STAT 870 Lecture 10

Initial Distributions

Meaning of unconditional expected values?

Markov property specifies only conditional probabilities; no way to deduce marginal distributions.

For every distribution π on S and transition matrix $\mathbf{P}$ there is a a stochastic process $X_0, X_1, \ldots$ with

$$P(X_0 = k) = \pi_k$$

and which is a Markov Chain with transition matrix $\mathbf{P}$. corresponding probability measure.

Notice Strong Markov Property proof used only conditional expectations.

Notation: π a probability on S. E^π and P^π are expected values and probabilities for chain with initial distribution π.

Summary of easily verified facts:

- For any sequence of states $i_0, \ldots, i_k$

$$P(X_0 = i_0, \ldots, X_k = i_k) = \pi_{i_0} \mathbf{P}_{i_0 i_1} \cdots \mathbf{P}_{i_{k-1} i_k}$$

- For any event A:

$$\mathbf{P}^\pi(A) = \sum_k \pi_k \mathbf{P}^k(A)$$

- For any bounded rv $Y = f(X_0, \ldots)$

$$\mathrm{E}^\pi(Y) = \sum_k \pi_k \mathrm{E}^k(A)$$

Recurrence and Transience

Now consider a transient state k, that is, a state for which

$$f_k = P^k(T_k < \infty) < 1$$

Note that $T_k = \min\{n > 0 : X_n = k\}$ is a stopping time. Let N_k be the number of visits to state k. That is

$$N_k = \sum_{n=0}^{\infty} 1(X_n = k)$$

Notice that if we define the function

$$f(x_0, x_1, \dots) = \sum_{n=0}^{\infty} 1(x_n = k)$$

then

$$N_k = f(X_0, X_1, \dots)$$

Notice, also, that on the event $T_k < \infty$

$$N_k = 1 + f(X_{T_k}, X_{T_k+1}, \dots)$$

and on the event $T_k = \infty$ we have

$$N_k = 1$$

In short:

$$N_k = 1 + f(X_{T_k}, X_{T_k+1}, \dots)1(T_k < \infty)$$

Hence

$$\begin{aligned}
&\mathbf{P}^k(N_k = r) \\
&\quad = \mathrm{E}^k\{P(N_k = r|\mathcal{F}_T)\} \\
&\quad = \mathrm{E}^k[P\{1 + f(X_{T_k}, X_{T_k+1}, \dots) \\
&\qquad\qquad \times 1(T_k < \infty) = r|\mathcal{F}_T\}] \\
&\quad = \mathrm{E}^k[1(T_k < \infty) \\
&\qquad\quad \times P^{X_{T_k}}\{f(X_0, X_1, \dots) = r - 1\}\Big] \\
&\quad = \mathrm{E}^k\Big\{1(T_k < \infty)P^k(N_k = r - 1)\Big\} \\
&\quad = \mathrm{E}^k\{1(T_k < \infty)\}\, P^k(N_k = r - 1) \\
&\quad = f_k P^k(N_k = r - 1)
\end{aligned}$$

It is easily verified by induction, then, that

$$\mathbf{P}^k(N_k = r) = f_k^{r-1} P^k(N_k = 1)$$

But $N_k = 1$ if and only if $T_k = \infty$ so

$$\mathbf{P}^k(N_k = r) = f_k^{r-1}(1 - f_k)$$

so N_k has (chain starts from k) Geometric dist'n, mean $1/(1-f_k)$. Argument also shows that if $f_k = 1$ then

$$P^k(N_k = 1) = P^k(N_k = 2) = \cdots$$

which can only happen if all these probabilities are 0. Thus if $f_k = 1$

$$P(N_k = \infty) = 1$$

Since

$$N_k = \sum_{n=0}^{\infty} 1(X_n = k)$$

$$\mathrm{E}^k(N_k) = \sum_{n=0}^{\infty} (\mathbf{P}^n)_{kk}$$

So: State k is transient if and only if

$$\sum_{n=0}^{\infty} (\mathbf{P}^n)_{kk} < \infty$$

and this sum is $1/(1 - f_k)$.

Proposition 1 *Recurrence (or transience) is a class property. That is, if i and j are in the same communicating class then i is recurrent (respectively transient) if and only if j is recurrent (respectively transient).*

Proof: Suppose i is recurrent and $i \leftrightarrow j$. There are integers m and n such that

$$(\mathbf{P}^m)_{ji} > 0 \quad \text{and} \quad (\mathbf{P}^n)_{ij} > 0$$

Then

$$\begin{aligned}
\sum_k (\mathbf{P}^k)_{jj} &\geq \sum_{k \geq 0} (\mathbf{P}^{m+k+n})_{jj} \\
&\geq \sum_{k \geq 0} (\mathbf{P}^m)_{ji} (\mathbf{P}^k)_{ii} (\mathbf{P}^n)_{ij} \\
&= (\mathbf{P}^m)_{ji} \left\{ \sum_{k \geq 0} (\mathbf{P}^k)_{ii} \right\} (\mathbf{P}^n)_{ij}
\end{aligned}$$

The middle term is infinite and the two outside terms positive so

$$\sum_k (\mathbf{P}^k)_{jj} = \infty$$

which shows j is recurrent.

A finite state space chain has at least one recurrent state:

If all states we transient we would have for each k $P(N_k < \infty) = 1$. This would mean $P(\forall k . N_k < \infty) = 1$. But for any ω there must be at least one k for which $N_k = \infty$ (the total of a finite list of finite numbers is finite).

