

Stat 536 Homework 2

September 17, 2008

1. Consider a single locus with two alleles. Suppose p_f is the frequency of the first allele in females and p_m is the frequency in males. How long does it take to reach HWE and what is the equilibrium allele frequency at HWE?
2. Now, suppose that the locus in question is on the X chromosome. Females have two X chromosomes and males have one. A daughter receives one of her mom's X chromosomes at random and her dad's only X chromosome. A son receives one of his mom's X chromosomes at random and his dad's Y chromosome. Determine whether the difference in allele frequencies decreases after one generation if the initial allele frequencies in the sexes are $p_f \neq p_m$ but HWE assumptions apply thereafter? (Note, p_m means the frequency of the allele on the X chromosomes in males.) Does the difference decrease to zero in one generation as in Question 1? Will the allele and genotype frequencies stay constant if $p_f = p_m$?

[Please note that autosomal loci are loci that are not on the X or Y chromosome. Sex-linked loci are on the X chromosome.]

3. When the allele frequencies at an autosomal locus differ by δ in the sexes, it can be shown that the genotype frequencies in the next generation diverge from HWE according to the following equations.

$$\begin{aligned}P_{11} &= p_1^2 - \frac{\delta^2}{4} \\P_{12} &= 2p_1p_2 + \frac{\delta^2}{2} \\P_{22} &= p_2^2 - \frac{\delta^2}{4}\end{aligned}\tag{1}$$

- (a) Derive formulae for the maximum likelihood estimates for p_1, p_2 , and δ and compute \hat{p}_1, \hat{p}_2 , and $\hat{\delta}$ for the data in the table below.

A_1A_1	A_1A_2	A_2A_2
1	17	82

- (b) Estimate the variance of these estimates using the information matrix.
- (c) Find 95% confidence intervals for p_1 .

4. Consider this blood type data taken from a sample of Kuwaiti individuals.

A	B	AB	O
29	23	14	35

- (a) Write an EM algorithm to estimate the allele frequencies \hat{p}_A, \hat{p}_B , and \hat{p}_O .

- (b) Use bootstrap to estimate the variance of these estimates.
- (c) Do the allele frequencies in this population differ from those in Korea, where $\hat{p}_A = 0.231$, $\hat{p}_B = 0.209$, and $\hat{p}_O = 0.560$ were estimated based on a sample of size $n = 253$?