COLUMBIA UNIVERSITY

DEPARTMENT OF BIOSTATISTICS

P 8149 - HUMAN POPULATION GENETICS

exercise sheet 4 (covers chapters 7 - 10)

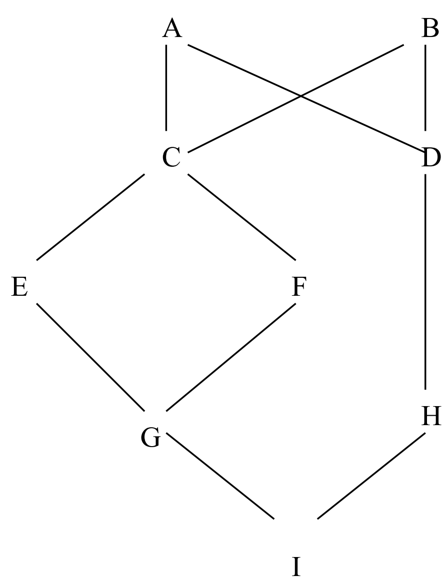
Date due: Monday Dec 8, 2025 (SUBBMIT ONLINE IN CANVAS)

Question 1
Suppose that a random-breeding population is sampled and the following genotype frequencies of a protein variant are found:

AA	*AB*	*BB*
42	76	448

What is the inbreeding coefficient indicated by these numbers?

Question 2



Calculate the inbreeding coefficient for individual I in the diagram above.

Question 3

Consider a population with inbreeding coefficient *f.* Let a bi-allelic locus have allele frequencies *p* and $q(=1-p)$, and genotype frequencies (P_{11}, P_{12}, P_{22}). Prove that equation (A) below implies both (B) and (C):

$$P_{11} = p^2 + pqf$$
$$P_{12} = 2pq(1-f)$$
$$P_{22} = q^2 + pqf$$

Question 4

In a diploid population of size 50, the frequencies of A_1 and A_2 are .7 and .3. What is the expected heterozygosity after ten generations?

Question 5

An allele has a frequency of .01 in a population.

(a) If genetic drift is the only force operating, what is the probability under random mating that the allele will ultimately be lost if the (diploid) population size is 50?

(b) What is the probability of lost if the size of the (diploid) population is 5000?

Question 6

(a) What is the probability of loss of a new mutation in the first generation if all families are of size 2?

(b) What is the probability of loss of a new mutation in the first generation if the family size follows a Poisson distribution of size 2?

(c) Explain the difference, if any, between your results in **a.** and **b.** above.

Question 7

	number of individuals		
	A_1A_1	A_1A_2	A_2A_2
subpopulation 1	10	180	810
subpopulation 2	250	500	250

(a) For the above subdivided population, calculate f_{is}, f_{it}, and f_{st}.

(b) Now assume that the two subpopulations are indistinguishable and calculate *f* for the combined population.

Question 8
Assume that one population has 5 alleles at a particular locus, each with a frequency of .2. A second population has five different alleles (i.e. different from the set of alleles in the first population) at the same locus, each with a frequency of .2. Calculate f_{st} for this locus and comment on your result.

Question 9
(a) What is the probability that two alleles coalesced 10 generations before the present in a diploid population of size $N = 50$?
(b) What is the expected time in a diploid population of size $N = 5$ that there are exactly 10 lineages?
(c) What is the expected time in a diploid population of size $N = 5$ that there are exactly 2 lineages?

Question 10
Consider a haploid population with size $2N$ in mutation-drift balance. Assuming an infinite-alleles model, use the coalescent approach to prove that the probability of ibd between two gene copies is

$$\overline{F}_{eq} = \frac{1}{1+4N\mu},$$

where μ is the mutation rate per gene per generation.

Question 11
Assume $h = .03$ for a mutation that is lethal when homozygous and $h = .4$ for a mildly deleterious mutation when homozygous ($s = .05$). In both cases, the mutation rate to the deleterious form is 10^{-5} .
(a) Assuming random mating, what are the equilibrium frequencies for the mutant allele in these two cases?
(b) Assuming $f = .1$ and that both mutants are now completely recessive, what are the equilibrium frequencies for the mutant allele in these two cases?

ANSWERS

Ans: Q1: .446; Q2: f = .0625; Q4: $H_{10} = .38$, Q7: (a) $f_{is} = 0, f_{it} = .19$, $f_{st} = .19$; (b) $f = .19$; Q8: $f_{st} = .11$; Q9: (a) 9.14 x 10^{-3}, (b) 1.00, (c)10.0; Q11: (a) $\tilde{q} = 3.3 \times 10^{-4}$ (first case), $\tilde{q} = 5.0 \times 10^{-4}$ (second case), (b) $\tilde{q} = 1.0 \times 10^{-4}$ (first case), $\tilde{q} = 2.0 \times 10^{-3}$ (second case)
