Aula 10

Roteiro

- Distribuição estacionária
- Tempo de chegada
- Distância de variação total
- Convergência
- Reversibilidade
- Passeios aleatórios
- Nascimento e morte

Cadeia de Markov

- S : espaço de estados da CM
- P : matriz de transição de estados
- X_t : v.a. que determina o estado da cadeia no instante de tempo t, para $t = 0,1,2,\ldots$
- Estamos interessados em
  - $P[\, X_t = s \,] = \pi_s(t)$
- Sabemos que para um estado incial, temos

$$\pi(t) = \pi(t-1)P = \pi(0)P^t$$

Mas para onde vai π(t)?

  - será que depende em $\pi(0)$?

Possível Convergência

- Observação 1: X_t não converge
  - X_t passeia pela cadeia para sempre, trocando de estado
- Observação 2: P[$X_t = s_i$] pode convergir
  - Prob. de encontrar X_t no estado s_i ou fração de vezes que X_t é igual ao estado s_i (intuitivamente)
- Estamos interessados em valores grandes de t

$$\lim_{t\to\infty}\pi(t)=\lim_{t\to\infty}\pi(0)P^t$$

  - mas temos que formalizar este limite, pois π é um vetor

Distribuição Estacionária

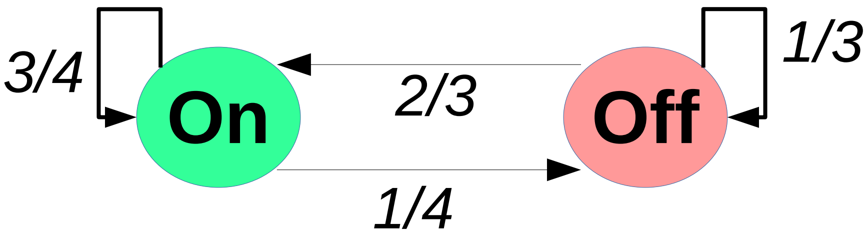
- π : um vetor de distribuição em uma CM com matriz de transição *P*
- Dizemos que π (vetor linha) é uma distribuição estacionária sse

$$\pi_s \geq 0 \qquad \sum_{s \in S} \pi_s = 1 \longleftarrow \text{Vetor de probabilidade}$$

$$\pi P = \pi \longrightarrow \pi_i = \sum_j \pi_j P_{ji} \quad \text{, para todo } i$$

- Ou seja, ao multiplicar π por *P* temos π de volta
  - “estacionou”, não temos mais evolução
- Repare que se $\pi(0)$ é uma distribuição estacionária, então $\pi(1) = \pi(0)P = \pi(0)$

Exemplo Modelo On-Off



$$P=\begin{pmatrix} .75 & .25 \\ .67 & .33 \end{pmatrix}$$

$$\pi=\begin{pmatrix} 8/11 & 3/11 \end{pmatrix}$$ ← Distribuição estacionária

- Verificando: πP = ?

$$\pi_1=8/11*3/4+3/11*2/3=8/11$$

$$\pi_2=8/11*1/4+3/11*1/3=3/11$$

- Se $\pi(0) = \pi$ então temos que $\pi(1) = \pi(0)P = \pi$
  - vetor de probabilidade está estacionado!

Tempo de Chegada

- T_{ij}: Tempo (em transições) necessário para sair de um estado s_i e chegar a outro estado s_j (*hitting time*)
  - número de transições T_{ij} é aleatório, $\tau_{ij} = E[\, T_{ij} \,]$

$$T_{ij} = min\{\, t \mid X_t = s_j \wedge X_0 = s_i \,\}$$

- Se $i = j$, temos τ_{ii}, chamado de tempo médio de retorno ao estado s_i
  - número médio de transições para sair e voltar ao estado s_i
- T_{ij} não depende de $\pi(t)$ pois estamos condicionando em estar em s_i

Tempo de Chegada

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer dois estados s_i e s_j, temos o seguinte:

$$P[T_{ij} < \infty] = 1$$

- Probabilidade de T_{ij} ser infinito é zero
- Não há chances de sair de s_i ficar circulando pela cadeia e nunca chegar a s_j

$$E[T_{ij}] = \tau_{ij} < \infty$$

- Valor esperado do tempo de retorno para qualquer estado s_i é finito
- Boa notícia!

