

# Redes complexas – MO804 (MC908)

## Homofilia

Prof. Dr. Ruben Interian

Instituto de Computação, UNICAMP

# Resumo

- 1 Objetivo
- 2 Homofilia
  - Homofilia em atributos categóricos
  - Homofilia em atributos numéricos
- 3 Aplicações e mecanismos subjacentes

## 1 Objetivo

## 2 Homofilia

- Homofilia em atributos categóricos
- Homofilia em atributos numéricos

## Objetivo

- Estudar o conceito de **homofilia** em redes complexas.
- Estudar as medidas de homofilia, suas causas e consequências para a estrutura da rede.

# Resumo

## 1 Objetivo

## 2 Homofilia

- Homofilia em atributos categóricos
- Homofilia em atributos numéricos

# Homofilia

**Homofilia** – tendência dos nós da rede de ter vínculos com aqueles que são semelhantes a eles mesmos.

# Homofilia

**Homofilia** – tendência dos nós da rede de ter vínculos com aqueles que são semelhantes a eles mesmos.

O que é **semelhante**? – Depende do tipo de rede complexa!

# Homofilia

**Homofilia** – tendência dos nós da rede de ter vínculos com aqueles que são semelhantes a eles mesmos.

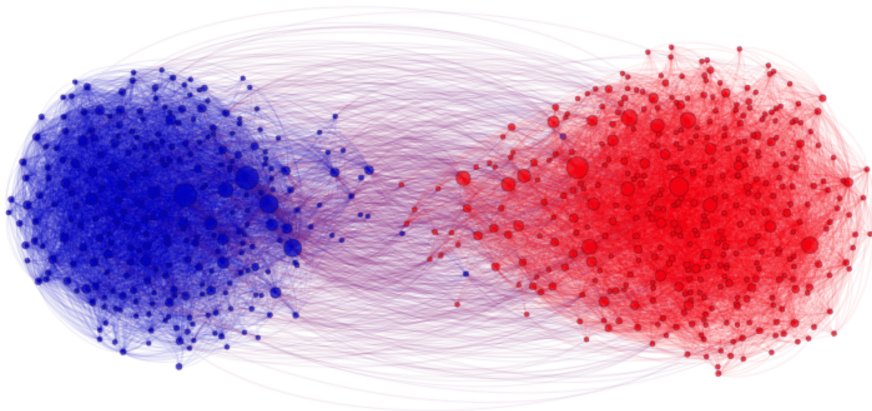
O que é **semelhante**? – Depende do tipo de rede complexa!

O conceito de homofilia surgiu na sociologia, e tem uma longa **história**:

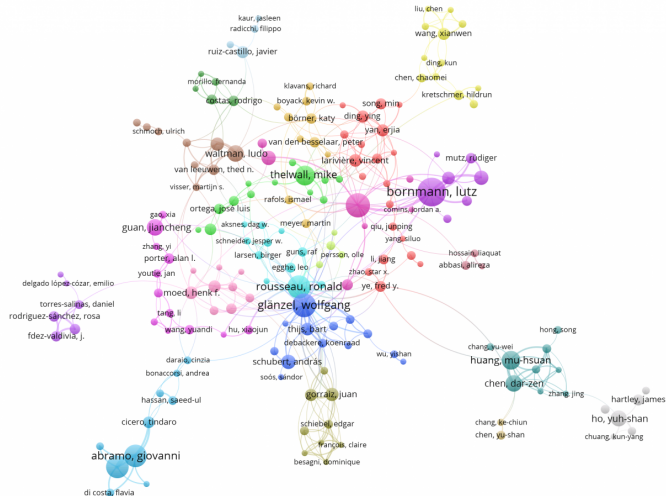
- **Platão:** *Similaridade leva a amizade.*
- **Aristóteles:** *Pessoas amam àqueles que são como eles.*



## Existe homofilia na rede de sites e blogs políticos?



## Existe homofilia na rede de colaborações ou citações entre cientistas?



## A complex network graph visualization showing a large number of nodes (approximately 100) connected by numerous edges. The nodes are colored red, green, and blue, suggesting different categories or clusters. The graph is highly interconnected, with many edges forming a dense web. The nodes are distributed across the frame, with a higher concentration of red nodes on the left and blue nodes on the right. Green nodes are scattered throughout the network. The edges are thin black lines connecting the nodes, creating a complex, tangled pattern.

## Existe homofilia na rede de amizades?



Seus amigos não são pessoas aleatórias! São **similares** a você em diversos aspectos:

- Idade, gênero, origem, cidade, time de futebol favorito, academia;
- Interesses, crenças, opiniões e muitas outras características.

Seus amigos não são pessoas aleatórias! São **similares** a você em diversos aspectos:

- Idade, gênero, origem, cidade, time de futebol favorito, academia;
- Interesses, crenças, opiniões e muitas outras características.

Os adjacentes de um vértice  $v$  não são vértices aleatórios! São **similares** a  $v$  em diversos aspectos.

# Homofilia

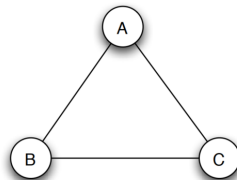
A homofilia é um dos **princípios universais** que caracterizam os vínculos em redes complexas.

## 17 / 95

# Homofilia

A **homofilia** é **fundamental** para explicar o surgimento de **novos vínculos**:

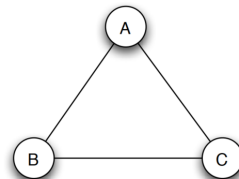
- Se duas pessoas se conhecem através de um amigo comum, podemos argumentar que está agindo o **fecho triádico**: um princípio intrínseco **da própria rede**.
- O **fecho triádico não consegue** explicar tudo: por exemplo, duas pessoas que se conhecem porque **compartilham interesses** ou crenças.



