Population genetics is based on statistical models:

“A model is an intentional simplification of a complex situation designed to eliminate extraneous detail in order to focus attention on the essentials of the situation” (Daniel L. Hartl).

Statistical modeling and inference:

- Define a model
- Explore properties
- Estimate model parameters from the data
- Test goodness of fit
- Refine Model

Concerns:

- Rules / parameters / quantities
- Summary stats / graphical data exploration / simulation
- Moments / maximum likelihood / Bayesian methods
- Compare estimators / heterogeneity / outliers
- Update parameters
## HW model: no change in frequencies

## Alt model; change in frequencies (molecular evolution)

<table>
<thead>
<tr>
<th>Agency</th>
<th>Genotype</th>
<th>Allele</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linkage</td>
<td>no</td>
<td>no</td>
<td>Creates disequilibrium among loci</td>
</tr>
<tr>
<td>Inbreeding</td>
<td>yes</td>
<td>no</td>
<td>Acts on all loci in genome; results in loss of heterozygosity</td>
</tr>
<tr>
<td>Assortative Mating</td>
<td>yes</td>
<td>no</td>
<td>Only acts on the locus subject to assortiment, and those loci linked to it</td>
</tr>
<tr>
<td>Migration</td>
<td>yes</td>
<td>yes</td>
<td>Depends on migration rate and frequency differences between populations</td>
</tr>
<tr>
<td>Mutation</td>
<td>yes</td>
<td>yes</td>
<td>Very very very slow</td>
</tr>
<tr>
<td>Natural Selection</td>
<td>yes</td>
<td>yes</td>
<td>Acts on the locus subject to selection, and those loci linked to it</td>
</tr>
<tr>
<td>Genetic Drift</td>
<td>yes</td>
<td>yes</td>
<td>Acts on all loci in the genome; results in loss of heterozygosity and loss of alleles</td>
</tr>
</tbody>
</table>

**Population Genetics 8:**

**transient verses equilibrium polymorphism**

**Note:** Many natural populations exhibit extensive genetic polymorphism.

**How do we explain this?**
Mutation pressure and selection can operate in opposite directions as a force for change in allele frequencies in populations. Note that effectiveness of both depends on the allele frequency. \( \Delta p \) is the change in allele frequency from one generation to the next. In this example, mutation and selection are acting in opposite directions as \( \Delta p \) is positive under mutation pressure and negative under selection pressure. Note that the values of \( \Delta p \) under both forces only become comparable when the allele frequency is low.

\[ \Delta p \text{ (mutation pressure)} = \Delta p \text{ (selection)} \]

Mutation - selection equilibrium

Attainment of the equilibrium allele frequency given selection and a variety of different mutation rates. Note that the time to equilibrium varies in addition to the actual equilibrium frequencies. Note that for realistic mutation rates, the equilibrium frequencies are quite low (freq of \( a \) allele \(< 0.05 \)). If we assume stronger selection pressure (\( s > 0.1 \)), the equilibrium point will be lower and the rate to equilibrium will be faster.
Mutation - selection equilibrium

1. Mutation pressure:

Let \( \mu = \) the mutation rate from \( A \rightarrow a \)
Let \( \nu = \) the mutation rate from \( a \rightarrow A \)

Let \( p_t = \) the frequency of \( A \) in the population in generation \( t \).
Let \( q_t = \) the frequency of \( a \) in the population in generation \( t \), with \( q_t = (1 - p_t) \).

\[
\Delta q = \frac{\mu}{1 - s q_t^2} - q_t \nu
\]

2. Natural selection against a deleterious recessive allele:

Remember from our earlier lecture:
\[ q_{t+1} = q_t - s q_t^2 / 1 - s q_t^2 \]

So for \( \Delta q \),
\[ \Delta q = q_{t+1} - q_t \]
\[ \Delta q = (q_t - s q_t^2 / 1 - s q_t^2) - q_t \]
\[ \Delta q = -s q_t^2 (1-q) / 1 - s q_t^2 \]

Mutation - selection equilibrium

\[ \Delta q \text{ (mutation pressure) = } \Delta q \text{ (selection)} \]
\[ p\mu - q\nu = s q_t^2 (1-q) / 1 - s q_t^2 \]

YUCK!

[Approximate and simplify]
\[ p\mu = s q_t^2 (1-q) / 1 - s q_t^2 \]
\[ p\mu = s q_t^2 (1-q) \]
\[ (1-q)\mu = s q_t^2 (1-q) \]
\[ \mu = s q_t^2 \text{ (approx.) -or- } q = \sqrt{\mu / s} \text{ (approx.)} \]

[Dominance: \( q = \mu / hs \text{ (approx.)} \)]
Effect of partial dominance on mutation-selection equilibrium. The fitness of genotypes AA, Aa, and aa are assumed to be 1, 1-hs, and 1-s respectively.