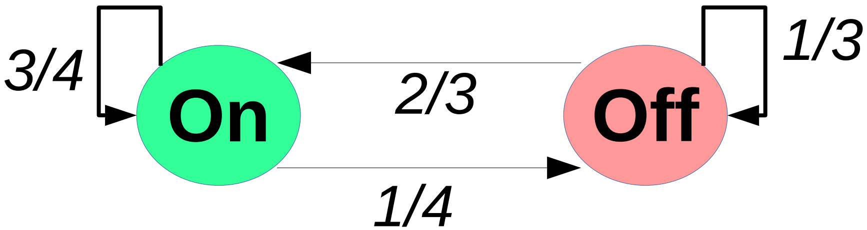
Distribuição Estacionária e Tempo de Retorno

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer estado s_i, temos a seguinte relação

$$\pi_i = \frac{1}{\tau_{ii}}$$

- Relação entre tempo médio de retorno e distribuição estacionária
- Conhecer um determina o outro!
- **Intuição:** na média a dinâmica visita s_i 1 vez a cada τ_{ii} passos

Exemplo



$$P = \begin{pmatrix} .75 & .25 \\ .67 & .33 \end{pmatrix}$$

$$\pi = \begin{pmatrix} 8/11 & 3/11 \end{pmatrix}$$

Distribuição estacionária

- Tempo médio de retorno do estado On (1) e Off (2) ?
- $\tau_{11} = 1/\pi_1 = 11/8 = 1.375$
- $\tau_{22} = 1/\pi_2 = 11/3 = 3.666$

Variação Total

- Como medir a distância entre dois vetores?
  - Precisamos disto para medir aproximação da distribuição estacionária
  - há muitas maneiras, uma delas é variação total
- Seja α e β dois vetores de probabilidade em *S*, a distância de variação total entre eles é dada por

$$d_{TV}(\alpha, \beta) = \frac{1}{2} \sum_k |\alpha_k - \beta_k|$$

d_{TV} tem valor entre 0 e 1

- Se α = β então $d_{TV}(\alpha_t, \beta) = 0$
  - e vice-versa

Exemplo e Convergência

- Vetores de probabilidade de mesma dimensão, n=5
- α = (0.1 0.2 0.3 0.3 0.1)
- β = (0.2 0.1 0.1 0.5 0.1)

$$d_{TV}(\alpha,\beta)=\frac{1}{2}\sum_k |\alpha_k-\beta_k|=\frac{1}{2}(0.1+0.1+0.2+0.2+0)=0.3$$

- Considere sequência de vetores α_1, α_2, ...
- Sequência de vetores α_t converge para β em variação total sse

$$\lim_{t\to\infty} d_{TV}(\alpha_t,\beta)=0$$

Convergência em CM

- Teorema: Para qualquer CM irredutível e aperiódica, para qualquer condição inicial $\pi(0)$, temos que

$$\lim_{t \to \infty} d_{TV}(\pi(t), \pi) = 0$$

π é uma distribuição estacionária da CM

- Além disso, π é única!
  - CM sempre converge para a mesma distribuição estacionária, independente da condição inicial
  - Também chamado de estado estacionário, ou equilíbrio da CM

Encontrando π

- Ótima notícia, mas como encontrar π ?

1) Método iterativo

$$\pi(t)=\pi(t-1)P=\pi(0)P^t$$

  - fazer iteração até critério de convergência

2) Método direto

$$\pi=\pi P$$

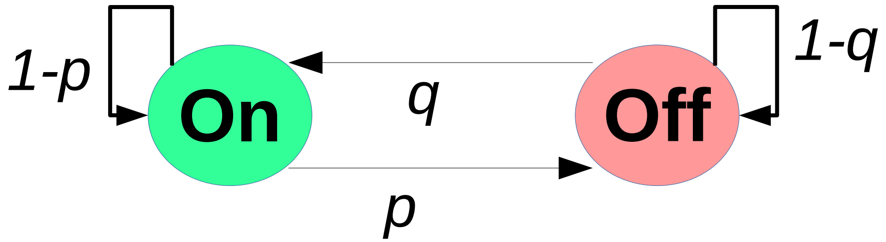
  - resolver sistema de equações, adicionando a equação

$$\sum_i \pi_i=1$$

3) Monte Carlo

  - Usar própria cadeia para gerar amostras para estimar π_i ou estimar τ_{ii} para todo s_i

Modelo On-Off



- Distribuição estacionária?
- Sistemas de equações

$$\pi_i = \sum_j \pi_j P_{ji} \qquad \sum_i \pi_i = 1$$

$$\pi_1 = \pi_1 P_{11} + \pi_2 P_{21} = (1-p)\pi_1 + q\pi_2$$
$$\pi_2 = \pi_1 P_{12} + \pi_2 P_{22} = p\pi_1 + (1-q)\pi_2$$
$$\pi_1 + \pi_2 = 1$$

