Redes Complexas
Aula 5

Roteiro

- Lei de potência
- Distribuição Zeta
- Propriedades
- Distribuição Zipf
- Exemplos

Aula passada

- Padrões de mixagem (*mixing patterns*)
- Assortatividade
- Correlação entre graus
- Similaridade entre vértices

Lei de Potência

- Qualquer função que tenha a forma

$$f(x) \sim c\,x^{a}$$

c, a são constantes

- f(x) evolui de acordo com uma potência
  - lei de potência (ou polinômio com grau real)
- Relação ocorre quando x é grande

$$f(x) \sim g(x) \longrightarrow \frac{f(x)}{g(x)} \rightarrow 1 \quad \text{quando} \quad x \rightarrow \infty$$

Invariância em Escala

- Propriedade de “*scale free*”
- Multiplicar o argumento por *k* mantém a *forma* da função

$$f(kx) \sim c(kx)^a \sim k^a c x^a = k^a f(x)$$

constante depende apenas de *k*

- Forma é preservada em qualquer escala
- Ex. a = 2, k = 3
  - f(3x) = 9 f(x)
  - evento 3 vezes maior tem 9 vezes mais impacto, independente do tamanho do evento (x)

Escala log-log

- Aplicar logaritmo aos dois lados da equação

$$f(x) \sim c\,x^{a}$$

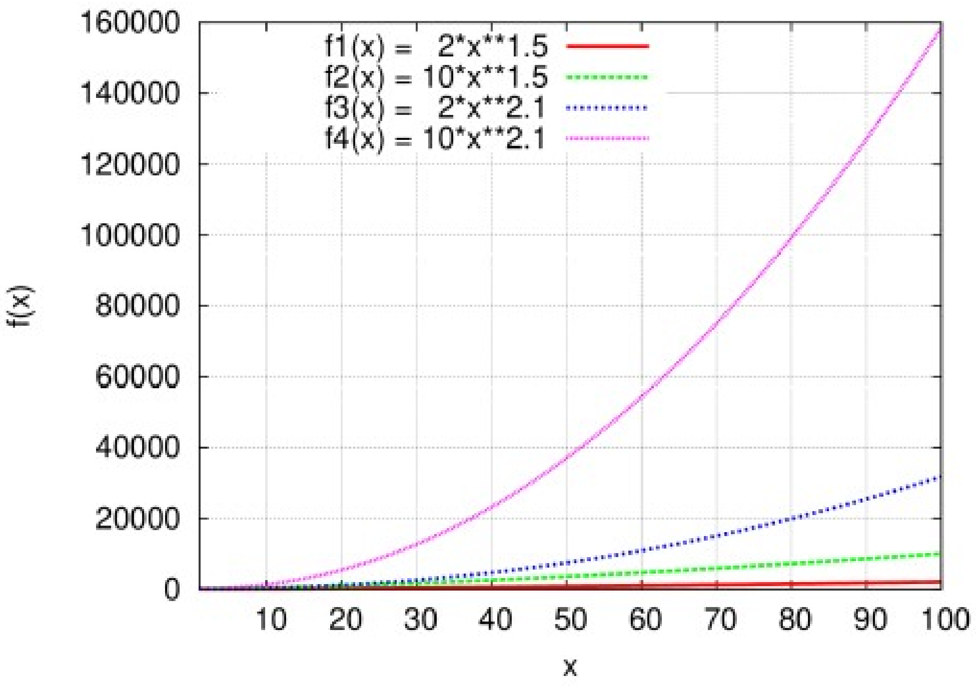
- Temos então

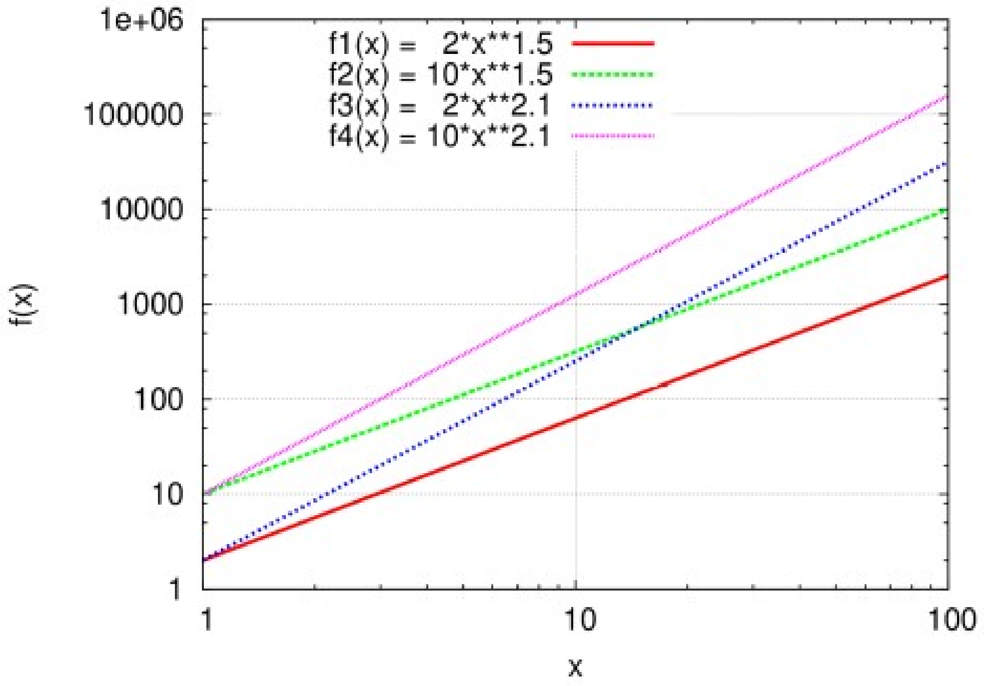
$$\log(f(x)) \sim \log(c\,x^{a}) = a \log x + \log c$$

- Relação linear entre log f(x) e log x com inclinação a
- Expoente determina inclinação da reta na escala log-log
- Invariância de escala
  - inclinação (forma) não depende de x

