Math 338, Problem Assignments, Spring 2008

**Week 9**

**1.** Exercise 5.1. (Page 14.)

**2.** Exercise 5.2. (Page 15.)

**3.** Use the transition probabilities of Exercise 5.4, but answer these questions instead:

(a) Write down the probability transition matrix for the chain.

(b) Determine the matrices of 2- and 3-step ahead transition probabilities. (This question applies the material of section 5.6.)

(c) Determine $\mathbb{P}(X_3 = 0), \mathbb{P}(X_3 = 1), \mathbb{P}(X_3 = 2)$. (Again, this uses material from section 5.6.)

(d) As written.

**4.** Exercise 5.5. (Page 16.)

**5.** Exercise 5.6. (Page 16.)

**6.** Suppose we decide to build a Markov chain model of yeast DNA. We collect some randomly selected fragments, sequence them and come up with the sequences

GGCCTATATAAAGAGAGCACTTCGGGAAAACCCTTT
ATAGAGACTATCTAGCTGCCGCACATCAGAGTTACCGCGCGGCAAT

Estimate a transition matrix for the Markov chain model.

**7.** A simple Markov model for a DNA sequence may be insufficient to capture the amount of statistical dependence between sites. Letting $X_1, X_2, \ldots$ denote the successive bases along a strand of DNA, the Markov assumption says

$$\mathbb{P}\left(X_{n+1}=x \;\middle|\; X_n=x_n, X_{n-1}=x_{n-1}, \ldots, X_1=x_1\right) = \mathbb{P}\left(X_{n+1}=x \;\middle|\; X_n=x_n\right).$$

It may be more accurate though to assume that

$$\mathbb{P}\left(X_{n+1}=x \;\middle|\; X_n=x_n, X_{n-1}=x_{n-1}, \ldots, X_1=x_1\right) = \mathbb{P}\left(X_{n+1}=x \;\middle|\; X_n=x_n, X_{n-1}=x_{n-1}\right). \tag{1}$$

This allows the bases at the two sites previous to site $n$ to affect the probability of what letter occurs at site $n$. Let us make the assumption that (1) holds. Define the step-ahead transition probabilities

$$a_{st,u} \triangleq \mathbb{P}\left(X_{n+1}=u \;\middle|\; X_n=t, X_{n-1}=s\right).$$

We can still compute the probability of a path relatively easily assuming (1). Show that

$$\begin{aligned}&\mathbb{P}\left(X_n=j_n, X_{n-1}=j_{n-1}, \ldots, X_1=j_1\right) = \\ &\quad \mathbb{P}\left(x_1=j_1, X_2=j_2\right) a_{j_1 j_2, j_3} a_{j_2 j_3, j_4} a_{j_3 j_4, j_5} \cdots a_{j_{n-2} j_{n-1}, j_n}.\end{aligned}$$

(Hint: start your calculation applying $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$ and property (1) to write

$$\begin{aligned}&\mathbb{P}\left(X_n=j_n, X_{n-1}=j_{n-1}, \ldots, X_1=j_1\right) \\ &= \mathbb{P}\left(X_n=j_n \;\middle|\; X_{n-1}=j_{n-1}, \ldots, X_1=j_1\right) \mathbb{P}\left(X_{n-1}=j_{n-1}, \ldots, X_1=j_1\right) \\ &= \mathbb{P}\left(X_n=j_n \;\middle|\; X_{n-1}=j_{n-1}, X_{n-2}=j_{n-2}\right) \mathbb{P}\left(X_{n-1}=j_{n-1}, \ldots, X_1=j_1\right)\end{aligned}$$

The first term can be expressed in terms of the step-ahead transition probabilities. Continue the calculation using this technique.)

**8.** Consider a three state Markov chain with transition probability matrix:

| | 0 | 1 | 2 |
|---|---|---|---|
| 0 | 1/4 | 1/4 | 1/2 |
| 1 | 1/2 | 1/4 | 1/4 |
| 2 | 1/4 | 1/2 | 1/4 |

Find the invariant distribution $\rho = (\rho_1, \rho_2, \rho_3)$. (Remember use the equation $\rho_1 + \rho_2 + \rho_3 = 1$.)