# Homofilia

A **homofilia** é **fundamental** para explicar o surgimento de novos vínculos:

- O próprio **fecho triádico** pode ser visto como um tipo específico de **homofilia**: o vínculo entre duas pessoas aparece pois eles **compartilham algo** (um amigo).
- O **fecho triádico** e **homofilia** podem agir juntos:  $B$  e  $C$  têm mais chance de se conhecer, mas se as arestas  $(A, B)$  e  $(A, C)$  existem,  $B$  e  $C$  podem ser *parecidos*.





# Homofilia

Já sabemos como avaliar o **fecho triádico**: usando os coeficientes de clusterização **local** (para um nó) e **global** (para toda a rede).

## Como medir a homofilia?

- Como **avaliar** em que medida uma rede apresenta comportamento homofílico?

# Homofilia

Já sabemos como avaliar o **fecho triádico**: usando os coeficientes de clusterização **local** (para um nó) e **global** (para toda a rede).

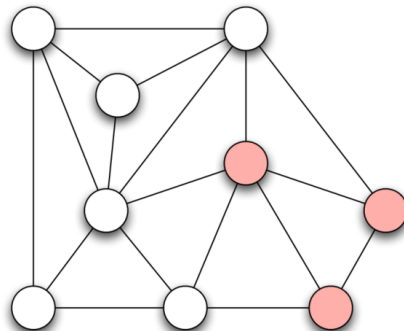
## Como medir a homofilia?

- Como **avaliar** em que medida uma rede apresenta comportamento homofílico?
- Para **medir a homofilia**, vamos supor **inicialmente** que existem grupos de nós que são semelhantes entre si (grupos com “algo em comum”).

# Homofilia

Seja  $G = (V, E)$  um grafo, onde  $V = \{v_1, \dots, v_n\}$  é o conjunto de vértices ou nós, e  $E \subseteq V \times V$  é o conjunto de arestas.

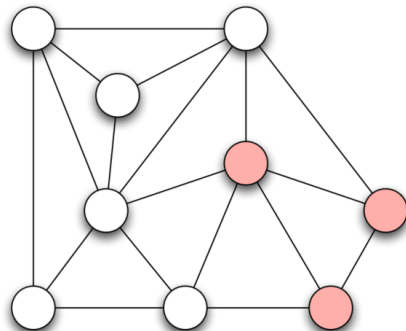
- Seja  $A = \{A_1, \dots, A_k\}$  uma coleção de  $k$  grupos de nós:  $A_i \subseteq V$  para  $i = 1, \dots, k$ .



# Homofilia

Seja  $G = (V, E)$  um grafo, onde  $V = \{v_1, \dots, v_n\}$  é o conjunto de vértices ou nós, e  $E \subseteq V \times V$  é o conjunto de arestas.

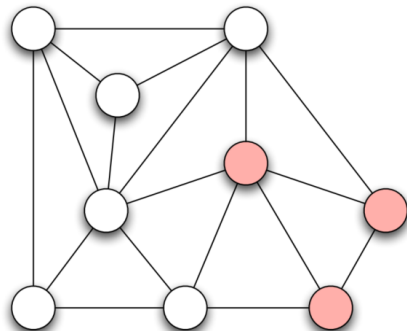
- Seja  $A = \{A_1, \dots, A_k\}$  uma coleção de  $k$  grupos de nós:  $A_i \subseteq V$  para  $i = 1, \dots, k$ .
- Os grupos  $A = \{A_1, \dots, A_k\}$  formam uma partição do conjunto  $V$ ?



# Homofilia

Seja  $G = (V, E)$  um grafo, onde  $V = \{v_1, \dots, v_n\}$  é o conjunto de vértices ou nós, e  $E \subseteq V \times V$  é o conjunto de arestas.

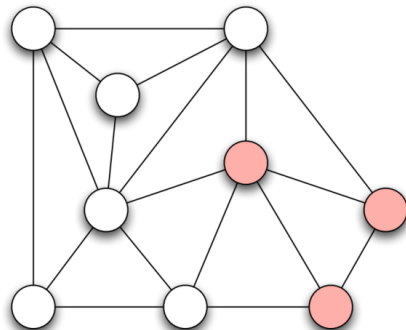
- Seja  $A = \{A_1, \dots, A_k\}$  uma coleção de  $k$  grupos de nós:  $A_i \subseteq V$  para  $i = 1, \dots, k$ .
- Os grupos  $A = \{A_1, \dots, A_k\}$  formam uma partição do conjunto  $V$ ?
  - **Caso sim**, cada nó pertence a exatamente um grupo. Analisamos este caso mais simples.
  - **Caso não**: alguns nós podem pertencer a mais de um grupo.



