

PoCS Assignment 4

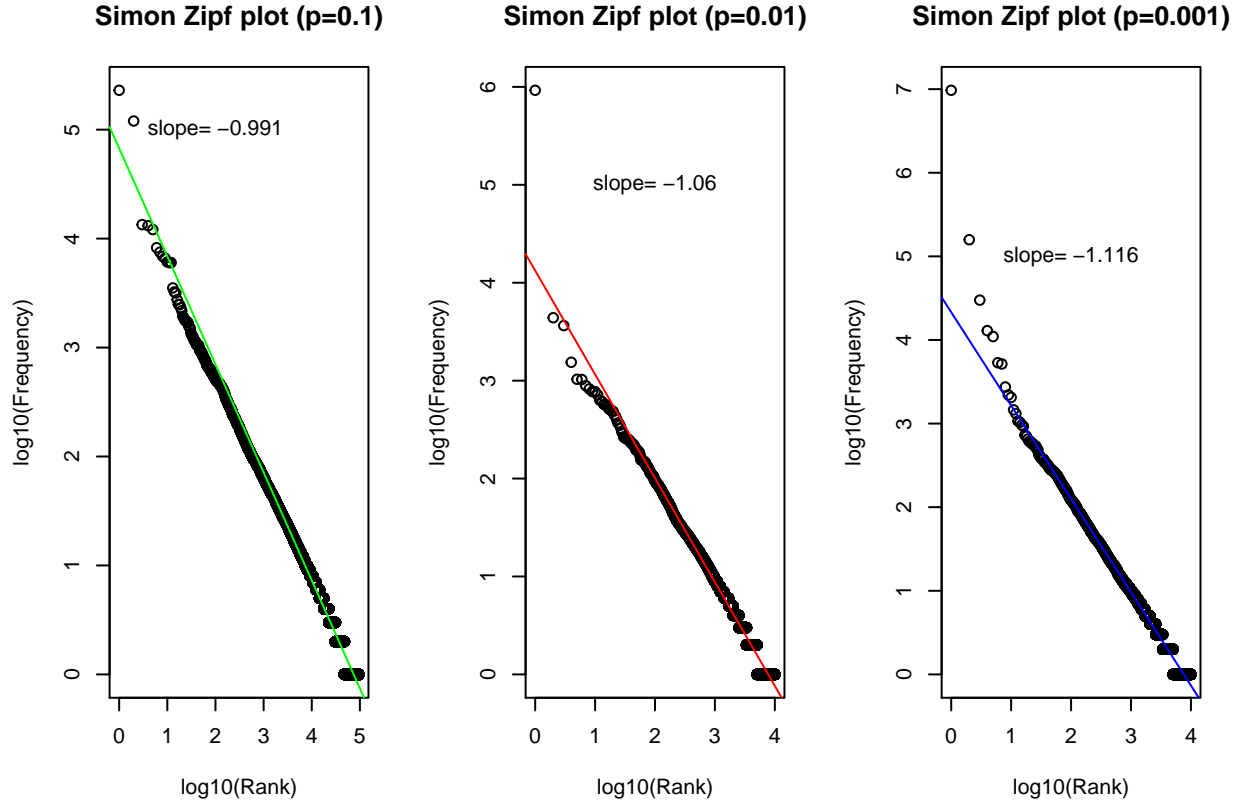
P. Alexander Burnham

September 25, 2018

PoCS Assignment github repo: <https://github.com/alexburn17/BurnhamPoCS>

Worked with: Yu Han, Edison & Kewang

Problem 1:



Problem 2:

a)

Given: $\frac{n_k}{n_{k+1}} = \frac{(k-1)(1-\rho)}{1+(1-\rho)k}$, $z = (1-\rho)$, $n_1 = \frac{\rho}{2-\rho}$

Gamma Function says: $\Gamma(k) = (k-1)!$

From Video:

Step 1: write difference equation for an idea of what is going on

$$n_k = \left[\frac{(k-1)(1-\rho)}{1+(1-\rho)k} \right]$$

$$n_{k-1} = \left[\frac{(k-2)(1-\rho)}{1+(1-\rho)(k-1)} \right]$$

$$n_{k-2} = \left[\frac{(k-3)(1-\rho)}{1+(1-\rho)(k-2)} \right]$$

etc.

