Markov Chains (Continued)

Let us now consider an irreducible Markov Chain with one step transition probability $\pi(x, y)$ that is periodic with period $d > 1$. Let us choose and fix a reference point $x_0 \in \mathcal{X}$. For each $x \in \mathcal{X}$ let $D_{x_0,x} = \{n : \pi^{(n)}(x_0, x) > 0\}$.

Lemma 3.10. If $n_1, n_2 \in D_{x_0,x}$ then d divides $n_1 - n_2$.

Proof: Since the chain is irreducible there is an m such that $\pi^{(m)}(x, x_0) > 0$. By the Chapman-Kolmogorov equations $\pi^{(m+n_i)}(x_0, x_0) > 0$ for $i = 1, 2$. Therefore $m + n_i \in D_{x_0} = D_{x_0,x_0}$ for $i = 1, 2$. This implies that d divides both $m + n_1$ as well as $m + n_2$. Thus d divides $n_1 - n_2$.

The residue modulo d of all the integers in $D_{x_0,x}$ are the same and equal some number $r(x)$, satisfying $0 \leq r(x) \leq d - 1$. By definition $r(x_0) = 0$. Let us define $\mathcal{X}_j = \{x : r(x) = j\}$. Then $\{\mathcal{X}_j : 0 \leq j \leq d-1\}$ is a partition of $\mathcal{X}$ into disjoint sets with $x_0 \in \mathcal{X}_0$.

Lemma 3.11. If $x \in \mathcal{X}$, then $\pi^{(n)}(x, y) = 0$ unless $r(x) + n = r(y)$ modulo d.

Proof: Suppose that $x \in \mathcal{X}$ and $\pi(x, y) > 0$. Then if $m \in D_{x_0,x}$ then $(m+1) \in D_{x_0,y}$. Therefore $r(x)+1 = r(y)$ modulo d. The proof can be completed by induction. The chain marches through $\{\mathcal{X}_j\}$ in a cyclical way from a state in $\mathcal{X}_j$ to one in $\mathcal{X}_{j+1}$

Theorem 3.12. Let $\mathcal{X}$ be irreducible and positive recurrent with period d. Then

$$\lim_{\substack{n\to\infty \\ n+r(x)=r(y) \text{ modulo } d}} \pi^{(n)}(x, y) = \frac{d}{m(y)}$$

Of course

$$\pi^{(n)}(x, y) = 0$$

unless $n + r(x) = r(y)$ modulo d.

Proof: If we replace π by $\tilde{\pi}$ where $\tilde{\pi}(x, y) = \pi^{(d)}(x, y)$, then $\tilde{\pi}(x, y) = 0$ unless both x and y are in the same $\mathcal{X}_j$. The restriction of $\tilde{\pi}$ to each $\mathcal{X}_j$ defines an irreducible aperiodic Markov chain. Since each time step under $\tilde{\pi}$ is actually d units of time we can apply the earlier results and we will get for $x, y \in \mathcal{X}_j$ for some j,

$$\lim_{k\to\infty} \pi^{(k\,d)}(x,y) = \frac{d}{m(y)}$$

We note that

$$\begin{aligned}
\pi^{(n)}(x,y) &= \sum_{1\le m\le n} f_m(x,y)\,\pi^{(n-m)}(y,y)\\
f_m(x,y) = P_x\{\tau_y = m\} &= 0 \ \text{ unless } \ r(x)+m = r(y) \ \text{ modulo } \ d\\
\pi^{(n-m)}(y,y) &= 0 \ \text{ unless } \ n-m = 0 \ \text{ modulo } \ d\\
\sum_m f_m(x,y) &= 1
\end{aligned}$$

The theorem now follows.

Suppose now we have a chain that is not irreducible. Let us collect all the transient states and call the set $\mathcal{X}_{tr}$. The complement consists of all the recurrent states and will be denoted by $\mathcal{X}_{re}$.

Lemma 3.13. If $x \in \mathcal{X}_{re}$ and $y \in \mathcal{X}_{tr}$, then $\pi(x,y) = 0$.

Proof: If x is a recuurrent state, and $\pi(x,y) > 0$, the chain will return to x infinitely often and each time there is a positive probability of visiting y. By the renewal property these are independent events and so y will be recurrent too.

