Recap: Stationary Distribution

- π is said to be a stationary distribution for the Markov chain with transition probabilities P if

$$\pi(y) = \sum_{x \in S} \pi(x) P(x, y), \quad \forall y \in S$$

- When π is stationary distribution,
  $\pi_0 = \pi \quad \Rightarrow \quad \pi_n = \pi, \ \forall n$
- If $\pi_n = \pi, \ \forall n$ then π is a stationary distribution
- For a finite chain: $P^T \pi = \pi$
- A stationary distribution always exists for a finite chain

Recap

- $N_n(y)$ – number of visits to y till n
- $G_n(x, y) = E_x[N_n(y)] = \sum_{m=1}^{n} P^m(x, y)$
  – expected number of visits to y till n
- $m_y = E_y[T_y]$ – mean return time to y

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{I_{[T_y<\infty]}}{m_y}, \quad w.p.1$$

$$\lim_{n\to\infty} \frac{G_n(x,y)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x, y) = \frac{\rho_{xy}}{m_y}$$

Recap: positive and null recurrent states

- y is positive recurrent if $m_y < \infty$
- y is null recurrent if $m_y = \infty$
- If x is positive recurrent and x leads to y, then y is positive recurrent
- In a closed irreducible set of recurrent states either all states are positive recurrent or all states are null recurrent
- A finite closed set has to have at least one positive recurrent state
- A finite chain cannot have null recurrent states

Recap: Existence of stationary distribution

- In any stationary distribution π, $\pi(y) = 0$ if y is transient or null recurrent
- An irreducible transient or null recurrent chain does not have a stationary distribution
- An irreducible positive recurrent chain has a unique stationary distribution: $\pi(y) = \frac{1}{m_y}$
- An irreducible chain has a stationary distribution iff it is positive recurrent
- For a non-irreducible chain, for each closed irreducible set of positive recurrent states, there is a unique stationary distribution concentrated on that set.
- All stationary distributions of the chain are convex combinations of these

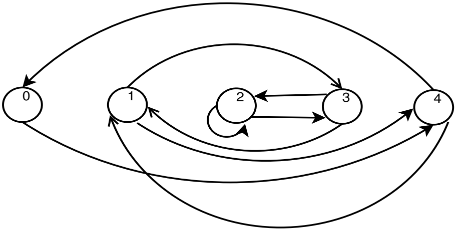
Recap: Periodic chains

- The period of a state x is $d_x = \gcd\{n \geq 1 \; : \; P^n(x,x) > 0\}$
- If x and y lead to each other, $d_x = d_y$
- In an irreducible chain, all states have the same period
- An irreducible chain is called aperiodic if the period is 1
- For an irreducible aperiodic positive recurrent chain, π_n converges to π, the unique stationary distribution, irrespective of what π_0 is.
- Also, for an irreducible, aperiodic, positive recurrent chain, $P^n(x,y)$ converges to $\frac{1}{m_y}$

Example

- Consider the umbrella problem

$$P = \begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1-p & p \\ 2 & 0 & 0 & 1-p & p & 0 \\ 3 & 0 & 1-p & p & 0 & 0 \\ 4 & 1-p & p & 0 & 0 & 0 \end{array}$$



- This is an irreducible, aperiodic positive recurrent chain

- We want calculate the probability of getting caught in a rain without an umbrella.
- This would be the steady state probability of state 0 multiplied by p
- We are using the fact that this chain converges to the stationary distribution starting with any initial probabilities.

$$P = \left[\begin{array}{c|ccccc} & 0 & 1 & 2 & 3 & 4 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1-p & p \\ 2 & 0 & 0 & 1-p & p & 0 \\ 3 & 0 & 1-p & p & 0 & 0 \\ 4 & 1-p & p & 0 & 0 & 0 \end{array}\right]$$