Step 2: Denominator sub in z for $1-\rho$

$$[(1+zk)(1+zk-1)(1+z)(k-2)\dots(1+2z)(1+z)] \frac{1}{1+z}$$

simplify...

$$z^k \left(\frac{1}{z} + k\right) \left(\frac{1}{z} + k - 1\right) \dots \left(\frac{1}{z} + 1\right) \left(\frac{1}{1+z}\right)$$

$$Denominator = \frac{z^k \Gamma(\frac{1}{z} + k - 1)}{\Gamma(\frac{1}{z} + 1)} * \frac{1}{1+z}$$

Step 3: Numerator sub in z for $1-\rho$ and $\Gamma(k)$ for $(k-1)!$

$$= (k-1)(1-\rho)(k-2)(1-\rho)(k-3)(1-\rho)\dots(1-\rho)$$

simplify...

$$= (1-\rho)^{(k-1)}(k-1)!$$

$$= z^{(k-1)}\Gamma(k)$$

Step 4: put numerator of denominator, multiply by n_1 and simplify

$$= \frac{z^{(k-1)}\Gamma(k)}{\frac{z^k \Gamma(\frac{1}{z} + k - 1)}{\Gamma(\frac{1}{z} + 1)} * \frac{1}{1+z}} * \frac{\rho}{2-\rho}$$

cross out z^k and move z^{-1} to denominator sub in: $\rho = 1 - z$

$$= \frac{\Gamma(k)}{z \frac{\Gamma(\frac{1}{z} + k - 1)}{\Gamma(\frac{1}{z} + 1)} * \frac{1}{1+z}} * \frac{(1-z)}{2-(1-z)}$$

$$= \frac{\Gamma(k)}{z} (1+z) \frac{\Gamma(\frac{1}{z} + 1)}{\Gamma(\frac{1}{z} + k - 1)} * \frac{(1-z)}{2-(1-z)}$$

simplify and sub ρ back in:

$$= \frac{\Gamma(k)\Gamma(\frac{1}{z} + 1)}{\Gamma(\frac{1}{z} + k - 1)} * \frac{\rho}{1-\rho}$$

Use beta function to derive:

$$\beta(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \frac{(x-1)!(y-1)!}{(x+y-1)!}$$

$$n_k = \frac{\rho}{1-\rho} * \beta(k, \frac{1}{1-\rho} + 1)$$

b)

$$n_k = \frac{\rho}{1-\rho} * \beta(k, \frac{1}{1-\rho} + 1)$$

use the fact that $\beta(k, \gamma) = k^{-\gamma}$

$$n_k = \frac{\rho}{1-\rho} * k^{-(\frac{1}{1-\rho}+1)}$$

$$\gamma = \frac{1}{1-\rho} + 1$$

$$\gamma = \frac{2-\rho}{1-\rho}$$

Problem 3:

$$\lim_{\gamma \rightarrow 0} \gamma = \frac{2-\rho}{1-\rho}$$

As ρ approaches 0, γ goes to 2.

$$\lim_{\gamma \rightarrow 1} \gamma = \frac{2-\rho}{1-\rho}$$

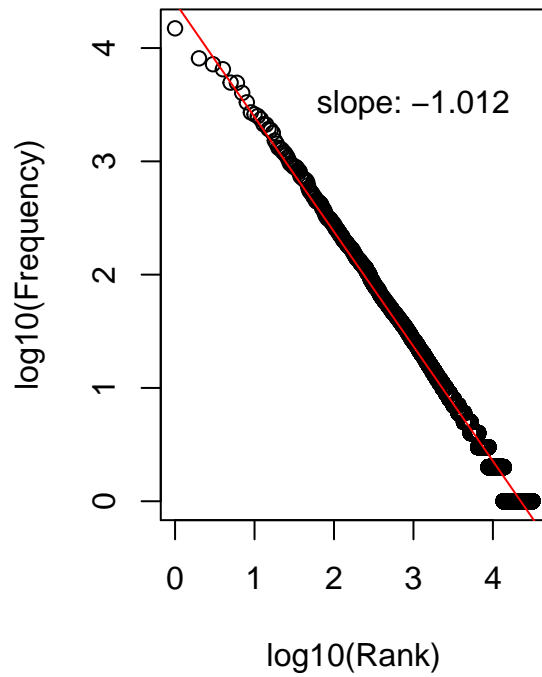
As ρ approaches 1, γ goes to ∞ .

