ANSC 446 / IB 416
Population Genetics
Exam 1, Sept. 19, 2008

**Name ______________________**

**(5 pages) Please show decimals rounded to 4 significant digits. Show your work or describe your logic to earn partial credit for incomplete answers.**

(9) 1. A survey in one state finds that the proportion of newborn infants who have the recessive autosomal disease cystic fibrosis is 1 in 2000.

(3) a) What is the estimated frequency of this disease allele?

q = P(a) = square root of 1/2000 = 0.02236

(3) b) What proportion of individuals would be carriers of the disease?

P(carriers) = 2pq = 2 X P(A) X P(a) = 2 X (1-0.02236) X (0.02236) = 0.04372

(3) c) Assuming a random-mating population, what proportion of matings would be between two carriers?

(P(carriers))^2 = 0.04372^2 = 0.001911

(15) 2. Assume that the following mtDNA sequences were found in four different individuals sampled from a population.

```
AATCGAGACTTTAGT
ATTCCAGATTTAAGC
ATTCCAGATTTAAGC
AATTGAGACTTTAGT
```

(3) a) How many sites are segregating? 6

(3) b) What proportion of nucleotide sites differ between the first and second sequences? 5/15 = 0.3333

(3) c) How many transitions are present between the first and second sequences? 2

(6) d) Estimate the population nucleotide diversity from this sample.

Three unique sequences: A (sequence 1), B (sequences 2, 3), C (sequence 4)
Frequencies P(A) = 0.25, P(B) = 0.5, P(C) = 0.25
$\pi ij$ for comparisons: A-B 5/15 = 0.3333, A-C 1/15 = 0.06666, B-C 6/15 = 0.4

$\pi$-hat (estimate) = (N/(N-1)) $\Sigma$ (pi X pj) X $\pi$ij
= 4/3 X [(2 X 0.25 X 0.5 X 0.3333) + (2 X 0.25 X 0.25 X 0.06666) + (2 X 0.5 X 0.25 X 0.4)]
= 4/3 X (.08333 + 0.008333 + 0.1)] = **0.2555**

(6) 3. A population of caracals was sampled to determine the weight of adult males. A normal distribution was present in which the arithmetic mean weight of the sampled caracals was 20 kg, with standard deviation of 5 kg.

Hint: draw a bell shaped curve

(3) a) What proportion of adult male caracals would be expected to weigh 29.8 kg or more?

29.8 is 9.8 kg above the mean. 9.8/5 = 1.96 standard deviations; 95% of values fall within 1.96 sd's of the mean, only 5% are more extreme than this, with half above and half below. So **0.025** of adult male caracals would be $\geq$ 29.8 kg.

(3) b) Among 200 adult male caracals, how many individuals would be expected to weigh between 15 and 25 kg?

This range of one standard deviation above and below the mean in a normal distribution includes 68% of samples; 0.68 X 200 = 136 adult male caracals.

(16) 4. Red-green color blindness in humans is an X-linked recessive trait present in 7% of males in the United States. Assuming Hardy Weinberg Equilibrium:

(3) a) What would be the expected frequency of the trait in females?

P(Xa) = 0.07; in females P(trait) = P(Xa)^2 = 0.0049

(3) b) What would be the expected frequency of carriers among females?

Female heterozygotes are carriers, with P(het) = 2 X P(Xa) X P($X_A$)
= 2 X(0.07) X (1-0.07) = 2 X 0.07 X 0.93 = 0.1302

(3) c) Alice is a daughter born to a father with normal vision and a mother who is a carrier. Using just this information, what is the probability that Alice is also a carrier?

Hint: draw Punnett square. P(carrier) = 0.5

(7) d) As an adult, Alice has four sons who all have normal vision.

Using information regarding her ancestors and descendants, what is the (posterior or Bayesian) probability that Alice is a carrier?

$P(X_A X_a) =$

$$\frac{P(4\ X_A \text{ sons}) \text{ given } X_A X_a) * P(X_A X_a)}{[(P(4\ X_A \text{ sons}) \text{ given } X_A X_a) * P(X_A X_a)] + [(P(4\ X_A \text{ sons}) \text{ given } X_A X_A) \times P(X_A X_A)}$$

$$= \frac{(.5)^4 * 0.5}{((.5)^4 * 0.5) + (1 * 0.5)}$$

= 0.03125 / (0.03125 + 0.5) = 0.05882

(3) 5. Four babies were born in a hospital on the same night, and their blood groups were later found to be O, A, B and AB. The four pairs of biological parents were:

O and O – baby O
AB and A – baby AB
A and A – baby A
B and O – baby B

Assign the four babies to their correct parents. Indicated next to parents

(2) 6. What is the difference between a Punnett square and a unit square?