Exemplo Fictício

Escala Linear → Escala Log-Log



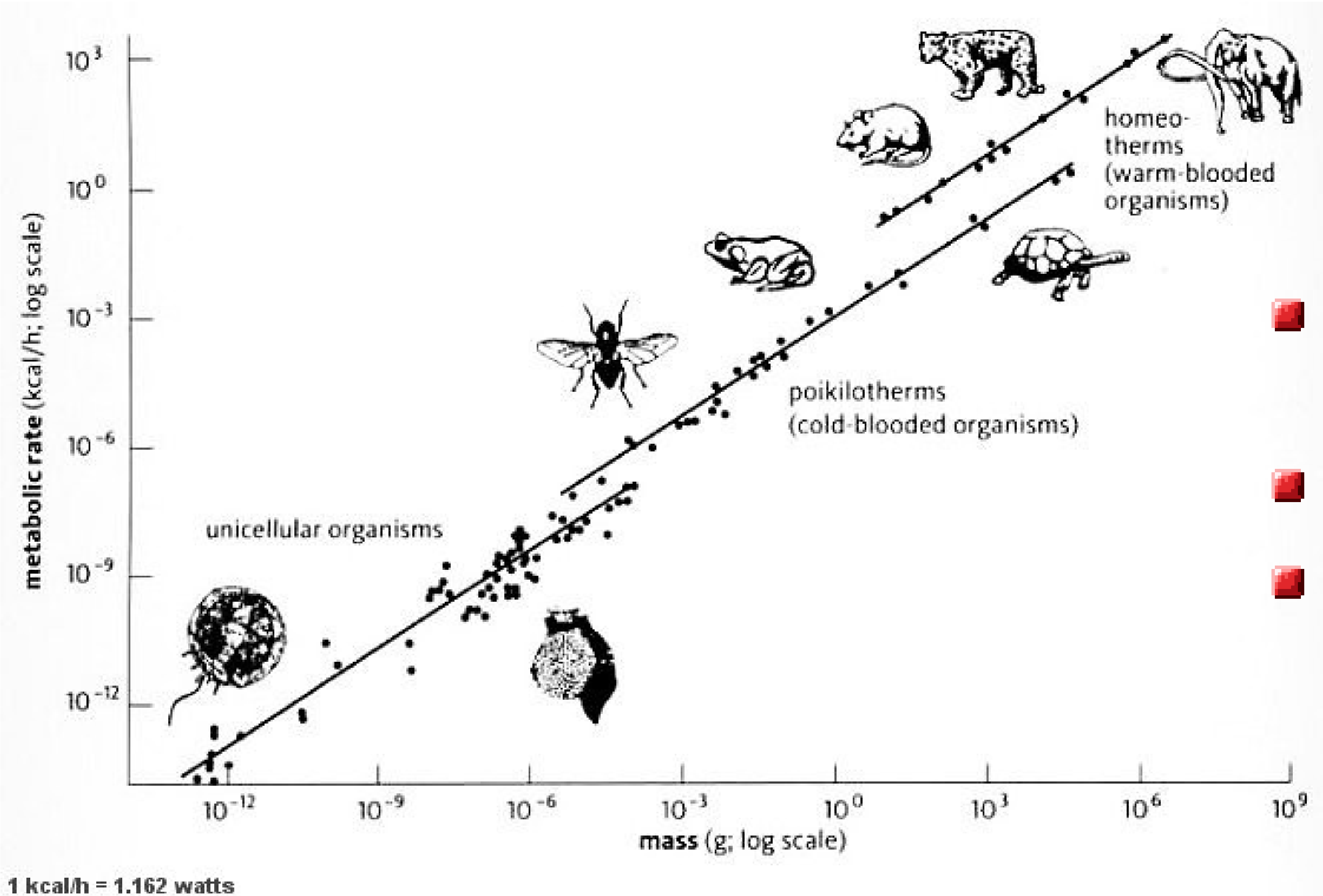


- f1 e f2: retas paralelas
- f3 inclinação maior que f1

- Mais fácil interpretar comportamento quando apresentado em escala log-log

Exemplo Real

- Muitos fenômenos naturais e artificiais seguem lei de potência

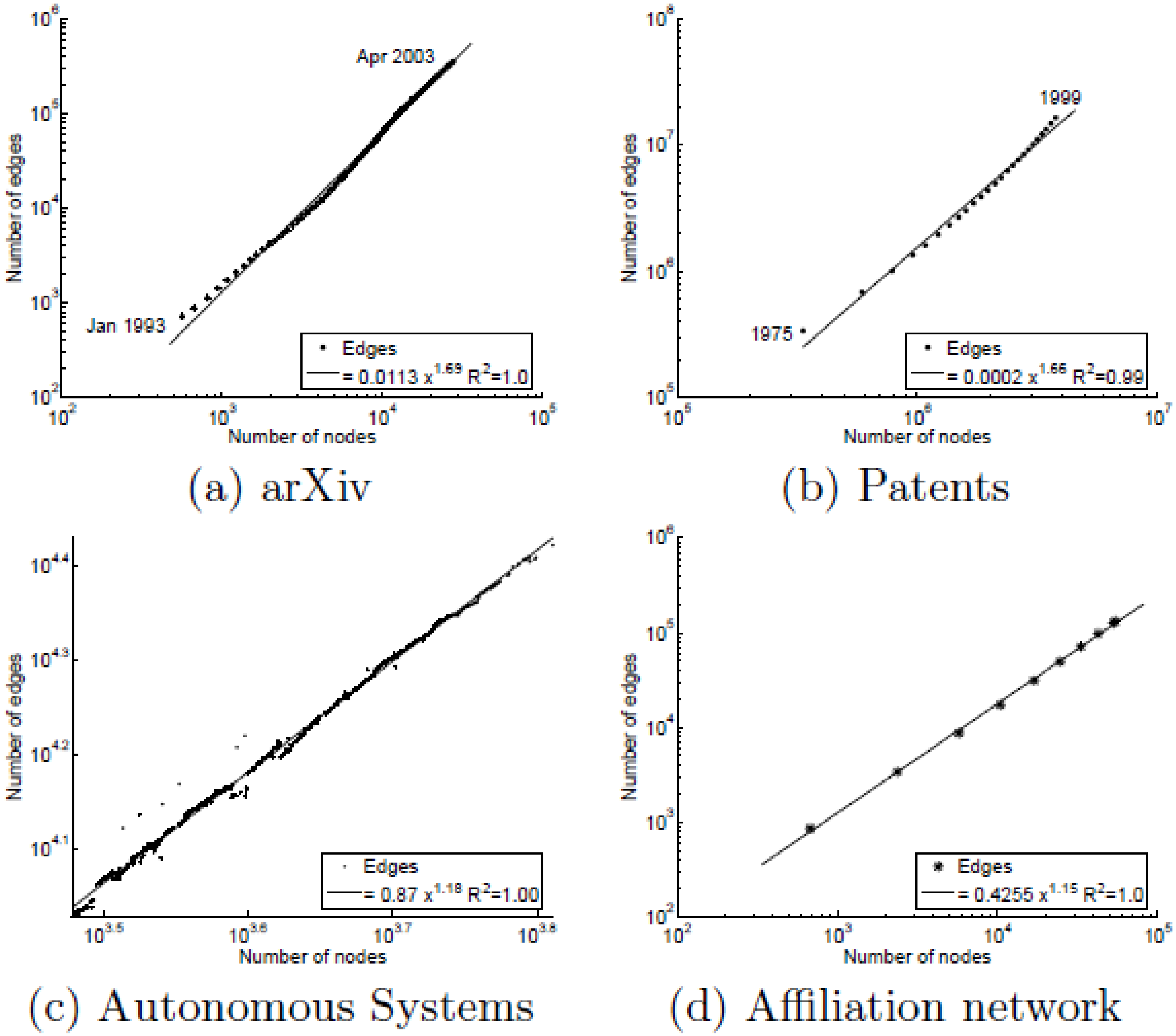


- R: taxa metabólica
- M : massa
- Lei de Kleiber

$$R \sim M^{3/4}$$

Tamanho das Redes

- Número de vértices x número de arestas
  - pontos correspondem a diferentes instâncias da rede



(a) arXiv (b) Patents

(c) Autonomous Systems (d) Affiliation network

- Taxa de crescimento da rede
- Arestas mais rápido que vértices

