Key ideas from lecture 1.

Evolution by Natural Selection as a syllogism* (Endler 1986)

1. If there is heritable variation \((h^2 > 0)\), and

2. the variants differ with respect to their expected abilities to survive and reproduce in the present environment \((S \neq 0)\), then

3. those heritable traits conferring enhanced success will tend to increase in frequency \((R \neq 0)\).

As we will see, \(R = h^2S\). The response to selection \((R)\) is equal to heritability \((h^2)\) times the selection differential \((S)\).

Is this testable idea? If so, it becomes science.

Better yet: Is the idea falsifiable? (See http://en.wikipedia.org/wiki/Falsifiability)

*I want to make the distinction between “natural selection,” and “evolution by natural selection.”
Lecture 2. A brief history of population genetics (from Provine’s book), and basic theoretical population genetics.

Do you have an intuitive understanding of $e = mc^2$, or $R = h^2S$?

Intuitive understanding of $R = h^2S$ requires unpacking $h^2$, which requires an understanding of variance, covariance and regression. That is where we are going.

More philosophy: in addition to Peters’ attack on Natural Selection, Karl Popper (famous for introducing the criterion of “falsifiability”) also hammered the theory:

“Darwinism is not a testable scientific theory, but a metaphysical research program.”

But he later changed his mind (Popper 1978):

“I have changed my mind about the testability and logical status of the theory; and I am glad to have an opportunity to make a recantation.”

Popper realized that the theory was testable (i.e. falsifiable).
Darwin’s views on variation prior to 1856

2 types of variation, both heritable
   “Sports”: large, discrete effect
   “Individual differences”: small continuous effects

Darwin’s view in 1856: *The Origin of Species*

a) Sports are too rare

b) Darwin concludes that selection operates on “individual differences.” Variation in these differences is generated every generation. Change is gradual. And there is a gradual radiation of species from a common ancestor (the tree of life.)

Reaction to *The Origin of Species*
   --Evolution as such accepted
   --Natural selection rejected.

As we discussed before, one of the main objections to Evolution by Natural Selection was that there was no known mechanism for heredity.
Darwin’s view on heredity in 1868. The provisional theory of pangenesis

a) gemmules circulate through the body.

b) gemmules change as the environment affects the body.

c) gemmules collect and multiply in reproductive organs, and pass on acquired characteristics.

d) gemmules have small, additive effects (due to blending).
The problem with blending inheritance

a) variation needed for natural selection is lost in the blender.

b) Darwin: Artificial selection works even with blending inheritance, because the favored types are isolated. No chance to blend.

The views of Thomas Henry Huxley ("Darwin’s Bulldog")

a) Accepts the mechanism of NS, but feels that Gradual change is an “unnecessary burden” for the theory. Can’t explain the fossil record.

b) Natural selection work on sports, which are more stable to blending inheritance

c) Evolution by NS is jumpy (saltational. Saltation means jump)

d) “Hopeful monsters” can reproduce.
Francis Galton (Darwin’s cousin, also wrote *Hereditary Genius*)

a) 1869: Galton agrees with Huxley that evolution by natural selection is saltational.

b) Galton tested the “provisional theory of pangenesis by transferring blood between rabbits.

--Galton concludes that the provisional theory is false. Darwin responds in *Nature* that he never claimed that gemmules circulated in the blood.

--Mivart’s (1870) thought experiment. Weismann’s mouse experiment. Both undermine the provisional theory.
1875: Galton’s theory of Heredity...

..gemmules don’t circulate. Gametes are concentrated in the reproductive organs, and they stay there.

1889: Publication of Galton’s *Natural Inheritance* (Galton 1889)

a) Galton begins the tradition of statistical analysis of correlations among characters, especially between parents and offspring.

b) Galton’s “Principle of Regression” (Galton 1886; Galton 1889)

--There is a regression (here: movement) of offspring towards the mean. Thus 2 extreme parents produce a less extreme set of progeny. Graph this.

--Regression to mediocrity means, to Galton, that natural selection must operate on “sports” and the response to selection must be manifested as discrete jumps.
Figure 1.1  The relationship between height of adult children and the average height of their parents (in inches). Circles denote average offspring heights for different 1-inch classes of midparent heights. The best linear fit is given by the solid line, while the dotted line is the expected pattern if mean offspring height were equal to midparent height. (Data from Galton 1889.)

Source: Galton_regressionTomedicrity.ppt
Karl Pearson and the Biometricians (Pearson of the Pearson product moment correlation).

--reject Galton’s logic for regression on the mean, but they continue Galton’s tradition of statistical analysis

--The Biometricians support gradual evolution

--Reject Mendelism... Why??

