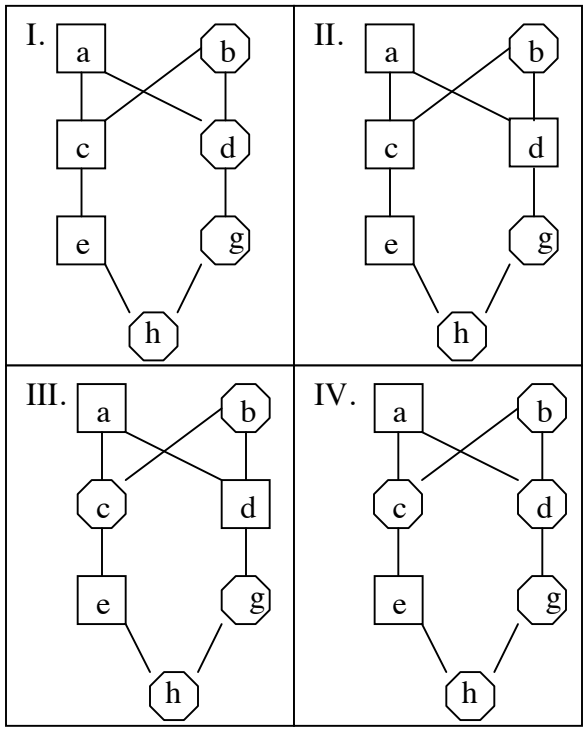
Homework #2 Answer Key

1. Lumped heterozygosity = $2\bar{p}\bar{q} - 2\,\mathrm{var}(p)$ where $\bar{p} = (.2 + .4)/2 = .3$, $\bar{q} = 1 - .3 = .7$, and $\mathrm{var}(p) = (.2^2 + .4^2)/2 - (.3)^2 = .01 \Rightarrow$ lumped heterozygosity $= 2(.3)(.7) - 2(.01) = .4$. The expected heterogyzosity if one random mating population $= 2\bar{p}\bar{q} = 2(.3)(.7) = .42$. $\Rightarrow$ observed heterozygosity < expected heterozygosity.
$\bar{H}_i(\text{observed}) = 2\bar{p}_i\bar{p}_j + 2\mathrm{cov}(p_i, p_j) \Rightarrow \bar{H}_{12}(\text{observed}) = 2\bar{p}_1\bar{p}_2 + 2\mathrm{cov}(p_1, p_2)$.
$\bar{p}_1 = (.2 + .6)/2 = .4$, $\bar{p}_2 = (.4 + .2)/2 = .3$,
$\mathrm{cov}(p_1, p_2) = [(.2)(.4) + (.6)(.2)]/2 - (.4)(.3) = -.02 \Rightarrow$
$\bar{H}_{12}(\text{observed}) = 2(.4)(.3) + 2(-.02) = .2$, which is less than $\bar{H}_{12}(\text{expected}) = 2(.4)(.3) = .24$
$\bar{H}_{13}(\text{observed}) = 2\bar{p}_1\bar{p}_3 + 2\mathrm{cov}(p_1, p_3)$, $\bar{p}_3 = 1 - \bar{p}_1 - \bar{p}_2 = .3$,
$\mathrm{cov}(p_1, p_3) = [(.2)(.4) + (.6)(.2)]/2 - (.4)(.3) = -.02 \Rightarrow \bar{H}_{13}(\text{observed}) = .2 <$
$\bar{H}_{13}(\text{expected}) = .24$. $\bar{H}_{23}(\text{observed}) = 2\bar{p}_2\bar{p}_3 + 2\mathrm{cov}(p_2, p_3)$,
$\mathrm{cov}(p_2, p_3) = [(.4)(.4) + (.2)(.2)]/2 - (.3)(.3) = .01 \Rightarrow$
$\bar{H}_{23}(\text{observed}) = 2(.3)(.3) + 2(.01) = .2 > \bar{H}_{23}(\text{expected}) = 2(.3)(.3) = .18$.
Overall observed heterozygosity = .2 + .2 + .2 = .6 < overall expected heterozygosity = $1 - (.4)^2 - (.3)^2 - (.3)^2 = .66$.

2. In general, the ratio of freq(recessive homozygote with inbreeding f) to freq(recessive homozygote with inbreeding f = 0) given a recessive allele of frequency q is $\frac{q^2 + fpq}{q^2} = 1 + f\frac{p}{q}$. If $f = .005$ and $q = .005$, then $p = 1 - .005 = .995$ and the ratio is $1 + (.005)(.995)/(.005) = 1.995$.

Even a relatively small amount of inbreeding can significantly amplify the frequency of a rare recessive condition relative to a randomly mating population. In this case, an inbreeding coefficient of just 0.5% would almost double the frequency of the recessive disease.

