

NAME _____
POPULATION GENETICS AND MICROEVOLUTIONARY THEORY
MIDTERM EXAMINATION

The following table may be useful in answering one or more of the questions. Use it as needed:

Probability	Chi-Square Test Statistic Value			
	df = 1	df = 2	df = 3	df=4
0.10	2.706	4.605	6.251	7.779
0.05	3.841	5.991	7.815	9.488
0.01	6.635	9.210	11.345	13.277

1. Briefly define (5 points each):

a) **Interval Mapping. A measured genotype approach in quantitative genetics in which a QTL is hypothesized to lie between two adjacent markers and the likelihood of the QTL being at various intermediate positions between the flanking markers is statistically evaluated.**

b) **Tree Scanning A measured genotype approach in quantitative genetics applicable to genomic regions with little to no recombination such that a haplotype tree can be estimated. Tree scanning partitions the haplotype tree into two or more mutually exclusive and exhaustive clades and then treats each clade as an “allele” in a genotypic analysis of phenotypic associations.**

c) **Environmental deviation What is left over of the phenotype of an individual that is not explicable by the grand mean the genetic model being used.**

2. A sample of individuals is surveyed tested at an autosomal locus with two alleles as follows:

Genotype	AA	Aa	aa	Total
Number	22	36	42	100

a. (1 point) Characterize this population by its genotypic frequencies.

Freq(AA)=22/100=0.22; Freq(Aa)=36/100=0.36; Freq(aa)=42/100=0.42

b. (2 points) Characterize the gene pool by the allele frequencies for A and a.

Freq(A)=(2*22+36)/200 = 0.4; Freq(a)=(36+2*42)/200 = 0.6

c. (4 points) Test the goodness of fit of this population to the Hardy-Weinberg expectations.

				Sum
H-W. Freq.	0.1600	0.4800	0.3600	1
Exp.	16.000	48.000	36.000	100
(o-e) ² /e	2.25	3	1	6.25

From the Chi-square table with 1 degree of freedom, the above is significant at the 5% level, so reject the null hypothesis of HW.

d. (4 points) Estimate the correlation of uniting gametes in this sample and list three distinct explanations to explain its value.

The correlation of uniting gametes is the same as f , so $f=1-0.36/0.48 = 0.25$.

$f > 0$ could be explained by

- 1) inbreeding**
- 2) assortative mating**
- 3) population stratification (Wahlund effect)**

3. a. (4 points) Suppose a deme is scored for 3 randomly chosen loci. The genotypic frequency results are as follows:

locus i	A_iA_i	A_ia_i	a_ia_i
1	.09	.42	.49
2	.36	.48	.16
3	.111	.378	.511

What do you conclude about this population's system of mating (assume no selection or drift)?

For locus 1, $p = 0.3$; $f=1-0.42/(2pq)=0$

For locus 2, $p = 0.6$; $f=1-0.48/(2pq)=0$

For locus 3, $p = 0.3$; $f=1-0.378/(2pq)=0.1$

The population is randomly mating at loci 1 and 2, and assortatively mating at locus 3. This is not inbreeding at locus 3 because inbreeding affects all loci.

b. (4 points) Same question as (a), but now assume the results were:

locus i	A_iA_i	A_ia_i	a_ia_i
1	.05	.50	.45
2	.2025	.5950	.2025
3	.45	.50	.05

For locus 1, $p = 0.3$; $f=1-0.5/(2pq)= -0.19$

For locus 2, $p = 0.5$; $f=1-0.595/(2pq)= -0.19$

For locus 3, $p = 0.7$; $f=1-0.5/(2pq)= -0.19$

The population is avoiding inbreeding (all loci equally affected and negative f)

4. (4 points) Linkage disequilibrium is monitored between two loci, each with two alleles, in an infinite-sized population, with no natural selection occurring. The statistic D is observed over several generations:

Generation	1	2	3	4
D	0.10	0.13	0.17	0.20

Suppose the genotype frequencies at each of the two loci involved in the linkage disequilibrium are determined, and f is calculated for each locus. Do you expect these f 's to be positive, negative, or zero?

Because D is increasing, there must be assortative mating associated with these two loci, so expect f to be positive for these two loci.

5. 100 isolated populations are established, each with an initial allele frequency of 0.2 for the A allele at an autosomal locus, and each with a constant variance effective size of 25 and each randomly mating.

a) (2 points) How many of these populations are expected to become fixed for the A allele under neutrality?

Since drift has no direction, the average allele frequency is expected to remain the same, so the expected number of populations fixed for A is $0.2(100)=20$

b) (2 points) The system of mating is now changed such that $f = -0.1$. How many of these populations are now expected to become fixed for the A allele under neutrality?

Drift operates on the gene pool level, so changes in system of mating have no effect, so $0.2(100)=20$

c) (2 points) The variance effective size of each population is increased to 2 billion. How many of these populations are now expected to become fixed for the A allele?