# Homofilia

**Primeiras ideias:** A homofilia de um nó  $v$  pode ser definida como a razão entre o número de vizinhos de  $v$  que estão no seu grupo  $A_i$ , e o grau de  $v$ :

$$h(v) = \frac{d_i(v)}{d(v)}$$

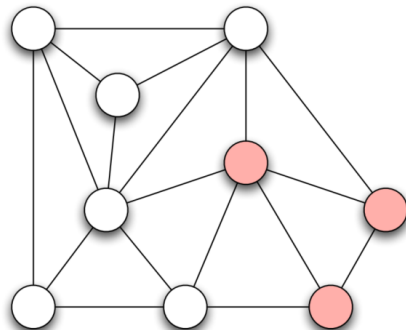


# Homofilia

**Primeiras ideias:** A homofilia de um nó  $v$  pode ser definida como a razão entre o número de vizinhos de  $v$  que estão no seu grupo  $A_i$ , e o grau de  $v$ :

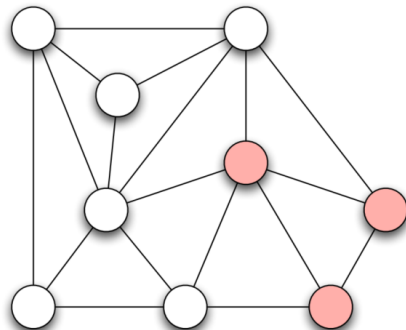
$$h(v) = \frac{d_i(v)}{d(v)}$$

- **Exemplo:**  $d_i(v) = 3$ ,  $d(v) = 5 \Rightarrow h(v) = 0.6$ .  
O valor  $h(v)$  está sempre no intervalo  $[0, 1]$ .



# Homofilia

Se o pertencimento dos nós aos grupos **não influencia nas arestas** de cada nó, então qual seria o valor **esperado** de  $h(v)$ ?

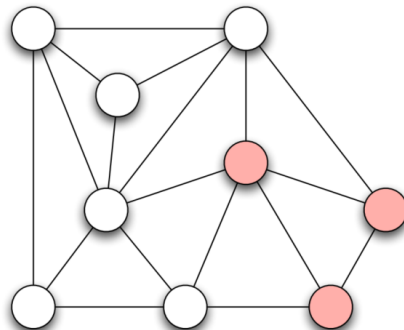


# Homofilia

Se o pertencimento dos nós aos grupos **não influencia nas arestas** de cada nó, então qual seria o valor **esperado** de  $h(v)$ ?

**Resposta:**

$$h(v) \approx \frac{|A_i|}{|V|} = w_i.$$



# Homofilia

**Problema:** Avaliar em que medida o nó  $v$  apresenta homofilia em suas conexões.

**Pergunta:** Qual é o valor a partir do qual  $h(v)$  teria um valor atipicamente alto?

# Homofilia

**Problema:** Avaliar em que medida o nó  $v$  apresenta homofilia em suas conexões.

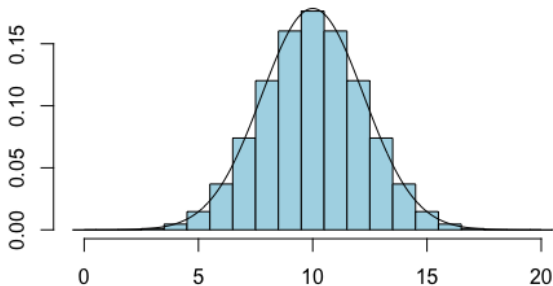
**Pergunta:** Qual é o valor a partir do qual  $h(v)$  teria um valor atipicamente alto?

**Possível solução:** analisar a **significância estatística** dos valores de  $d_i(v)$  e  $d(v)$ .

**Procedimento:**

- Fixar um **nível de significância**  $\alpha$  (por exemplo,  $\alpha = 5\%$ ).
- Fazer um teste estatístico sob a **hipótese nula** = não há homofilia ( $\frac{d_i(v)}{d(v)} \approx w_i$ ).
- O teste estatístico (binomial) calcula a chance (**p-valor**) de observar, do total de  $d(v)$  vizinhos, o número de vizinhos do mesmo grupo que seja  $\geq d_i(v)$ .
- **Comportamento homofílico:** se o **p-valor** é menor do que  $\alpha$ .

# Homofilia



$$Pr(d_i = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

# Homofilia

**Exemplo.** Vamos supor que temos uma rede grande na qual a metade (50%) dos nós é vermelha, e a metade dos nós é azul. Seja um vértice **v** que é vermelho.

# Homofilia

**Exemplo.** Vamos supor que temos uma rede grande na qual a metade (50%) dos nós é vermelha, e a metade dos nós é azul. Seja um vértice  $v$  que é vermelho.

- Se  $v$  tem 3 vizinhos, dos quais 2 são vermelhos, então se não há homofilia, a chance de termos 2 ou mais (2 ou 3) vermelhos é de  $0.5 = 50\%$ .  
Podemos concluir que de fato não temos homofilia para esse nó.

# Homofilia

**Exemplo.** Vamos supor que temos uma rede grande na qual a metade (50%) dos nós é vermelha, e a metade dos nós é azul. Seja um vértice  $v$  que é vermelho.

- Se  $v$  tem 3 vizinhos, dos quais 2 são vermelhos, então se não há homofilia, a chance de termos 2 ou mais (2 ou 3) vermelhos é de  $0.5 = 50\%$ .  
Podemos concluir que de fato não temos homofilia para esse nó.
- Se  $v$  tem 30 vizinhos, dos quais 20 são vermelhos, então se não há homofilia, a chance de termos 20 ou mais vermelhos é de  $0.04937 = 4.937\%$ .  
Com um nível de significância de  $\alpha = 5\%$ , podemos rejeitar a hipótese que não há homofilia, e concluir que de fato há homofilia para  $v$ .

# Homofilia

**Exemplo.** Vamos supor que temos uma rede grande na qual a metade (50%) dos nós é vermelha, e a metade dos nós é azul. Seja um vértice  $v$  que é vermelho.

- Se  $v$  tem 3 vizinhos, dos quais 2 são vermelhos, então se não há homofilia, a chance de termos 2 ou mais (2 ou 3) vermelhos é de  $0.5 = 50\%$ .  
Podemos concluir que de fato não temos homofilia para esse nó.
- Se  $v$  tem 30 vizinhos, dos quais 20 são vermelhos, então se não há homofilia, a chance de termos 20 ou mais vermelhos é de  $0.04937 = 4.937\%$ .  
Com um nível de significância de  $\alpha = 5\%$ , podemos rejeitar a hipótese que não há homofilia, e concluir que de fato há homofilia para  $v$ .