The set of recurrent states $\mathcal{X}_{re}$ can be divided into one or more equivalence classes accoeding to the following procedure. Two recurrent states x and y are in the same equivalence class if $f(x,y) = P_x\{\tau_y < \infty\}$, the probability of ever visiting y starting from x is positive. Because of recurrence if $f(x,y) > 0$ then $f(x,y) = f(y,x) = 1$. The restriction of the chain to a single equivalence class is irreducible and possibly periodic. Different equivalence classes could have different periods, some could be positive recurrent and others null recurrent. We can combine all our observations into the following theorem.

Theorem 3.14. If y is transient then $\sum_n \pi^{(n)}(x,y) < \infty$ for all x. If y is null recurrent (belongs to an equivalence class that is null recurrent) then $\pi^{(n)}(x,y) \to 0$ for all x, but $\sum_n \pi^{(n)}(x,y) = \infty$ if x is in the same equivalence class or $x \in \mathcal{X}_{tr}$ with $f(x,y) > 0$. In all other cases $\pi^{(n)}(x,y) = 0$ for all $n \geq 1$. If y is positive recurrent and belongs to an equivalence class with period d with $m(y) = E^{P_y}\{\tau_y\}$, then for a nontransient x, $\pi^{(n)}(x,y) = 0$ unless x is in the same equivalence class and $r(x) + n = r(y)$ modulo d. In such a case,

$$\lim_{\substack{n\to\infty \\ r(x)+n=r(y) \text{ modulo } d}} \pi^{(n)}(x,y) = \frac{d}{m(y)}$$

If x is transient then

$$\lim_{\substack{n\to\infty \\ n=r \text{ modulo } d}} \pi^{(n)}(x,y) = f(r,x,y)\frac{d}{m(y)}$$

where

$$f(r,x,y) = P_x\{X_{kd+r} = y \quad \text{for some} \quad k \geq 0\}$$

Proof: The only statement that needs an explanation is the last one. The chain starting from a transient state x may at some time get into a positive recurrent equivalence class $\mathcal{X}_j$ with period d. If it does, it never leaves that class and so gets absorbed in that class. The probability of this is $f(x,y)$ where y can be any state in $\mathcal{X}_j$. However if the period d is greater than 1, there will be cyclical subclasses $C_1, \cdots, C_d$ of $\mathcal{X}_j$. Depending on which subclass the chain enters and when, the phase of its future is determined. There are d such possible phases. For instance, if the subclasses are ordered in the correct way, getting into C_1 at time n is the same as getting into C_2 at time $n+1$ and so on. $f(r,x,y)$ is the probability of getting into the equivalence class in a phase that visits the cyclical subclass containing y at times n that are equal to r modulo d.

Example (Random Walk): If $\mathcal{X} = Z^d$, the integral lattice in R^d, a random walk is a Markov chain with transition probability $\pi(x,y) = p(y-x)$ where $\{p(z)\}$ specifies the probability distribution of a single step. We will assume for simplicity that $p(z) = 0$ except when $z \in F$ where F consists of the $2\,d$ neighbors of 0 and $p(z) = \frac{1}{2d}$ for each $z \in F$. For $\xi \in R^d$ the characteristic function of $\hat{p}(\xi)$ of $p(\cdot)$ is given by

$\frac{1}{d}(\cos\xi_1 + \cos\xi_2 + \cdots + \cos\xi_d)$. The chain is easily seen to irreducible, but periodic of period 2. Return to the starting point is possible only after an even number of steps.

$$\begin{aligned}\pi^{(2n)}(0,0) &= (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} [\hat{p}(\xi)]^{2n} d\xi \\ &\simeq \frac{C}{n^{\frac{d}{2}}}\end{aligned}$$

To see this asymptotic behavior let us first note that the integration can be restricted to the set where $|\hat{p}(\xi)| \geq 1-\delta$ or near the 2 points $(0,0,\cdots,0)$ and $(\pi,\pi,\cdots,\pi)$ where $|\hat{p}(\xi)| = 1$. Since the behaviour is similar at both points let us concentrate near the origin.

$$\frac{1}{d}\sum_{j=1}^{d}\cos\xi_j \leq 1 - c\sum_j \xi_j^2 \leq \exp[-c\sum_j \xi_j^2]$$

for some $c > 0$ and

$$[\frac{1}{d}\sum_{j=1}^{d}\cos\xi_j]^{2n} \leq \exp[-2\,n\,c\sum_j \xi_j^2]$$

and with a change of variables the upper bound is clear. We have a similar lower bound as well. The random walk is recurrent if $d = 1$ or 2 but transient if $d \geq 3$.