A Punnett square is used by biologists to determine the probability of offspring having a particular genotype when **two individuals** mate.

A unit square is a square with all sides equaling one, which is used to show frequencies of alleles and genotypes in a **population**.

(9) 7. Coat color in horses is determined by multiple alleles. A complete Black horse (C1 black horse with black mane and tail) is dominant to a Bay horse (C2 brown horse with black legs, mane and tail) and a Mahogany Bay. A Bay is dominant to a Mahogany Bay (C3 brown horse with black roots, legs, mane, and tail). Your sample has 1000 horses (Black, Bay, and Mahogany Bay).

| Color | Observed Number |
|---|---|
| Black | 360 |
| Bay | 150 |
| Mahogany Bay | 490 |

Estimate the allele frequencies for C1, C2, and C3.

Hint: draw a unit square

P(C3) = square root of 490/1000 = 0.7

P(C2) = (square root of 640/1000) – P(C3) = 0.8-0.7 = 0.1

P(C1) = 1 – P(C2) – P(C3) = 1 – 0.7 – 0.1 = 0.2

(5) 8. An AFLP marker in xantusid lizards was found to be heterozygous in 80 lizards and homozygous in 10 lizards.

Estimate the effective number of alleles at this marker.

AE = 1 / (1-(80/90)) = 1/(1-0.8888) = 1/.1111 = 9 alleles

(16) 9. A survey of MN blood type frequencies was conducted using samples from 400 Navaho in New Mexico. The phenotypic results were:

336 M
48 MN
16 N

(3) a. What is the frequency of the M allele?

P(M allele) = (336 + (48/2)) / 400 = 0.9

(3) b. What is the frequency of the N allele?

P(N allele) = (16 + (48/2)) / 400 = 0.1

(6) c. What are the expected genotypic frequencies under Hardy-Weinberg equilibrium?

MM = 0.9^2 = 0.81

MN = 2 X 0.9 X 0.1 = 0.18

NN = 0.1^2 = 0.01

(4) d. Using a chi-square ($\chi 2$) test, are the observed genotypes in the sampling consistent with Hardy-Weinberg equilibrium?

Expected: MM 0.81 X 400 = 324, MN = 0.18 X 400 = 72, NN = 0.01 X 400 = 4

Σ ((O-E)^2/E) = [(336-324)^2/324] + [(48-72)^2/72] + [(16-4)^2/4]

= 0.4444 + 8 + 36 = 44.44

Since 44.44 > 3.84, we conclude that the population significantly differs from Hardy Weinberg expectations

| Potentially useful chi square critical values. | |
|---|---|
| Degrees of freedom | P value = .05 |
| 1 | 3.84 |
| 2 | 5.99 |
| 3 | 7.81 |
| 4 | 9.49 |

4) 10. Give the best definition for the following terms:

(2) a) Synonymous mutation

A nucleotide substitution in a codon that does not alter the amino acid coded

(2) b) Transversion (mutation)

A substitution mutation between a purine and a pyrimidine, or vice versa

(15) 11. Two populations of cattle were sampled and found to have the following allele frequencies for two SNP sites:

| | Site 1 | | Site 2 | |
|---|---|---|---|---|
| | G | C | G | A |
| Ten Asian cattle | .20 | .80 | .30 | .70 |
| Forty African cattle | .60 | .40 | .35 | .65 |

(3) a) Estimate the mean allele frequency of G at Site 1.

Mean P(G) = (0.2 X 10/50) + (0.6 X 40/50) = 0.04 + 0.48 = 0.52

(8) b) Calculate genetic identity and its three components for Site 1.

$J_{xi}$ = 0.2^2 + 0.8^2 = 0.04 + 0.64 = 0.68

$J_{yi}$ = 0.6^2 + 0.4^2 = 0.36 + 0.16 = 0.52

$J_{xiyi}$ = (0.2 X 0.6) + (0.8 X 0.4) = 0.12 + 0.32 = 0.44

I = 0.44/((0.68 X 0.52)^0.5)) = 0.44/0.5946 = 0.7399

(4) c) Estimate Nei's standard genetic distance between the two cattle populations at Site 1.

D = -ln (I) = -ln (0.7399) = 0.3011