Nesses exemplos,  $h(v) = \frac{2}{3}$  nos dois casos!

Mas no primeiro caso, **não há informação suficiente** para dizer que o valor de  $h(v)$  é atípico, e no segundo exemplo há informação suficiente para isso.

# Homofilia

## Em resumo:

Um nó apresenta **homofilia**, quando o valor  $h(v) = \frac{d_i(v)}{d(v)}$  é significativamente maior do que o **esperado** para aquele nó:  $\frac{|A_i|}{|V|}$ .

Um nó apresenta **heterofilia**, quando o valor  $h(v) = \frac{d_i(v)}{d(v)}$  é significativamente menor do que o **esperado** para aquele nó:  $\frac{|A_i|}{|V|}$ .

A avaliação do que é **significativamente maior ou menor** pode ser feita usando testes estatísticos.

# Homofilia de uma rede

## Homofilia de uma rede:

- Avaliar quão longe está o **número de arestas entre vértices de diferentes grupos** do valor esperado.

# Homofilia de uma rede

**(1) Quão longe está a quantidade de arestas entre vértices de diferentes grupos do seu valor esperado?**

- Vamos supor que temos uma rede com dois grupos,  $X$  e  $Y$ , onde:

$$\frac{|X|}{|V|} = w_x, \quad \frac{|Y|}{|V|} = w_y.$$

# Homofilia de uma rede

**(1) Quão longe está a quantidade de arestas entre vértices de diferentes grupos do seu valor esperado?**

- Vamos supor que temos uma rede com dois grupos,  $X$  e  $Y$ , onde:

$$\frac{|X|}{|V|} = w_x, \quad \frac{|Y|}{|V|} = w_y.$$

- Se uma aresta é colocada aleatoriamente no grafo, então a chance de:
  - Escolher dois vértices de  $X$  é  $w_x^2$ .
  - Escolher dois vértices de  $Y$  é  $w_y^2$ .
  - Escolher vértices de grupos diferentes é  $2w_x w_y$ .

# Homofilia de uma rede

**(1) Quão longe está a quantidade de arestas entre vértices de diferentes grupos do seu valor esperado?**

- Vamos supor que temos uma rede com dois grupos,  $X$  e  $Y$ , onde:

$$\frac{|X|}{|V|} = w_x, \quad \frac{|Y|}{|V|} = w_y.$$

- Se uma aresta é colocada aleatoriamente no grafo, então a chance de:
  - Escolher dois vértices de  $X$  é  $w_x^2$ .
  - Escolher dois vértices de  $Y$  é  $w_y^2$ .
  - Escolher vértices de grupos diferentes é  $2w_x w_y$ .
- Veja que  $(w_x + w_y)^2 = w_x^2 + 2w_x w_y + w_y^2 = 1$ .

# Homofilia de uma rede

Se o pertencimento dos nós aos grupos **não influencia nas arestas**, então a proporção **esperada** de arestas entre nós de grupos diferentes (arestas inter-grupo) é  $2w_x w_y$ .

# Homofilia de uma rede

Se o pertencimento dos nós aos grupos **não influencia nas arestas**, então a proporção **esperada** de arestas entre nós de grupos diferentes (arestas inter-grupo) é  $2w_x w_y$ .

Se a proporção de arestas inter-grupo é significativamente menor do que  $2w_x w_y$ , então há evidência de homofilia.

# Homofilia de uma rede

Se o pertencimento dos nós aos grupos **não influencia nas arestas**, então a proporção **esperada** de arestas entre nós de grupos diferentes (arestas inter-grupo) é  $2w_x w_y$ .

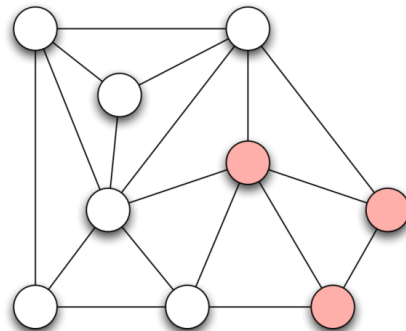
Se a proporção de arestas inter-grupo é significativamente menor do que  $2w_x w_y$ , então há evidência de homofilia.

**Equivalentemente:**

Se a proporção de arestas intra-grupo é significativamente maior do que  $w_x^2 + w_y^2$ , então há evidência de homofilia.

# Homofilia de uma rede: Exemplo

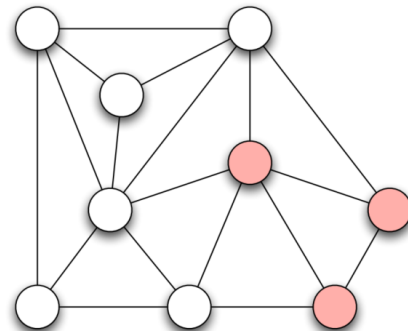
- Quantas arestas inter-grupo e intra-grupo há na rede?



# Homofilia de uma rede: Exemplo

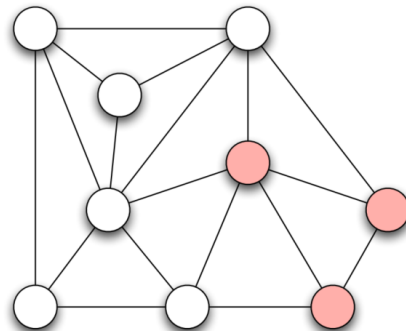
- Quantas arestas inter-grupo e intra-grupo há na rede?

