

SEGREGATION, LINKAGE, AND ESTIMATION

SEGREGATION AND HARDY-WEINBERG

READING: Nielsen & Slakin, pp. 5-20

– Idea behind Hardy-Weinberg: “simplest” scenario.

– Generic assumptions: "HARDY-WEINBERG" CONDITIONS

- 1) Random mating
- 2) No differences in viability or fecundity of the genotypes
- 3) No mutation
- 4) No migration
- 5) Infinite population size
- 6) Hermaphrodites, or equal gene frequencies in the two sexes
- 7) Non-overlapping generations (e.g. annual plants, bivoltine insects).

– Benefits & drawbacks of Hardy-Weinberg conditions.

–**Question:** Does Mendelian inheritance change genotype frequencies?

ASEXUAL INHERITANCE

- Mixed population with two clones named "1" and "2"
- Number (count) of strain i in generation t is $N_i(t)$, $i = 1, 2$
- Asexual inheritance simple: offspring genotypes are exact copies of their parents
- Suppose each individual in the pop has $W(t)$ kids in generation t , regardless of genotype:

$$N_i(t+1) = W(t)N_i(t) \quad i = 1, 2$$
- What happens?:

– Consider the fraction (or frequency) of individuals that are of type 1: $p_1(t)$

- In generation t :

$$p_1(t) = \frac{N_1(t)}{N_1(t) + N_2(t)}$$

- In generation $t + 1$:

$$\begin{aligned}
 p_1(t+1) &= \frac{N_1(t+1)}{N_1(t+1) + N_2(t+1)} \\
 &= \frac{W(t)N_1(t)}{W(t)N_1(t) + W(t)N_2(t)} \quad [\text{cancel } W(t)] \\
 &= \frac{N_1(t)}{N_1(t) + N_2(t)} \\
 &= p_1(t)
 \end{aligned}$$

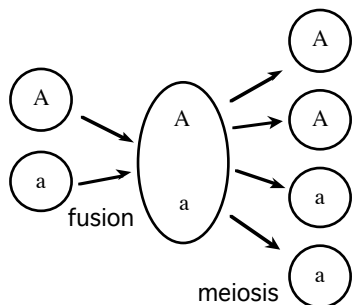
- Numerical Example: Suppose $N_1(t) = 100$, $N_2(t) = 300$, $W(t) = 2$ (2 offspring/individual)

$$p_1(t) = \frac{100}{100 + 300} = 0.25; \quad p_1(t+1) = \frac{2 \cdot 100}{2 \cdot 100 + 2 \cdot 300} = \frac{200}{800} = 0.25 = p_1(t)$$

- Same result holds with any number of asexually reproducing genotypes
- Observations:
 - 1) When different genotypes reproduce equally well, their relative frequencies do not change;
 - 2) Population size can fluctuate, but no change in genotypic frequencies.
 - 3) Note the method of investigation: **"Recursion equation"**
 - 4) Relative frequencies, rather than numbers, are often best way to think about problems of population composition.
- **Practice Exercises:**
 - 1) What happens if $W(t) = 3, 0.5,$ or 10 in the last example?
 - 2) Show that $p_2(t+1) = p_2(t)$.
 - 3) Show that, for any t , $p_2(t) = 1 - p_1(t)$, $p_1(t) = 1 - p_2(t)$, and $p_1(t) + p_2(t) = 1$.

A SINGLE HAPLOID LOCUS WITH TWO ALLELES

- Similar to asexuals, except there's a brief diploid sexual phase (e.g. fungi)
- Life cycle:



- Consider haploid population with two genotypes, A and a . Let p be the frequency of A and $1 - p \equiv q$ be the frequency of the a genotype.
- Mating is random: genotypes of the two mates are independent of one another;
- Frequencies of diploid genotypes are:

$$AA: p^2 \quad aa: (1-p)^2 = q^2 \quad Aa + aA: p(1-p) + (1-p)p = 2p(1-p) = 2pq$$

- These are the "familiar" Hardy-Weinberg proportions
- What about the frequencies of A and a in the offspring?
 - All offspring of AA are A 's, aa are a 's, while half of Aa 's are A and half are a
 - Use a prime (') to denote "next generation"
 - Get

$$A: p' = p^2 + (1/2) \cdot 2pq = p^2 + pq = p(p+q) = p[p + (1-p)] = p$$

$$a: q' = 1 - p' = 1 - p = q$$

Punch line: Genotype frequencies do not change.

