5. Markov Processes.

A stochastic process in discrete time is just a sequence \( \{X_j : j \geq 0\} \) of random variables with values in some \((\mathcal{X}, \mathcal{F})\) defined on a probability \((\Omega, \Sigma, P)\). It can also be specified by prescribing, in a self consistent manner, the joint distribution of \(\{X_0, X_1, X_2, \ldots, X_n\}\) for every \(n\). A convenient way of doing it is by specifying the the distribution \(p_0(dx_0)\) of \(X_0\) and the conditional distributions \(p_n(x_0, x_1, \ldots, x_{n-1}; dx_n)\) of \(X_n\) given \(X_0, \ldots, X_{n-1}\). \(\Omega\) can be the product space \(\mathcal{X}^\infty\), i.e. the space of sequences with values in \(\mathcal{X}\). There is a canonical \(P\) on the natural \(\sigma\)-field \(\mathcal{F}_\infty\) on \(\Omega\). There is also the sub-\(\sigma\)-fields \(\mathcal{F}_n\) generated by \(x_0, x_1, \ldots, x_n\). The canonical \(P\) will equal \(p_0\) on \(\mathcal{F}_0\) and the conditional distribution of on \(\mathcal{F}_n\) given \(\mathcal{F}_{n-1}\) will be given by \(p_n(x_0, x_1, \ldots, x_{n-1}; dx_n)\). In the special case when \(p_n(x_0, x_1, \ldots, x_{n-1}; dx_n) = \pi_n(x_{n-1}, dx_n)\), for \(n \geq 1\), depends only on \(x_{n-1}\), the process is called a Markov Process. Of course when they are just \(p_n(dx_n)\) and do not depend on any \(x_i\) for \(0 \leq i \leq n-1\) we have independent random variables and \(P\) is the product measure. If, in the Markov case, \(\pi_n(\cdot, \cdot)\) is the same \(\pi(\cdot, \cdot)\) for all \(n \geq 1\), it is called a Markov process with stationary transition probabilities.

A simple example is to take \(\mathcal{X}\) to be a countable set. Then \(p_0\) is just the set of probabilities \(p_0(x) = P[X_0 = x]\), and

\[
P[X_n = y|X_0 = x_0, \ldots, X_{n-1} = x_{n-1}] = \pi_n(x, y)
\]

are the transition probabilities which in the stationary case is independent of \(n\). It is natural to consider \((\Omega, \mathcal{F}_n, \mathcal{F}_\infty, P)\). There are some natural martingales. For simplicity we limit ourselves to the stationary case.

**Theorem.** For any function \(f\) on \(\mathcal{X}\) let us define

\[
(\pi f)(x) = \sum_y \pi(x, y)f(y) = E[f(X_n)|X_{n-1} = x]
\]

Then

\[
Z_n = f(X_n) - f(X_0) - \sum_{j=0}^{n-1} (\pi f - f)(X_j)
\]
is a martingale with respect to \((\Omega, \mathcal{F}_n, P)\).

**Proof:** Let us compute \(E[Z_n|\mathcal{F}_{n-1}]\).

\[
E[Z_n|\mathcal{F}_{n-1}] = E[f(X_n)|\mathcal{F}_{n-1}] - \sum_{j=0}^{n-1} (\pi f - f)(X_j) \\
= (\pi f)(X_{n-1}) - \sum_{j=0}^{n-1} (\pi f - f)(X_j) \\
= f(X_{n-1}) - \sum_{j=0}^{n-2} (\pi f - f)(X_j) \\
= Z_{n-1}
\]

**Remark.** If we replace the definition \((\pi f)(x) = \sum_y \pi(x,y)f(y)\) with

\[
(\pi f)(x) = \int f(y)\pi(x,dy)
\]

then the theorem is true for Markov processes on any state space. For simplicity we will assume that we have a countable state space.

Martingales are a useful tool in studying Markov Processes. Let us look at some examples.

1. Let \(A \subset \mathcal{X}\). Define

\[
\tau_A = \inf\{j : X_j \in A\}
\]

is the first hitting time of \(A\). It is possible that \(X_j\) never hits \(A\) in which case we take \(\tau_A = \infty\). We wish to calculate for \(\lambda > 0\),

\[
(5.1) \quad \phi_\lambda(x) = E[e^{-\lambda \tau_A}|X_0 = x]
\]