The stationary distribution satisfies $\pi^T P = \pi^T$

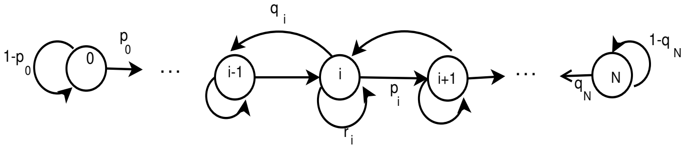
$$\begin{aligned} \pi(0) &= (1-p)\pi(4) \\ \pi(1) &= (1-p)\pi(3) + p\pi(4) \;\;\Rightarrow\;\; \pi(3) = \pi(1) \\ \pi(2) &= (1-p)\pi(2) + p\pi(3) \\ \pi(3) &= (1-p)\pi(1) + p\pi(2) \;\;\Rightarrow\;\; \pi(2) = \pi(1) \\ \pi(4) &= \pi(0) + p\pi(1) \;\;\Rightarrow\;\; \pi(4) = \pi(1) \end{aligned}$$

This gives $4\pi(1) + (1-p)\pi(1) = 1$ and hence

$$\pi(i) = \frac{1}{5-p} \; i = 1, 2, 3, 4 \quad \text{and} \quad \pi(0) = \frac{1-p}{5-p}$$

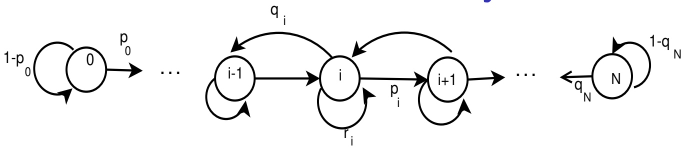
Birth-Death chains

- The following is a finite birth-death chain



- We assume $p_i, q_i > 0, \forall i$.
- Then the chain is irreducible, positive recurrent
- If we assume $r_i > 0$ at least for one i, it is also aperiodic
- We can derive a general form for its stationary probabilities

birth-death chains – stationary distribution



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

$$\begin{aligned}
\pi(0) &= \pi(0)(1-p_0) + \pi(1)q_1 \\
&\Rightarrow \ \pi(1)q_1 - \pi(0)p_0 = 0 \\
\pi(1) &= \pi(0)p_0 + \pi(1)(1-p_1-q_1) + \pi(2)q_2 \\
&\Rightarrow \pi(1)q_1 - \pi(0)p_0 = \pi(2)q_2 - \pi(1)p_1 \\
&\Rightarrow \pi(2)q_2 - \pi(1)p_1 = 0 \\
\pi(2) &= \pi(1)p_1 + \pi(2)(1-p_2-q_2) + \pi(3)q_3 \\
&\Rightarrow \pi(2)q_2 - \pi(1)p_1 = \pi(3)q_3 - \pi(2)p_2 = 0
\end{aligned}$$

- Thus we get

$$\begin{aligned}\pi(1)q_1 - \pi(0)p_0 &= 0 \Rightarrow \pi(1) = \frac{p_0}{q_1}\,\pi(0) \\ \pi(2)q_2 - \pi(1)p_1 &= 0 \Rightarrow \pi(2) = \frac{p_1}{q_2}\,\pi(1) = \frac{p_0 p_1}{q_1 q_2}\,\pi(0)\end{aligned}$$

- Iterating like this, we get

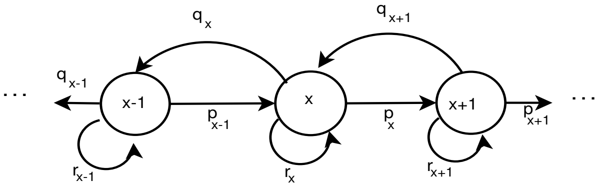
$$\pi(n) = \eta_n\,\pi(0), \quad \text{where} \quad \eta_n = \frac{p_0 p_1 \cdots p_{n-1}}{q_1 q_2 \cdots q_n}, \; n = 1, 2, \cdots, N$$