3. The four possible first-cousin pedigrees are shown at right. Note that pedigrees I and II contain successive males in all lines leading to offspring h and so have inbreeding coefficients $f = 0$. In pedigree III, the only chain that contributes to inbreeding is "e-c-**b**-d-g" since the chain through "a" contains successive males. The inbreeding coefficient for this case is $f = 1\cdot(1/2)\cdot(1/2)\cdot 1\cdot(1/2) = 1/8$. In case IV, there are two feasible chains: "e-c-**b**-d-g" and "e-c-**a**-d-g". (Note that neither contains successive males.) Inbreeding



via the first chain is 1·(1/2) ·(1/2) · ·(1/2)·(1/2) = 1/16 and 1·(1/2) ·1 ·(1/2)·(1/2) = 1/8 via the second chain. Summing these gives $f = 3/16$.

4. Case 1:

$H_I = (.5 + .3)/2 = .4$

$\overline{H}_S$: $H_{S,1} = 2p_1q_1 = 2(.5)(.5) = .5$; $H_{S,2} = 2p_2q_2 = 2(.5)(.5) = .5 \Rightarrow \overline{H}_S = (.5+.5)/2 = .5$

H_T: $q_1 = q_2 = 0.5 \Rightarrow \overline{q} = \big(.5 + .5\big)/2 = .5 \Rightarrow H_T = 2\overline{p}\overline{q} = 2(.5)(.5) = .5.$

So $F_{IS} = \dfrac{\overline{H}_S - H_I}{\overline{H}_S} = (.5 - .4)/.5 = .2$; $F_{ST} = \dfrac{H_T - \overline{H}_S}{H_T} = (.5 - .5)/.5 = 0$;

$F_{IT} = \dfrac{H_T - H_I}{H_T} = (.5 - .4)/.5 = .2.$

Case 2:

$H_I = (.5 + .42)/2 = .46$

$\overline{H}_S$: $H_{S,1} = 2p_1q_1 = 2(.5)(.5) = .5$; $H_{S,2} = 2p_2q_2 = 2(.7)(.3) = .42 \Rightarrow \overline{H}_S = (.5+.42)/2 = .46$

H_T: $q_1 = 0.5, q_2 = 0.3 \Rightarrow \overline{q} = \big(.5 + .3\big)/2 = .4 \Rightarrow H_T = 2\overline{p}\overline{q} = 2(.6)(.4) = .48$

So $F_{IS} = \dfrac{\overline{H}_S - H_I}{\overline{H}_S} = (.46 - .46)/.46 = 0$; $F_{ST} = \dfrac{H_T - \overline{H}_S}{H_T} = (.48 - .46)/.48 = .0417$;

$F_{IT} = \dfrac{H_T - H_I}{H_T} = (.48 - .46)/.48 = .0417$

5. Subpopulation 1: $p_1 = p_2 = \cdots = p_5 = 1/5$, $p_6 = p_7 = \cdots = p_{10} = 0$

Subpopulation 2: $p_1 = p_2 = \cdots = p_5 = 0$, $p_6 = p_7 = \cdots = p_{10} = 1/5$

Averages: $\overline{p}_1 = \overline{p}_2 = \cdots = \overline{p}_{10} = 1/10$

$$H_{S,1} = H_{S,2} = 1 - \sum_{i=1}^{10} p_i^2 = 1 - 5(1/5)^2 = 4/5 \Rightarrow \overline{H}_S = 4/5.$$

$$H_T = 1 - \sum_{i=1}^{10} \overline{p}_i^2 = 1 - 10\big(1/10\big)^2 = 9/10 \Rightarrow G_{ST} = (.9 - .8)/.9 = .111$$

On the surface, this value of G_{ST} is not particularly large despite the two subpopulations not sharing even a single allele! The morale of the exercise is that any particular value of F_{ST} (G_{ST}) has limitations as an indicator of population substructuring. One must also consider the number and distribution of alleles on which a given fixation index computation is based.

6. $$N_e\big(\text{diploid}\big) = \frac{4N_fN_m}{N_f + N_m} = \frac{4(5)(1)}{5+1} = 3\tfrac{1}{3}$$

$$N_e\big(\text{haplo - diploid}\big) = \frac{9N_fN_m}{2N_f + 4N_m} = \frac{9(1)(10)}{2(1) + 4(10)} = 2.14$$

7. Exact: $N_e = \frac{1}{2}\left\{1-\left[(1-1/10)(1-1/100)(1-1/20)(1-1/200)\right]^{1/4}\right\}^{-1} = 11.9;$

Approximate: $N_e \approx \frac{4}{1/5+1/50+1/10+1/100} = 12.1.$ The exact and approximate values of N_e are obviously very close to one another!

8. $N_e(\text{mtDNA}) = \frac{N_{ef}}{2} = \frac{100}{2} = 50;$

$N_e(\text{Y chrom}) = \frac{N_m}{2} = \frac{10}{2} = 5;$

$N_e(\text{automsomal}) = \frac{4N_{ef}N_{em}}{N_{ef}+N_{em}} = \frac{4(100)(10)}{100+10} = 36.4$