All finite populations ultimately go to fixation, so $0.2(100)=20$

6. a) (4 points) A deme of inbreeding effective size 20 has an initial $\bar{F}(0)=0.1$. What is $\bar{F}(1)$ and $\bar{F}(2)$?

Use: $\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\bar{F}(t-1)$ to get $F(1)=1/40 + (39/40)(0.1) = 0.1225$

Use: $\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\bar{F}(t-1)$ to get $F(2)=1/40 + (39/40)(0.1225) = 0.1444$

6. b) (4 points) A deme of inbreeding effective size 100 has an initial $\bar{F}(0)=0.1$. What is $\bar{F}(1)$ and $\bar{F}(2)$?

Use: $\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\bar{F}(t-1)$ to get $F(1)=1/200 + (199/200)(0.1) = 0.1045$

Use: $\bar{F}(t) = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)\bar{F}(t-1)$ to get $F(2)=1/200 + (199/200)(0.1045) = 0.1090$

7. a) (2 points) A neutral pseudogene has a nucleotide mutation rate of 10^{-9} . What is the rate of nucleotide substitution in this pseudogene in a population with variance effective size 100?

Neutral rate of substitution = neutral mutation rate = 10^{-9}

b) (2 points) Suppose now the variance effective size is changed to 1,000,000. What is the rate of nucleotide substitution?

Neutral rate of substitution = neutral mutation rate = 10^{-9}

c) (2 points) Suppose now that nucleotide mutation rate is 10^{-8} and the variance effective size is 1000. What is the rate of nucleotide substitution?

Neutral rate of substitution = neutral mutation rate = 10^{-8}

8. The times (in years) to the most recent common ancestor (TMRCA) for several X-linked loci were determined as follows:

Locus	1	2	3	4	5	6	7	8	9	10
TMRCA	90,000	120,000	100,000	80,000	110,000	75,000	105,000	125,000	95,000	100,000

a) (5 points) Given that there is one generation per year, what is the long term inbreeding effective size of this population?

The Average coalescence time for an X-linked locus from the above data is 100,000 years.

The expected coalescence time for an X-linked locus is $3N_{ef}$, so estimate N_{ef} by $100,000/3 = 33,333$.

b) (3 points) What is the expected coalescence time for an autosomal locus in this population?

The expected coalescence time for an autosomal locus is $4N_{ef}$, so $(4/3)100,000 = 133,333$ years.

c) (4 points) Assuming that the reproductive properties of both sexes are identical and that the sex ratio is 50:50, what is the expected coalescence time for mitochondrial DNA and for the Y chromosome?

The expected coalescence time for a haploid, unisexual locus in a population with 50:50 sex ratio and identical reproductive properties for both sexes is $N_{ef} = 33,333$ years.

9. (5 points) What is the expected value of a Molecule Genetic Distance? Briefly define all terms used.

The expected value of a molecule genetic distance is always $2\mu t$ where μ is the mutation rate at the locus (nucleotide) and t is the time to the common ancestral molecule of the two molecules being compared.

10. (5 points) The expected heterozygosity at a neutral locus is determined to be 0.6. What is the probability that two randomly chosen copies of this gene will experience coalescence before mutation given that coalescence or mutation has occurred?

Prob(coal. before mutation | coal. or mutation) = 1 - Prob(mutation before coal. | coal. or mutation) = 1 - Expected Heterozygosity = 1 - (0.6) = 0.4

11. (8 points) Two populations exchange gametes at a rate of 0.005 every generation. Given that the two populations were initially fixed for different alleles at an autosomal locus, what is the expected difference (assuming neutrality and infinite population size in each deme) in the allele frequency between the demes at generation 10 and at generation 100? What are the expected differences in allele frequency at equilibrium?

From lecture, $d_n = d_0(1-2m)^n$ Here, $d_0 = 1$, so $d_n = (1-2m)^n = (0.99)^n = 0.9044$ for $n=10$ and 0.3660 for $n=100$

At equilibrium, the expected difference is 0.

12. Population 1 has a frequency of the A allele of 0.3 at an autosomal locus, and population 2 has a frequency of 0.7 for A. Both populations have the same number of individuals.

a) (6 points) What is f_{st} ?

Expected Heterozygosity Pop. 1 = $2(.3)(.7) = 0.42$; Expected Heterozygosity Pop. 2 = $2(.7)(.3) = 0.42$

The average expected heterozygosity within populations is therefore 0.42

Pooling the two demes together yields a pooled p of 0.5, so the Expected Heterozygosity in the Pooled Population is $2(.5)(.5) = 0.5$.

$$f_{st} = (0.5 - 0.42) / (.5) = 0.16$$

b) (6 points) Each population is determined to have a size of 100 ideal individuals (e.g., self-compatible, randomly mating hermaphrodites, etc.). What is the variance effective size of the total population consisting of these two demes (assume there are no other demes)?

$$N_{evT} = \frac{nN}{1 - f_{st}}$$

Here, $n=2$, $N=100$, so the total variance effective size = $200 / (.84) = 238.1$