- With $\eta_0 = 1$, we get $\pi(0)\sum_{j=0}^{N} \eta_j = 1$ and hence

$$\pi(0) = \frac{1}{\sum_{j=0}^{N} \eta_j} \quad \text{and} \quad \pi(n) = \eta_n\,\pi(0), \; n = 1, \cdots, N$$

- Note that this process is applicable even for infinite chains with state space $\{0, 1, 2, \cdots\}$ (but there may not be a solution)

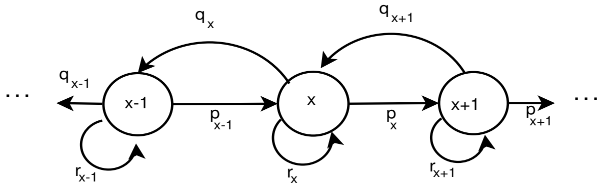
- Consider a birth-death chain



- The chain may be infinite or finite
- Let $a, b \in S$ with $a < b$. Assume $p_x, q_x > 0$, $a < x < b$.
- Define

$$U(x) = P_x[T_a < T_b], \ a < x < b, \quad U(a) = 1, \ U(b) = 0$$

- We want to derive a formula for $U(x)$
- This can be useful, e.g., in the gambler's ruin chain



$$
\begin{aligned}
U(x) &= P_x[T_a < T_b] = Pr[T_a < T_b | X_0 = x] \\
&= \sum_{y=x-1}^{x+1} Pr[T_a < T_b | X_1 = y] \; Pr[X_1 = y | X_0 = x] \\
&= U(x-1)q_x + U(x)r_x + U(x+1)p_x \\
&= U(x-1)q_x + U(x)(1 - p_x - q_x) + U(x+1)p_x
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \quad & q_x[U(x) - U(x-1)] = p_x[U(x+1) - U(x)] \\
\Rightarrow \quad & U(x+1) - U(x) = \frac{q_x}{p_x} \, [U(x) - U(x-1)]
\end{aligned}
$$

$$
\begin{aligned}
U(x+1) - U(x) &= \frac{q_x}{p_x}\,[U(x) - U(x-1)] \\
&= \frac{q_x}{p_x}\,\frac{q_{x-1}}{p_{x-1}}\,[U(x-1) - U(x-2)] \\
&= \frac{q_x q_{x-1} \cdots q_{a+1}}{p_x p_{x-1} \cdots p_{a+1}}\,[U(a+1) - U(a)]
\end{aligned}
$$

$$
\text{Let} \quad \gamma_y = \frac{q_y q_{y-1} \cdots q_{a+1}}{p_y p_{y-1} \cdots p_{a+1}}, \; a < y < b, \;\; \gamma_a = 1
$$

Now we get

$$
U(x+1) - U(x) = \frac{\gamma_x}{\gamma_a}[U(a+1) - U(a)]
$$

$$U(x+1) - U(x) = \frac{\gamma_x}{\gamma_a} \left[U(a+1) - U(a)\right]$$

- By taking $x = b-1,\ b-2, \cdots, a+1,\ a$,

$$\begin{aligned}
U(b) - U(b-1) &= \frac{\gamma_{b-1}}{\gamma_a} \left[U(a+1) - U(a)\right] \\
U(b-1) - U(b-2) &= \frac{\gamma_{b-2}}{\gamma_a} \left[U(a+1) - U(a)\right] \\
&\vdots \\
U(a+1) - U(a) &= \frac{\gamma_a}{\gamma_a} \left[U(a+1) - U(a)\right]
\end{aligned}$$

- Adding all these we get

$$\frac{1}{\gamma_a} \left[U(a+1) - U(a)\right] \sum_{x=a}^{b-1} \gamma_x = U(b) - U(a) = 0 - 1$$

$$\Rightarrow \quad U(a) - U(a+1) = \frac{\gamma_a}{\sum_{x=a}^{b-1} \gamma_x}$$

- Using these, we get

$$\begin{aligned} U(x) - U(x+1) &= \frac{\gamma_x}{\gamma_a}[U(a) - U(a+1)] \\ &= \frac{\gamma_x}{\gamma_a} \frac{\gamma_a}{\sum_{x=a}^{b-1} \gamma_x} = \frac{\gamma_x}{\sum_{x=a}^{b-1} \gamma_x} \end{aligned}$$