- Haploid vs. asexuals
 - 1) Haploids can have mating type-linked genes
 - 2) When there's more than 1 locus, haploid genotypes can recombine

A SINGLE DIPLOID LOCUS WITH TWO ALLELES

- **genotype** frequencies vs. **gene** or **allele** frequencies

– We can determine the **allele** frequencies from the genotype frequencies:

- Freq. of **allele** $A = p = \frac{N_A}{N_A + N_a}$; Freq. of **allele** $a = q = 1 - p = \frac{N_a}{N_A + N_a}$.

- Freq. of genotype $AA = P_{AA} = \frac{N_{AA}}{N_{AA} + N_{Aa} + N_{aa}}$, etc.
- Then $p = P_{AA} + 1/2 P_{Aa}$, $q = P_{aa} + 1/2 P_{Aa}$.
- *Practice exercises:*
 - (1) Show this last result using numbers of alleles. NOTE: $N_A = 2N_{AA} + N_{Aa}$, etc.
 - (2) Verify the above formulas.
 - (3) Show that $P_{Aa} = 1 - P_{AA} - P_{aa}$. (Hint: Show that $P_{AA} + P_{Aa} + P_{aa} = 1$.)
- Note: cannot do converse, i.e. compute genotype frequencies from allele frequencies, without making additional assumptions.
- A given p is consistent with many values of P_{AA} and P_{Aa}
 - e.g. $P_{AA}:P_{Aa}:P_{aa} = 1/2:0:1/2$ and $1/3:1/3:1/3$ both correspond to $p = 1/2$.
- *Practice exercises:* Verify this example. Find other genotype frequencies corresponding to $1/2$.
- **Questions:**
 - Does Mendelian inheritance change genotype frequencies?
 - If so, what happens to the genotype frequencies?
- **Two Answers:**
 - 1) Straightforward-but-tedious way: Consider all possible matings.
 - 2) Easier way
 - use intuition: random mating = random union of haploid gametes.
 - if intuition is correct, then:
 - frequencies of A and a among the gametes produced = frequencies in mating adults: p and $q = 1 - p$. As with the haploid results, random union of gametes implies:
$$P'_{AA} = p^2, P'_{Aa} = 2p(1-p) = 2pq, P'_{aa} = (1-p)^2 = q^2$$

(the notorious HARDY-WEINBERG PROPORTIONS!)
- Note: $p' = P_{AA} + 1/2 P_{Aa} = p^2 + 1/2 2p(1-p) = p(p+1-p) = p!$ (as we you might expect).
- *Practice exercise:* Show that $q' = q$.
- Is our intuition correct?
 - Consider the more tedious way. (see **HANDOUT I.1: Random mating = Random Union of Gametes**)

– Yes, intuition is correct.

• **Hardy-Weinberg Proportions for multiple alleles:**

– Above results generalize to multiple alleles (named "A₁", "A_s", "A₃", ...):

$$p'_3 = p_3, P'_{33} = p_3^2, P'_{13} = 2p_1p_3, \text{ etc.}$$

• **Predictions:**

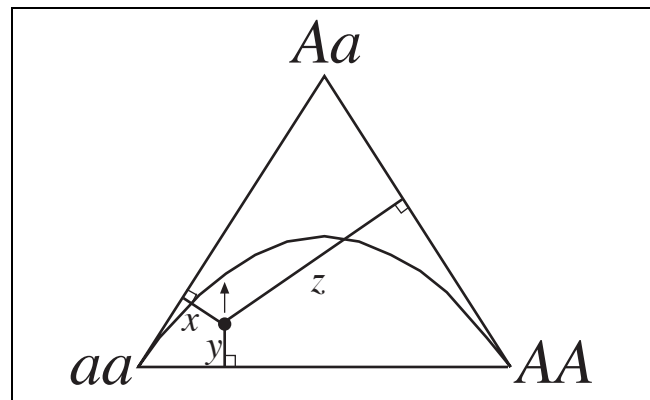
If the Hardy-Weinberg assumptions hold:

- (1) Allele frequencies do not change;
- (2) Allele frequencies equal those in the reference population, p_0 and $q_0 = 1 - p_0$;
- (3) Genotype frequencies may change;
- (4) Genotype frequencies will not change *after* the first generation;
- (5) Genotype frequencies after the first generation may be found from reference allele frequencies: $P_{AA} = p_0^2$, $P_{Aa} = 2p_0q_0$, and $P_{aa} = q_0^2$.

• **de Finetti Diagram and Hardy-Weinberg Dynamics**

– $x = P_{AA}$, $y = P_{Aa}$, $z = P_{aa}$

– parabolic curve represents Hardy-Weinberg equilibria for different allele frequencies.



• **Implications:**

- (A) Neutral Stability:
- (B) Variation Maintained
- (C) Parameter Number Reduced
- (D) Null Model

TWO COMPLICATIONS

(1) Two sexes with unequal allele frequencies

- Let frequency of A be p_f in female gametes, p_m in male gametes,
 - Frequency of a is $q_f \equiv 1 - p_f$ in females, etc.
- A mating table for gametes:

		Female Gametes:		
		A	a	
		frequency	p_f	q_f
Male Gametes:	A	p_m	$p_m p_f$ (AA)	$p_m q_f$ (Aa)
	a	q_m	$q_m p_f$ (Aa)	$q_m q_f$ (aa)

- Offspring (Both sexes)
 - $AA : P'_{AA} = p_m p_f$ $Aa : P'_{Aa} = p_m q_f + p_f q_m$ $aa : P'_{aa} = q_m q_f$
- Biological Conclusion: Genotype and allele frequencies within each sex have become the same.

(2) Sex Linkage: What if genes are not autosomally inherited?

- In mammals & many other animals, $XX = \text{female}$, $XY = \text{male}$
- see **HANDOUT I.2: Sex linkage and Hardy-Weinberg**
- General Conclusions:

(1) Weighted average of allele frequencies remains constant through time:

$$\frac{1}{2} p_m(t) + \frac{1}{2} p_f(t) = \frac{1}{2} p_m(t-1) + \frac{1}{2} p_f(t-1) = \dots = \frac{1}{2} p_m(0) + \frac{1}{2} p_f(0)$$

(2) Allele frequencies in both sexes ultimately approach this average, regardless of sex ratio.

(3) Allele frequencies always oscillates toward the equilibrium [unless $p_f(0) = p_m(0)$

].

(4) Magnitude of the difference in allele frequencies in sexes decreases by 1/2 each generation: $|p_f(t+1) - p_m(t+1)| = \frac{1}{2} |p_f(t) - p_m(t)|$

- Some biological consequences of sex linkage: If the frequency of allele A is the same in both sexes, i.e., $p_f = p_m = p$, then among the offspring:

<u>Females</u>		<u>Males</u>	
<i>AA</i>	p^2	<i>AY</i>	p
<i>Aa</i>	$2p(1-p)$	<i>aY</i>	$1-p$
<i>aa</i>	$(1-p)^2$		

- Recessive genes are expressed much more frequently in the heterogametic sex
 - E.g., if *A* codes for a recessive trait and $p = 0.01$ in both sexes, then the ratio of affected male to female offspring (*AY:AA*) will be $\frac{1}{100} : \frac{1}{10,000}$ or 100:1.

– Other genetic sex-determining systems:

- *XX-XO* (some spiders, beetles) Works like *XX-XY*.
- *ZW-ZZ* (birds, lepidoptera) Same except reverse male/female labels
- Haplo-diploids (hymenoptera) All loci are sex-linked .
- Principles of sex-linkage remain the same: only the labels have been changed!