Recap: Stationary Distribution of markov chain

- 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
- The expected (limiting) fraction of time spent in state y for a chain starting in x is

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

- $m_y = E_y[T_y]$ – mean return time to y.
  It is the expected time to return to y for a chain starting in y.

Recap

- **Theorem**: For a recurrent state y

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

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

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

Recap: positive and null recurrent states

- A recurrent state y is called **null recurrent** if $m_y = \infty$.
- y is called **positive recurrent** if $m_y < \infty$
- The limiting fraction of time spent by the chain in transient and null recurrent states is zero.
- **Theorem:** Let x be positive recurrent and let x lead to y. Then y is positive recurrent.

- Thus, in a closed irreducible set of recurrent states, if one state is positive recurrent then all are positive recurrent.
- Hence, in the partition: $S_R = C_1 + C_2 + \cdots$, each C_i is either wholly positive recurrent or wholly null recurrent.
- We next show that a finite chain cannot have any null recurrent states.

- Let C be a finite closed set of recurrent states.
- Suppose all states in C are null recurrent. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) = 0, \quad \forall x, y \in C$$

- Since C is closed, $\sum_{y \in C} P^m(x,y) = 1, \ \forall m, \ \forall x \in C$.
- Thus we get

$$1 = \frac{1}{n} \sum_{m=1}^{n} \sum_{y \in C} P^m(x,y) = \sum_{y \in C} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y), \ \forall n$$

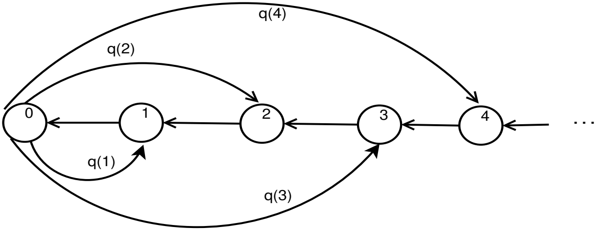
$$\Rightarrow \quad 1 = \lim_{n\to\infty} \sum_{y \in C} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) = \sum_{y \in C} \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) = 0$$

where we could take the limit inside the sum because C is finite.

- If C is a finite closed set of recurrent states then all states in it cannot be null recurrent.
- Actually what we showed is that any closed finite set must have at least one positive recurrent state.
- Hence, in a finite chain, every closed irreducible set of recurrent states contains only positive recurrent states.
- Hence, a finite chain cannot have a null recurrent state.

Example of null recurrent chain

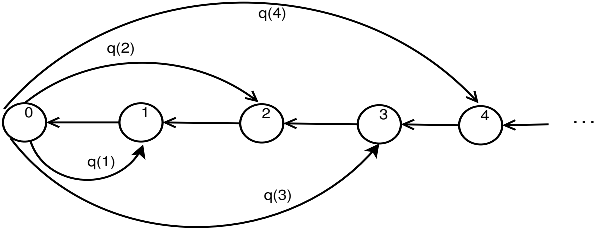
- Consider the chain with state space $\{0, 1, \cdots\}$ given by



- Here, $q(k) \geq 0, \forall k$ and $\sum_{k=1}^{\infty} q(k) = 1$.
- The chain is irreducible.
- So, we want to know whether it is transient or recurrent.
- We can calculate ρ_{00} to test this.

Example of null recurrent chain

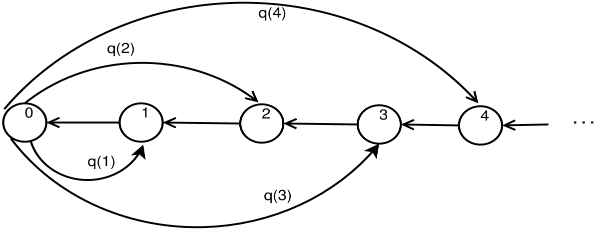
- Consider the chain with state space $\{0, 1, \cdots\}$ given by



We have
$$P_0[T_0 = j + 1] = q(j), \; j = 1, 2, \cdots$$
(Note that $P_0[T_0 = 1] = 0$)
$$P_0[T_0 < \infty] = \sum_{j=2}^{\infty} P_0[T_0 = j] = \sum_{j=1}^{\infty} q(j) = 1$$

- So, the chain is recurrent.
- We want to know whether it is positive recurrent or null recurrent

Example of null recurrent chain

- 

$$m_0 = \sum_{j=2}^{\infty} j \; P_o[T_0 = j] = \sum_{j=2}^{\infty} j \; q(j-1) = \sum_{j=1}^{\infty} (j+1) q(j)$$

