EVE102 Practice Problems 1

Graham Coop

These study questions are to help you prepare for the exams. These, and questions given in the homework, are the sorts of questions that you can expect to see. Indeed, some of these questions are old exam questions. Your answers should not to be handed in and no key will be posted. If you have questions come to see your TA or instructor during office hours or try asking your peers for help in the Chat Room on Piazza.

Unless otherwise stated assume that loci are biallelic and autosomal.
For full credit on the exam you must show your work, and state the asumptions you have to make (if any).

Question 1. You are studying a codominant flower colour polymorphism. Skipping through a meadow of flowers you compile the following data:

red	pink	white
200	100	200

A) What frequencies would you expect at this locus under Hardy-Weinberg equilibrium?
B) Calculate the inbreeding coefficient at this locus.
C) Name two distinct processes that could lead to the deviation you see, and describe how they would result in a deficit of heterozygotes.

Question 2. The colour and shape of a a species of beetles wings are controled by two distinct polymorphisms (with alleles big/small and red/yellow respectively). In a museum collection you estimate the frequency of the four haplotypes to be:

big/red	big/yellow	small/red	small/yellow
0.69	0.00	0.09	0.22

This collection is from 60 years ago. In present day populations you estimate the frequencies of the haplotypes to be:

0.5452 0.1448 0.2348 0.0752

A) Assuming one generation per year, what is the recombination fraction between these loci?
B Qualitatively how would your answer change if you determined that crossing over only occurred in females and not in males?

Question 3. Charles the 2^{nd} of Spain was the offspring of a uncle-niece marriage. **A)** What were the relatedness coefficients (r_0, r_1, r_2) of his parents?
B) What was Charles's inbreeding coefficient from this inbreeding loop?
In answering this question ignore the other deeper inbreeding in his pedigree.

Question 4. What are the relatedness coefficients of the X chromosome between:
A) Two male full siblings?
B) Two female full siblings?
C) What is the probability that a female offspring of a full sib mating is homozygous by descent at a locus on her X chromosome?

Question 5. You are studying the wing spot polymorphism in a butterfly species. From crosses in the lab you find that the presence of wing spots is determined by a dominant allele.
You collect 100 butterflies, 84 of them have the wing spots. What is the frequency of the wing-spot allele?

Question 6. An allele has frequency of 0.001 in the population. What the probability that both you and your first (full) cousin are heterozygote for the allele?

Question 7. The kinship coefficient of the parents is the inbreeding coefficient of the offspring. Explain, with reference to the weighting of relatedness coefficients in the inbreeding coefficient, why the inbreeding coefficient is the probability that a locus is homozygous by descent.

Question 8. In terms of identity by descent, explain why multiple inbreeding loops in an individual's pedigree lead to higher levels of inbreeding.

Question 9. Based on museum samples, from $\sim$ 1800, you estimate that the average heterozygosity in Northern Elephant Seals was 0.0304 across many loci. Based on further samples you estimate that in 1960 this had dropped to 0.011. Elephant Seals have a generation time of 8 years.
What effective population size do you estimate is consistent with this drop?

Question 10. Assume that at a locus, where allele 1 has frequency p. What is the probability that two 1/2 sibs are both heterozygotes?

Question 11. **A)** Why are large populations expected to harbor more neutral variation?
B) What is the effective population size? Is it usually higher or lower than the census population size?

Question 12. You are providing genetic counseling to a couple. They are first cousins. The man is a heterozygote carrier for a rare recessive disease allele, the carrier status for the woman is unknown. The frequency of the allele is 0.00001 in the population.
A) If they have a child what is the probability that the child will be homozygote for the disease allele?
B) How would your answer change if they were 2nd cousins? (2nd cousins share a pair of great grandparents).

Question 13. You sequence a genomic region of a species of Baboon. Out of 100 thousand basepairs, on average, 200 differ between each pair of sequences. Assume a per base mutation rate of 1×10^{-8} and a generation time of ten years.
A) What is the effective population size of these Baboons?
B) What is the average coalescent time (in years) of a pair of sequences in this species?

Question 14. You find a pair of sites where Neanderthal and human populations had a fixed difference at both loci (all neanderthals had one allele and all humans had the alternate allele). You find that in present day

populations the frequency of haplotype with the Neanderthal allele at both sites is 2.24%. Based on independent information you know that the initial Neanderthal admixture proportion was 5%. The marginal frequency of the Neanderthal allele has not changed from this initial frequency to be present day, at either locus. The recombination fraction between your loci is 0.0005. How many generations back do you estimate that the Neanderthal admixture occurred?

Question 15. In a species of lemurs you estimate the allele frequency to be 20%. In a particular sub-population you estimate that the allele frequency is 10%. In this population only 9% of individuals are heterozygote. What is F_{IT}, F_{ST}, and F_{IS} for this population?

Question 16. Under a model of genetic drift why does the expected level of heterozygosity decrease by a factor of $1 - 1/(2N)$ every generation? Explain your answer in terms of the probability of identity by descent.

Question 17. Why is the historical rate of genetic drift often much higher than the census size of the population might suggest? Name two distinct processes that can contribute to this mismatch in your answer.