$$E \sim V^{\alpha}$$

> 1

Distribuição de Lei de Potência

- Lei de potência aplicada a distribuição de probabilidade

$$f_X(x) \sim c\,x^{-a}$$

expoente negativo

- onde X é uma v.a. e $f_X(x)$ sua função de probabilidade, $f_X(x) = P[X = x]$
- X pode ser discreta ou contínua, limitada ou ilimitada superiormente
  - se contínua, $f_X(x)$ é a função densidade

Distribuição Zeta

- Seja X uma v.a. discreta ilimitada superiormente
- X assume valores 1, 2, ...

$$p_k = P[X = k] \sim c k^{-a}$$

- Determinar constante c de normalização

$$p_k = c k^{-a}$$

- Temos então

$$\sum_{k>0} p_k = 1 \longrightarrow c = 1 / \sum_{k>0} k^{-a} = 1 / z(a)$$

Função Zeta de Riemann

- z(a) é a função zeta de Riemann

$$z(a)=\sum_{k>0} k^{-a}$$

Converge para qualquer valor de a?

- Para a=1 temos a série harmônica ← Não converge
- Para a=2 temos a série quadrática ← Converge para $\pi^2/6$
- Quando a>1, z(a) converge
- Logo função de probabilidade definida apenas quando a>1

Momentos da Zeta

- Calcular momentos da distribuição zeta
  - n-ésimo momento é dado por $E[X^n]$
- Primeiro momento é valor esperado (n=1)

Como calcular n-ésimo momento?

- Aplicar definição

$$E[X^n] = \sum_{k>0} k^n p_k = \sum_{k>0} k^n c k^{-a} =$$

$$= \frac{1}{z(a)} \sum_{k>0} k^{-(a-n)} = \frac{z(a-n)}{z(a)}$$

Momentos da Zeta

- n-ésimo momento é dado por

$$E[X^n] = z(a-n)/z(a)$$

Converge sempre?

