Fourth Homework Assignment

Write the solution to each question on a single page. We will not grade the solutions, but if you turn them in by Friday, we will give you feedback before the exam, which is Wednesday, January 29, 2014.

Question 1. (20 = 10 + 10 points). Consider a branching process in which a node can have up to \( k \) children, for some positive constant integer \( k \). The potential children are numbered, and each is born with probability \( p \) independent of the others.

(a) What is the critical value above which a population has positive probability to survive forever?

(b) What is the average size of the population as a function of \( p \)?

Question 2. (20 = 5 + 5 + 5 + 5 points). Patterson et al. (Nature, 2006) analyzed a single genome from a human, a chimpanzee, and a gorilla. A simple model for the divergence of the three species is that human and chimp diverged from a common ancestral population \( T_1 \) generations ago, and that this ancestral population diverged from the population ancestral to all three species \( T_1 + T_2 \) generations ago. For simplicity, we can assume that the ancestral populations had the same size, \( N_e \). Patterson et al. classified mutations into those found only in one species \( (n_H = 28, 504, n_C = 28, 495, n_G = 38, 677) \) or shared by two \( (n_{HC} = 8, 561, n_{HG} = 1, 302, n_{CG} = 1, 430) \).

You can assume that the mutations are far enough apart on the genome that each comes from a different genealogy, that is an independent realization of the coalescent process.

(a) How can you explain the qualitative pattern here? What are the possible tree topologies?

(b) Ignoring gorillas, how many mutations distinguish humans from chimp? Similarly, how many distinguish human from gorilla, ignoring chimp, and chimp from gorilla, ignoring human?

(c) What are the expected numbers of these mutations, assuming a mutation rate \( \mu \)? Use this relationship to estimate \( \mu T_2 \).

(d) What is the expected number of mutations shared by human and gorilla, or by chimp and gorilla? Use this to estimate \( N_e \mu \) and \( \mu T_1 \).