Hardy Weinberg

Population Genetics Modeling
Using mathematics to model the behavior of alleles in populations.
Core model: Hardy Weinberg Law

Start with Ideal Population (Keep things simple st…)
Hardy-Weinberg Assumption
1. All individuals have equal rates of survival and equal reproductive success. (i.e. no selection.)
2. No new alleles are created or converted from existing alleles by mutation. (i.e. no mutation.)
3. Individuals do not migrate into or out of population (no migration)
4. Population is infinitely large (or large enough that sampling error doesn't alter allele frequencies from one generation to next)
5. Individuals mate randomly
6. Alleles studied are on autosomes.

Hardy Weinberg Law

- Allele frequencies do not change over generations.
- Relationships between allelic and genotypic frequencies can be described by the equation $p^2 + 2pq + q^2 = 1$ (Hardy-Weinberg Equation)
- If genotypic frequencies are disturbed one generation, they will return to equilibrium after one generation of random breeding.

- Significance of HW: changes in allele frequency are the genetic basis of evolution.

Value of Hardy-Weinberg model

1. Estimating p and q for alleles with complete dominance

2. Testing whether populations are in Hardy-Weinberg Equilibrium

3. Evaluate factors that cause populations to evolve - analyzing assumption

---

Value of Hardy-Weinberg Model

1. Estimating allelic and genotypic frequencies

Reminder for H-W equilibrium
$p^2 + 2pq + q^2 = 1$, where p is freq of dominant allele
and q is freq of recessive allele.

p and q can be estimated if you assume the population fits the HW model

$q = \sqrt{\text{freq of homozygous recessives}}$

$p = 1-q$

Example 1

About 99% of Native Americans are Rh +, 95% of African Americans are Rh positive, 85% of European American are Rh+.

- – What is the frequency of the Rh+ allele among Native Americans?
- – What percentage of African Americans with Rh+ are heterozygous for the Rh factor

Example 2

- Cystic Fibrosis affects about 1 in every 2500 births among Europeans Americans and affects about 1 in 15,000 births in African Americans. What is the frequency of Cystic Fibrosis carriers (heterozygotes) in the two populations?

Value of Hardy-Weinberg Model

2. Testing if populations are in equilibrium

Why test for H-W equilibrium?
  1. Some of H-W assumptions must not be true for pop.
  2. Population may be evolving

How to Test
1. Use Chi Square
2. Only when you can distinguish heterozygotes (Codominance or partial dominance)
3. Compare actual counts of homozygotes, heterozygotes
4. Lose an extra degree of freedom

Example

1. A sample of 300 plants from a population was examined for electrophoretic mobility of alcohol dehydrogenase enzyme. The genotypes are determined by two alleles, the fast allele and the slow allele. 7 plants were found to be FF, 106 were found to be FS and 187 were found to be SS. Use Chi-square to test whether the population is in Hardy-Weinberg Equilibrium.

Value of Hardy-Weinberg Model

3. Evaluate factors that cause populations to evolve

	Assumption	Exception	Effect
1.	Individuals must randomly mate	Inbreeding /Outbreeding	Decrease or increase in heterozygosity
2.	No migration	Migration	Homogenize different Populations
3.	No mutation	Source of all Variation	Increase Heterozygosity
4.	Genes are autosomal	Sex Linkage	Disrupts return to equilibrium
5.	**No selection**	Directional Selection	Reduce variation
		Disruptive Selection	Increase Variation
		Stabilizing Selection	Increase Variation
6.	**Population must be very large**	Genetic Drift	Reduce Variation (tend to loose rare alleles)

Types of Selection

Copyright © 2006 Pearson Prentice Hall, Inc.
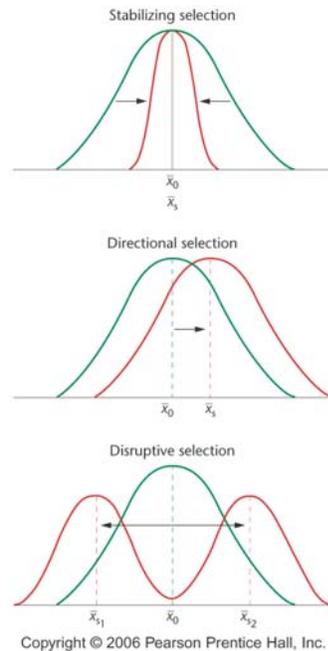
Selection and Hardy Weinberg

- Selection results when the genotypes have different fitness (survival and reproductive success)

- Goals
  - Predict affect on p and q when fitness is known
  - Estimate fitness when p and q change because of selection.