5 arestas inter-grupo e 13 arestas intra-grupo.



# Homofilia de uma rede: Exemplo

- Quantas arestas inter-grupo e intra-grupo há na rede?  
5 arestas inter-grupo e 13 arestas intra-grupo.
- Qual é o número esperado dessas arestas se não houvesse homofilia?



# Homofilia de uma rede: Exemplo

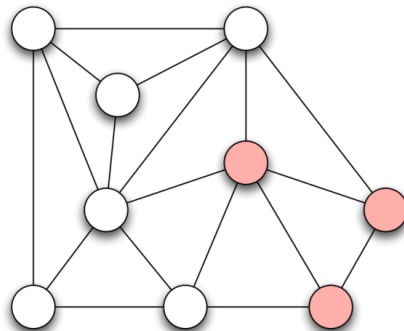
- Quantas arestas inter-grupo e intra-grupo há na rede?

5 arestas inter-grupo e 13 arestas intra-grupo.

- Qual é o número esperado dessas arestas se não houvesse homofilia?

$$2w_x w_y = 2 \frac{2}{3} \frac{1}{3} = \frac{4}{9}$$

⇒ o esperado é  $18 \frac{4}{9} = 8$  arestas inter-grupo.



## Homofilia de uma rede: Exemplo

- Quantas arestas inter-grupo e intra-grupo há na rede?

5 arestas inter-grupo e 13 arestas intra-grupo.

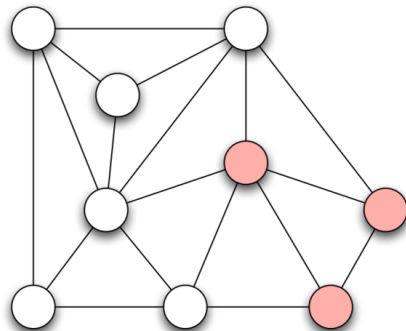
- Qual é o número esperado dessas arestas se não houvesse homofilia?

$$2w_x w_y = 2\frac{2}{3}\frac{1}{3} = \frac{4}{9}$$

⇒ o esperado é  $18\frac{4}{9} = 8$  arestas inter-grupo.

$$w_x^2 + w_y^2 = \frac{2^2}{3} + \frac{1^2}{3} = \frac{5}{9}$$

⇒ o esperado é  $18\frac{5}{9} = 10$  arestas intra-grupo.



## Homofilia de uma rede: Exemplo

- Quantas arestas inter-grupo e intra-grupo há na rede?

5 arestas inter-grupo e 13 arestas intra-grupo.

- Qual é o número esperado dessas arestas se não houvesse homofilia?

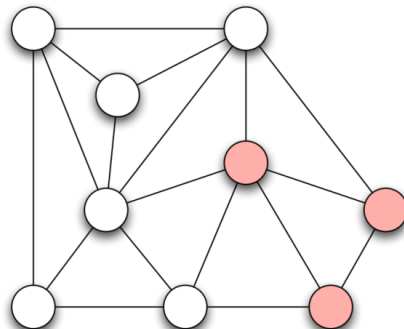
$$2w_x w_y = 2\frac{2}{3}\frac{1}{3} = \frac{4}{9}$$

⇒ o esperado é  $18\frac{4}{9} = 8$  arestas inter-grupo.

$$w_x^2 + w_y^2 = \frac{2^2}{3} + \frac{1^2}{3} = \frac{5}{9}$$

⇒ o esperado é  $18\frac{5}{9} = 10$  arestas intra-grupo.

- **Comparando** os valores, observamos que há alguma evidência de homofilia na rede.



# Homofilia de uma rede

## Homofilia **vs** fecho triádico

É possível uma rede **não ter triângulos**, e apresentar alto grau de homofilia!

# Homofilia de uma rede

## Pendências:

**Pergunta 1:** A diferença entre os números real e esperado de arestas intra-grupo (inter-grupo) na rede é **significativa, suficiente** para afirmar que há homofilia?

# Homofilia de uma rede

## Pendências:

**Pergunta 1:** A diferença entre os números real e esperado de arestas intra-grupo (inter-grupo) na rede é **significativa, suficiente** para afirmar que há homofilia?

**Pergunta 2:** Seria possível criar um **indicador simples** (um número) para a homofilia de uma rede?

# Homofilia de uma rede

## Pendências:

**Pergunta 1:** A diferença entre os números real e esperado de arestas intra-grupo (inter-grupo) na rede é **significativa, suficiente** para afirmar que há homofilia?

**Pergunta 2:** Seria possível criar um **indicador simples** (um número) para a homofilia de uma rede?

**Pergunta 3:** O que fazer se temos **mais de 2 grupos**?

# Homofilia de uma rede

## Pendências:

**Pergunta 1:** A diferença entre os números real e esperado de arestas intra-grupo (inter-grupo) na rede é **significativa, suficiente** para afirmar que há homofilia?

**Pergunta 2:** Seria possível criar um **indicador simples** (um número) para a homofilia de uma rede?

**Pergunta 3:** O que fazer se temos **mais de 2 grupos**?

To be continued... (**modularidade**)

# Resumo

- 1 Objetivo
- 2 **Homofilia**
  - Homofilia em atributos categóricos
  - Homofilia em atributos numéricos
- 3 Aplicações e mecanismos subjacentes

# Homofilia: atributos categóricos vs numéricos

**Até agora vimos atributos categóricos:** grupos de nós, dentro dos quais há nós semelhantes entre si.

- **Exemplos:** pessoas que falam o mesmo idioma, gostam do mesmo time de futebol.

