PoCS Assignment 3

P. Alexander Burnham September 18, 2018

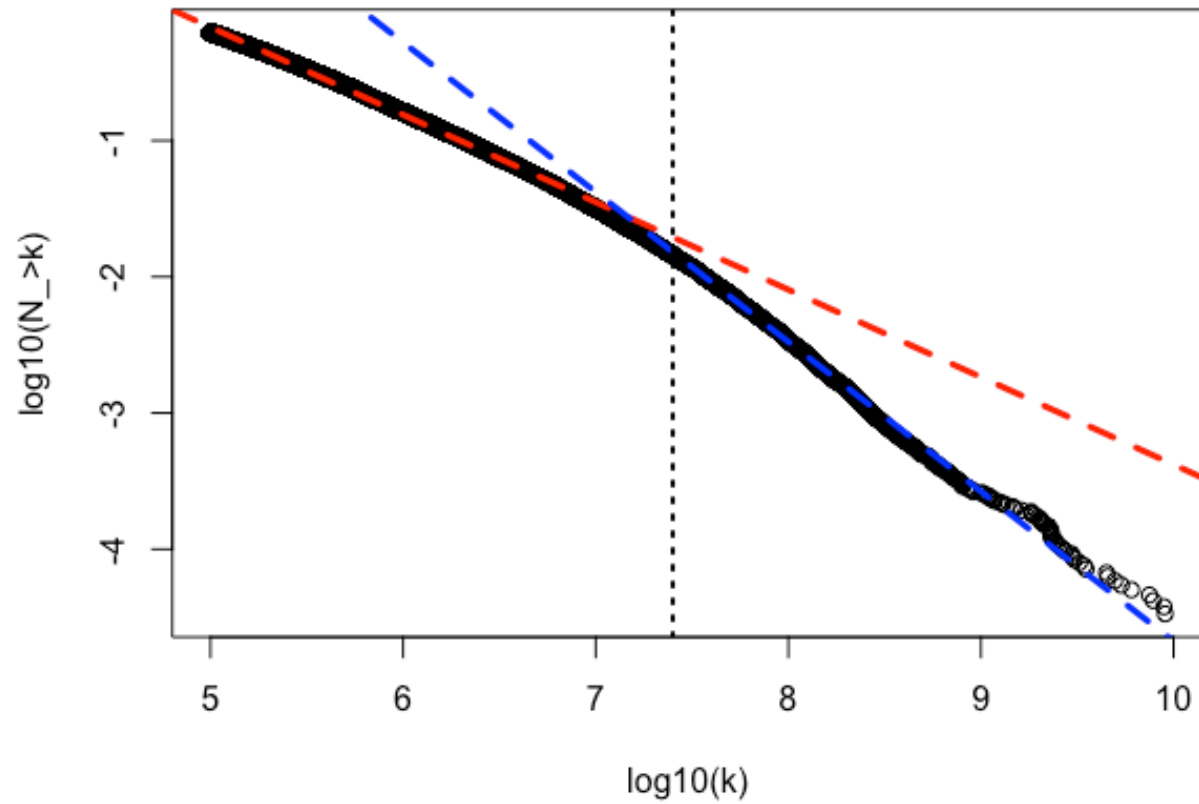PoCS Assignment github repo: https://github.com/alexburn17/BurnhamPoCS

Worked with: Sam and Edison

Problem 1:

ccdf in log-log space
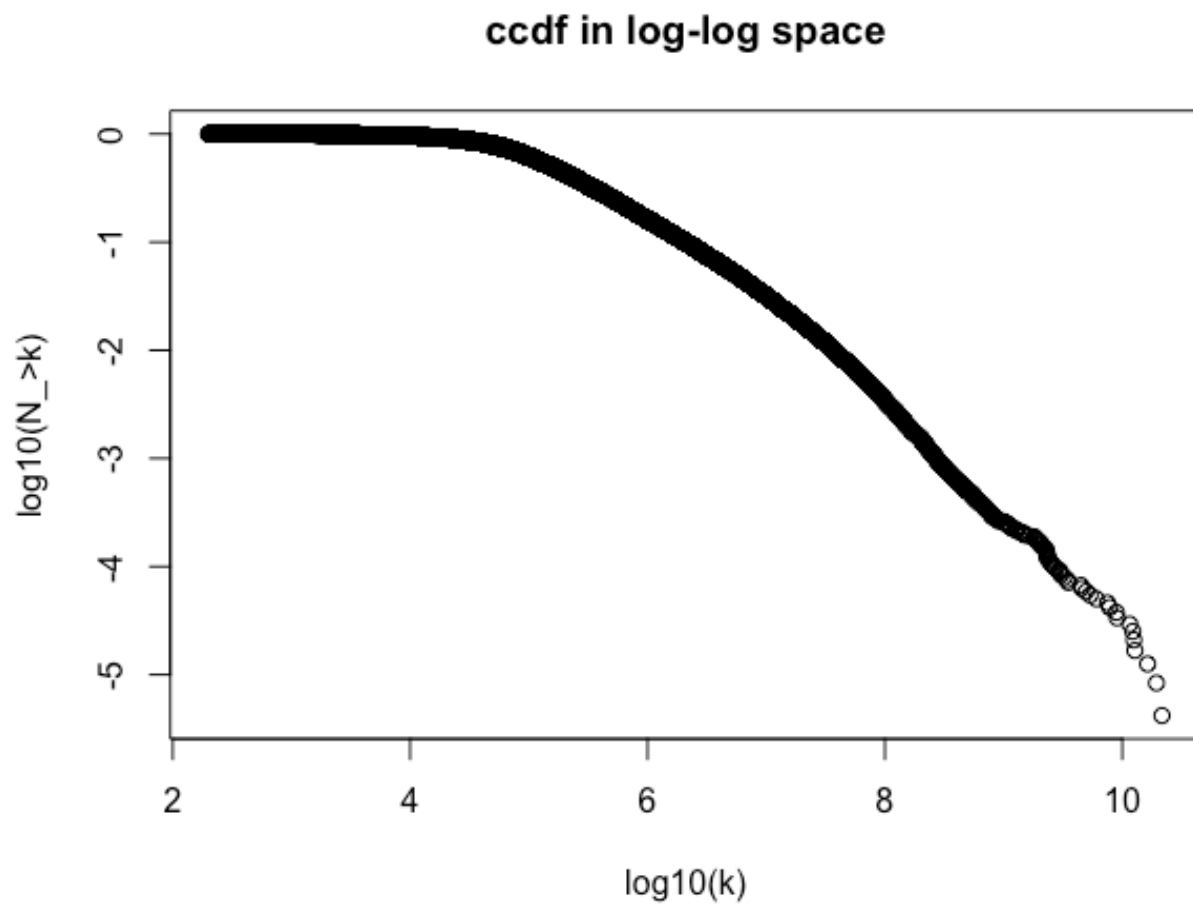