

NAME _____

POPULATION GENETICS AND MICROEVOLUTIONARY THEORY
FINAL EXAMINATION

(Write your name on every page. One point will be deducted for every page without your name!)

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) Net Reproductive Rate

The expected number of half-offspring born to a zygote (or new-born, in many cases).

Could add that it is a function of age-specific survivorship and age-specific mating success and fecundity.

b) Mutualism

An interaction between individuals that is of mutual benefit to both.

c) R1 element

A type of retrotransposon that specially inserts into the 28S ribosomal DNA on the X chromosome of *Drosophila*. Could add that it is one of the necessary (but not sufficient) conditions for *aa* in *Drosophila mercatorum*.

d) Unit of Selection

The level of genetic organization that allows the prediction of the genetic response to selection.

e) Coarse-Grained Heterogeneity

A coarse-grained environment occurs when an individual remains in a single environment throughout its lifetime, but the environment varies between demes occupying different spatial locations or across generations.

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

Genotype	AA	Aa	aa	Total
Number	310	591	99	1000

a. (6 points) Given random mating at this locus and the allele frequency of *A* in the gene pool produced by the parents of these individuals being 0.6, what are the zygote-to-adult relative viabilities of the genotypes, setting the relative viability of *Aa* to 1?

adult freq. \propto (zygotic freq.) \times viability, so viability \propto adult freq/zygotic freq.

The zygotic freq. are 0.36 for AA, 0.48 for Aa, and 0.16 for aa under rm with p=.6 (1 pt)

Hence, the viabilities are 0.310/.36=0.861, 0.591/0.48=1.231 and 0.099/0.16=0.619 for AA, Aa, and aa (3 pts)

The relative viabilities are 0.861/1.231 = 0.7 for AA, 1 for Aa, and 0.619/1.231=0.5 for aa (2 pts)

b. (6 points) Same problem as a), but now assume there is positive assortative mating at this locus such that $f=0.1$.

The zygotic freq. are 0.384 for AA, 0.432 for Aa, and 0.184 for aa under $f=0.1$ with p=.6 (1 pt)

Hence, the viabilities are 0.310/.384=0.807, 0.591/0.432=1.368 and 0.099/0.184=0.538 for AA, Aa, and aa (3 pts)

The relative viabilities are 0.807/1.368 = 0.59 for AA, 1 for Aa, and 0.538/1.368=0.39 for aa (2 pts)

3. (6 points) An effectively infinite-sized, random mating population is fixed for the a allele at an autosomal locus. A mutation to A then occurs, with the mean number of offspring being 1 for aa , 1.05 for Aa and 1.3 for AA . The variance in the number of offspring is 1 for aa , 1.1 for Aa , and 0.9 for AA . What is the probability of fixation of A ?

Only the mean and variance of offspring of aa and Aa are relevant under random mating when A is rare, since once A gets common it will be driven to fixation by selection (1 pt). Hence, the probability of fixation is the probability of A survives these initial generations of rarity in the population (1 pt). The probability that A survives is

$$\Pr(A \text{ survives}) = \frac{2s}{1 + s + \frac{1}{s}} \quad (1 \text{ pt})$$

In this case, $s=0.05$ and $\frac{1}{s} = 20$ (2 pts), so prob. A survives = $0.1/1.1 = 0.09$ (1 pts).

4. (6 points) A trait y is related to fitness by the function $w(y)=y^2-15y+14$. Assuming that the fundamental theorem of natural selection is valid in this case, what condition(s) is(are) necessary for the average value of y at equilibrium to be at its optimal value [i.e., $w(y_{eq})$ is maximized]?

The relevant equation is $\bar{w}_{eq} = w(x_{eq}) + \frac{1}{2}w''(x_{eq})\sigma_{eq}^2(x)$, although in this case the trait is indicated by y and not x (2 pts).

The second derivative of $w(y)$ is 2 (2 pts), so the only way trait y is optimized at equilibrium is to have $\sigma_{eq}^2(y) = 0$, that is, there is no phenotypic variance in trait y at equilibrium (2 pts).

5. (12 points): The coding region of a gene is sequenced in several individuals of species 1 and a haplotype tree is estimated from these sequence data. One individual of the closely-related species 2 is sequenced to root the haplotype tree. The following table summarizes the results:

Mutational class	Polymorphic Tip	Polymorphic Interior	Fixed Interspecific
Nonsynonymous	15	3	7
Synonymous	5	6	14

Test the null hypothesis of neutrality with a contingency chi-square test that uses what should be the most informative categories, keeping in mind that contingency chi-squares require that the expected numbers in all cells ≥ 5 . If you reject the null hypothesis of neutrality, describe the type of selection indicated by the data.

In the 2x3 table above, the expected values in the polymorphic interior cells are 4.5, so pooling is necessary to perform a contingency chi-square (2 points).