- para a – n > 1 ou seja, a > n + 1
  - infinito caso contrário (não converge)
- Exemplo: a = 2.3, E[X] existe? Var[X] existe?
  - Var[X] = E[X^2] – E[X]^2 , função do segundo momento
- E[X] está definido, pois 2.3 > 1 + 1
- E[X^2] é infinito, pois 2.3 < 2 + 1
- Var[X] é infinita!

Invariância em Escala

- Razão entre probabilidades depende apenas da razão entre magnitudes
- Não depende do tamanho do evento (k)
- Matematicamente, seja *b* uma constante positiva

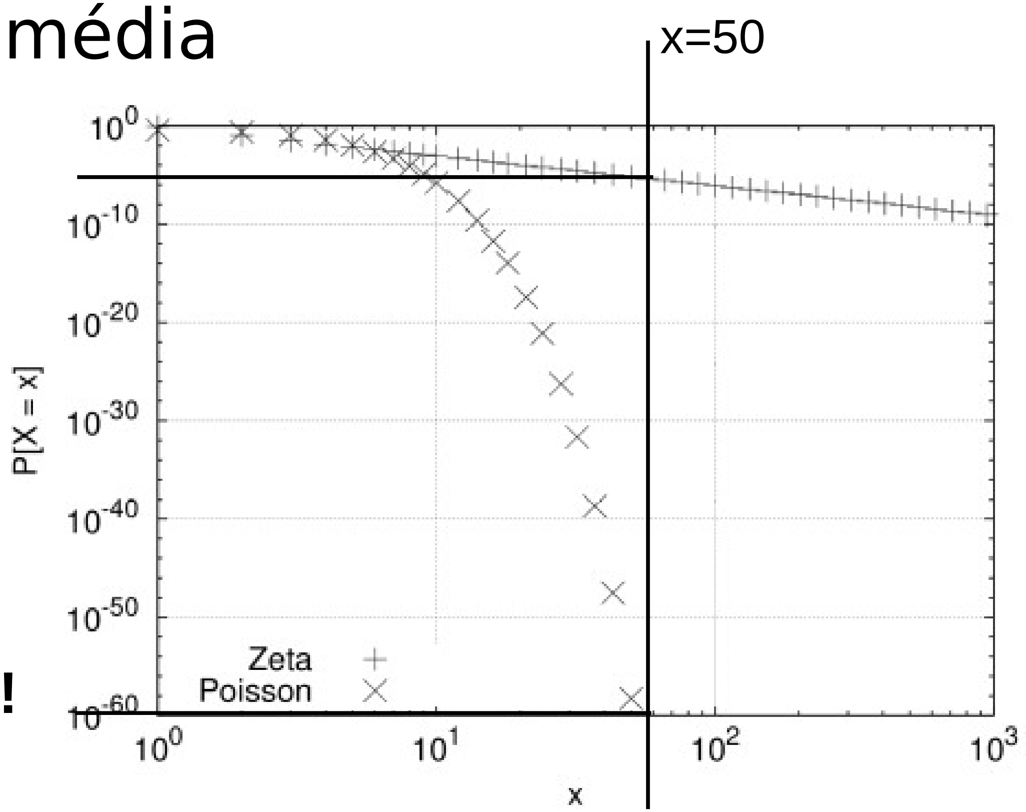
$$\frac{p_{bk}}{p_k} = \frac{c\left(bk\right)^{-a}}{ck^{-a}} = b^{-a}$$

Não depende de *k* (escala), apenas de *b* (fator multiplicativo)

Cauda Pesada

- *Heavy tail* ou *long tail* ou *fat tail*
- Probabilidade decresce devagar com x
  - segue lei de potência, ao invés de lei exponencial
- Eventos “enormes” podem ocorrer com probabilidade não desprezível
  - muito maiores do que a média
- Comparação entre Zeta (a = 3) e Poisson
- Mesmo valor esperado, E[X] = 5
- Poisson tem cauda exponencial

10^{50} vezes mais provável!