# Homofilia: atributos categóricos vs numéricos

**Até agora vimos atributos categóricos:** grupos de nós, dentro dos quais há nós semelhantes entre si.

- **Exemplos:** pessoas que falam o mesmo idioma, gostam do mesmo time de futebol.

**Se temos atributos numéricos:** cada nó  $v$  possui um valor numérico  $x_v$  associado. Se o valor do atributo é semelhante, os vértices são semelhantes.

- **Exemplos:** idade de uma pessoa, salário ou renda.
- Outra definição de ter “algo em comum”: ter um atributo igual ou próximo ao atributo de outro nó.

## Homofilia: atributos numéricos

**Coeficiente de assortatividade** (numérico):

$$r = \frac{\frac{1}{m} \sum_{(u,v)} (x_u - \mu)(x_v - \mu)}{\sigma^2}.$$



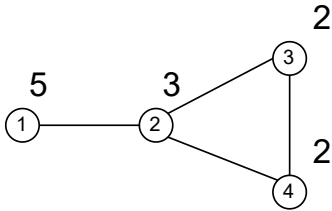
## Homofilia: atributos numéricos

Seja um grafo  $G(V, E)$ , com um conjunto de  $m$  arestas  $E$ .

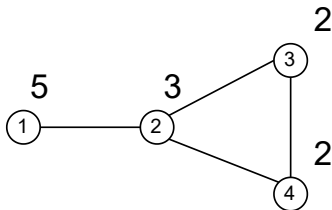
- Existe um valor  $x_i$ , associado ao nó  $i$ .
- Para cada aresta  $(u, v)$ , queremos avaliar a relação entre o valor  $x_u$  e o valor  $x_v$ :  
Existe uma tendência dos nós de ter vínculos com seus semelhantes?

Sejam as variáveis aleatórias  $X$  e  $Y$  que indicam os **valores associados** aos **dois extremos** de uma aresta aleatória do grafo.

## Homofilia: atributos numéricos

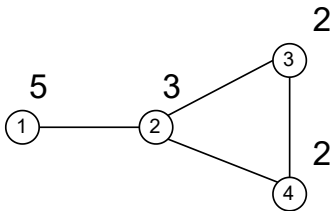


## Homofilia: atributos numéricos



| X |           | Y |
|---|-----------|---|
| 5 | (1) — (2) | 3 |
| 3 | (2) — (3) | 2 |
| 3 | (2) — (4) | 2 |
| 2 | (3) — (4) | 2 |

## Homofilia: atributos numéricos



| X |           | Y |
|---|-----------|---|
| 5 | (1) — (2) | 3 |
| 3 | (2) — (3) | 2 |
| 3 | (2) — (4) | 2 |
| 2 | (3) — (4) | 2 |

**Importante:** Duplicar as arestas. Os vértices podem ser escolhidos em ordem inversa.  
(Ao duplicar as arestas, as variáveis aleatórias  $X$  e  $Y$  serão iguais.)

## Homofilia: atributos numéricos

Existe uma relação entre as variáveis aleatórias  $X$  e  $Y$ ?

## Homofilia: atributos numéricos

**Existe uma relação** entre as variáveis aleatórias **X** e **Y**?

- **Coeficiente de correlação de Pearson:** avalia o grau da correlação entre duas variáveis aleatórias.

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

## Homofilia: atributos numéricos

Existe uma relação entre as variáveis aleatórias  $X$  e  $Y$ ?

- **Coeficiente de correlação de Pearson:** avalia o grau da correlação entre duas variáveis aleatórias.

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

- No nosso caso,  $X$  e  $Y$  possuem a mesma distribuição, então  $\mu_X = \mu_Y$  e  $\sigma_X = \sigma_Y$ .

## Homofilia: atributos numéricos

Existe uma relação entre as variáveis aleatórias  $X$  e  $Y$ ?

- **Coeficiente de correlação de Pearson:** avalia o grau da correlação entre duas variáveis aleatórias.

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

- No nosso caso,  $X$  e  $Y$  possuem a mesma distribuição, então  $\mu_X = \mu_Y$  e  $\sigma_X = \sigma_Y$ .
- Covariância:  $\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[(X - \mu)(Y - \mu)]$ .

## Homofilia: atributos numéricos

Existe uma relação entre as variáveis aleatórias  $X$  e  $Y$ ?

- **Coeficiente de correlação de Pearson:** avalia o grau da correlação entre duas variáveis aleatórias.

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

- No nosso caso,  $X$  e  $Y$  possuem a mesma distribuição, então  $\mu_X = \mu_Y$  e  $\sigma_X = \sigma_Y$ .
- Covariância:  $\text{cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[(X - \mu)(Y - \mu)]$ .

$$r = \frac{\frac{1}{m} \sum_{(u,v)} (x_u - \mu)(x_v - \mu)}{\sigma^2}$$

## Homofilia: atributos numéricos

$$r = \frac{\frac{1}{m} \sum_{(u,v)} (x_u - \mu)(x_v - \mu)}{\sigma^2}$$

**Coeficiente de assortatividade** (numérico) calculado para toda a rede com atributos associados aos nós. Varia de -1 a 1.

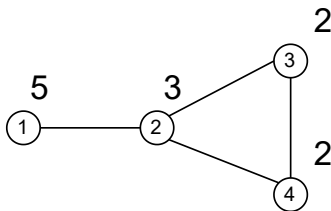
## Homofilia: atributos numéricos

$$r = \frac{\frac{1}{m} \sum_{(u,v)} (x_u - \mu)(x_v - \mu)}{\sigma^2}$$

**Coeficiente de assortatividade** (numérico) calculado para toda a rede com atributos associados aos nós. Varia de -1 a 1.