The pooling that should have the highest power is to pool polymorphic interior with fixed interspecific to yield the young vs. old classification (4 points). The resulting table is:

	Synonymous	Non-Synonymous	total
young	5	15	20
old	20	10	30
total	25	25	50

	expected 1	expected 2
young	10	10
old	15	15

chi-square 8.333333333
 p value 0.003892417 Or <.01 from Table, df=1

1 pt. table, 1 pt. expected values, 1 pt. chi-square, 1 pt. p<.01 (need to have df right).

There is an excess of synonymous mutations in the old, evolutionarily successful category, (or an excess of nonsynonymous mutations in the young, unproven category) which implies conservative selection that maintains the amino acid structure. (2 points).

If the student used the McDonald-Kreitman test, they don't get the 4 pts for choosing the most informative pooling, but they can get some credit if they implemented it correctly as follows:

	Synonymous	Non-Synonymous	total
Polymorphic	11	18	29
Fixed	14	7	21
Total	25	25	50

	expected 1	expected 2
Polymorphic	14.5	14.5
Fixed	10.5	10.5

chi-square 4.022988506
 p value 0.04488412 Or <.05 from Table, df=1

1 pt. table, 1 pt. expected values, 1 pt. chi-square, 1 pt. p<.01 (need to have df right).

There is an excess of synonymous mutations in the old, evolutionarily successful category, (or an excess of nonsynonymous mutations in the young, unproven category) which implies conservative selection that maintains the amino acid structure. (2 points).

6. The following fitnesses were estimated for the genotypes at an autosomal locus found in a deme:

Genotype:	AA	Aa	aa
Fitness:	0.8	1	1.2

The frequency of A in the gene pool is 0.2, and the system of mating is avoidance of inbreeding with $f=0.1$.

a) (6 points) What are the average excesses for the phenotype of fitness for the A and a alleles?

The expected genotype frequencies, t_{ij} , with $f=0.1$ and $p=0.2$ are 0.024, 0.352, and 0.624 respectively. (1 pt). From this, get the average fitness as 1.12 (1 pt)

Then use $a_i = \frac{t_{ii}}{p_i} g_{ii} + \sum_{j \neq i} \frac{\frac{1}{2} t_{ij}}{p_i} g_{ij} = \sum_j t(ij|i) g_{ij}$ to get $a_A = -0.144$ and $a_a = 0.036$ (4 pts)

b) (2 points) What is the change in the frequency of A in the next generation due to selection in this deme?

Use $\Delta p = \frac{p}{\bar{w}} a_A$ (1 pt), to get -0.026 (1 pt).

c) (4 points) A second deme exists for which these alleles are neutral and for which the system of mating is inbreeding with $f=0.2$. The frequency of A in this second deme is 0.7. Gene flow exists between the 2 demes such that 10% of the gametes are exchanged between them every generation. Given this additional information, what is the frequency of A in the next generation in deme 1?

The relevant new info is that $m=0.1$ and $p_2=0.7$. (1 pt). Then use the equation $\Delta p_1 = \frac{p_1}{\bar{w}} a_A + m(p_2 - p_1)$ = -0.026 + 0.05 = 0.024 (2 pts.) Hence, the frequency of A in deme 1 is 0.2+0.024=0.224 (1 pt).

7. A locus is sequenced in species 1 and it is found that the statistic $D = \frac{! \#}{\sqrt{\text{Var}(! \#)}} = -4$ is significantly different from zero.

a) (6 points) Assuming the infinite sites model is appropriate, and that the mutation rate at this locus is 10^{-5} , provide reasonable evolutionary explanation(s) for this result.

This is Tajima's D statistic, and given the infinite sites model, it is an appropriate statistic. A significant negative value has two possible evolutionary explanations:

- 1) **positive selection leading to a selective sweep in the recent past (3 pts)**
- 2) **a bottleneck or founder event in the recent past (3 pts)**

b) (6 points) An outgroup species is now sequenced such that one can infer which polymorphic nucleotide states in species 1 are ancestral versus derived. This allows the calculation of S_i , the number of derived variants found i times in the sample of n sequences. Then the statistic $!_H = \sum_{i=1}^{n-1} \frac{2S_i i^2}{n(n-1)}$ is substituted for $!$ in the D statistic given above, and the resulting new statistic remains significantly negative. Provide reasonable evolutionary explanation(s) for this result.

This is the Fay and Wu modification, and it is not sensitive to bottleneck or founder events (3 pts), so this implies that there has been positive selection at this locus (3 pts).

8. Members of a species encounter each other at random and engage in competitive interactions. The competitive fitness interaction matrix for the genotypes at a locus with two alleles, A and a , with random mating, is:

	AA	Aa	aa
AA	0.3	0.4	1
Aa	1.6	0.89	1.5
aa	2.1	3	3.1

a). (3 points) Is the A allele protected from loss?

When A is rare, the relevant contrast is Aa competing with aa versus aa competing with aa ; that is 3 versus 3.1 Since $3 < 3.1$, A is NOT protected from loss.

Alternative, $w_{\bar{a}}$ is about 3.1 when A is rare, so the average excess of A is about $3 - 3.1 = -0.1 < 0$, so A is not protected.

b) (3 points) Is the a allele protected from loss?