- Substituindo $\pi_2 = 1 - \pi_1$ na primeira equação, temos

$$\pi_1 = (1-p)\pi_1 + q(1-\pi_1) \rightarrow \pi_1 = \frac{q}{p+q}$$

$$e \quad \pi_2 = \frac{p}{p+q}$$

Reversibilidade

- Uma classe bem especial de CM
  - usada por muitos algoritmos, incluindo MCMC
- Reversível no tempo: evolução de X_t é o mesmo se t vai para frente ou para trás
- Uma CM é dita reversível para a distribuição de probabilidade π sse

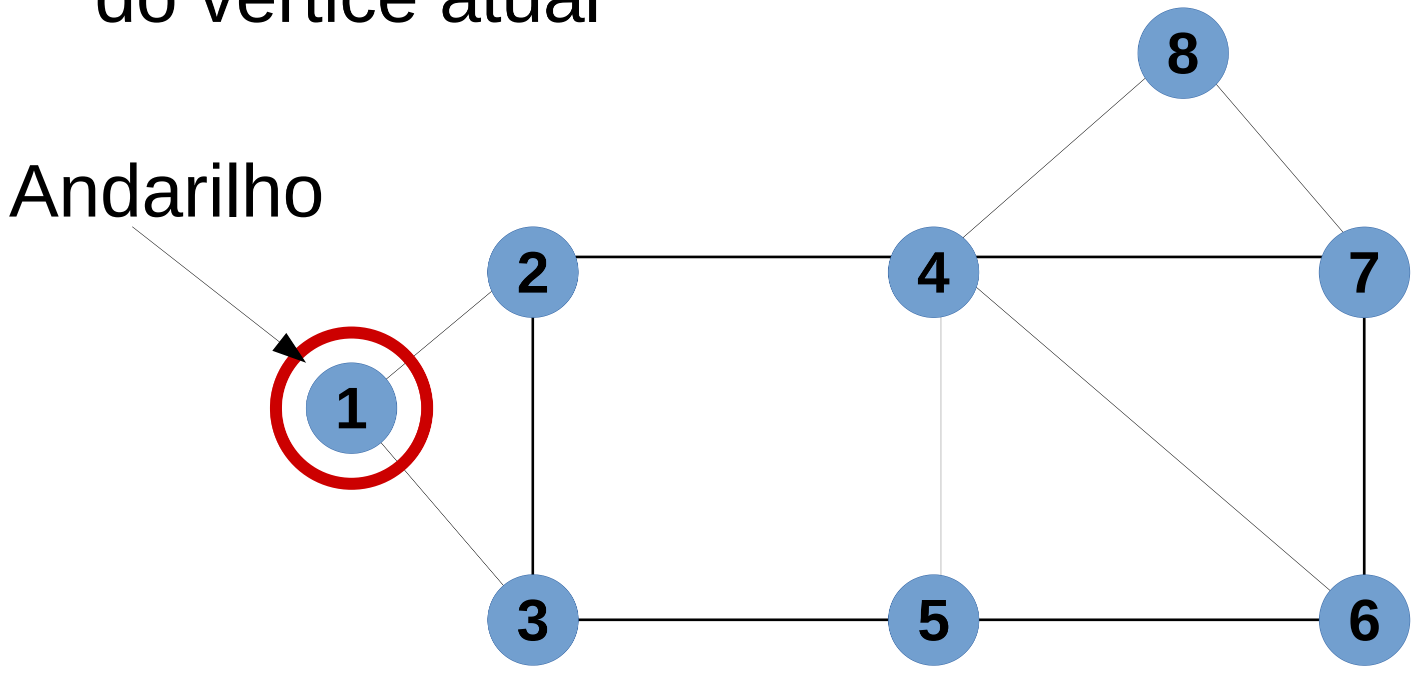
$$\pi_i P_{ij} = \pi_j P_{ji}$$

Massa de prob fluindo de s_i para s_j = Massa de prob fluindo de s_j para s_i

- Se π existe então π é também distribuição estacionária
- Uma forma ainda mais forte de equilíbrio!

Passeios Aleatórios

- Seja G=(V, E) um grafo não direcionado
- Considere um andarilho que passeia pelo grafo de forma aleatória, sem preferência e sem memória
  - escolhe uniformemente próximo vértice entre vizinhos do vértice atual



- $p_{12} = ½$, $p_{13} = ½$, $p_{21} = 1/3$, $p_{23} = 1/3$, $p_{24} = 1/3$, ...