Exercise 3.11. If the distribution $p(\cdot)$ is arbitrary, determine when the chain is irreducible and when it is irreducible and aperiodic.

Exercise 3.12. If $\sum_z zp(z) = m \neq 0$ conclude that the chain is transient by an application of the strong law of large numbers.

Exercise 3.13. If $\sum_z zp(z) = m = 0$, and if the covariance matrix, $\sum_z z_i z_j p(z) = \sigma_{i,j}$, is nondegenerate show that the transience or recurrence is determined by the dimension as in the case of the nearest neighbor random walk.

Exercise 3.14. Can you make sense of the formal calculation

$$\sum_n \pi^{(n)}(0,0) = \sum_n (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} [\hat{p}(\xi)]^n d\xi$$

$$\begin{aligned} &= (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} \frac{1}{(1-\hat{p}(\xi))} d\xi \\ &= (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} \mathcal{R}eal\ \mathcal{P}art\Big[\frac{1}{1-\hat{p}(\xi)}\Big] d\xi \end{aligned}$$

to conclude that a necessary and sufficient condition for transience or recurrece is the convergence or divergence of the integral

$$\int_{\mathbf{T}^d} \mathcal{R}eal\ \mathcal{P}art\Big[\frac{1}{1-\hat{p}(\xi)}\Big] d\xi$$

with an integrand

$$\mathcal{R}eal\ \mathcal{P}art\Big[\frac{1}{1-\hat{p}(\xi)}\Big]$$

that is seen to be nonnegative ?

Hint: Consider instead the sum

$$\begin{aligned} \sum_{n=0}^{\infty} \rho^n \pi^{(n)}(0,0) &= \sum_n (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} \rho^n [\hat{p}(\xi)]^n d\xi \\ &= (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} \frac{1}{(1-\rho\hat{p}(\xi))} d\xi \\ &= (\frac{1}{2\pi})^d \int_{\mathbf{T}^d} \mathcal{R}eal\ \mathcal{P}art\Big[\frac{1}{1-\rho\hat{p}(\xi)}\Big] d\xi \end{aligned}$$

for $0 < \rho < 1$ and let $\rho \to 1$.

Example: The Queue Problem: In the example of customers arriving, except in the trivial cases of $p_0 = 0$ or $p_0 + p_1 = 1$ the chain is irreducible and aperiodoc. Since the service rate is at most 1 if the arrival rate $m = \sum_j j\, p_j > 1$, then the queue will get longer and by an application of law of large numbers it is seen that the queue length will become infinite as time progresses. This is the transient behavior of the queue. If $m < 1$, one can expect the situation to be stable and there should be an asymptotic distribution for the queue length. If $m = 1$, it is the borderline case and one should probably expect this to be the null recurent case. The actual proofs are not hard. In time n the actual number of customers served is at most n because the queue may sometomes be empty. If $\{\xi_i : i \geq 1\}$ are the number of new customers arriving at time i and X_0 is the initial number in the queue, then the number X_n in the queue at time n satisfies $X_n \geq X_0 + (\sum_{i=1}^n \xi_i) - n$ and if $m > 1$, it follows from the law of large numbers that $\lim_{n\to\infty} X_n = +\infty$, thereby establishing transience. To prove positive recurrence when $m < 1$ it is sufficient to prove that the equations

$$\sum_x q(x)\pi(x,y) = q(y)$$

has a nontrivial nonnegative solution such that $\sum_x q(x) < \infty$. We shall proceed to show that this is indeed the case. Since the equation is linear we can alaways normalize tha solution so that $\sum_x q(x) = 1$. By iteration

$$\sum_x q(x)\pi^{(n)}(x,y) = q(y)$$

for every n. If $\lim_{n\to\infty} \pi^{(n)}(x,y) = 0$ for every x and y, because $\sum_x q(x) = 1 < \infty$, by the bounded convergence theorem the right hand side tends to 0 as $n \to \infty$. therefore $q \equiv 0$ and is trivial. This rules out the transient and the null recurrent case. In our case $\pi(0,y) = p_y$ and $\pi(x,y) = p_{y-x+1}$ if $y \geq x-1$ and $x \geq 1$. In all other cases $\pi(x,y) = 0$. The equations for $\{q_x = q(x)\}$ are then

$$q_0 p_y + \sum_{x=1}^{y+1} q_x p_{y-x+1} = q_y \quad \text{for} \quad y \geq 0 \tag{3.6}$$