- Um valor próximo de 1 indica **forte correlação positiva**: os nós tendem a se conectar a outros com valores semelhantes (homofilia).
- O valor zero indica que não há homofilia.
- Um valor próximo de -1 indica **forte correlação negativa**: os nós tendem a se conectar a outros com valores diferentes (heterofilia).

## Homofilia: atributos numéricos

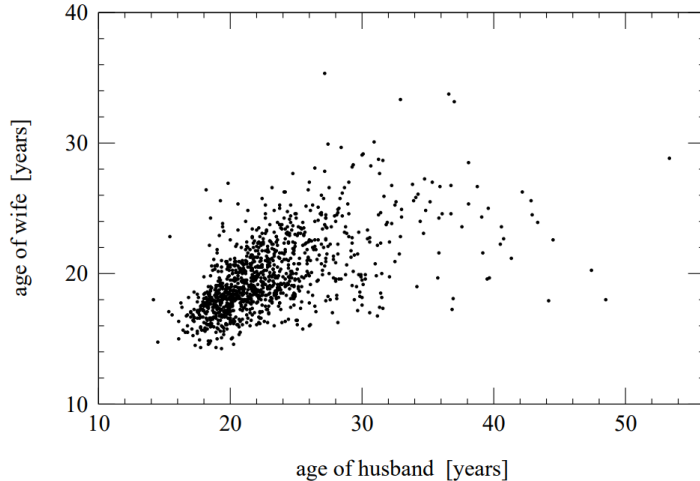


| X |           | Y |
|---|-----------|---|
| 5 | (1) — (2) | 3 |
| 3 | (2) — (3) | 2 |
| 3 | (2) — (4) | 2 |
| 2 | (3) — (4) | 2 |

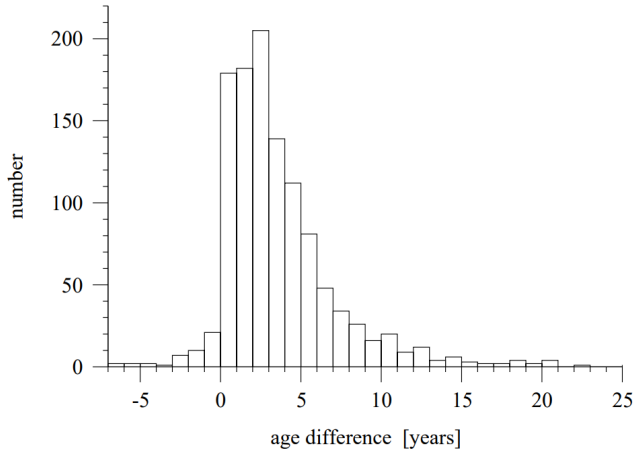
**Média e variância** das variáveis  $X$  e  $Y$ :  $\mu = \frac{11}{4}$ ,  $\sigma = \frac{15}{16}$ .

**Coeficiente de assortatividade:**  $r = 0,2$ .

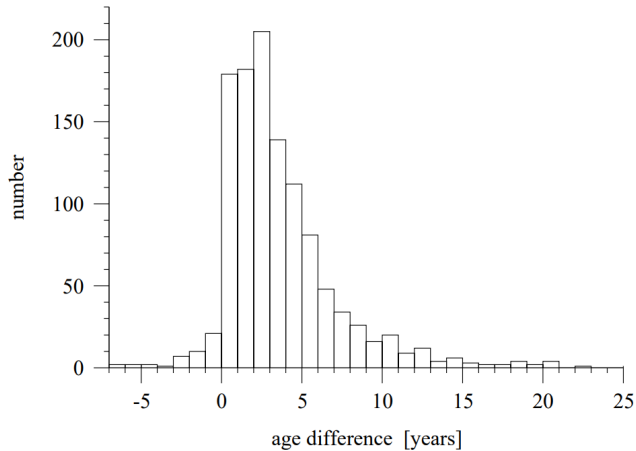
## Homofilia: atributos numéricos



## Homofilia: atributos numéricos



## Homofilia: atributos numéricos



$$r = 0,574$$

## Homofilia: atributos numéricos

**Assortatividade de grau:** tendência dos nós de se conectar a outros cujo grau é semelhante.

- O valor numérico neste caso é o grau do vértice.
- A rede apresenta **assortatividade**, se os nós tendem a se conectar a outros cujo grau é semelhante (numericamente). **Exemplo**: redes sociais.
- Rede **dissassortativa**: nós de alto grau tendem a se conectar com nós de baixo grau. **Exemplo**: World Wide Web.

# Resumo

## 1 Objetivo

## 2 Homofilia

- Homofilia em atributos categóricos
- Homofilia em atributos numéricos

### 3 Aplicações e mecanismos subjacentes

## Aplicações

## O mundo não é tão simples como nestes modelos!

Um nó ou uma rede pode **apresentar** homofilia **em relação a um atributo** dos vértices, e **não apresentar** homofilia, ou mostrar heterofilia, **em relação a outro**.

## Aplicações

## O mundo não é tão simples como nestes modelos!

Um nó ou uma rede pode **apresentar** homofilia **em relação a um atributo** dos vértices, e **não apresentar** homofilia, ou mostrar heterofilia, **em relação a outro**.

- **Exemplo:** a mesma rede de indivíduos **apresenta homofilia** pela **empresa** na qual cada um é empregado, e **não apresenta** homofilia pelo **time de futebol** favorito de cada indivíduo.