Fitness

(w) is a mathematic measure of relative fitness

w values will be between 1 and 0

Max fitness, w = 1

Min fitness, w = 0

In principle each genotype can have its own fitness value

w_{AA} = fitness of homozygous AA

w_{Aa} = fitness of heterozygotes Aa

w_{aa} = fitness of homozygous aa

If fitness's are known, they can be used to predict how allelic and genotypic frequencies will change the next generation

Effect of Selection

$$q_1 = \frac{\text{\# heterozygotes} + 2(\text{\# homozygote aa})}{2\ (\text{\# homozygotes AA} + \text{\# heterozygotes Aa} + \text{\# homozygotes aa})}$$

$$q_1 = \frac{2pq + 2(q^2)}{2(p^2 + 2pq + q^2)}$$

Adding fitness values

$$q_1 = \frac{2pq(w_{Aa}) + 2(q^2)\ (w_{aa})}{2(p^2(w_{AA}) + 2pq(w_{Aa}) + q^2(w_{aa}))}$$

If all have the same fitness value (w=1), then it returns to original
(q and p don't change if there is not selection)

If fitness values are different, than q1 may differ from q

Squirrel Example

Assume on Grand Island there is a population of squirrels. The coat color phenotype is controlled by a single gene with two alleles, the dominant B allele for grey fur and the recessive b allele for black squirrels. In the current generation about 40% of the squirrels on the island are black.

Teenage boys are aggressive predators of squirrels. Because boys can spot black squirrels easily, black squirrels have 30% the survival and reproductive rate of grey squirrels. Predict the frequency of black squirrels in the next generation.

To Solve Squirrel Problem

Freq of black squirrel in next generation is q_1^2

$$q_1 = \frac{2pq(w_{Aa}) + 2(q^2)(w_{aa})}{2(p^2(w_{AA}) + 2pq(w_{Aa}) + q^2(w_{aa}))}$$

To solve equation you need to know

Allele frequency in current generation
p_0 and q_0

Fitness values for each genotype
w_{BB}
w_{Bb}
w_{bb}

Estimating fitness based on changes in p and q

Typically, the fitness of different genotypes cannot be observed directly, instead it is estimated based on changes in p and q over generations.

For example assume that homozygous recessive has reduced fitness compared to other genotypes.

$$q1 = \frac{2pq(w_{Aa}) + 2(q2)(w_{aa})}{2(p^2(w_{AA}) + 2pq(w_{Aa}) + q^2(w_{aa}))}$$

Solve for w_{aa} assuming W_{AA} and W_{Aa} are = 1

$waa = 1 - [(q-q_1)/(q^2(1-q_1)]$

Example

The phenotypic frequency of black squirrels in 25% in one generation, but after heavy predation by boys it decreases to 21% in the next generation. What is the relative fitness of black squirrels?

Need q and q_1

Genetic Drift

- Definition - Changes in allele frequency due to sampling error of gametes during fertilization.

Consider this example.
- Suppose you had a population of 100 in which p and q were both equal to 0.5. That population would generate 1000's of gametes, 50% of the gametes would have an A allele and 50% would have an a allele.

- To generate the next generation of 100, you would need 200 gametes to fuse during fertilization. If there was no sampling error than you would expect 100 of these gametes would have an A allele and 100 of these gametes would have an a allele. Without sampling error, p and q would remain at 0.5.

- However, because of sampling error, the chances that the 200 gametes would be exactly 50% A and 50% a is not likely. For example it would be unlikely that if gametes are randomly sampled 105 might have the A allele and 95 might have the a allele. With this level of sampling error p = 52.5 and q = 47.5.

Effect of population size

Suppose you followed 5 populations with 100 individuals
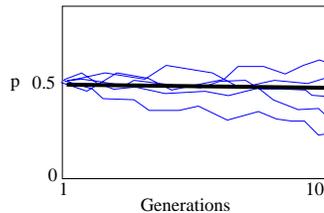
Suppose you followed 5 populations with 10 individuals
(Note Allele Fixation)
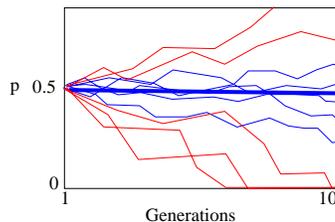
Evaluating Genetic Drift

- Sample error is always occurring. The magnitude of the change depends on the population size. The greater the population the smaller the expected change due to drift.

- Range of genetic drift changes.

  Sample error generates a normal distribution (bell curved shaped)

  Standard error (square root of variance) is a measure of range of distribution

  p $\pm$ 2 SE corresponds to 95% of the values in the distribution

  Std Error $s = \sqrt{pq/2N}$

  95% confidence interval

Example

- Suppose that a population of 64 individuals had a p value of 0.6 in one generation but a p value of 0.5 the next generation. Would it be reasonable to believe that genetic drift could account for this change in p?