Passeios Aleatórios

- Andarilho induz uma cadeia de Markov sobre o grafo G
  - forma de colocar “movimento” no grafo
- X_t : vértice onde andarilho se encontra no tempo t
- Matriz de transição de probabilidade

$$P = D^{-1} A$$

  - D^{-1} é a matriz diagonal com o inverso do grau de cada vértice
  - A é a matriz de adjacência de G
    - A[i,j] = 1 sse i e j são vizinhos
- Outra forma: $P_{ij} = 1/d_i$ se (i, j) são vizinhos, 0 c.c.
  - onde d_i é o grau do vértice i

Passeios Aleatórios

- Qual a distribuição estacionária do andarilho ?

$$\pi_i = \frac{d_i}{K} \qquad K = \sum_i d_i = 2m \longleftarrow$$

m é o número de arestas em G

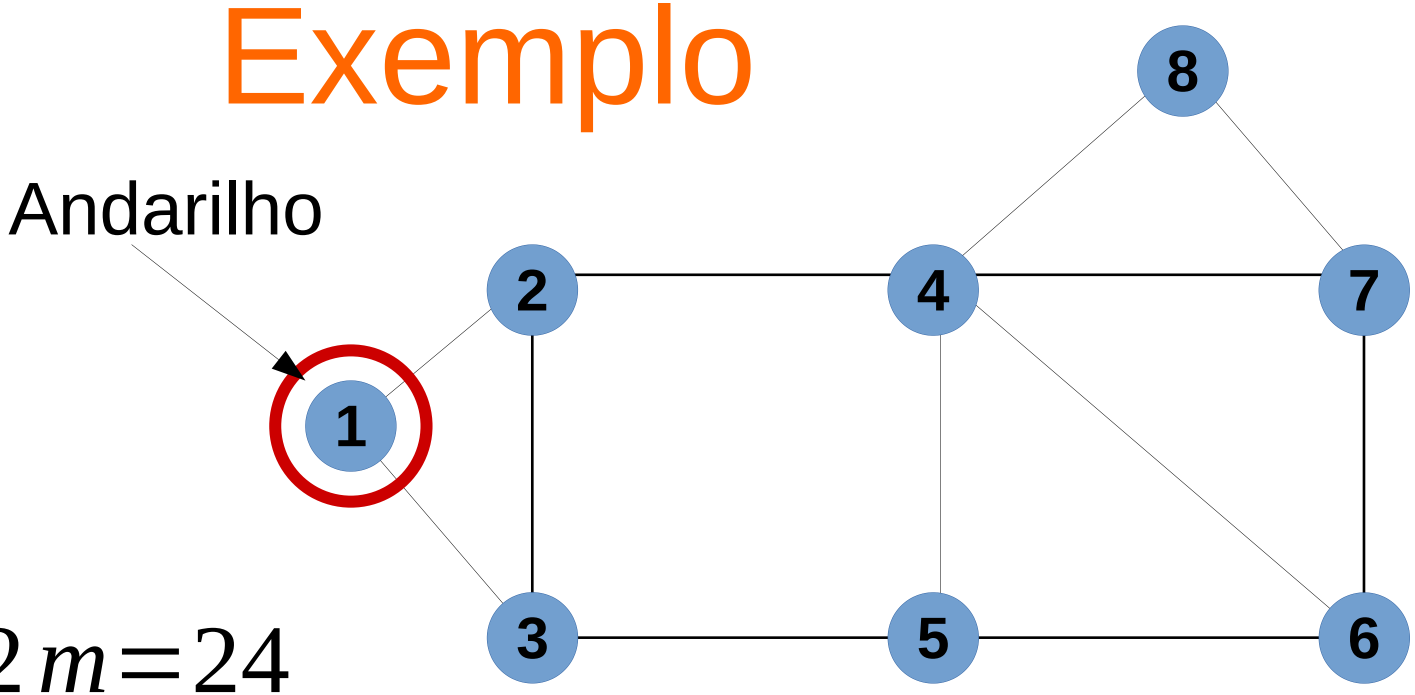
- **Incrível:** só depende do grau do vértice *i*
  - *K* é constante de normalização
  - vértices com mesmo grau tem a mesma probabilidade
- Além disso, CM é reversível

$$\pi_i P_{ij} = \pi_j P_{ji} \longleftarrow$$

Podemos verificar

  - tanto faz andarilho andar para frente ou para trás
  - equilíbrio mais forte para esta CM