- So, $m_0 = \infty$ if the $q(\cdot)$ distribution has infinite expectation. For example, if $q(k) = \frac{c}{k^2}$
- Then state 0 is null recurrent. Implies chain is null recurrent

- Suppose π is a stationary distribution.
- Then $\pi(y) = 0$ if y is transient or null recurrent
- We prove this as follows

$$
\begin{aligned}
\pi(y) &= \sum_x \pi(x) P^m(x,y) \;\; \forall m \\
\Rightarrow \;\; \pi(y) &= \frac{1}{n}\sum_{m=1}^{n}\sum_x \pi(x) P^m(x,y) = \sum_x \pi(x)\, \frac{1}{n}\sum_{m=1}^{n} P^m(x,y) \\
\Rightarrow \;\; \pi(y) &= \lim_{n\to\infty} \sum_x \pi(x)\, \frac{1}{n}\sum_{m=1}^{n} P^m(x,y)
\end{aligned}
$$

- The proof is complete if we can take the limit inside the sum

Bounded Convergence Theorem

- **Bounded Convergence Theorem**:
  1. Suppose $a(x) \geq 0, \ \forall x \in S$ and $\sum_x a(x) < \infty$.
  2. Let $b_n(x), \ x \in S$ be such that $|b_n(x)| \leq K, \ \forall x, n$ and

$$\lim_{n \to \infty} b_n(x) = b(x), \forall x \in S.$$

  Then

$$\lim_{n \to \infty} \sum_{x \in S} a(x) \ b_n(x) = \sum_{x \in S} a(x) \ \lim_{n \to \infty} b_n(x) = \sum_{x \in S} a(x) \ b(x)$$

- **Bounded Convergence Theorem**: Suppose $a(x) \geq 0, \ \forall x \in S$ and $\sum_x a(x) < \infty$. Let $b_n(x), \ x \in S$ be such that $|b_n(x)| \leq K, \ \forall x, n$ and suppose $\lim_{n\to\infty} b_n(x) = b(x), \forall x \in S$. Then

$$\lim_{n\to\infty} \sum_{x\in S} a(x) \ b_n(x) = \sum_{x\in S} a(x) \ \lim_{n\to\infty} b_n(x) = \sum_{x\in S} a(x) \ b(x)$$

- We had

$$\pi(y) = \lim_{n\to\infty} \sum_x \pi(x) \ \frac{1}{n} \sum_{m=1}^{n} P^m(x, y)$$

- We have

$$\pi(x) \geq 0; \ \ \sum_x \pi(x) = 1; \ \ 0 \leq \frac{1}{n} \sum_{m=1}^{n} P^m(x, y) \leq 1, \forall x$$

- Hence, if y is transient or null recurrent, then

$$\pi(y) = \sum_x \pi(x) \ \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x, y) = 0$$

- In any stationary distribution π, we would have $\pi(y) = 0$ is y is transient or null recurrent.
- Hence an irreducible transient or null recurrent chain would not have a stationary distribution.
- The null recurrent chain we considered earlier is an example of a Markov chain that does not have a stationary distribution.

- **Theorem** An irreducible positive recurrent chain has a unique stationary distribution given by
$$\pi(y) = \frac{1}{m_y}, \ \forall y \in S$$
- Suppose π is such that $\pi(y) = \sum_x \pi(x) P(x, y)$. Then

$$\begin{aligned}
\pi(y) &= \sum_x \pi(x)\, P^m(x,y), \ \forall m \\
\Rightarrow \quad \pi(y) &= \sum_x \pi(x)\, \frac{1}{n} \sum_{m=1}^{n} P^m(x,y), \ \forall n \\
\Rightarrow \quad \pi(y) &= \lim_{n\to\infty} \sum_x \pi(x)\, \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) \\
\Rightarrow \quad \pi(y) &= \sum_x \pi(x)\, \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P^m(x,y) \\
&= \sum_x \pi(x)\, \frac{1}{m_y} = \frac{1}{m_y}
\end{aligned}$$

- To complete the proof, we need to show $\sum_y \frac{1}{m_y} = 1$.
- We skip this step in the proof.
- The theorem shows that an irreducible positive recurrent chain has a unique stationary distribution
- Corollary: An irreducible finite chain has a unique stationary distribution
- Corollary: An irreducible chain has a stationary distribution if and only if it is positive recurrent

