563 HW#6 Mathematical Genetics

(1) Consider a 2-allele WF(N) model; alleles A, a.

Let A_t = # A's in the population in generation t

and $X_t = \frac{A_t}{N}$ = fraction of A's - - - - - gen t

(a) Calculate $Var(A_{t+1} \mid A_t = k)$. (Be specific about what you are using.)

(b) Explain why $Var(A_{t+1} \mid X_t = \frac{k}{N}) = Var(A_{t+1} \mid A_t = k)$.

(c) What is $Var(A_{t+1} \mid X_t = x)$?

(d) Find $Var(X_{t+1} \mid X_t = x)$. (Your answer should involve $\frac{1}{N}$.)

(e) Use (d) to define the variance effective size, $N_e^{(v)}$.

(f) Use the equation in (e) to compute $N_e^{(v)}$ for a WF model in which the population size N is chosen at random each generation with

$$N = \begin{cases} N_1 & \text{with probability } 2/5 \\ N_2 & \text{with probability } 3/5. \end{cases}$$

(If you use something besides (e) to derive your answer, you will get no credit.)

(2) SNP data:

		Site 1	site 2	site 3	site 4
	a	T	G	C	G
	b	T	G	T	G
sequence	c	A	A	T	G
	d	T	G	C	G
	e	T	G	C	T

(a) Calculate Watterson's $\hat{\theta}_W$ and Tajima's $\hat{\theta}_T$ (est's of θ).

(b) Compute Tajima's D statistic. Note: the denominator will involve $\widehat{Var}\left(D_n - \frac{S_n}{a_n}\right) = c_n S_n + d_n(S_n - 1) \cdot S_n$

where $c_n = \frac{1}{a_n}\left[\frac{n+1}{3(n-1)} - \frac{1}{a_n}\right]$

and $d_n = \frac{1}{a_n^2 + b_n}\left[\frac{2(n^2+n+3)}{9n(n-1)} - \frac{n+2}{n a_n} + \frac{b_n}{a_n^2}\right]$.

(c) Does your value of D lead you to accept or reject the null hypothesis of a standard WF? Why?