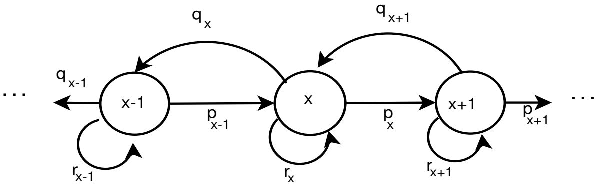
- Putting $x = b-1,\ b-2, \cdots, y$ in the above

$$\begin{aligned} U(b-1) - U(b) &= \frac{\gamma_{b-1}}{\sum_{x=a}^{b-1} \gamma_x} \\ U(b-2) - U(b-1) &= \frac{\gamma_{b-2}}{\sum_{x=a}^{b-1} \gamma_x} \\ &\vdots \\ U(y) - U(y+1) &= \frac{\gamma_y}{\sum_{x=a}^{b-1} \gamma_x} \end{aligned}$$

- Adding these we get

$$U(y) - U(b) = U(y) = \frac{\sum_{x=y}^{b-1} \gamma_x}{\sum_{x=a}^{b-1} \gamma_x},\ a < y < b$$

- We are considering birth-death chains



- We have derived, for $a < y < b$,

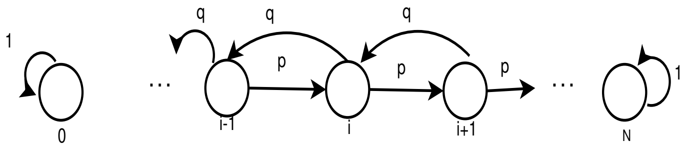
$$U(y) = P_y[T_a < T_b] = \frac{\sum_{x=y}^{b-1} \gamma_x}{\sum_{x=a}^{b-1} \gamma_x}, \quad \gamma_x = \frac{q_x q_{x-1} \cdots q_{a+1}}{p_x p_{x-1} \cdots p_{a+1}}$$

- Hence we also get

$$P_y[T_b < T_a] = \frac{\sum_{x=a}^{y-1} \gamma_x}{\sum_{x=a}^{b-1} \gamma_x}$$

- Suppose this is a Gambler's ruin chain:

$$p_x = p, q_x = q, \ \forall x$$

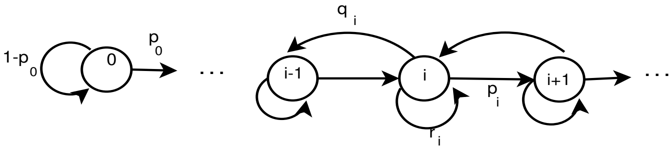


- Then, $\gamma_x = \left(\frac{q}{p}\right)^x$
- Hence, for a Gambler's ruin chain we get, e.g.,

$$P_i[T_N < T_0] = \frac{\sum_{x=0}^{i-1} \gamma_x}{\sum_{x=0}^{N-1} \gamma_x} = \frac{\left(\frac{q}{p}\right)^i - 1}{\left(\frac{q}{p}\right)^N - 1}$$

- This is the probability of gambler being successful

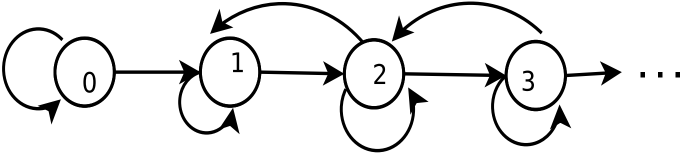
- Consider the following chain over $\{0, 1, \cdots\}$



- This is an infinite irreducible birth-death chain
- We want to know whether the chain is transient or recurrent etc.
- We can use the earlier analysis for this too.

$$\begin{aligned}P_1[T_0 < T_n] &= \frac{\sum_{x=1}^{n-1} \gamma_x}{\sum_{x=0}^{n-1} \gamma_x}, \ \forall n > 1 \\ &= \frac{\sum_{x=0}^{n-1} \gamma_x - \gamma_0}{\sum_{x=0}^{n-1} \gamma_x} = 1 - \frac{1}{\sum_{x=0}^{n-1} \gamma_x}\end{aligned}$$

- Consider this chain started in state 1.



$$[T_0 < T_n] \subset [T_0 < T_{n+1}], \;\; n = 2, 3, \cdots$$

since the chain cannot hit $n+1$ without hitting n.