When a is rare, the relevant contrast is Aa competing with AA versus AA competing with AA ; that is 0.4 versus 0.3 Since $0.4 > 0.3$, a is protected from loss.

Alternative, $w_{\bar{a}}$ is about .3 when a is rare, so the average excess of a is about $.4 - .3 = 0.1 > 0$, so a is protected.

9. A species lives in an environment that changes gradually from one state to a second state over 10 km.

a) (2 points) A study was made at a genetically variable locus in this species over a transect that straddles the spatial change, including the documentation of a change in selective values over the transect and the

amount of gene flow, to yield the following statistic: $c = d\sqrt{\frac{m}{b!}} = 6$ km. Is the spatial heterogeneity an ecotone or a gradient with respect to this locus.

Since $c < 10$, this is a GRADIENT.

b) (3 points) It is observed that the body color of the species changes from light brown to black in the two different environments, with the color change occurring gradually over the 10 km long transition zone. Is this an example of a genetic cline?

There is a phenotypic cline, but no information is given about its genetic basis, so the correct answer is INSUFFICIENT INFORMATION TO ANSWER THE QUESTION, or UNKNOWN, or MAYBE.

10. Consider a Levene model of a population with an island-model of dispersal with $m=0.1$ across three environments, each with random mating within, such that the relative fitnesses for the genotypes determined by an autosomal locus with 2 alleles are:

Genotype \ Env.	AA	Aa	Aa
1	0.8	1	1.1
2	1.1	1	1.1
3	1.2	1	0.7

a). (4 points) Assume soft selection in which the portions of the total population that come from the three environments are 0.24, 0.61, and 0.15 respectively. Is the polymorphism protected?

**For soft selection, first look at $v_i < 1 - m = 0.9$ and $w_i < 1 - m = 0.9$
Here, $v_1 = 0.8 < 0.9$, and $w_3 = 0.7 < 0.9$, so protected.**

b). (4 points) Is your answer to a) sensitive to the exact proportions of the population that come from the three environments, given that all environments provide individuals to the total population? Why or why not?

No, the conditions for protection above do not depend on the c 's because gene flow is sufficiently restricted relative to the fitness in particular environments such that protection will always occur.

c). (4 points) Now assume a hard selection model in which the proportions of zygotes entering the three environments are 0.3, 0.3 and 0.4 respectively. Is the polymorphism protected?

**For hard selection, first look at $v_i = (1 - m)v_i + m \sum_i z_i v_i / 1 - m$ and $w_i = (1 - m)w_i + m \sum_i z_i w_i / 1 - m$
 $v_1 = 0.825 < 0.9$, and $w_3 = 0.724 < 0.9$, so protected.**

11. (6 points) Consider the following life history table:

Age	l_x	$m_x b_x$
0	1	0
1	0.8	0.5
2	0.5	1
3	0.3	0.5
4	0.1	0.5
5	0	0

What is the approximate value of the Malthusian parameter?

Use $r = \frac{1}{\bar{T}} \ln R_0$ where $\bar{T} = \frac{\sum_{x=0}^{\max \text{ age}} l_x m_x b_x}{\sum_{x=0}^{\max \text{ age}} l_x b_x}$ and $R_0 = \sum_{x=0}^{\max \text{ age}} l_x m_x b_x$ (2 pts).

$R_0=1.1$ (1 pt), and $\bar{T} = 2.05/1.1 = 1.86$ (1 pt), so $r = 0.049$ (2 pts).

12. A population is nearly fixed for the allele A. The neutral mutation rate at this locus is 10^{-5} , and the mutation rate at this locus to deleterious, loss of function alleles is 10^{-4} . All loss of function alleles are lethal when homozygous.

a) (3 points) Assuming the population is randomly mating and that the heterozygotes for a loss of function allele suffer a 1% fitness loss relative to individuals with two functional alleles, what is the equilibrium frequency for the class of loss of function alleles?

The relevant mutation rate is 10^{-4} and $s = 0.01$ (1 pt), and the appropriate equation is $q_{eq} = \frac{\mu}{hs}$ (1 pt).

Hence, the equilibrium frequency for the loss of function class is, since $h=0.01$, 0.01 (1 pt)

b) (3 points) Assuming the population is inbreeding with $f=0.01$ and that heterozygotes for a loss of function allele have identical fitness to individuals with two functional alleles, what is the equilibrium frequency for the class of loss of function alleles?

The relevant mutation rate is 10^{-4} and $s = 1$ (1 pt), and the appropriate equation is $q_{eq} = \frac{\mu}{fs}$ (1 pt).

Hence, the equilibrium frequency for the loss of function class is, since $f=0.01$, 0.01 (1 pt)

NAME _____

13. (10 points) Briefly describe a model in which senescence can evolve due to a) genetic drift, and a model in which senescence can evolve due to b) natural selection.

a), mutations that cause deleterious effects of late age of onset can be effectively neutral, so they have a finite probability of fixation, as described by the neutral theory.

b), mutations that have beneficial effects early in age but deleterious effects later in age often have a higher Malthusian parameter, and hence are selected for.