Problem 4:

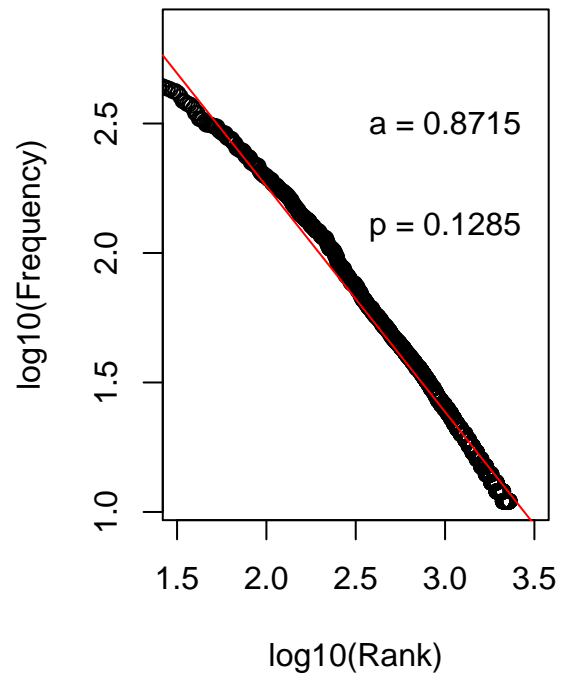
a)

b)

Zipf Plot of Ulysses



Zipf Plot of Ulysses (Trimmed)



c)

Problem 5:

a)

Given: $P_k = ck^{-\gamma}$

Step 1: find c

$$\int_{k_{min}}^{\infty} ck^{-\gamma} = 1$$

$$c \frac{1}{-\gamma + 1} k^{(-\gamma+1)} \Big|_{k_{min}}^{\infty} = 1$$

$$c \frac{1}{\gamma - 1} k_{min}^{(-\gamma+1)} = 1$$

find c:

$$c = (\gamma - 1) k_{min}^{(\gamma-1)}$$

Step 2: find k_{max}

$$\sum_{min k_{max}}^{\infty} ck^{-\gamma} = \frac{1}{N}$$

$$\int_{k_{max}}^{\infty} ck^{-\gamma} = \frac{1}{N}$$

$$\int_{k_{max}}^{\infty} k^{-\gamma} = \frac{1}{cN}$$

Find antiderivative

$$\frac{1}{\gamma - 1} k_{max}^{-\gamma+1} = \frac{1}{cN}$$

multiply both sides by $(\gamma - 1)$

$$k_{max}^{-\gamma+1} = \frac{\gamma - 1}{cN}$$

$$k_{max} = \left(\frac{\gamma - 1}{c} \right)^{\left(\frac{1}{1-\gamma} \right)} \frac{1}{N^{\left(\frac{1}{1-\gamma} \right)}}$$

Step 3: sub in C and simplify:

$$k_{max} = \left(\frac{\gamma - 1}{(\gamma - 1)k_{min}^{(\gamma-1)}} \right)^{\left(\frac{1}{1-\gamma} \right)} \frac{1}{N^{\left(\frac{1}{1-\gamma} \right)}}$$

cancel terms:

$$k_{max} = k_{min} \frac{1}{N^{\left(\frac{1}{1-\gamma} \right)}}$$

Answer:

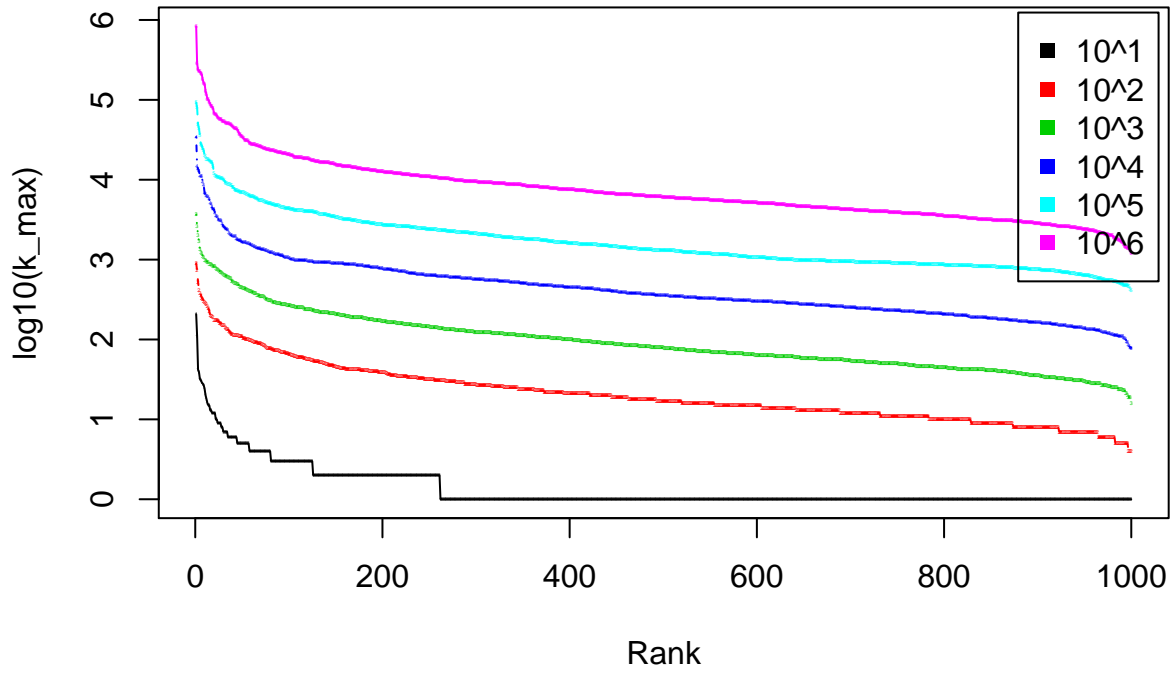
$$k_{max} = k_{min} N^{\left(\frac{1}{\gamma-1} \right)}$$

b)

Problem 6:

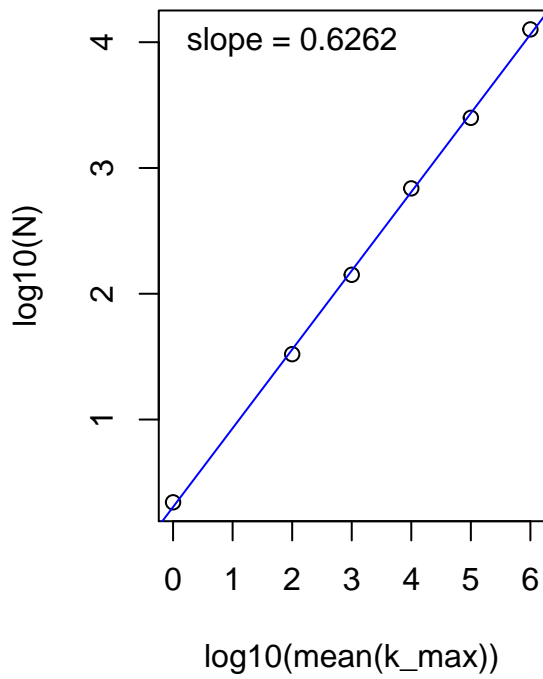
a)

K_max by N

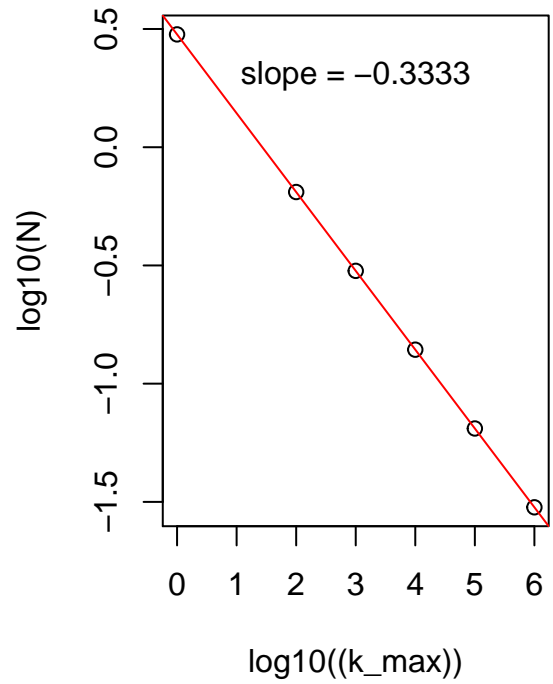


b)

K_max estimate by N



K_max theoretical by N



My theoretical value and empirical values do not match up. The one I calculated by sampling a powerlaw distribution has a positive slope of around 0.6. This makes sense as the max k should increase as the sample size increases. However, my derivation of the theoretical max k has a negative slope of -0.3.