

POWER LAWS E O FENÔMENO RICO-FICA-MAIS-RICO

Redes Sociais e Econômicas

Prof. André Vignatti

POWER LAWS E FENÔMENO “RICO FICA MAIS RICO”

popularidade - fenômeno caracterizado por **desequilíbrios extremos**

- **ex:** quase todo mundo conhece **somente pessoas do seu círculo social**, alguns alcançam uma **visibilidade maior**, e muito poucos têm **reconhecimento global**
- **ex:** alguns *livros, filmes, música...* são **extremamente populares**, mas a **GRANDE maioria** é conhecida por um **público pequeno**

como **quantificar** estes **desequilíbrios**? **por que eles surgem?**

POWER LAWS E FENÔMENO “RICO FICA MAIS RICO”

vamos estudar modelos básicos de **comportamento da rede** para obter respostas

é difícil medir quantas pessoas conhecem o **Brad Pitt** ou a **Anitta**

assim, vamos nos **concentrar na Web**, pois é **mais fácil de medir**

POWER LAWS E FENÔMENO “RICO FICA MAIS RICO”

mais especificamente: **número de links que apontam** para uma página Web

- **in-links**: número de links que *apontam para uma página*

lembre-se: a **Web é só um exemplo** de um fenômeno que **acontece em outras situações**

HIPÓTESE: A DISTRIBUIÇÃO NORMAL

uma pergunta, feita nos primórdios da Web:

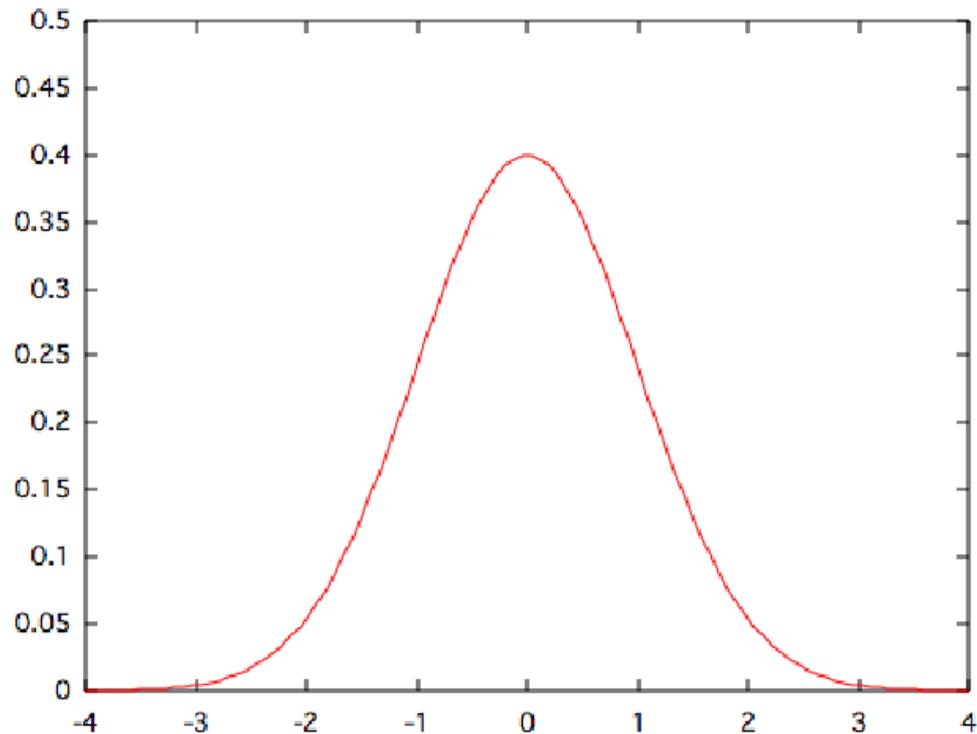
em função de k , qual a fração de páginas da Web tem k in-links?

qual seria uma possível resposta?

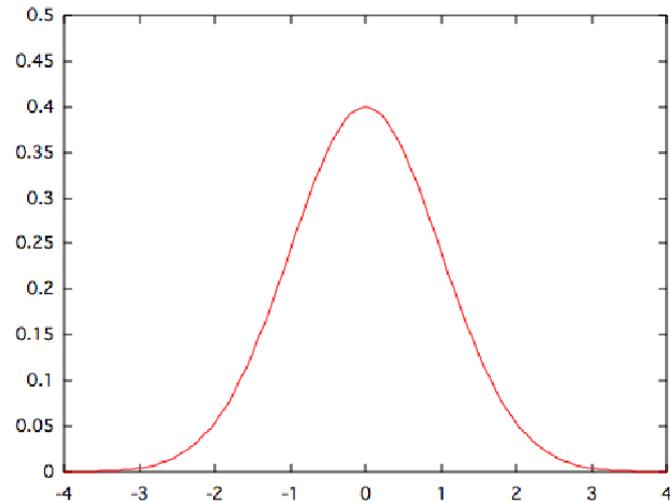
- uma tentativa natural seria a **distribuição normal (ou gaussiana)**

HIPÓTESE: A DISTRIBUIÇÃO NORMAL

- altura
- peso
- tempo de viagem
- jogada de dados
- ...