Distribuição de Zipf

- Variação de distribuição de Zeta
  - proposta por George Zipf em 1935 para caracterizar distribuição de palavras em textos
- Valores de X são limitados superiormente por n (parâmetro do modelo)

$$p_k = c\,k^{-a}$$

- Constante de normalização?

$$\sum_{k=1}^{n} p_k = 1$$

$$c = 1 / \sum_{k=1}^{n} k^{-a} = 1 / H_{n,a}$$

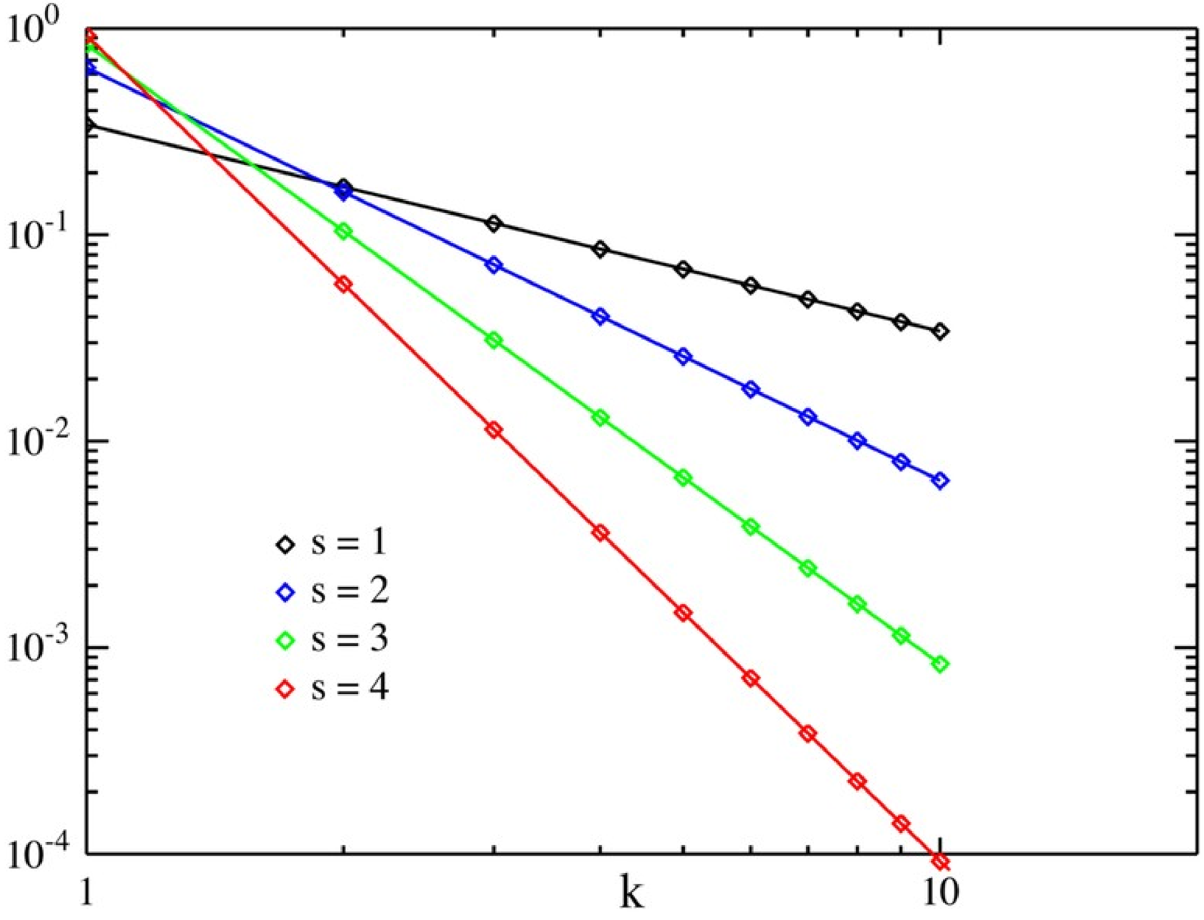
Distribuição de Zipf

- $H_{n,a}$ é o n-ésimo número harmônico generalizado

$$H_{n,a} = \sum_{k=1}^{n} k^{-a}$$

- Para qualquer n finito, $H_{n,a}$ converge
  - diferença fundamental para Zeta
- Momentos da Zipf?
  - aplicar definição, todos definidos