Exemplo



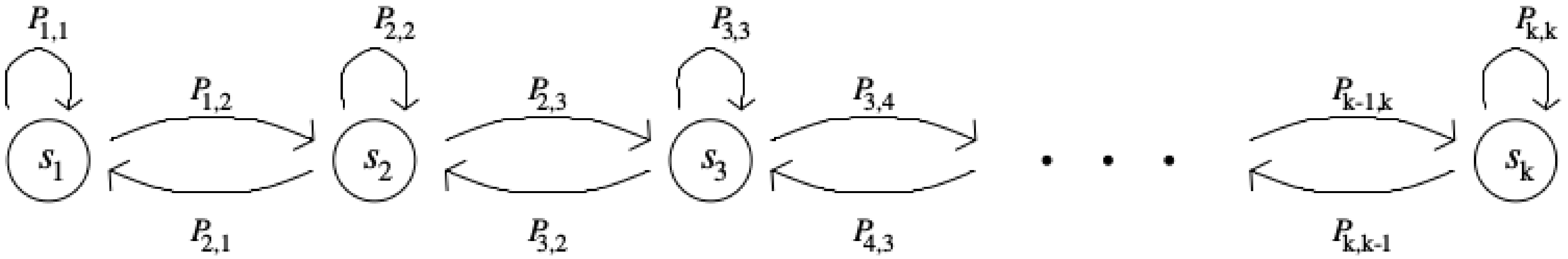
$$K = \sum_i d_i = 2m = 24$$

$$\pi = \left(\frac{2}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{5}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{3}{24} \quad \frac{2}{24} \right)$$

- Independente de onde andarilho inicia, distribuição de sua posição converge para π
  - prob de encontrar o andarilho no vértice *i*

Processo de Nascimento e Morte (*Birth-Death*)

- Generalização do modelo On-Off para mais estados
  - CM muito utilizada, várias aplicações práticas
- K estados: 1, …, K
- Transições apenas entre estados vizinhos (com possível *loops*)
  - matriz *P* é tridiagonal

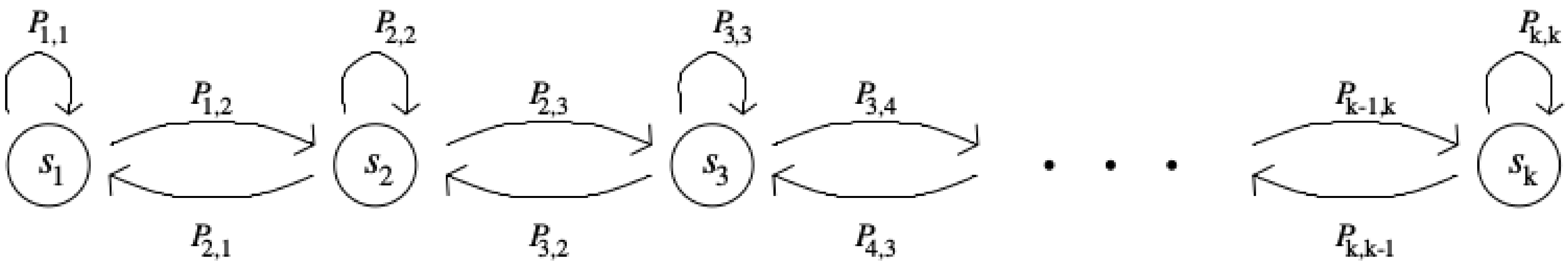


- Em geral, $p_{i,i+1}$ e $p_{i+1,i}$ possuem alguma lei de formação
  - modelo geralmente é bem regular

Figueiredo 2021

Processo de Nascimento e Morte (*Birth-Death*)

- Qual é a distribuição estacionária?