HIPÓTESE: A DISTRIBUIÇÃO NORMAL



a distribuição normal é meio que "mágica", porque a vemos muito na natureza

mas há uma explicação para isso: **Teorema do Limite Central**

HIPÓTESE: A DISTRIBUIÇÃO NORMAL

Teorema do Limite Central (grosseiramente): para qualquer sequência de *pequenas quantidades aleatórias independentes*, no limite, sua soma (ou média) será distribuída de acordo com a distribuição normal

IMPORTANTE: à medida que se **afasta da média**, a probabilidade de enxergar um valor **decrece exponencialmente**

HIPÓTESE: DISTRIBUIÇÃO NORMAL

mas **por que a distribuição normal?**

- **ex:** ao fazer **repetidas medidas numa quantidade física fixa**, e plotar o **gráfico da soma de determinados resultados**, ela será aproximadamente como a **distribuição normal**

como **aplicar em páginas Web?**

- se cada página decide **aleatoriamente** com qual outra **página vai se linkar**, então a **soma dos in-links** é a soma de **quantidades aleatórias independentes** \Rightarrow distribuição normal

POWER LAWS

MAS quando as pessoas **mediram a Web**, descobriram algo **totalmente diferente**

a **fração de páginas Web** que têm em k -links é aproximadamente **proporcional** à $1/k^2$

por que isso é **diferente da distribuição normal?**

- $1/k^2$ decresce **muito mais devagar** à medida que k aumenta
- então páginas com muitos in-links são **muito mais comuns que esperávamos** pela distribuição normal

POWER LAWS

ex: $k = 1000$

- $1/k^2$ é “um em 1 milhão”
- uma função que decresce exponencialmente, como $1/2^k$, teria um valor **MUITO** pequeno

a função em k que **decrece de uma potência fixa** (no caso, **2**) é chamada de **Power Law (Lei da Potência)**

POWER LAWS

a popularidade tem **desequilíbrios extremos**: poucas páginas Web são muito populares (têm muitos in-links)

serve para **outros domínios**:

- a fração dos **números de telefone** que recebem k chamadas por dia: $1/k^2$
- a fração de **livros comprados** por k pessoas: $1/k^3$
- a fração de **trabalhos científicos** que recebem k citações: $1/k^3$

conclusão: assim como a distribuição normal aparece nas *ciências naturais*, as distribuições power law aparecem em *contextos em que há “popularidade”*

POWER LAWS

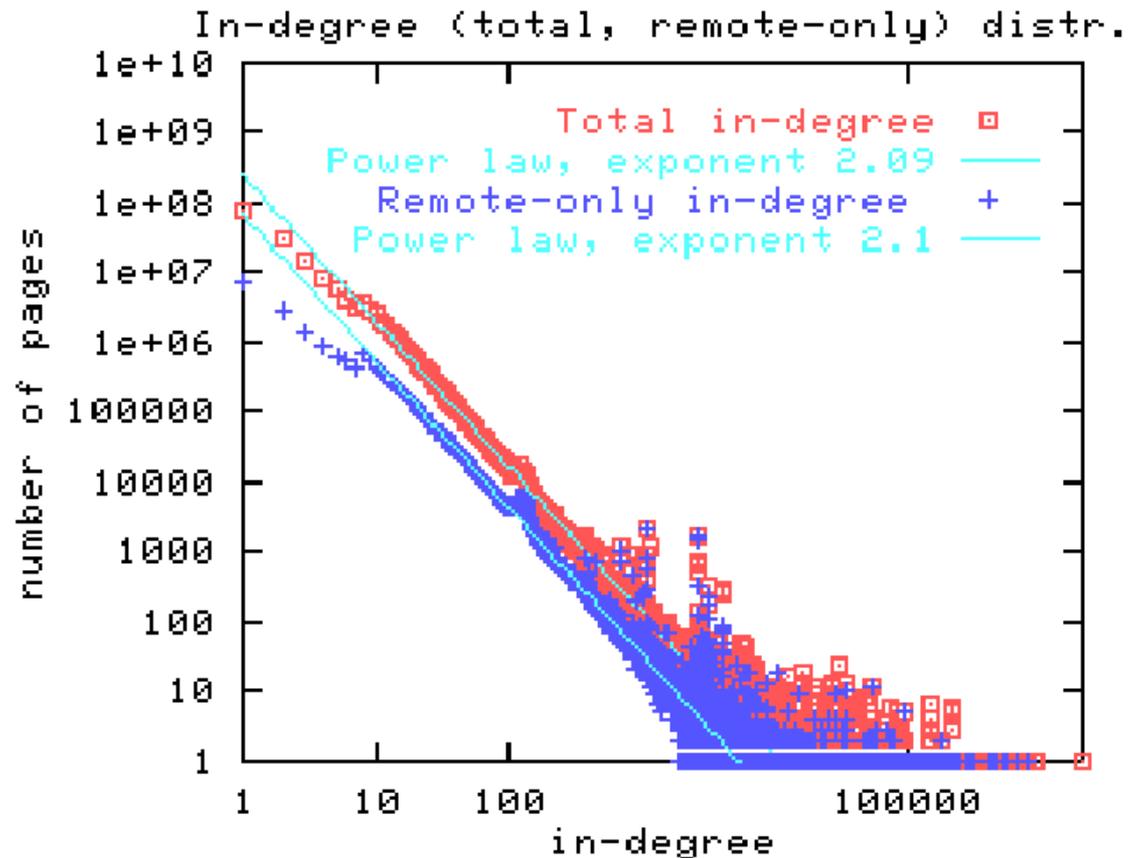
como **saber** se dados tem distribuição **power law**?

- seja $f(k)$ - a **fração de itens** que têm valor k
- queremos saber se $f(k) = a/k^c$ onde a e c são **constantes**
- tirando o log de ambos os lados da equação:

$$\log f(k) = \log a - c \log k$$

se temos de distribuição **power law**, obtemos uma **linha reta** no gráfico com eixos **log-log**

A DISTRIBUIÇÃO POWER-LAW DOS IN-LINKS DE PÁGINAS WEB



Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, and Janet Wiener. **Graph structure in the Web**. In Proc. 9th International World Wide Web Conference, 2000