Multiplying equation 3.6 by z^n and summimg from 1 to ∞, we get

$$q_0\, P(z) + \frac{1}{z}\, P(z)\, [Q(z) - q_0] = Q(z)$$

where $P(z)$ and $Q(z)$ are the generating functions

$$\begin{aligned} P(z) &= \sum_{x=0}^{\infty} p_x z^x \\ Q(z) &= \sum_{x=0}^{\infty} q_x z^x \end{aligned}$$

We can solve for Q to get

$$\begin{aligned} \frac{Q(z)}{q_0} &= P(z)\Big[1 - \frac{P(z)-1}{z-1}\Big]^{-1} = P(z)\sum_{k=0}^{\infty}\Big[\frac{P(z)-1}{z-1}\Big]^k \\ &= P(z)\sum_{k=0}^{\infty}\Big[\sum_{j=1}^{\infty} p_j(1 + z + \cdots + z^{j-1})\Big]^k \end{aligned}$$

is a power series in z with nonnegative coefficients. If $m < 1$, we can let $z \to 1$ to get

$$\frac{Q(1)}{q_0} = \sum_{k=0}^{\infty} \Big[\sum_{j=1}^{\infty} j\, p_j \Big]^k = \sum_{k=0}^{\infty} m^k = \frac{1}{1-m} < \infty$$

proving positive recurrence.

The case $m = 1$ is a little bit harder. The calculations carried out earlier are still valid and we know in this case that there exists $q(x) \geq 0$ such that each $q(x) < \infty$ for each x, $\sum_x q(x) = \infty$, and

$$\sum_x q(x)\, \pi(x, y) = q(y)$$

In other words the chain admits an infinite invariant measure. Such a chain cannot be positive recurrent. To see this we note

$$q(y) = \sum_x \pi^{(n)}(x, y) q(x)$$

and if the chain were positive recurrent

$$\lim_{n\to\infty} \pi^{(n)}(x, y) = \tilde{q}(y)$$

would exist and $\sum_y \tilde{q}(y) = 1$. By Fatou's lemma

$$q(y) \geq \sum_x \tilde{q}(y) q(x) = \infty$$

giving us a contradiction. To decide between transience and null recurrence a more detiled investigation is needed. We will outline a general procedure.

Suppose we have a state x_0 that is fixed and would like to calculate $F_{x_0}(\ell) = P_{x_0}\{\tau_{x_0} \leq \ell\}$. If we can do this, then we can answer questions about transience, recurrence etc. If $\lim_{\ell\to\infty} F_{x_0}(\ell) < 1$ then the chain is transient and otherwise recurrent. In the recurrent case the convergence or divergence of

$$E^{P_{x_0}}\{\tau_{x_0}\} = \sum_\ell [1 - F_{x_0}(\ell)]$$

determines if it is positive or null recurrent. If we can determine

$$F_y(\ell) = P_y\{\tau_{x_0} \le \ell\}$$

for $y \neq x_0$, then for $\ell \geq 1$

$$F_{x_0}(\ell) = \pi(x_0, x_0) + \sum_{y \neq x_0} \pi(x_0, y) F_y(\ell - 1)$$

We shall outline a procedure for determining for $\lambda > 0$,

$$U(\lambda, y) = E_y[\exp[-\lambda \tau_{x_0}]]$$

Clearly $U(x) = U(\lambda, x)$ satisfies

$$U(x) = e^{-\lambda} \sum_y \pi(x, y) U(y) \quad \text{for} \quad x \neq x_0 \tag{3.7}$$

and $U(x_0) = 1$. One would hope that if we solve for these equations then we have our U. This requires uniqueness. Since our U is bounded in fact by 1, it is sufficient to prove uniqueness within the class of bounded solutions of equation 3.7. We will now establish that any bounded solution U of equation 3.7 with $U(x_0) = 1$, is given by

$$U(y) = U(\lambda, y) = E_y[\exp[-\lambda \tau_{x_0}]]$$

Let us define $E_n = \{X_1 \neq x_0, X_2 \neq x_0, \cdots, X_{n-1} \neq x_0, X_n = x_0\}$. Then we will prove, by induction, that for any solution U of equation (3.7), with $U(\lambda, x_0) = 1$,

$$U(y) = \sum_{j=1}^{n} e^{-\lambda j} P_y\{E_j\} + e^{-\lambda n} \int_{\tau_{x_0} > n} U(X_n)\, dP_y \tag{3.8}$$

