A	A
A	AA	AA
a	Aa	Aa

= 50% AA
50% Aa

Biological Fitness

Suppose survival rates to adulthood $\overbrace{w_{AA}, w_{Aa}, w_{aa}}^{\text{fitness}}$ depend on the genotype. Then Hardy-Weinberg equilibrium does not hold.

Def. The mean fitness $w_t = \underbrace{p_t^2}_{P_{AA}} w_{AA} + \underbrace{2p_tq_t}_{P_{Aa}} w_{Aa} + \underbrace{q_t^2}_{P_{aa}} w_{aa}$ of next generation

Case 1: AA

Juvenile frequency: p^2

Survival rate: w_{AA}

Adult frequency: $\frac{p^2 w_{AA}}{w_t}$

$\longrightarrow$

	AA	Aa	aa
Juvenile frequencies	p^2	$2pq$	q^2
Relative survival rates	w_{AA}	w_{Aa}	w_{aa}
Relative adult frequencies	p^2w_{AA}	$2pqw_{Aa}$	q^2w_{aa}
Adult frequencies	$\frac{p^2w_{AA}}{w_t}$	$\frac{2pqw_{Aa}}{w_t}$	$\frac{q^2w_{aa}}{w_t}$

Can write a difference equation.

note: index, not power

Frequency of allele A at time $t+1$: $P_{t+1} = P_{AA}^{t+1} + \dfrac{P_{Aa}^{t+1}}{2}$

$$\Rightarrow \quad P_{t+1} = \frac{p_t^2 w_{AA}}{w_t} + \frac{p_t q_t w_{Aa}}{w_t}$$

$$P_{t+1} = \frac{p_t}{w_t}\left[p_t w_{AA} + q_t w_{Aa}\right]$$

$$P_{t+1} = \frac{p_t}{w_t}\left[p_t w_{AA} + (1-p_t) w_{Aa}\right]$$

$$P_{t+1} = \frac{1}{w_t}\left[p_t^2 w_{AA} + p_t(1-p_t) w_{Aa}\right]$$

Aside: if $w_{AA} = w_{Aa} = w_{aa} = 1$, then $w_t = 1$, so $p_{t+1} = p_t$, and we get Hardy-Weinberg equilibrium.

Normalize by survival rates

WLOG, let $w_{AA} = 1-s$, $w_{Aa} = 1$, $w_{aa} = 1-r$,

where $r, s < 1$ but r, s can be negative,

which ensures that $w_{AA}, w_{Aa}, w_{aa} > 0$.

Then $w_t = p_t^2(1-s) + 2p_t q_t + q_t^2(1-r) = 1 - p_t^2 s - (1-p_t)^2 r$

$$\Rightarrow \quad p_{t+1} = \frac{p_t\left[p_t(1-s) + (1-p_t)\right]}{1 - p_t^2 s - (1-p_t)^2 r} = \frac{p_t(1-p_t s)}{1-p_t^2 s-(1-p_t)^2 r} = f(p_t)$$

$$p_{t+1} = \frac{p_t(1-p_t s)}{1-p_t^2 s - (1-p_t)^2 r} = f(p_t)$$

$\bar{p} = 0, \quad 1, \quad \frac{r}{r+s}$

only "a" present — only "A" present — both alleles present

$$f'(p) = \frac{(1-s)p^2 + 2(1-s)(1-r)p(1-p) + (1-r)(1-p)^2}{(1-p^2 s - r + 2rp - rp^2)^2}$$

Note: $f'(p) > 0$ for $0 \le p \le 1$ and $r, s < 1$

$f'(0) = \frac{1-r}{(1-r)^2} = \frac{1}{1-r}$ $\qquad$ $f'(1) = \frac{1}{1-s}$

When $r < 0$, $|f'(0)| < 1$, so 0 is stable. $\qquad$ When $s < 0$, $|f'(1)| < 1$, so 1 is stable.

$f'\left(\frac{r}{r+s}\right) = \frac{2rs - r - s}{rs - r - s}$ $\qquad$ $\left(\bar{p} = \frac{r}{r+s}\right)$

For $0 < \frac{r}{r+s} < 1$, either $r, s > 0$ or $r, s < 0$ $\Big\}$ $rs > 0$.

If $r, s < 0$, then $C = rs - r - s > 0$, so $f'(\bar{p}) = \frac{rs + C}{C} > 1$, so unstable.

If $r, s \in (0,1)$, then $rs < r + s$, so $2rs < r + s$ (since $f'(\bar{p}) > 0$)

$\Rightarrow f'(\bar{p}) = \frac{r + s - 2rs}{r + s - rs} < 1$, so stable.

In this last case, $w_{Aa} > w_{aa}$ and $w_{Aa} > w_{AA}$, so the heterozygote has an advantage, and so both alleles remain present.