## Mecanismos subyacentes

Os nós podem ter **diversos atributos** e preferências. O surgimento de novos vínculos na rede pode depender de **vários fatores ao mesmo tempo**.

- Alguns atributos são **imutáveis** ou **estáticos**: cidade ou país de origem, raça ou etnia. Time de futebol?

## Mecanismos subyacentes

Os nós podem ter **diversos atributos** e preferências. O surgimento de novos vínculos na rede pode depender de **vários fatores ao mesmo tempo**.

- Alguns atributos são **imutáveis** ou **estáticos**: cidade ou país de origem, raça ou etnia. Time de futebol?
- Outros atributos são **mutáveis** ou **dinâmicos**: idade, interesses, comportamentos, ideologias, crenças, opiniões.

## Seleção

No caso dos atributos **imutáveis** ou **estáticos**:

- A tendência dos indivíduos de se associar com seus semelhantes é chamada de **seleção**: nos escolhemos ter vínculos com pessoas com os mesmos atributos.

## Seleção

No caso dos atributos **imutáveis** ou **estáticos**:

- A tendência dos indivíduos de se associar com seus semelhantes é chamada de **seleção**: nos escolhemos ter vínculos com pessoas com os mesmos atributos.
- Às vezes, a seleção é mais **explícita** ou **intencional**: em um grupo, podemos escolher para conversar com uma pessoa que vem da mesma cidade ou gosta do mesmo time.

## Seleção

No caso dos atributos **imutáveis** ou **estáticos**:

- A tendência dos indivíduos de se associar com seus semelhantes é chamada de **seleção**: nos escolhemos ter vínculos com pessoas com os mesmos atributos.
- Às vezes, a seleção é mais **explícita** ou **intencional**: em um grupo, podemos escolher para conversar com uma pessoa que vem da mesma cidade ou gosta do mesmo time.
- Às vezes, a seleção é **implícita** ou **involuntária**: ambientes mais homogêneos (empresa, escola, universidade) incentivam o encontro com pessoas com os mesmos atributos.

## Atributos determinam vínculos

No caso dos atributos **mutáveis** ou **dinâmicos** (comportamentos, ideologias, crenças):

- A relação entre características individuais e a criação de vínculos com pessoas semelhantes já **não é unilateral**.

## Atributos determinam vínculos

No caso dos atributos **mutáveis** ou **dinâmicos** (comportamentos, ideologias, crenças):

- A relação entre características individuais e a criação de vínculos com pessoas semelhantes já **não é unilateral**.
- A **seleção** continua operando: **atributos** continuam **influenciando** o surgimento dos vínculos:

atributos  $\Rightarrow$  vínculos

## Vínculos determinam atributos

Porém, o processo **inverso** também ocorre:

- As pessoas podem **mudar** os seus atributos e comportamentos para estar mais alinhados aos atributos e comportamentos dos seus amigos;
- Países tendem a **alinhar** suas características produtivas aos interesses dos seus parceiros comerciais. **Brasil produz mais soja porque a China consome mais soja!**

# Vínculos determinam atributos

Porém, o processo **inverso** também ocorre:

- As pessoas podem **mudar** os seus atributos e comportamentos para estar mais alinhados aos atributos e comportamentos dos seus amigos;
- Países tendem a **alinhar** suas características produtivas aos interesses dos seus parceiros comerciais. **Brasil produz mais soja porque a China consome mais soja!**

Nesse processo inverso (influência), os vínculos determinam os atributos:

**vínculos  $\Rightarrow$  atributos**

## Mecanismos subyacentes

Suponha que **uma rede apresenta homofilia**: nós compartilham características com os seus adjacentes.

- **Exemplo 1:** Eu posso **criar vínculos** econômicos (e.g., com outros países) devido às minhas características, mas eu posso **mudar as minhas características** por causa dos meus vínculos!
- **Exemplo 2:** As pessoas **se adaptam** para serem mais parecidas aos seus amigos, e também **procuraram amigos** que são mais parecidos a eles.

## Mecanismos subyacentes

Entender o que é causa e o que é consequência é importante!

Em que medida os nós **se adaptaram** para serem mais parecidos aos seus adjacentes, **ou procuraram vínculos** com nós que são **mais parecidos** a eles?

## Mecanismos subyacentes

Entender o que é causa e o que é consequência é importante!

Em que medida os nós **se adaptaram** para serem mais parecidos aos seus adjacentes, **ou procuraram vínculos** com nós que são **mais parecidos** a eles?

- Parece o dilema do ovo e da galinha!

# Mecanismos subjacentes

Entender o que é causa e o que é consequência é importante!

Em que medida os nós **se adaptaram** para serem mais parecidos aos seus adjacentes, **ou procuraram vínculos** com nós que são **mais parecidos** a eles?

- Parece o dilema do ovo e da galinha!

Embora pareça difícil responder esta questão, tendo **dados dinâmicos** dos vínculos em uma rede de interação, é possível responder essa pergunta objetivamente:

- Os atributos se adaptaram à estrutura da rede, ou a estrutura da rede se adaptou aos atributos?

# Aplicações

Entender o que é causa e o que é consequência é importante!

**Intervenções para reduzir o uso de drogas:** alunos usam drogas com maior probabilidade se seus amigos o fazem. Vamos supor que vai ser criada uma intervenção (um programa) que estimula alunos a não usarem drogas.

- Se a homofilia no uso de drogas é causada pela **influência**, o programa pode ter um **efeito muito maior**, fazendo com que seus amigos também deixem de (ou não comecem a) usar drogas!

## Material bibliográfico

M. Newman. *Mixing patterns in networks* (2002).

David Easley and Jon Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*, pp 77-88.

# Dúvidas?