- If π^1 and π^2 are stationary distributions, then so is $\alpha\pi^1 + (1-\alpha)\pi^2$ (easily verified)
- Let C be a closed irreducible set of positive recurrent states.
  Then there is a unique stationary distribution π that satisfies $\pi(y) = 0, \ \forall y \notin C$.
- Any other stationary distribution of the chain is a convex combination of the stationary distributions concentrated on each of the closed irreducible sets of positive recurrent states.
- This answers all questions about existence and uniqueness of stationary distributions

- Consider an irreducible positive recurrent chain.
- It has a unique stationary distribution and $\frac{1}{n}\sum_{m=1}^{n} P^m(x,y)$ converges to $\pi(y)$.
- The next question is convergence of π_n

$$\lim_{n\to\infty} \pi_n(y) = \lim_{n\to\infty} \sum_x \pi_0(x)\ P^n(x,y) =?$$

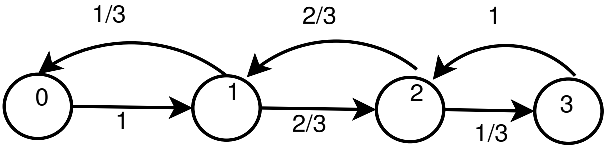
- If $P^n(x,y)$ converges to $g(y)$ then that would be the stationary distribution and π_n converges to it
- But, $\frac{1}{n}\sum_{m=1}^{n} a_m$ may have a limit though $\lim_{n\to\infty} a_n$ may not exist.
  For example, $a_n = (-1)^n$

- Consider a chain with transition probabilities

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- One can show $\pi^T = \left[\frac{1}{8}\ \frac{3}{8}\ \frac{3}{8}\ \frac{1}{8}\right]$
- However, P^n goes to different limits based on whether n is even or odd

- The chain is the folowing



$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 2/3 & 0 & 1/3 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad P^2 = \begin{bmatrix} 1/3 & 0 & 2/3 & 0 \\ 0 & 7/9 & 0 & 2/9 \\ 2/9 & 0 & 7/9 & 0 \\ 0 & 2/3 & 0 & 1/3 \end{bmatrix}$$

- We can return to a state only after even number of time steps
- That is why P^n does not go to a limit
- Such a chain is called a periodic chain

- We define period of a state x as

$$d_x = \gcd\{n \geq 1 \ : \ P^n(x,x) > 0\}$$

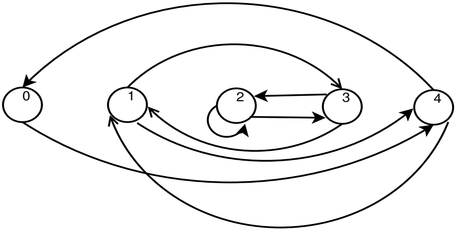
- If $P(x,x) > 0$ then $d_x = 1$
- If x leads to y and y leads to x, then $d_x = d_y$
- Let $P^{n_1}(x,y) > 0$, $P^{n_2}(y,x) > 0$.
  Then $P^{n_1+n_2}(x,x) > 0 \ \Rightarrow \ d_x$ divides $n_1 + n_2$.
- For any n s.t. $P^n(y,y) > 0$, we get $p^{n_1+n+n_2}(x,x) > 0$
- Hence, d_x divides n for all n s.t. $P^n(y,y) > 0 \Rightarrow d_x \leq d_y$
- Similarly, $d_y \leq d_x$ and hence $d_y = d_x$
- All states in an irreducible chain have the same period.
- If the period is 1 then chain is called aperiodic

- The extra condition we need for convergence of π_n is aperiodicity
- For an aperiodic, irreducible, positive recurrent chain, there is a unique stationary distribution and π_n converges to it irrespective of what π_0 is.
- An aperiodic, irreducible, positive recurrent chain is called an ergodic chain

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 to 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.

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

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

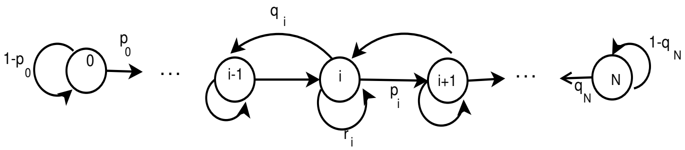
$$
\begin{aligned}
\pi(0) &= (1-p)\pi(4) \\
\pi(1) &= (1-p)\pi(3) + p\pi(4) \;\;\Rightarrow\;\; \pi(3) = \pi(1) \\
\pi(2) &= (1-p)\pi(2) + p\pi(3) \\
\pi(3) &= (1-p)\pi(1) + p\pi(2) \;\;\Rightarrow\;\; \pi(2) = \pi(1) \\
\pi(4) &= \pi(0) + p\pi(1) \;\;\Rightarrow\;\; \pi(4) = \pi(1)
\end{aligned}
$$