By letting $n \to \infty$ we would obtain

$$U(y) = \sum_{j=1}^{\infty} e^{-\lambda j} P_y\{E_j\} = E^{P_y}\{e^{-\lambda \tau_{x_0}}\}$$

because U is bounded and $\lambda > 0$.

$$\begin{aligned}
&\int_{\tau_{x_0}>n} U(X_n)\, dP_y \\
&\quad = e^{-\lambda} \int_{\tau_{x_0}>n} [\sum_y \pi(X_n\,, y)\, U(y)]\, dP_y \\
&\quad = e^{-\lambda} P_y\,\{E_{n+1}\} + e^{-\lambda} \int_{\tau_{x_0}>n} [\sum_{y \neq x_0} \pi(X_n\,, y)\, U(y)]\, dP_y \\
&\quad = e^{-\lambda} P_y\,\{E_{n+1}\} + e^{-\lambda} \int_{\tau_{x_0}>n+1} U(X_{n+1})\, dP_y
\end{aligned}$$

completing the induction argument. In our case, if we take $x_0 = 0$ and try $U_\sigma(x) = e^{-\sigma\, x}$ with $\sigma > 0$, for $x \geq 1$

$$\begin{aligned}
\sum_y \pi(x,y) U_\sigma(y) &= \sum_{y \geq x-1} e^{-\sigma\, y} p_{y-x+1} \\
&= \sum_{y\geq 0} e^{-\sigma\,(x+y-1)} p_y \\
&= e^{-\sigma\, x} e^{\sigma} \sum_{y \geq 0} e^{-\sigma\, y} p_y = \psi(\sigma) U_\sigma(x)
\end{aligned}$$

where

$$\psi(\sigma) = e^{\sigma} \sum_{y\geq 0} e^{-\sigma\, y} p_y$$

Let us solve $e^\lambda = \psi(\sigma)$ for σ which is the same as solving $\log \psi(\sigma) = \lambda$ for $\lambda > 0$ to get a solution $\sigma = \sigma(\lambda) > 0$. Then

$$U(\lambda, x) = e^{-\sigma(\lambda)\, x} = E^{P_x}\{e^{-\lambda \tau_0}\}$$

We see now that recurrence is equivalent to $\sigma(\lambda) \to 0$ as $\lambda \downarrow 0$ and positive recurrence to $\sigma(\lambda)$ being differentiable at $\lambda = 0$. The function $\log \psi(\sigma)$ is convex and its slope at the origin is $1 - m$. If $m > 1$ it dips below 0 initially for $\sigma > 0$ and then comes back up to 0 for some positive σ_0 before turning positive for good. In that situation $\lim_{\lambda \downarrow 0} \sigma(\lambda) = \sigma_0 > 0$ and that is transience. If $m < 1$ then $\log \psi(\sigma)$ has a positive slope at the origin and $\sigma'(0) = \frac{1}{\psi'(0)} = \frac{1}{1-m} < \infty$. If $m = 1$, then $\log \psi$ has zero slope at the origin and $\sigma'(0) = \infty$. This concludes the discussion of this problem.

Exercise 3.15. **The Urn Problem.** We now turn to a discussion of the urn problem.

$$\pi(p,q\,;p+1,q)=\frac{p}{p+q}\quad\text{and}\quad\pi(p,q\,;p,q+1)=\frac{q}{p+q}$$

and π is zero otherwise. In this case the equation

$$F(p,q)=\frac{p}{p+q}F(p+1,q)+\frac{q}{p+q}F(p,q+1)\ \ \text{for all } p,q$$

which will play a role later, has lots of solutions. In particular, $F(p,q)=\frac{p}{p+q}$ is one and for any $0<x<1$

$$F_x(p,q)=\frac{1}{\beta(p,q)}x^{p-1}(1-x)^{q-1}$$

where

$$\beta(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

is a solution as well. The former is defined on $p+q>0$ where as the latter is defined only on $p>0,q>0$. Actually if p or q is initially 0 it remains so for ever and there is nothing to study in that case. If f is a continuous function on $[0,1]$ then

$$F_f(p,q)=\int_0^1 F_x(p,q)f(x)\,dx$$

is a solution and if we want we can extend F_f by making it $f(1)$ on $q=0$ and $f(0)$ on $p=0$. It is a simple exercise to verify