Then if \(x \in A\) then \(\phi_\lambda(x) = 1\). Moreover for \(x \notin A\) it is easy to see that

\[
\phi_\lambda(x) = e^{-\lambda} \sum_y \pi(x,y)\phi_\lambda(y)
\]
Clearly $0 \leq \phi_\lambda(x) \leq 1$. We will show that the only bounded solution of
\[(5.2) \quad F(x) = e^{-\lambda} \sum_y \pi(x, y) F(y)\]
for $x \notin A$ with $F(x) = 1$ for $x \in A$ is given by (5.1). Let $F(x)$ be a solution of (5.2). Define
\[Z_n = e^{-\lambda n} F(X_n)\]
Then with $\Sigma_n = \sigma\{X_0, X_1, \ldots, X_n\}$,
\[
E[Z_{n+1} | \Sigma_n] = e^{-\lambda (n+1)} E[F(X_{n+1} | \Sigma_n)] \\
= e^{-\lambda (n+1)} \sum_y \pi(X_n, y) F(y) \\
= e^{-\lambda n} F(X_n) \\
= Z_n
\]
provided $X_n \notin A$. One can rewrite this as
\[
E[Z_{n+1} - Z_n | \sigma_n] = \begin{cases} 
0 & \text{if } X_n \notin A \\
e^{-\lambda n} G(X_n) & \text{if } X_n \in A.
\end{cases}
\]
with
\[G(x) = e^{-\lambda} \sum_y \pi(X_n, y) F(y) - F(x)\]
Therefore
\[Z_n - Z_0 - \sum_{j=0}^{n-1} e^{-\lambda n} G(X_j) 1_A(X_j)\]
is a martingale. Let $\tau_A$ is a stopping time and for $n \leq \tau_A$ $X_n \notin A$ and $G(X_n) = 0$. Therefore $\{Z_n\}$ is bounded uniformly until $\tau_A$ even if $\tau_A$ itself can be large. Doob’s stopping theorem applies and
\[
E[e^{-\lambda \tau_A}] = E[e^{-\lambda \tau_A} F(X_{\tau_A})] = E[Z_{\tau}] = E[Z_0] = F(x)
\]
**Example.** Consider the random walk on $\mathbb{Z}$ where $\pi(x, x \pm 1) = \frac{1}{2}$. If one starts from 0, and $\tau$ is the first time $\pm k$ is reached calculate $E[e^{-\lambda \tau}]$. Solve the equation
\[F(x) = e^{-\lambda} \left[ \frac{1}{2} F(x-1) + \frac{1}{2} F(x+1) \right]\]
for $|x| \leq k - 1$, with $F(x) = 1$ for $|x| \geq k$. One can isolate $[-k, k]$. Need to solve

$$F(x - 1) + F(x + 1) - 2e^{\lambda}F(x) = 0$$

with $F(\pm k) = 1$. Solve the quadratic

$$\rho^2 - 2e^{\lambda}\rho + 1 = 0$$

with roots

$$\rho_{\pm} = e^{\lambda} \pm \sqrt{e^{2\lambda} - 1} = e^{\pm \theta}$$

where $\theta = \log[e^{\lambda} + \sqrt{e^{2\lambda} - 1}]$. The solution is seen to be

$$F(x) = \frac{e^{\theta}x + e^{-\theta}x}{e^{\theta}k + e^{-\theta}k}$$

and

$$F(0) = [\cosh(\theta k)]^{-1}$$

**Exercise.** Start from $x > 0$. Show that sooner or later 0 is reached. Calculate $E[e^{-\lambda}\tau]$ where $\tau$ is the first time 0 is reached.

**Exercise.** What happens when

$$p = \pi(x, x - 1) > \frac{1}{2} > \pi(x, x + 1) = q = 1 - p$$

and when

$$p = \pi(x, x - 1) < \frac{1}{2} < \pi(x, x + 1) = q = 1 - p$$

**Example.** A game is being played where the probability is $\frac{1}{2}$ for each of two players to win any one round. It is agreed that the first person to win $k$ rounds will be the winner. They put equal amounts to make a kitty for the winner to take. Unfortunately the game is interrupted before either player can win $k$ rounds. It stops when player A needs to win $a$ more rounds, and player B needs $b$ rounds. $1 \leq a \leq k, 1 \leq b \leq k$. What is the ”fair” way to divide the kitty between the two players?

Let $u(a, b)$ be the proportion of the kitty that player A should get in a fair division when he needs $a$ rounds and player B needs $b$ rounds. Since the
game is fair neither player can expect to gain or lose by playing an extra
game.

\[ u(a, b) = \frac{1}{2} u(a - 1, b) + \frac{1}{2} u(a, b - 1) \]

\[ u(0, b) = 1 \text{ if } b > 0 \text{ and } u(a, 0) = 0 \text{ if } a > 0. \]

Solution is

\[ u(a, b) = \frac{1}{2^{a+b-1}} \sum_{a+b-1 \geq r \geq a} \binom{a + b - 1}{r} \]

You can verify that this is a solution. Can you show directly?