PBG 200A Notes

Sam Fleischer

October 24, 2016

1 Questions from the Notes - my answers in magenta

1. You are investigating a locus with three alleles, A, B, and C, with allele frequencies p_A, p_B, and p_C. What fraction of the population is expected to be homozygotes under Hardy-Weinberg? The probability an individual is homozygote i is p_i^2 for $i = A, B, C$. So the total probability an individual is homozygote is $p_A^2 + p_B^2 + p_C^2$.

2. What are r_0, r_1, and r_2 for $\frac{1}{2}$ sibs? (share exactly one parent)? Parent i has alleles A_{i1} and A_{i2} for $i = 1, 2, 3$. If parent 1 mates with parent 2, their child has the following genotype with the following probabilities:

$$\begin{aligned} A_{11}A_{21} &: \frac{1}{4} \\ A_{12}A_{21} &: \frac{1}{4} \\ A_{11}A_{22} &: \frac{1}{4} \\ A_{12}A_{22} &: \frac{1}{4} \end{aligned}$$

and if parent 1 mates with parent 3 their child has the following genotype with the following probabilities:

$$\begin{aligned} A_{11}A_{31} &: \frac{1}{4} \\ A_{12}A_{31} &: \frac{1}{4} \\ A_{11}A_{32} &: \frac{1}{4} \\ A_{12}A_{32} &: \frac{1}{4} \end{aligned}$$

Then here are the possibilities:

$A_{11}A_{21},\ A_{11}A_{31}$: 1 IBD	$A_{11}A_{21},\ A_{12}A_{31}$: 0 IBD
$A_{11}A_{21},\ A_{11}A_{32}$: 1 IBD	$A_{11}A_{21},\ A_{12}A_{32}$: 0 IBD
$A_{12}A_{21},\ A_{11}A_{31}$: 0 IBD	$A_{12}A_{21},\ A_{12}A_{31}$: 1 IBD
$A_{12}A_{21},\ A_{11}A_{32}$: 0 IBD	$A_{12}A_{21},\ A_{12}A_{32}$: 1 IBD
$A_{11}A_{22},\ A_{11}A_{31}$: 1 IBD	$A_{11}A_{22},\ A_{12}A_{31}$: 0 IBD
$A_{11}A_{22},\ A_{11}A_{32}$: 1 IBD	$A_{11}A_{22},\ A_{12}A_{32}$: 0 IBD
$A_{12}A_{22},\ A_{11}A_{31}$: 0 IBD	$A_{12}A_{22},\ A_{12}A_{31}$: 1 IBD
$A_{12}A_{22},\ A_{11}A_{32}$: 0 IBD	$A_{12}A_{22},\ A_{12}A_{32}$: 1 IBD

Thus $r_0 = \frac{8}{16} = \frac{1}{2}$, $r_1 = \frac{8}{16} = \frac{1}{2}$, and $r_2 = 0$.

3. Consider a biallelic locus where allele 1 is at frequency p and two individuals who have IBD allele sharing probabilities r_0, r_1, r_2. What is the overall probability that these two individuals are both homozygous for allele 1? One individual is homozygote for allele 1 with probability p^2, so it must be $r_2 p^2$.

2 Population Genetics

2.1 Why study it?

It is key to our thinking as evolutionary biologists. Darwin lacked a mechanistic basis for gradual evolution. Mendelian genetics went unnoticed in this time unfortunately. Pop gen is the extension of Mendelian genetics to evolving populations. Quant Gen is the extension to phenotypic evoltion. Modern Synthesis in the 1930s. This is the basis of modern evolutionary thinking to this day.

2.2 What is it?

- Genetic basis for evolutionary change.
- study of genetic variantion within and between populations and species
- basis of "micro"-evolutionary thought

2.3 What is theoretical Population Genetics?

- Interplay between
  - mutation, assortive mating, migration, drift, recombination, selection
  - how these "forces" shape polymorphism and divergence

2.4 Why are population genetics models useful?

- support or discount verbal models
- intuition
- evolution is fundamentally statiscial
- Mendelian inheritance, segregation, recombination provide a powerful framework

2.5 What is empirical Population Genetics?

- lots and lots of data!
- DNA sequencing is getting CHEAP!

2.6 What is evolution?

- descent with modification
  - The process of descent - each individual has two parents - we can talk about the pedigree (family tree) of an individual
  - Modification only happens when there is genetic variantion
  - Mutations cause multiple alleles (ploymorphism)
  - Changes in phylogeny is actually substitution over many generations

2.7 DNA Sequencing

- provides unbiased description of genetic variation
- A "locus" is a location. The allele is the genetic information at that locus.
- Diploid individuals have two haplotypes (one genotype)
- At each locus, individuals are either homo- or heterozygous.
- Some changes in base pairs don't change the amino acids produced (synonymous) but some do (non-syn)
- we can talk about the frequency of an allele, the number of segregating sites, etc.

- How much variation is there?
  - In humans, 0.1% diversity
  - In D. melanogaster, 1% diversity
  - Vast amount of genetic diversity in every species.. why is there so much variation?

2.8 Example - Theory

Remember $p_A = f_{AA} + \frac{1}{2} f_{Aa}$. GIVEN the frequency of the different genotypes, you don't NEED the Hardy-Weinberg Equilibrium.

The HW Equilibrium assumes

- no migration
- $\boxed{\text{random mating}}$ ← most important
- no selection
- infinite population
- equal allele frequency in males and females
- no mutation

Supposing $f_{AA} = 0.07$, $f_{aA} = 0.4$, and $f_{aa} = 0.53$, we can calculate p_A and p_a using

$$p_A = f_{AA} + \frac{1}{2} f_{Aa} \qquad\qquad p_a = f_{aa} + \frac{1}{2} f_{Aa}$$

then note that p_A^2 is remarkably close to f_{AA} even though we never assumed HW equilibrium exists. If we start with frequencies which deviate from HW, it only takes a single generation of random mating to restore HW.

2.9 Example - Kuru outbreak in the Fore people

Fore people of Papua New Guinea practiced ritual funereal cannibalism (till the 1950s). Resistance to Kuru has a genetic basis. $f_{\text{Met}} = \frac{1}{30}\left(4 + \frac{1}{2} \cdot 23\right) = 52\%$. We expect the heterozygotes to have frequency $2 \times 0.52 \times 0.48$. This is not what we see though. The counts are post-outbreak, so there is deviation from HW equilibrium. A single generation of random mating would restore.

2.10 Clarification

Alleles being Identical By Descent does NOT mean it is not Identical. Question 3 is based on this fact

$$\begin{aligned}
\mathbb{P}[\text{two homozygote } AA \mid 0 \text{ IBD}] &= p^4 \\
\mathbb{P}[\text{two homozygote } AA \mid 1 \text{ IBD}] &= p^3 \\
\mathbb{P}[\text{two homozygote } AA \mid 2 \text{ IBD}] &= p^2
\end{aligned}$$

and thus the total probability is $r_0 p^4 + r_1 p^3 + r_2 p^2$.