$$\lim_{\substack{p,q\to\infty\\ \frac{p}{q}\to x}} F_f(p,q)=f(x)$$

for any continuous f on $[0,1]$. We will show that the ratio $\xi_n=\frac{p_n}{p_n+q_n}$ which is random, stabilizes asymptotically (i.e. has a limit) to a random variable ξ and if we start from p,q the distribution of ξ is the Beta distribution on $[0,1]$ with density

$$F_x(p,q)=\frac{1}{\beta(p,q)}x^{p-1}(1-x)^{q-1}$$

Suppose we have a Markov Chain on some state space $\mathcal{X}$ with transition probability $\pi(x,y)$ and $U(x)$ is a bounded function on $\mathcal{X}$ that solves

$$U(x) = \sum_y \pi(x,y)\, U(y)$$

Such functions are called (bounded) Harmonic functions for the Chain. Consider the random variables $\xi_n = U(X_n)$ for such an harmonic function. ξ_n are uniformly bounded by the bound for U. If we denote by $\eta_n = \xi_n - \xi_{n-1}$ an elementary calculation reveals

$$\begin{aligned} E^{P_x}\{\eta_{n+1}\} &= E^{P_x}\{U(X_{n+1}) - U(X_n)\} \\ &= E^{P_x}\{E^{P_x}\{U(X_{n+1}) - U(X_n)\}|\mathcal{F}_n\}\} \end{aligned}$$

where $\mathcal{F}_n$ is the σ-field generated by $X_0, \cdots, X_n$. But

$$E^{P_x}\{U(X_{n+1}) - U(X_n)\}|\mathcal{F}_n\} = \sum_y \pi(X_n, y)[U(y) - U(X_n)] = 0$$

A similar calculation shows that

$$E^{P_x}\{\eta_n \eta_m\} = 0$$

for $m \neq n$. If we write

$$U(X_n) = U(X_0) + \eta_1 + \eta_2 + \cdots + \eta_n$$

this is an orthogonal sum in $L_2[P_x]$ and because U is bounded

$$E^{P_x}\{|U(X_n)|^2\} = |U(x)|^2 + \sum_{i=1}^n E^{P_x}\{|\eta_i|^2\} \leq C$$

is bounded in n. Therefore $\lim_{n\to\infty} U(X_n) = \xi$ exists in $L_2[P_x]$ and $E^{P_x}\{\xi\} = U(x)$. Actually the limit exists almost surely and we will show it when we discuss martingales later. In our example if we take $U(p,q) = \frac{p}{p+q}$, as remarked earlier, this is a harmonic function bounded by 1 and therefore

$$\lim_{n\to\infty} \frac{p_n}{p_n + q_n} = \xi$$

exists in $L_2[P_x]$. Moreover if we take $U(p,q) = F_f(p,q)$ for some continuous f on $[0,1]$, because $F_f(p,q) \to f(x)$ as $p,q \to \infty$ and $\frac{p}{q} \to x$, $U(p_n,q_n)$ has a limit as $n \to \infty$ and this limit has to be $f(\xi)$. On the other hand

$$\begin{aligned} E^{P_{p,q}}\{U(p_n,q_n)\} &= U(p_0,q_0) = F_f(p_0,q_0) \\ &= \frac{1}{\beta(p_0,q_0)} \int_0^1 f(x) x^{p_0-1}(1-x)^{q_0-1}\,dx \end{aligned}$$

giving us

$$E^{P_{p,q}}\{f(\xi)\} = \frac{1}{\beta(p,q)} \int_0^1 f(x)\, x^{p-1}(1-x)^{q-1}\,dx$$

thereby identifying the distribution of ξ under $P_{p,q}$ as the Beta distribution with the right parameters.

