

COLUMBIA UNIVERSITY

DEPARTMENT OF BIOSTATISTICS

P8149 – HUMAN POPULATION GENETICS

Exercise Sheet 3 (Model Answers)

Question 1 [3 marks]

The mean fitness is

$$\bar{w} = P_{AA}w_{AA} + P_{Aa}w_{Aa} + P_{aa}w_{aa}$$

where $0 \leq P_{AA}, P_{Aa}, P_{aa} \leq 1$. Since the fitnesses are standardized, $0 \leq w_{AA}, w_{Aa}, w_{aa} \leq 1$.

Therefore, $w_{AA}P_{AA} \leq P_{AA}$, $w_{Aa}P_{Aa} \leq P_{Aa}$ and $w_{aa}P_{aa} \leq P_{aa}$, so that

$$\bar{w} = P_{AA}w_{AA} + P_{Aa}w_{Aa} + P_{aa}w_{aa} \leq P_{AA} + P_{Aa} + P_{aa} \leq 1$$

Question 2 [1+2=3 marks]

(a) Because of random mating, the genotypic frequencies are

$$(P_{AA}, P_{Aa}, P_{aa}) = (.2^2, 2 \times .2 \times .8, .8^2) = (.04, .32, .64).$$

The unstandardized frequencies in the adult population are

$$(1 \times .04, .5 \times .32, .25 \times .64) = (.04, .16, .16).$$

The standardized frequencies are

$$(P_{AA}, P_{Aa}, P_{aa}) = \left(\frac{.04}{.04 + .16 + .16}, \frac{.16}{.04 + .16 + .16}, \frac{.16}{.04 + .16 + .16} \right) = \left(\frac{1}{9}, \frac{4}{9}, \frac{4}{9} \right).$$

(b) The observed frequencies of the three genotypes are, respectively,

$$P_{AA} = \frac{165}{165 + 562 + 339} = .155,$$

$$P_{Aa} = \frac{562}{165 + 562 + 339} = .527,$$

$$P_{aa} = 1 - .155 - .527 = .318.$$

The observed frequencies of the alleles A and a are, respectively,

$$p = P_{AA} + \frac{1}{2}P_{Aa} = .155 + \frac{1}{2}(.527) = .418$$

$$q = 1 - p = 1 - .418 = .582$$

Therefore. the expected genotype frequencies for AA, Aa, and aa are, respectively,

$$p^2 = .418^2 = .175$$

$$2pq = 2(.418)(.582) = .486$$

$$q^2 = (.582)^2 = .339$$

The absolute viabilities then are

$$w_{AA}^* = \frac{Obs}{Exp} = \frac{.155}{.175} = .886,$$

$$w_{Aa}^* = \frac{.527}{.486} = 1.084,$$

$$w_{aa}^* = \frac{.318}{.339} = .938,$$

and the relative viabilities are

$$w_{AA} = \frac{.886}{1.084} = .817,$$

$$w_{Aa} = 1,$$

$$w_{aa} = \frac{.938}{1.084} = .865.$$

Question 3 [2+1=3 marks]

For population 1,

genotype	AA	AS	SS	AC	SC	CC
fitness	.9	1.0	.2	.9	.71	1.31
Freq.	.81	.09	.0025	.09	.005	.0025

The average fitness of the population is

$$\bar{w} = (.9)(.81) + (1.0)(.09) + + (1.31)(.0025) = .907325$$

The average excesses are

$$\begin{aligned}\alpha_A &= p_A w_{AA} + p_S w_{AS} + p_C w_{AC} - \bar{w} \\ &= (.9)(.9) + (.05)(1.0) + (.05)(.9) - .907325 \\ &= -.002325\end{aligned}$$

$$\begin{aligned}\alpha_S &= p_S w_{SS} + p_A w_{AS} + p_C w_{SC} - \bar{w} \\ &= (.05)(.2) + (.9)(1.0) + (.05)(.71) - .907325 \\ &= .038175\end{aligned}$$

$$\begin{aligned}\alpha_C &= p_C w_{CC} + p_A w_{AC} + p_S w_{SC} - \bar{w} \\ &= (.05)(1.31) + (.9)(.9) + (.05)(.71) - .907325 \\ &= .003675\end{aligned}$$

The change in allele frequencies are

$$\begin{aligned}\Delta p_A &= \frac{p_A}{\bar{w}} \alpha_A = \frac{.9}{.907325} (-.002325) = -.0023062 \\ \Delta p_S &= \frac{p_S}{\bar{w}} \alpha_S = \frac{.05}{.907325} (.038175) = .00210371 \\ \Delta p_C &= \frac{p_C}{\bar{w}} \alpha_C = \frac{.05}{.907325} (.003675) = .00020252\end{aligned}$$

Similarly, for population 2,

$$\begin{aligned}\bar{w} &= .90895 \\ \alpha_A &= -.00395, \alpha_S = .02205, \alpha_C = .02255 \\ \Delta p_A &= -.0036938, \Delta p_S = .00121294, \Delta p_C = .00248088\end{aligned}$$

Similarly, for population 3,

$$\begin{aligned}\bar{w} &= .91835 \\ \alpha_A &= -.01335, \alpha_S = -.01635, \alpha_C = .05415 \\ \Delta p_A &= -.0109027, \Delta p_S = -.0008902, \Delta p_C = .01179289\end{aligned}$$

(c) Yes, the initial frequencies do affect the course of adaptive evolution.

Question 4 [2 marks]

Mutation rate per gene per generation = $2 \times 10^{-6} \times 36 = 7.2 \times 10^{-5}$

Question 5 [2 marks]

To find the equilibrium frequency of allele a, we set the change in the allele frequency of allele A to zero:

$$\Delta p = v\tilde{q} - u(1 - \tilde{q}) = 0 \Rightarrow \tilde{q} = \frac{u}{u + v}$$

For $u = 10v$,

$$\tilde{q} = \frac{10v}{10v + v} = \frac{10}{11} \approx .91.$$

For $u = 100v$,

$$\tilde{q} = \frac{100v}{100v + v} = \frac{100}{101} \approx .99$$

Question 6 [1+2=3 marks]

The observed heterozygosity is

$$H_0 = \frac{10}{20} = .5$$

The allele frequencies are

$$p = \text{freq}(A_1) = \frac{10(2) + 10}{40} = \frac{3}{4},$$

$$q = \text{freq}(A_2) = 1 - \frac{3}{4} = \frac{1}{4}.$$

Therefore, the expected heterozygosity is

$$H_e = 1 - .75^2 - .25^2 = .375$$

(a) The observed heterozygosity is

$$H_0 = \frac{20}{20} = 1$$

Denoting the frequency of allele A_i by p_i ($i = 1, 2, \dots, 40$), we have $p_i = 1/40$.

Therefore, the expected heterozygosity is

$$\begin{aligned} H_e &= 1 - \sum_{i=1}^{40} p_i^2 \\ &= 1 - \sum_{i=1}^{40} \left(\frac{1}{40} \right)^2 \\ &= 1 - 40 \left(\frac{1}{40} \right)^2 \\ &= 1 - \frac{1}{40} \\ &= .975 \end{aligned}$$

Question 7 [1+1+2=4 marks]

(a) Number of segregating sites = 8

(b) Nucleotide polymorphism = $\frac{8}{15} = .53$

(c) Nucleotide diversity = $\frac{\sum_{i < j} T_{ij}}{\binom{n}{2}} = \frac{4 + 6 + 6 + 6 + 4 + 6 + 4 + 4}{\binom{5}{2}} = 4.0$