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

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

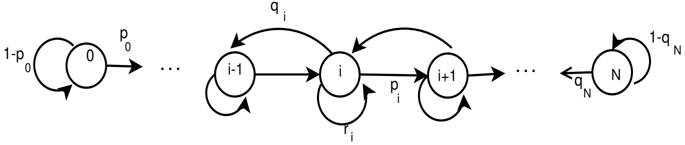
Birth-Death chains

- The following is a finite birth-death chain



- We assume $0 < p_i, q_i < 1, \forall i$.
- Then the chain is irreducible, positive recurrent
- It is also aperiodic
- We can derive a general form for its stationary probabilities

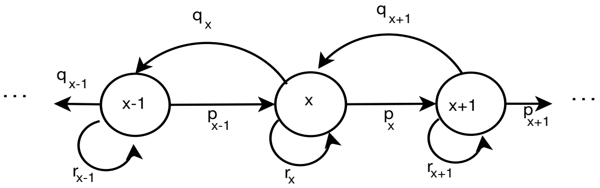
birth-death chains – stationary distribution



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

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

- For any x, the relevant part of the chain is



- We get

$$\pi(x) = \pi(x-1)p_{x-1} + \pi(x)(1 - p_x - q_x) + \pi(x+1)q_{x+1}$$

- This gives us the general recurrence

$$\pi(x+1)q_{x+1} - \pi(x)p_x = \pi(x)q_x - \pi(x-1)p_{x-1} = 0$$

- Thus we get

$$\pi(1)q_1 - \pi(0)p_0 \;=\; 0 \;\Rightarrow\; \pi(1) = \frac{p_0}{q_1}\,\pi(0)$$
$$\pi(2)q_2 - \pi(1)p_1 \;=\; 0 \;\Rightarrow\; \pi(2) = \frac{p_1}{q_2}\,\pi(1) = \frac{p_0 p_1}{q_1 q_2}\,\pi(0)$$

- Iterating like this, we get

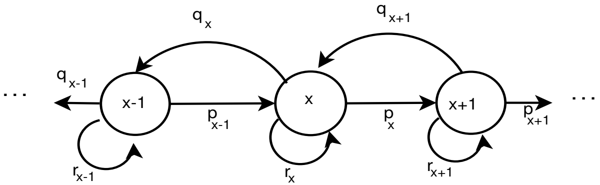
$$\pi(n) = \eta_n\,\pi(0), \quad \text{where } \; \eta_n = \frac{p_0 p_1 \cdots p_{n-1}}{q_1 q_2 \cdots q_n}, \; n = 1, 2, \cdots, N$$

- 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)

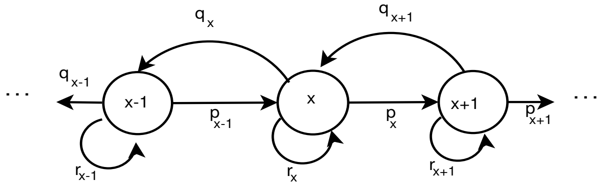
- Consider a birth-death chain



- 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



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

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

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

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

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

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

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

- Using these, we get

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

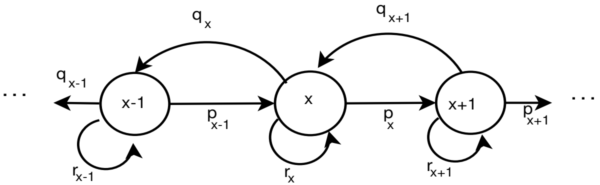
- Putting $x = b-1,\ b-2, \cdots, y$ in the above

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

- 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 $a < y < b$,

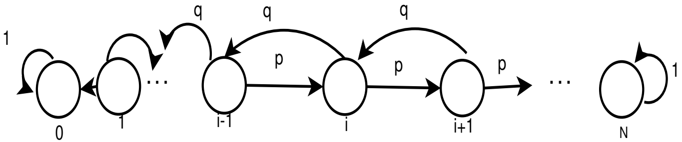
$$U(y) = P_y[T_a < T_b] = \frac{\sum_{x=y}^{b-1} \gamma_x}{\sum_{x=a}^{b-1} \gamma_x}, \quad \gamma_x = \frac{q_x q_{x-1} \cdots q_{a+1}}{p_x p_{x-1} \cdots p_{a+1}}$$

- 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