The symbol \( h \) is the amount of dominance in the heterozygote genotype. Note, that even a small amount of dominance (\( h = 0.01 \)) reduced the equilibrium frequency of the recessive allele. The reason is that when \( q \), the freq of the recessive allele is small, the majority of those alleles are in the heterozygote configuration, and even a small amount of selection on the heterozygotes leads to a major reduction in its equilibrium frequency as compared with full dominance.

Note that for reasonable values of \( \mu \), \( h \), and \( s \), the equilibrium frequencies are < 0.01. This means that mutation selection equilibria is not sufficient to explain low frequency detrimental alleles in populations where those alleles have frequencies > 0.01.

- **Mutation - selection equilibrium**

- **Genetic drift**

<table>
<thead>
<tr>
<th>( N_e )</th>
<th>Population Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Ne = 100</td>
</tr>
<tr>
<td>1000</td>
<td>Ne = 1000</td>
</tr>
<tr>
<td>10000</td>
<td>Ne = 10000</td>
</tr>
<tr>
<td>50000</td>
<td>Ne = 50000</td>
</tr>
</tbody>
</table>
Selection (push to fixation) – drift (high probability of loss)

Drift alone: probability of fixation of a new mutant $= 1/2N_e$

probability of loss of a new mutant $= 1-(1/2N_e)$

Selection + Drift: probability of fixation depends on interaction of $s$ and $N_e$.

$N_es > 1$: beneficial allele more likely to be fixed than under drift alone

$N_es < 1$: beneficial allele is fixed with probability close to its frequency in the population
Then fate of a beneficial recessive allele ($A_1$) is not always predictable under the combined effects of directional selection and genetic drift. If there is no genetic drift (left: $N_e = \infty$), the fate of the recessive allele ($A_1$) is always determined by selection. When there is drift (right: $N_e < \infty$) the fate of the recessive allele ($A_1$) is not necessarily determined by selection; hence a deleterious allele can be fixed in a population.

Note that $N_e > 1$ does not guarantee that an allele is going to be fixed; it simply indicates that (as a long term average) the frequency that it is fixed will be greater than the frequency under genetic drift alone.

### Selection - drift

$A$ is beneficial

$a$ is deleterious

<table>
<thead>
<tr>
<th>System</th>
<th>$N_e$</th>
<th>$s$</th>
<th>Initial frequency</th>
<th>Fixation of $A$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drift alone</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>1%</td>
</tr>
<tr>
<td>Selection + Drift</td>
<td>100</td>
<td>0.1</td>
<td>0.01</td>
<td>10%</td>
</tr>
<tr>
<td>Selection alone</td>
<td>infinity</td>
<td>0.1</td>
<td>0.01</td>
<td>100%</td>
</tr>
</tbody>
</table>

$W_{AA} = 1$

$W_{Aa} = 1$

$W_{aa} = 1$ or 0.9

- This sort of polymorphism is always transient!
- Selection causes the pop of fixation to be > chance alone
- For selection to dominate $N_e s >> 1$
Selection - drift

Because drift disturbs the allele frequencies each generation, frequencies in any one generation will not be in equilibrium. However, the long-term average will be the equilibrium frequencies. The polymorphism in both of the above cases is not transient.
Polymorphisms under balancing selection ($s_1 = 0.1$ and $s_2 = 0.3$) is transient if drift effects are strong. ($N_e = 50$).

Combined effects of mutation, selection and drift on polymorphism. When drift is weak polymorphism is not transient, but when drift is strong the polymorphism is transient, but recurring due to mutation.
Very high mutation rate (0.01) results in only a small shift in the long term average allele frequency under overdominant selection.
An explanation for population variation based on natural selection:

Very high mutation rate (0.01) results in only a small shift in the long term average allele frequency under overdominant selection.

Mutation - drift equilibrium

• Ignored until 1960’s
• “Neutral theory of molecular evolution”
• Transient polymorphism
• Fundamental to discipline of molecular evolution
Sources of polymorphism in populations

- Mutation-selection-drift equilibrium (long-term or transient)
- Selection-drift (transient)
- Overdominance-drift equilibrium (long-term or transient)
- Mutation-drift (transient, but important)

Drift constantly disturbs equilibrium.
- The strength of the disturbance depends on the effective population size ($N_e$).
- If strong enough, the disturbances can push the frequency to fixation.
- We don’t expect to see persistent equilibrium in populations with low $N_e$.

- For realistic values of $\mu$, $h$ and $s$, the equilibrium point is generally very low ($p < 0.01$).
  - As an explanation for natural polymorphisms $> 0.01$, the balance between mutation and selection is not satisfactory.

- Overdominance can explain population polymorphisms with frequencies $> 0.01$.
  - A cost in fitness makes it unlikely that it can be invoked as a common mechanism in natural populations.
  - We will return the notion of the cost of selection later in this course.