- Also, $1 \leq T_2 < T_3 < \cdots < T_n$ and $T_n \geq n$.
- Hence $[T_0 < \infty]$ is same as $[T_0 < T_n, \;\; \text{for some } n]$

- Consider this chain started in state 1.

$$[T_0 < T_n] \subset [T_0 < T_{n+1}], \quad n = 2, 3, \cdots$$

since the chain cannot hit $n+1$ without hitting n.

- Also, $1 \leq T_2 < T_3 < \cdots < T_n$ and $T_n \geq n$.
- Hence $[T_0 < \infty]$ is same as $[T_0 < T_n, \ \text{ for some } n]$

$$\begin{aligned} P_1[T_0 < T_n, \ \text{ for some } n] &= P_1\left(\cup_{n>1} [T_0 < T_n]\right) \\ &= P_1\left(\lim_{n\to\infty}[T_0 < T_n]\right) \\ &= \lim_{n\to\infty} P_1\left([T_0 < T_n]\right) \\ &= \lim_{n\to\infty} \ 1 - \frac{1}{\sum_{x=0}^{n-1} \gamma_x} \\ \Rightarrow \quad P_1[T_0 < \infty] &= 1 - \frac{1}{\sum_{x=0}^{\infty} \gamma_x} \end{aligned}$$

- **Theorem**: The chain is recurrent iff $\sum_{x=0}^{\infty} \gamma_x = \infty$

Proof

- Supoose chain is recurrent. Since it is irreducible,

$$P_1[T_0 < \infty] = 1 \;\; \Rightarrow \;\; \sum_{x=0}^{\infty} \gamma_x = \infty$$

- Suppose $\sum_{x=0}^{\infty} \gamma_x = \infty \Rightarrow \; P_1[T_0 < \infty] = 1$

$$\begin{aligned} P_0[T_0 < \infty] &= P(0,0) \; + \; P(0,1) \; P_1[T_0 < \infty] \\ &= P(0,0) + P(0,1) = 1 \end{aligned}$$

- Implies state 0 is recurrent and hence the chain is recurrent because it is irreducible.
- Note that we have used the fact that the chain is infinite only to the right.

- The chain is transient if $\sum_{x=0}^{\infty} \gamma_x < \infty$
- Let $p_x = p, q_x = q \Rightarrow \quad \gamma_x = \left(\frac{q}{p}\right)^x$

$$\text{Transient if} \quad \sum_{x=0}^{\infty} \left(\frac{q}{p}\right)^x < \infty \quad \Leftrightarrow \quad q < p$$

$$\text{Recurrent if} \quad \sum_{x=0}^{\infty} \left(\frac{q}{p}\right)^x = \infty \quad \Leftrightarrow \quad q \geq p$$

- Intuitively clear
- This chain with $q < p$ is an example of an irreducible chain that is wholly transient

- We know the chain is recurrent if $\sum_{x=0}^{\infty} \left(\frac{q}{p}\right)^x = \infty$
- When will this chain be positive recurrent?
- We know that an irreducible chain is positive recurrent if and only if it has a stationary distribution.
- We can check if it has a stationary distribution
- The earlier equations that we derived earlier hold for this infinite case also.