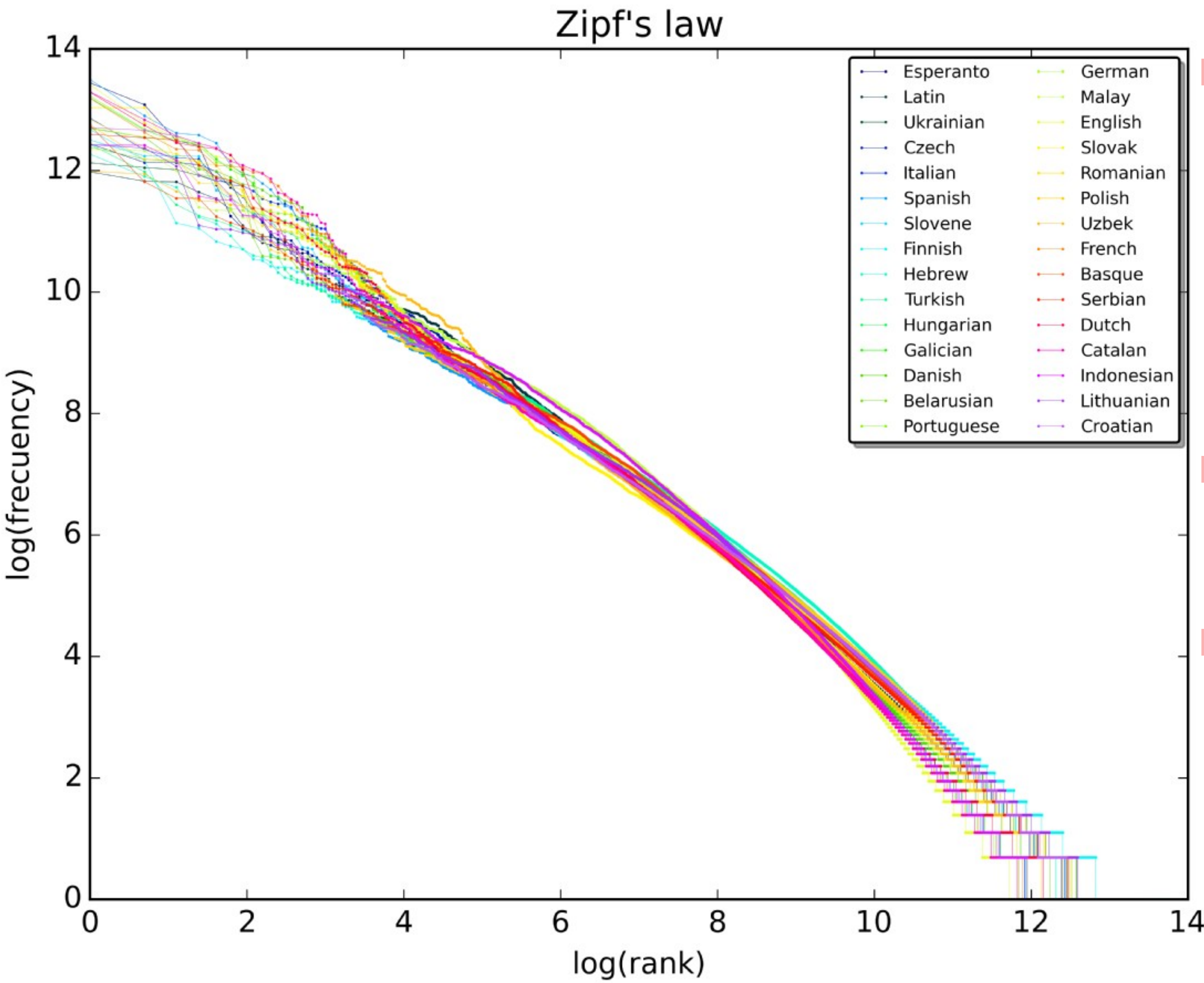
Exemplo da Zipf



- Escala log-log (k versus p_k)

- n = 10
- Diferentes valores para *a* (s na figura)
- Maior s, menor é a cauda, pois distribuição *decresce* mais rapidamente
- Peso da cauda é inversamente proporcional ao expoente

Wikipedia (real)



- Frequência de palavras no Wikipedia em diferentes idiomas
- Idiomas são parecidos
- Muito bem representado por distribuição de Zipf