Bateson and the Mendelians

--Reject gradualism: **selective advantage of slight differences is too small** to be important, and swamped out by cross-fertilization (blended out).

--Accept saltation, following Bateson’s study of the relationship between the phenotype and environment.

Accept Mendelism.

NOTE: both camps (Biometricians and Mendelians) claim Galton as one of their own.
WANTED: A demonstration that selection on variants having small differences in fitness could result in gradual change in reasonably short periods of time.

Such a demonstration came from simple mathematical models, which led to the synthesis of Darwinism, Mendelism and Biometrics. See (Fisher 1918) for the key original insights. The synthesis is now known as the MODERN SYNTHESIS.
Theoretical population genetics.

Assume no selection, no genetic drift (ie, infinite population size), no mutation, no migration, no inbreeding...

1. At time 1:

Let freq. of the allele 1 be $p_t$ ($p$ at time t)
Let the frequency of allele 2 be $q_t$, which is equal to $(1-p_t)$

Assuming random mating (panmixia) in an infinite population (no drift),

The frequency of 11 genotypes is __________ fill in

The frequency of 12 genotypes is __________ fill in

And the freq. of 22 genotypes is __________ fill in
2. At time 2 (the next generation), the frequency of the 1 allele is equal to

\[
p_{t+1} = \frac{\text{number of 1 alleles}}{\text{total number of alleles}}
\]

\[
p_{t+1} = \frac{2N(p_t^2 + 2p_t q_t / 2)}{2N(p_t^2 + 2p_t q_t + q_t^2)}, \text{ where } N = \text{number of individuals}
\]

\[
p_{t+1} = \frac{p_t^2 + p_t q_t}{1}
\]

\[
p_{t+1} = \frac{p_t^2 + p_t q_t}{1} = p_t^2 + p_t(1 - p_t)
\]

\[
p_{t+1} = p_t^2 + p_t - p_t^2 = p_t
\]

So the frequency of the 1 allele did not change. It is at equilibrium, which is independent of the frequency of the 1 allele at time \( t \). The equilibrium is called the Hardy-Weinberg equilibrium. How to graph this? The equilibrium is neutrally stable. What does that mean?
The Tweak: relax the assumption of no selection

Now we attach variables for the fitness of each genotype, thereby relaxing the assumption of no selection

Let  \( W_{11} = \) expected fitness relative reproductive success of genotype 11
  \( W_{12} = \) expected fitness relative reproductive success of genotype 12
  \( W_{22} = \) expected fitness relative reproductive success of genotype 22  
(\text{Note the “expected”})

Where the values for \( W \) are generally between 0 and 1.

As before, let \( p = \) freq of allele 1 (say, for example, the allele for \text{red} flowers)  
And let \( q \) be the freq of allele 2 (say, for example, the allele for \text{white} flowers)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>11</th>
<th>12</th>
<th>22</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before selection</td>
<td>( p^2 )</td>
<td>( 2pq )</td>
<td>( q^2 )</td>
</tr>
<tr>
<td>After selection</td>
<td>( p^2W_{11} )</td>
<td>( 2pqW_{12} )</td>
<td>( q^2W_{22} )</td>
</tr>
</tbody>
</table>
As above, we have (where the $t$ subscripts are dropped: $p = p_t$)

$$p_{t+1} = \frac{\text{number of 1 alleles}}{\text{total number of alleles}}$$

$$p_{t+1} = \frac{2N( \quad )}{2N( \quad )}, \text{ Fill in the equations.}$$

$$p_{t+1} = \frac{p(pW_{11} + qW_{12})}{p(pW_{11} + qW_{12}) + q(pW_{12} + qW_{22})} \quad \text{Q: what are the terms in parentheses?}$$

$$p_{t+1} = \frac{p(pW_{11} + qW_{12})}{W}$$

Where $\bar{W} = p^2W_{11} + 2pqW_{12} + q^2W_{22}$ is the mean fitness of the population. Convince yourself that this statement is true. Why is $\bar{W}$ the mean fitness?

Now create an excel file that “iterates”

$$p_{t+1} = \frac{p(pW_{11} + qW_{12})}{\bar{W}} \quad \text{for 200 generations. For starters, set } W_{11} = W_{12} = 1, \text{ and } W_{22} = 0.9. \text{ Set } p \text{ at time 1 equal to } 0.01.$$
Graph the output. You should see that allele 1 increases rapidly in frequency, and go nearly to fixation by generation 200.

Here are my results (file: PopGen_Sim1_2006.xls)

Note that there is rapid gradual change in the frequency of the allele. This solved the argument between the Biometricians and the Mendelians (page 5). Who won?

How do I write my own simulation in excel?
note: See online handout for help with variance, covariance and regression.

Literature Cited

   —. 1889. Natural Inheritance, Macmillan.