- Assumindo que CM é reversível, temos $\pi_i P_{ij} = \pi_j P_{ji}$
- Vamos definir π como possível candidato, e assumir algum valor para π_1. Temos então

$$\pi_2 = \frac{\pi_1 P_{12}}{P_{21}} \qquad \pi_3 = \frac{\pi_2 P_{23}}{P_{32}} = \frac{\pi_1 P_{12} P_{23}}{P_{21} P_{32}} \qquad \cdots \qquad \pi_i = \frac{\pi_1 \prod_{j=1}^{i-1} P_{j,j+1}}{\prod_{j=1}^{i-1} P_{j+1,j}}$$

Processo de Nascimento e Morte (*Birth-Death*)

- Mas quanto vale π_1?

$$1=\sum_{i=1}^{K}\pi_i=\sum_{i=1}^{K}\frac{\pi_1\prod_{j=1}^{i-1}P_{j,j+1}}{\prod_{j=1}^{i-1}P_{j+1,j}}=\pi_1\sum_{i=1}^{K}\frac{\prod_{j=1}^{i-1}P_{j,j+1}}{\prod_{j=1}^{i-1}P_{j+1,j}}$$

- Logo,

$$\pi_1=\left(\sum_{i=1}^{K}\frac{\prod_{j=1}^{i-1}P_{j,j+1}}{\prod_{j=1}^{i-1}P_{j+1,j}}\right)^{-1}$$

- Modelo BD é reversível, e π é sua distrib estacionária

Exemplo da Fila

- Considere uma fila onde a cada instante tempo chega ou sai um objeto da fila
  - chegada com prob *p*, saída com prob *q*, p+q < 1
  - fila tem capacidade para armazenar K objetos
- Modelo BD com K+1 estados (fila vazia + K)
  - $P_{i,i+1}$ = p , i=0,…., K-1 ; $P_{i+1,i}$ = q i=1,…., K, ; $P_{i,i}$ = 1-p-q
- Estado estacionário?

$$\pi_i = \frac{\pi_0 \prod_{j=1}^{i} P_{j,j+1}}{\prod_{j=1}^{i} P_{j+1,j}} = \pi_0 \left(\frac{p}{q}\right)^i \qquad \pi_0 = \left(\sum_{i=0}^{K} \left(\frac{p}{q}\right)^i\right)^{-1} = \frac{1-p/q}{1-(p/q)^{K+1}}$$

, para $p < q$

Exemplo da Fila

- K=10, p=0.3 (chegada), q=0.4 (saída)
- Aplicando as fórmulas temos

$$\pi_0 = \frac{1-0.3/0.4}{1-(0.3/0.4)^{10+1}} = 0.261 \qquad \pi_i = \pi_0 \left(\frac{0.3}{0.4}\right)^i$$

- Probabilidade da fila estar vazia (em estado estacionário)?

$$\pi_0 = 0.261$$

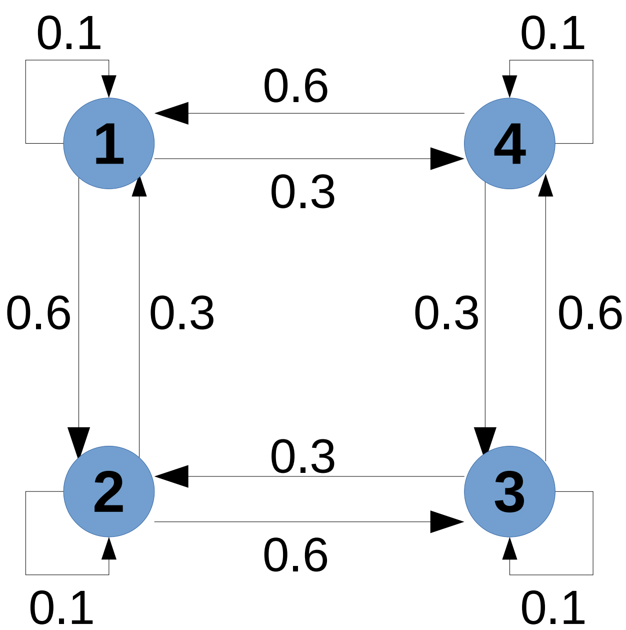
- Probabilidade da fila estar cheia (em estado estacionário)?

$$\pi_K = 0.015$$

- Valor esperado do tamanho da fila (em estado estacionário)?

$$\sum_{i=0}^{K} i\pi_i = \ldots$$

Outro Exemplo



- CM é irredutível? **Sim!**
- CM é aperiódica? **Sim!**
- Qual é sua distribuição estacionária?
  - por inspeção, sem fazer conta

  π_i = ¼ para *i* = 1, 2, 3, 4
- Podemos verificar facilmente!

- CM é reversível? **Não!**

$$\pi_1 p_{12} = \frac{1}{4}\frac{6}{10} = \frac{3}{20} \neq \pi_2 p_{21} = \frac{1}{4}\frac{3}{10} = \frac{3}{40}$$

- **Intuição:** processo gira mais em uma direção (anti-horário)
  - direção do tempo (para frente ou para trás) determina direção mais girada, não sendo reversível