- We derived earlier the equations that a stationary distribution of this chain (if it exists) has to satisfy

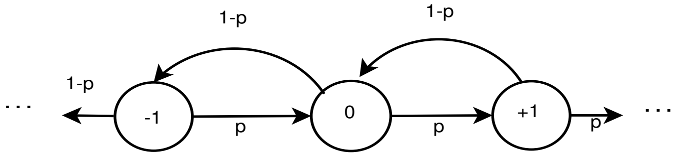
$$\pi(n) = \eta_n \, \pi(0), \quad \text{where} \quad \eta_n = \frac{p_0 p_1 \cdots p_{n-1}}{q_1 q_2 \cdots q_n}, \; n = 1, 2, \cdots ,$$

- Setting $\eta_0 = 1$, we get $\pi(0) \sum_{j=0}^{\infty} \eta_j = 1$
- Hence stationary distribution exists iff $\sum_{j=0}^{\infty} \eta_j < \infty$
- Let $p_x = p, q_x = q$

$$\sum_{j=0}^{\infty} \eta_j = \sum_{j=0}^{\infty} \left(\frac{p}{q} \right)^j < \infty \quad \Leftrightarrow \quad p < q$$

- Thus in this special case, the chain is
  - transient if $p > q$; recurrent if $p \le q$
  - positive recurrent if $p < q$
  - null recurrent if $p = q$

- This analysis can handle chains which are infinite in one direction
- Consider the following random walk chain



- The state space here is $\{\cdots, -1, 0, +1, \cdots\}$
- The chain is irreducible and periodic with period 2
- $P^{2n}(0,0) = \ ^{2n}C_n p^n (1-p)^n$.
- We can look at the limit of $\frac{1}{n}\sum_n P^{2n}(0,0)$
- We can show that the chain is transient if $p \neq 0.5$ and is recurrent if $p = 0.5$.

- In general, determining when an infinite chain is positive recurrent is difficult.
- The method we had works only for birth-death chains over non-negative integers.
- There is a useful general theorem.

  Foster's Theorem
  Let P be the transition probabilities of a homogeneous irreducible Markov chain with state space S. Let $h : S \to \Re$ with $h(x) \geq 0$ and
    - $\sum_{k \in S} P(i,k)h(k) < \infty \; \forall i \in F$ and
    - $\sum_{k \in S} P(i,k)h(k) \leq h(i) - \epsilon \; \forall i \notin F$

  for some finite set F and some $\epsilon > 0$. Then the Markov chain is positive recurrent
- The h here is called a Lyapunov function.
- We will not prove this theorem

- Let $\{X_n,\ n \geq 0\}$ be an irreducible markov chain on a finite state space S with stationary distribution π.
- Let $r : S \to \Re$ be a bounded function.
- Suppose we want $E[r(X)]$ with respect to the stationary distribution π ($E[r(X)] = \sum_{j \in S} r(j)\pi(j)$)
- Let $N_n(j)$ be as earlier. Then

$$\frac{1}{n} \sum_{m=1}^{n} r(X_m) = \frac{1}{n} \sum_{j \in S} N_n(j) r(j)$$

$$\Rightarrow \lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} r(X_m) = \sum_{j \in S} r(j) \lim_{n \to \infty} \frac{N_n(j)}{n} = \sum_{j \in S} r(j)\pi(j)$$

- For this to be true for infinite S, we need some extra conditions

MCMC Sampling

- Consider a distribution over (finite) S: $\pi(x) = \frac{b(x)}{Z}$
- Since this is a distribution, $Z = \sum_{x \in S} b(x)$
- We assume, we can efficiently calculate $b(x)$ for any x but computation of Z is intractable or computationally expensive
  E.g., the Boltzmann distribution: $b(x) = e^{-E(x)/KT}$
- We want $E[g(X)]$ w.r.t. distribution π (for any g)

$$E[g(X)] = \sum_x g(x)\ \pi(x) \approx \frac{1}{n} \sum_{i=1}^{n} g(X_i), \quad X_1, \cdots X_n \sim \pi$$

- One way to generate samples is to design an ergodic markov chain with stationary distribution π
  – MCMC sampling