Example: Let $\mathcal{X}$ be the set of integers. Assume that transitions from x are possible only to $x-1, x$, and $x+1$. The transition matrix $\pi(x,y)$ appears as a tridiagonal matrix with $\pi(x,y) = 0$ unless $|x-y| \leq 1$. For simplicity let us assume that $\pi(x,x), \pi(x,x-1)$ and $\pi(x,x+1)$ are positive for all x. The chain is then irreducible and aperiodic. Let us try to solve for

$$U(x) = P_x\{\tau_0 = \infty\}$$

that satisfies the equation

$$U(x) = \pi(x,x-1)\,U(x-1) + \pi(x,x)\,U(x) + \pi(x,x+1)\,U(x+1)$$

for $x \neq 0$ with $U(0) = 0$. The equations decouple into a set for $x > 0$ and a set for $x < 0$. If we denote by $V(x) = U(x+1) - U(x)$ for $x \geq 0$, then we always have

$$U(x) = \pi(x,x-1)\,U(x) + \pi(x,x)\,U(x) + \pi(x,x+1)\,U(x)$$

so that

$$\pi(x,x-1)\,V(x-1) - \pi(x,x+1)\,V(x) = 0$$

or

$$\frac{V(x)}{V(x-1)} = \frac{\pi(x,x-1)}{\pi(x,x+1)}$$

and therefore

$$V(x) = V(0) \prod_{i=1}^{x} \frac{\pi(i, i-1)}{\pi(i, i+1)}$$

and

$$U(x) = V(0)\Big[1 + \sum_{y=1}^{x-1} \prod_{i=1}^{y} \frac{\pi(i, i-1)}{\pi(i, i+1)}\Big]$$

If the chain is to be transient we must have for some choice of $V(0)$, $0 \le U(x) \le 1$ for all $x > 0$ and this will be possible only if

$$\sum_{y=1}^{\infty} \prod_{i=1}^{y} \frac{\pi(i, i-1)}{\pi(i, i+1)} < \infty$$

which then becomes a necessary condition for

$$P_x\{\tau_0 = \infty\} > 0$$

for $x > 0$. There is a similar condition on the negative side

$$\sum_{y=1}^{\infty} \prod_{i=1}^{y} \frac{\pi(-i, -i+1)}{\pi(-i, -i-1)} < \infty$$

Transience needs at least one of the two series to converge. Actually the converse is also true. If, for instance the series on the positive side converges then we get a function $U(x)$ with $0 \le U(x) \le 1$ and $U(0) = 0$ that satisfies

$$U(x) = \pi(x, x-1)\, U(x-1) + \pi(x, x)\, U(x) + \pi(x, x+1)\, U(x+1)$$

and by iteration one can prove that for each n,

$$U(x) = \int_{\tau_0 > n} U(X_n)\, dP_x \le P\{\tau_0 > n\}$$

so the existence of a nontrivial U implies transience.

Exercise 3.16. Determine the conditions for positive recurrence.

Exercise 3.17. If we replace the set of integers by the set of nonnegative integers, [$\pi(0, y)$ is nonzero only for $y = 0$ or 1] and such processes are called birth and death processes. Work out the conditions in that case.

Exercise 3.18. In the special case of a birth and death process with $\pi(0,1) = \pi(0,0) = \frac{1}{2}$, and for $x \geq 1$, $\pi(x,x) = \frac{1}{3}$, $\pi(x,x-1) = \frac{1}{3} + a_x$, $\pi(x,x+1) = \frac{1}{3} - a_x$ with $a_x = \frac{\lambda}{x^\alpha}$ for large x, find conditions on positive α and real λ for the chain to be transient, null recurrent and positive recurrent.

Exercise 3.19. The notion of a Markov Chain makes sense for a finite chain $X_0, \cdots, X_n$. Formulate it precisely. Show that if the chain $\{X_j : 0 \leq j \leq n\}$ is Markov so is the reversed chain $\{Y_j : 0 \leq j \leq n\}$ where $Y_j = X_{n-j}$ for $0 \leq j \leq n$. Can the transition probabilities of the reversed chain be determined by the transition probabilities of the forward chain? If the forward chain has stationary transition proabilities does the same hold true for the reversed chain? What if we assume that the chain has a finte invariant probability distribution and we initialize the chain to start with an initial distribution which is the invariant distribution?

Exercise 3.20. Consider the simple chain on nonnegative integers with the following transition probailities. $\pi(0,x) = p_x$ for $x \geq 0$ with $\sum_{x=0}^\infty p_x = 1$. For $x > 0$, $\pi(x,x-1) = 1$ and $\pi(x,y) = 0$ for all other y. Determine conditions on $\{p_x\}$ in order that the chain may be transient, null recurrent or positive recurrent. Determine the invariant probability measure in the positive recurrent case.

Exercise 3.21. Show that any null recurrent equivalence class must necessarily contain an infinite number of states. In patricular any Markov Chain with a finite state space has only transient and positive recurrent states and moreover the set of positive recurrent states must be non empty.