Bíblia (real)

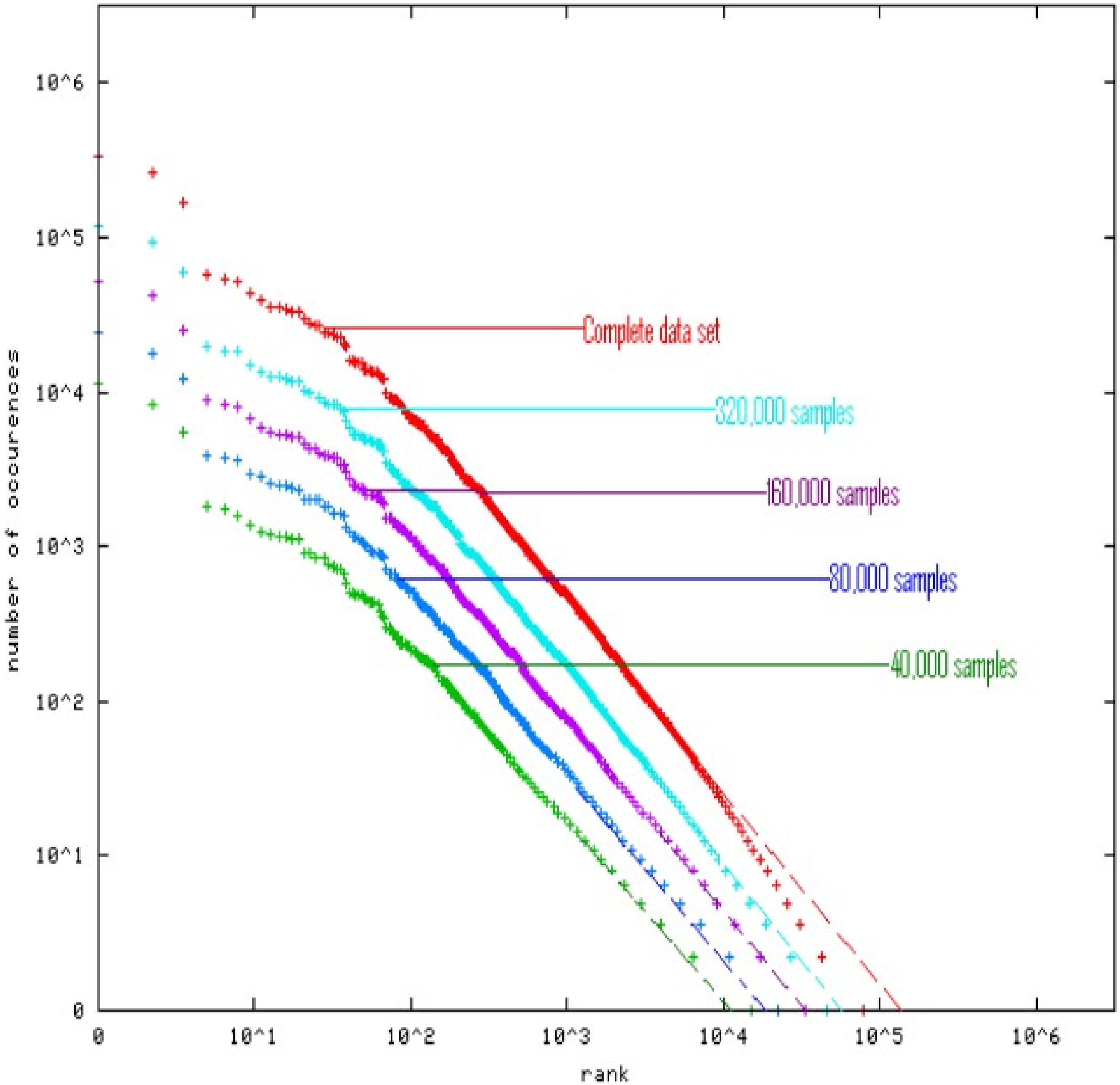


Figure 4: Word frequency in the Bible (including both the Old and the New Testament).

- Frequência de palavras na Bíblia para diferente número de amostras
- amostras aleatórias de palavras (preservam a forma)