Statistics 251/551 spring 2013
2013: Solutions to sheet 1

[1] (5 points) *Suppose $X_0, X_1, \dots$ is a Markov chain with state space $\mathcal{S}$ and transition probabilities $P(i,j)$ for $i,j \in \mathcal{S}$. Define $Y_n = X_{2n}$. Show that $Y_0, Y_1, \dots$ is also a Markov chain, with transition probability matrix P^2. Hint: If info stands for some information regarding $Y_0, Y_1, \dots, Y_{n-1}$, calculate $\mathbb{P}_\mu\{Y_{n+1} = j \mid Y_n = i, \textit{info}\}$ by conditioning on the value that X_{2n+1} takes.*

I was hoping for something like

$$\begin{aligned}
&\mathbb{P}_\mu\{Y_{n+1} = j \mid Y_n = i, \textit{info}\} \\
&\quad = \mathbb{P}_\mu\{X_{2n+2} = j \mid X_{2n} = i, \textit{info}\} \\
&\quad = \sum_{k\in\mathcal{S}} \mathbb{P}_\mu\{X_{2n+1} = k \mid X_{2n} = i, \textit{info}\}\times \\
&\qquad\qquad \mathbb{P}_\mu\{X_{2n+2} = j \mid X_{2n+1} = k, X_{2n} = i, \textit{info}\} \\
&\quad = \sum_{k\in\mathcal{S}} P(i,k)P(k,j) \qquad \text{by Markov property and definition of } P.
\end{aligned}$$

The last sum equals the (i,j)th element of P^2.

Some of you just assumed that the Markov property held for more than one step into the future after time $2n$, which is essentially what you were asked to show. In fact, it is true that

$$\mathbb{E}_\mu(F(X_{n+1}, X_{n+2}, \dots) \mid X_n = i, H) = \mathbb{E}_i(F(X_1, X_2, \dots))$$

if F is any function that is applied to future behavior of the chain and H is any information involving only $X_0, X_1, \dots, X_{n-1}$. A formal justification of this fact, starting from the "one step ahead" Markov property, would involve conditioning on X_{n+1} then on X_{n+2}, ... Compare with the comments following Chang Definition 1.3.

[2] (5 points) *For the hhh vs. thh coin tossing game, as described in the first lecture:*

(i) Which states are accessible from which states?

(ii) Which states communicate?

(iii) What is the period for each state?

We mostly ignored the confusion over how to define the period of a state like *nix*. (The most popular answer was $\textit{period}(\textit{nix}) = +\infty$, which has some claim to mathematical precision.) Some of you only considered getting from i to j in one step when deciding whether $i \rightsquigarrow j$. Some of you declared TH to have period 2 even though it communicates with state T, which has period 1.

[3] (5 points) *Let $\mathcal{S}$ be a finite set for which there is a nonnegative weight $w_{i,j}$ assigned to each distinct pair i, j from $\mathcal{S}$. Suppose $W_i = \sum_{j \neq i} w_{i,j}$ is strictly positive for each i. Define a transition matrix P by $P(i,j) = w_{i,j}/W_i$ for $i \neq j$, and $P(i,i) = 0$. Show that $\pi_i = W_i/W$, where $W = \sum_i W_i$, is a stationary distribution.*

There was much confusion caused by the notation $\sum_{j \neq i} w_{i,j}$. Is the sum over all pairs (i,j) for which $i \neq j$? Or is it over j for fixed i? Or over i for fixed j? The only clue was the fact that the left-hand side of the definition involved i, so that there could be no summation over i. It would be much clearer to write something like $\sum_{j \in \mathcal{S}} \mathbb{I}\{j \neq i\} w_{i,j}$. Even better, define $w_{i,i} = 0$ then write $\sum_{j \in \mathcal{S}} w_{i,j}$. In fact there was no need to require $w_{i,i}$ to be zero for the argument to work; it would be OK to allow $P(i,i) > 0$.

There was more confusion over the meaning of $w_{i,j}$. I intended that the weight be associated with (an undirected edge) $\{i,j\}$ without distinguishing between $\{i,j\}$ and $\{j,i\}$. With hindsight I should have posed the problem as: w is a symmetric $N \times N$ matrix of nonnegative weights, where N = the size of $\mathcal{S}$. I could have allowed $w[i,i]$ to be nonzero. Then $W_i := \sum_{j \in \mathcal{S}} w[i,j]$ and $W = \sum_{i \in \mathcal{S}} W_i$ and

$$\sum_{i \in \mathcal{S}} \pi_i P(i,j) = \sum_{i \in \mathcal{S}} (W_i/W)(w[i,j]/W_i).$$

The W_i factors cancel, leaving $1/W$ times $\sum_{i \in \mathcal{S}} w[i,j] = \sum_{i \in \mathcal{S}} w[j,i] = W_j$.

Students who tried to work with matrix notation to show $\pi P = \pi$, where π is the $1 \times N$ row vector of $\pi_i's$, often got bogged down in messy matrix expressions, which I found hard to read.