Assignment 2.

Given September 9, due September 23.
Last revised, September 10.

Objective: Conditioning and Markov chains.

1. Suppose that $\mathcal{F}$ and $\mathcal{G}$ are two algebras of sets and that $\mathcal{F}$ adds information to $\mathcal{G}$ in the sense that any $\mathcal{G}$ measurable event is also $\mathcal{F}$ measurable: $\mathcal{G} \subset \mathcal{F}$. Suppose that the probability space Ω is discrete (finite or countable) and that $X(\omega)$ is a variable defined on Ω (that is, a function of the random variable ω). The conditional expectations (in the modern sense) of X with respect to $\mathcal{F}$ and $\mathcal{G}$ are $Y = E[X \mid \mathcal{F}]$ and $Z = E[X \mid \mathcal{G}]$. In each case below, state whether the statement is true or false and explain your answer with a proof or a counterexample.

a. $Z \in \mathcal{F}$.

b. $Y \in \mathcal{G}$.

c. $Z = E[Y \mid \mathcal{G}]$.

d. $Y = E[Z \mid \mathcal{F}]$.

2. Let Ω be a discrete probability space and $\mathcal{F}$ a $\sigma-$algebra. Let $X(\omega)$ be a (function of a) random variable with $E[X^2] < \infty$. Let $Y = E[X \mid \mathcal{F}]$. The variance of X is $\text{var}(X) = E[(X - \overline{X})^2]$, where $\overline{X} = E[X]$.

a. Show directly from the (modern) definition of conditional expectation that

$$E[X^2] = E[(X - Y)^2] + E[Y^2] \; . \tag{1}$$

Note that this equation also could be written

$$E[X^2] = E\left[(X - E[X \mid \mathcal{F}])^2\right] + E\left[(E[X \mid \mathcal{F}])^2\right] \; .$$

.

b. Use this to show that $\text{var}(X) = \text{var}(X - Y) + \text{var}(Y)$.

c. If we interpret conditional expectation as an orthogonal projection in a vector space, what theorem about orthogonality does part a represent?

d. We have n independent coin tosses with each equally likely to be H or T. Take X to be the indicator function of the event that the first toss is H. Take $\mathcal{F}$ to be the algebra generated by the number of H tosses in all. Calculate each of the three quantities in (1) from scratch and check that the equation holds. Both of the terms on the right are easiest to do using the law of total probability, which is pretty obvious in this case.

3. (Bayesian identification of a Markov chain) We have a state space of size m and two $m \times m$ stochastic matrices, Q, and R. First we pick one of the matrices, choosing Q with probability f and R with probability $1-f$. Then we use the chosen matrix to run a Markov chain X, starting with $X(0) = 1$ up to time T.

a. Describe the probability space Ω appropriate for this situation.

b. Let $\mathcal{F}$ be the algebra generated by the chain itself, without knowing whether Q or R was chosen. Find a formula for $P(Q \mid \mathcal{F})$ (which would be $P(Q \mid X = x)$ in classical notation). Though this formula might be ugly, it is easy to probram.

4. Suppose we have a 3 state Markov chain with transition matrix

$$P = \begin{pmatrix} .6 & .2 & .2 \\ .3 & .5 & .2 \\ .1 & .2 & .7 \end{pmatrix}$$

and suppose that $X(0) = 1$. For any $t > 0$, the algebras $\mathcal{F}_t$ and $\mathcal{G}_t$ are as in the notes, and $\mathcal{H}_t$ is the algebra generated by $X(s)$ for $t \leq s \leq T$.

a. Show that the probability distribution of the first t steps conditioned on $\mathcal{G}_{t+1}$ is the same as that conditioned on $\mathcal{H}_{t+1}$. This is a kind of backwards Markov property: a forward Markov chain is a backward Markov chain also.

b. Calculate $P(X(3) = 2 \mid \mathcal{G}_4)$. This consists of 3 numbers.