Infinite state space chain may have all states transient:

The chain X_n satisfying $X_{n+1} = X_n + 1$ on the integers has all states transient.

More interesting example:

- Toss a coin repeatedly.
- Let X_n be X_0 plus the number of heads minus the number of tails in the first n tosses.
- Let p denote the probability of heads on an individual trial.

$X_n - X_0$ is a sum of n iid random variables Y_i where $P(Y_i = 1) = p$ and $P(Y_i = -1) = 1 - p$.

SLLN shows X_n/n converges almost surely to $2p - 1$. If $p \neq 1/2$ this is not 0.

In order for X_n/n to have a positive limit we must have $X_n \to \infty$ almost surely so all states are visited only finitely many times. That is, all states are transient. Similarly for $p < 1/2$ $X_n \to -\infty$ almost surely and all states are transient.

Now look at $p = 1/2$. The law of large numbers argument no long shows anything. I will show that all states are recurrent.

Proof: We show $\sum_n (\mathbf{P}^n)_{00}$ and show the sum is infinite. If n is odd then $(p_n)_{00} = 0$ so we evaluate

$$\sum_m (\mathbf{P}^{2m})_0 0$$

Now

$$(\mathbf{P}^{2m})_0 0 = \binom{2m}{m} 2^{-2m}$$

According to Stirling's approximation

$$\lim_{m\to\infty} \frac{m!}{m^{m+1/2} e^{-m} \sqrt{2\pi}} = 1$$

Hence

$$\lim_{m\to\infty} \sqrt{m} (\mathbf{P}^{2m})_0 0 = \frac{1}{\sqrt{\pi}}$$

Since

$$\sum \frac{1}{\sqrt{m}} = \infty$$

we are done.

Mean return times

Compute expected times to return. For $x \in S$ let T_x denote the hitting time for x.

Suppose x recurrent in **irreducible** chain (only one communicating class).

Derive equations for expected values of different T_x.

Each T_x is a certain function f_x applied to $X_1, \ldots$. Setting $\mu_{ij} = \mathrm{E}^i(T_j)$ we find

$$\mu_{ij} = \sum_k \mathrm{E}^i(T_j 1(X_1 = k))$$

Note that if $X_1 = x$ then $T_x = 1$ so

$$\mathrm{E}^i(T_j 1(X_1 = j)) = \mathbf{P}_{ij}$$

For $k \neq j$

$$T_x = 1 + f_x(X_2, X_3, \ldots)$$

and, by conditioning on $X_1 = k$ we find

$$\mathrm{E}^i(T_j 1(X_1 = k)) = \mathbf{P}_{ik}\left\{1 + \mathrm{E}^k(T_j)\right\}$$

This gives

$$\mu_{ij} = 1 + \sum_{k \neq j} \mathbf{P}_{ik}\mu_{kj} \tag{1}$$

Technically, I should check that the expectations in (1) are finite. All the random variables involved are non-negative, however, and the equation actually makes sense even if some terms are infinite. (To prove this you actually study

$$T_{x,n} = \min(T_x, n)$$

deriving an identity for a fixed n, letting $n \to \infty$ and applying the monotone convergence theorem.)

Here is a simple example:

$$\mathbf{P} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

The identity (1) becomes

$$\begin{aligned}
\mu_{1,1} &= 1 + \frac{1}{2}\mu_{2,1} + \frac{1}{2}\mu_{3,1} \\
\mu_{1,2} &= 1 + \frac{1}{2}\mu_{3,1} \\
\mu_{1,3} &= 1 + \frac{1}{2}\mu_{2,1} \\
\mu_{2,1} &= 1 + \frac{1}{2}\mu_{3,1} \\
\mu_{2,2} &= 1 + \frac{1}{2}\mu_{1,2} + \frac{1}{2}\mu_{3,2} \\
\mu_{2,3} &= 1 + \frac{1}{2}\mu_{1,3} \\
\mu_{3,1} &= 1 + \frac{1}{2}\mu_{2,1} \\
\mu_{3,2} &= 1 + \frac{1}{2}\mu_{1,2} \\
\mu_{3,3} &= 1 + \frac{1}{2}\mu_{1,3} + \frac{1}{2}\mu_{2,3}
\end{aligned}$$

Seventh and fourth show $\mu_{2,1} = \mu_{3,1}$. Similar calculations lead to $\mu_{ii} = 3$ and, for $i \neq j$ $\mu_{i,j} = 2$.

Example: The coin tossing Markov Chain with $p = 1/2$ shows that the situation can be different when S is infinite. The equations above become:

$$\begin{aligned}
m_{0,0} &= 1 + \frac{1}{2}m_{1,0} + \frac{1}{2}m_{-1,0} \\
m_{1,0} &= 1 + \frac{1}{2}m_{2,0}
\end{aligned}$$

and many more.

Some observations:

You have to go through 1 to get to 0 from 2 so

$$m_{2,0} = m_{2,1} + m_{1,0}$$

Symmetry (switching H and T):

$$m_{1,0} = m_{-1,0}$$

The transition probabilities are **homogeneous**:

$$m_{2,1} = m_{1,0}$$

Conclusion:

$$\begin{aligned} m_{0,0} &= 1 + m_{1,0} \\ &= 1 + 1 + \frac{1}{2} m_{2,0} \\ &= 2 + m_{1,0} \end{aligned}$$

Notice that there are **no** finite solutions!

Summary of the situation:

Every state is recurrent.

All the expected hitting times m_{ij} are infinite.

All entries $\mathbf{P}^n_{ij}$ converge to 0.

Jargon: The states in this chain are null recurrent.