Mathematical Genetics, Fall 2017, HW 4

Instructions: Do all problems and show your work.

1. Suppose that a model (that is not a standard WF) leads to a coalescent that is not the standard Kingman coalescent. It has mutations appearing according to the usual rate $\theta/2$ Poisson process along each lineage, and pairwise coalescence of lineages at the usual rate of 1, but it also allows for three-way coalescences of lineages at rate 0.1 per triple. (A three-way coalescence means that, at some time in the past, three of the individuals represented in the lineages had a common parent.)

(a) For a sample of size $n = 3$, compute the probability that all three individuals are genetically identical.

(b) Find the numerical value of this probability when $\theta = 0.2$.

2. For Kingman's coalescent with infinite-sites mutation, let K_i be the number of mutations that appear while there are exactly i lineages (i.e., between times T_{i+1} and T_i).

(a) Show that

$$P(K_i = k) = \left(\frac{i-1}{\theta + i - 1}\right)\left(\frac{\theta}{\theta + i - 1}\right)^k, \quad k = 0, 1, 2, \ldots$$

(Think about the sequence of events (coalescences and mutations) that must occur, as you go back in time, to get the desired event.) What type of probability distribution (name) does K_i have?

(b) Compute $P(K_3 = 2, K_2 = 6)$. What important property are you using?

3. We use a coalescent to study the genealogy of a sample of size $n = 4$ from a Wright-Fisher model with population size $N = 10^6$ and mutation probability (per individual, per generation) $u = 5 \times 10^{-8}$.

(a) Find the expected number of generations until the most recent common ancestor of the sample.

(b) Find the mean and variance of the number of mutations in the sample.