- Suppose $\{X_n\}$ is a an irreducible, aperiodic positive recurrent Markov chain with stationary dist $\pi(x) = \frac{b(x)}{Z}$
- Then we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} g(X_m) = \sum_{x} g(x)\pi(x)$$

- hence, if we can design a Markov chain with a given stationary distribution, we can use that to calculate the expectation.
- We can also use the chain to generate samples from distribution π

- $\{X_n\}$: Markov chain with stationary dist $\pi(x) = \frac{b(x)}{Z}$
  We can approximate the expectation as

$$\sum_x g(x)\pi(x) \approx \frac{1}{n}\sum_{i=1}^{n} g(X_{M+i})$$

  Where M is large enough to assume chain is in steady state
- When we take sample mean, $\frac{1}{n}\sum_{i=1}^{n} Z_i$, we want Z_i to be uncorrelated
- We can, for example, use

$$\sum_x g(x)\pi(x) \approx \frac{1}{n}\sum_{i=1}^{n} g(X_{M+Ki})$$

- For all these, we need to design a Markov chain with π as stationary distribution

- Let $Q = [q(i,j)]$ be the transition probability matrix of an irreducible Markov chain over S.
- Q is called the proposal distribution
- We start with arbitray X_0 and generate $X_{n+1},\ n = 0, 1, 2, \cdots$, iteratively as follows
  - If $X_n = i$, we generate Y with $Pr[Y = k] = q(i,k)$
  - Let the generated value for Y be j. Set

$$X_{n+1} = \begin{cases} j & \text{with probability } \alpha(i,j) \\ X_n & \text{with probability } 1 - \alpha(i,j) \end{cases}$$

- $\alpha(i,j)$ is called the acceptance probability
- We want to choose $\alpha(i,j)$ to make X_n an ergodic markov chain with stationary probabilities π

- The stationary distribution π satisfies (with transition probabilities P)

$$\pi(y) = \sum_x \pi(x)\ P(x, y), \ \ \forall y \in S$$

- Suppose there is a distribution $g(\cdot)$ that satisfies

$$g(y)\ P(y, x) = g(x)\ P(x, y), \ \ \forall x, y \in S$$

  This is called detailed balance
- Summing both sides above over x give

$$g(y) = \sum_x g(y)\ P(y, x) = \sum_x g(x) P(x, y), \ \ \forall y$$

- Thus if $g(\cdot)$ satisfies detailed balance, then it must be the stationary distribution
- Note that it is not necessary for a stationary distribution to satisfy detailed balance

- Any stationary distribution has to satisfy

$$\pi(y) = \sum_x \pi(x)\ P(x, y), \ \ \forall y \in S$$

- If I can find a π that satisfies

$$\pi(x)P(x, y) = \pi(y)P(y, x), \ \ \forall x, y \in S, \ x \neq y$$

  that would be the stationary distribution
- This is called detailed balance

- Recall our algorithm for generating $X_n, \ n = 0, 1, \cdots$
- Start with arbitrary X_0 and generate X_{n+1} from X_n
  - If $X_n = i$, we generate Y with $Pr[Y = k] = q(i, k)$
  - Let the generated value for Y be j. Set

$$X_{n+1} = \begin{cases} j & \text{with probability } \ \alpha(i,j) \\ X_n & \text{with probability } \ 1 - \alpha(i,j) \end{cases}$$

- Hence the transition probabilities for X_n are

$$\begin{aligned} P(i,j) &= q(i,j)\,\alpha(i,j), \quad i \neq j \\ P(i,i) &= q(i,i) + \sum_{j \neq i} q(i,j)\,(1 - \alpha(i,j)) \end{aligned}$$

- $\pi(i) = b(i)/Z$ is the desired stationary distribution
- So, we can try to satisfy

$$\begin{aligned} \pi(i)\ P(i,j) &= \pi(j)\ P(j,i), \ \ \forall i, j, i \neq j \\ \text{that is,} \quad b(i)q(i,j)\,\alpha(i,j) &= b(j)q(j,i)\,\alpha(j,i) \end{aligned}$$