Lista de Exercícios

Disciplina de Redes Complexas - PESC - COPPE - UFRJ

Vinícius W. Salazar, Prof. Daniel R. Figueiredo

Novembro de 2019

Questão 1

1.1) $q_k = (\frac{1-q}{n})^k$, pois as combinações de palavras com k letras é dada por n^k, e a chance de teclar alguma tecla que não seja o espaço é de $\frac{1-q}{n}$, se substituirmos n por $\frac{1-q}{n}$ temos q_k. Esse valor descresce monotonicamente em k pois $0 \leq \frac{1-q}{n} \leq 1$, logo se elevarmos esse valor a k, temos que $q_k \to 0 \mid k \to \infty$

1.2) (n^k), pois esse é o número de combinações possíveis de uma palavra de k letras em um alfabeto de tamanho n. Por exemplo, se temos $k = 1$, a chance é de $n^1 = n$ pois temos n possíveis palavras. Para $k = 2$, teremos $n \times n$, $k = 3$, $n \times n \times n$, e assim por diante.

1.3) Tomemos $k = 1$. O nosso ranqueamento terá n^1 palavras, e o valor do ranqueamento de $k = 2$ será de $n^1 + 1$. Para $k = 3$, esse valor será $n^1 + n^2 + 1$. Logo o valor a primeira palavra com k letras será $n^1 + \ ... + \ n^k + 1$, podemos anota-lo como

$$R_k = (\sum_{k=1}^{k} \prod_{k=1}^{k} n) + 1$$

.

1.4) k

1.5) Esse valor seria igual a q_k.

Questão 2

Dada a seguinte função de densidade para uma distribuição de Pareto:

$$f_X(x) = \frac{\alpha x_0^\alpha}{x^{\alpha-1}}, \ a > 0, \ x_0 > 0$$

A função de *likelihood* da distribuição de Pareto para uma amostra $X = (x_1, ..., x_n)$ é dada por:

$$L(x_1, ..., x_n \mid \alpha) = \prod_{i=1}^{n} f_X(x_i) = \prod_{i=1}^{n} \frac{\alpha x_0^\alpha}{x_i^{a+1}}$$

que ainda pode ser simplificada para:

$$\alpha^n x^{n\alpha} \prod_{i=1}^{n} \frac{1}{x_i^{\alpha+1}}$$

Sendo sua função *log*:

$$\ell(x_1, ..., x_n \mid \alpha) = n \ln(\alpha) + n\alpha \ln(x) - (\alpha + 1) \sum_{i=1}^{n} \ln$$

Para um dado α, fazemos a derivada de ℓ em função de α igual a 0:

$$\frac{\partial \ell(\alpha)}{\partial \alpha} = \frac{n}{\alpha} + n\ln(x) - \sum_{i=1}^{n} \ln(x_i) = 0$$

Encontrando a raíz, temos $\hat{\alpha}$ como uma v.a. condicionada em nossas amostras, temos o seu valor como:

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(x_i) - n\ln(\hat{x})}$$

.

Questão 3

O coeficiente de clusterização de um vértice v_i é dado por

$$C_i = \frac{2e_i}{d_i(d_i-1)}$$

, aonde e_i é o número de arestas entre os vizinhos de v_i e d_i é o grau de v_i. Se no modelo $G(n,p)$ as arestas aparecem com probabilidade p, temos

$$e_i = p \times \frac{d_i(d_i-1)}{2}$$

, aonde p é a probabilidade de um par acontecer e o segundo termo é o número de vizinhos do vértice i com grau d_i.

Se substituirmos e_i na equação do coeficiente de clusterização, temos:

$$C = p \times \frac{d_i(d_i-1)}{d_i(d_i-1)} = p$$

Questão 8

8.1) Em um grafo do modelo Watts-Strogatz, o coeficiente de clusterização $C(0) \mid p = 0$ não depende de N, somente de K. Sendo que $C_i = \frac{E_i}{\binom{d_i}{2}}$, $C(0) = \frac{3(K-1)}{2(2K-1)} \approx \frac{3}{4}$.

A distância média $\ell(0)$ depende de N e de K:

$$\ell = \frac{\sum_{u,v \in V} d(u,v)}{\binom{N}{2}}$$

quanto mais o grafo cresce, se não houverem "atalhos", a distância média vai crescer também. Para um K genérico,

$$\ell(0) \approx \frac{N}{4K}$$

(exibe um crescimento linear com N).

8.2) O modelo WS é similar ao modelo G(n,p) e possui propriedades estruturais equivalentes. No entanto, difere quanto a clusterização local e distância. Enquanto no G(n,p) estas são dadas, respectivamente, por $C = p$ e $\ell \approx \frac{\log(n)}{\log(d)}$. Enquanto no G(n,p) teremos um valor esperado de grau $E[D] = (n-1)p$, no modelo WS esse valor esperado será $E[D] = 2K$. Logo dado N e K, $m = NK$ e a densidade p será de $p = \frac{2K}{N-1} \approx \frac{2K}{N}$, logo $C(1) \approx \frac{2K}{N}$, decrescendo linearmente com N, e $\ell(1) = \frac{\log(N)}{log(2K)}$, crescendo logaritmicamente com N.

Com isso, a distância média decresce rapidamente a medida que $p \to 1$, enquanto a clusterização decresce mais devagar. Quando $p = 0$, $\ell(0) \approx \frac{N}{2K}$ e para $p = 1$, $\ell(1) = \frac{\log(N)}{log(2K)}$.

Como os valores de $E[D]$, C e ℓ serão parecidos no G(n,p) e no WS(N,K) onde $p = 1$, o G(n,p) serve como uma boa aproximação. Além disso, o modelo G(n,p) vai possuir distribuição de graus em cauda pesada, enquanto o WS vai ter uma distribuição mais aproximada da binomial.