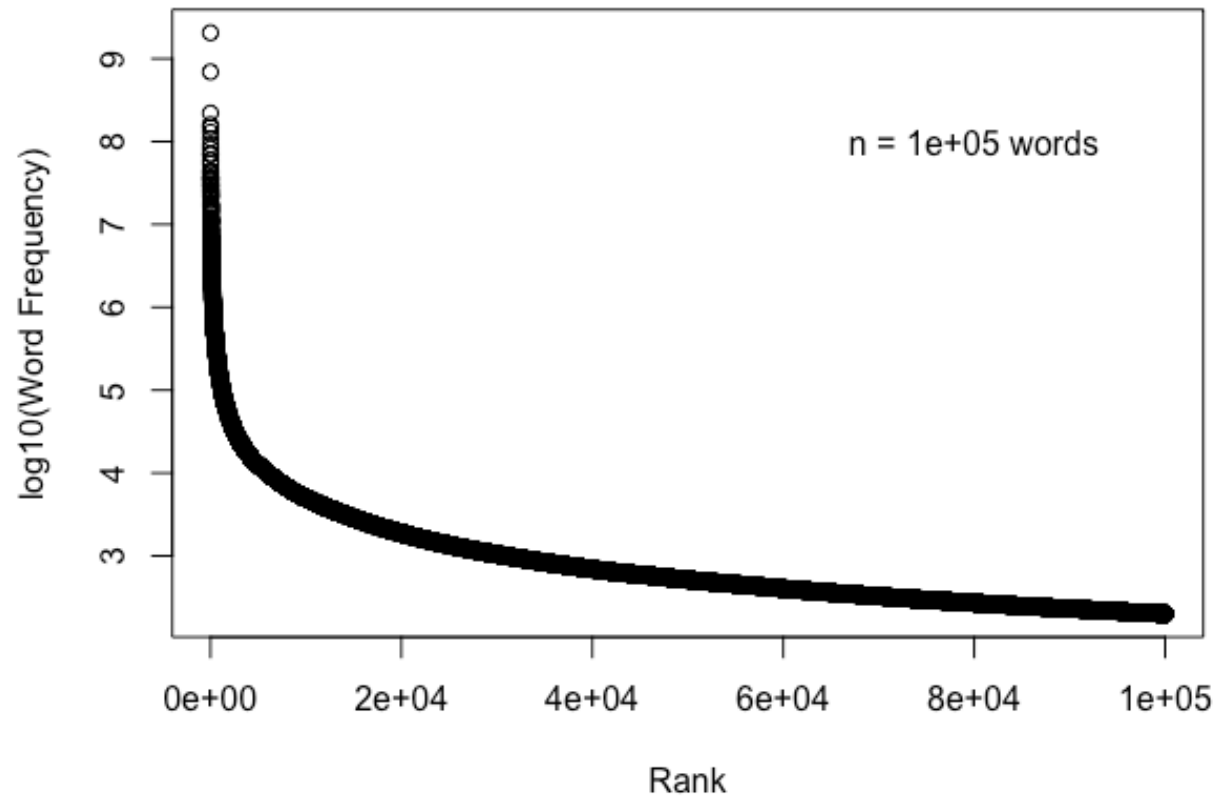
Problem 2:

Trimmed CCDF - Two Scaling Regimes

Gamma values after subtracting one and taking absolute values with 95% CI

- γ_{red} = 1.6423224, 95% CI [1.6421033 1.6425415]
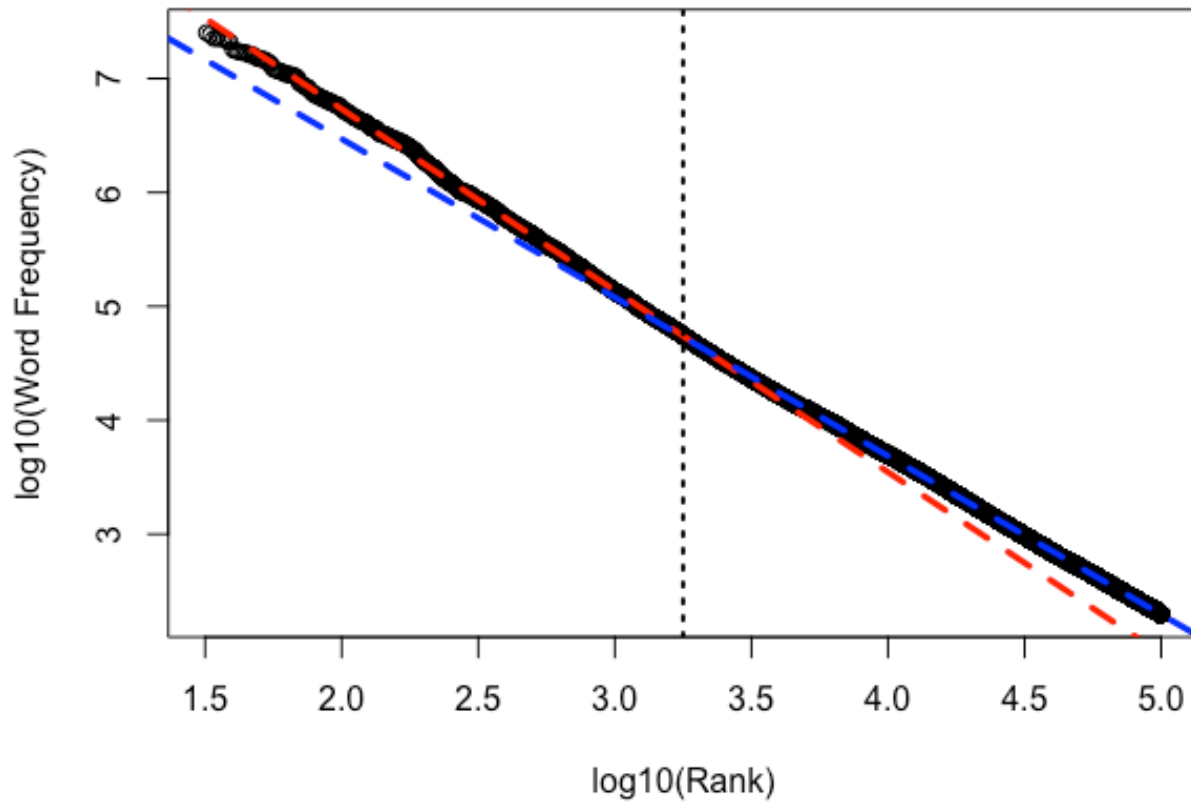- γ_{blue} = 2.099329, 95% CI [2.09735, 2.101307]

Problem 3:

Word Frequency by Rank

Problem 4:

Word Frequency by Rank Trimmed (log-log)

Alpha values after taking absolute value with 95% CI

- α_{red} = 1.593627, 95% CI [1.591714, 1.59554]
- α_{blue} = 1.391344, 95% CI [1.391236, 1.391452]

Problem 5:

Measurements for γ and α for 2 scaling regimes from empirical data

- γ_{red} = 1.6423224, 95% CI [1.6421033 1.6425415]
- γ_{blue} = 2.099329, 95% CI [2.09735, 2.101307]

- α_{red} = 1.593627, 95% CI [1.591714, 1.59554]
- α_{blue} = 1.391344, 95% CI [1.391236, 1.391452]

$$\alpha_{expected} = \frac{1}{1 - \gamma_{red}}$$

$$\alpha_{expected} = 1.556851$$

Very close to γ_{red} of 1.64

$$\alpha_{expected} = \frac{1}{1 - \gamma_{blue}}$$

$$\alpha_{expected} = 0.9096458$$

Far from α_{blue} of 1.39

This seems to indicate that the scaling regime in the red region holds true to Zipf's Law however, the second half of the regime (blue region) varied from this relationship. This could be due to sampling only 100,000 observations from a data set of 14 million and from eye-balling regions of different scalings.

Problem 6:

$$N(0, 2k, 2n) = \frac{2n}{n+k} = \frac{2n}{n-k}$$

1) Sterling's approximation: $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n$

$$\binom{2n}{n+k} = \frac{2n!}{(n+k)!(n-k)!} = \frac{\sqrt{2\pi 2n}(\frac{2n}{e})^{2n}}{\sqrt{2\pi(n+k)}(\frac{n+k}{e})^{n+k}\sqrt{2\pi(n-k)}(\frac{n-k}{e})^{n-k}}$$

2) Simplify this expression (cancel $\sqrt{2\pi}$ from numerator and denominator):

$$= \frac{\sqrt{2\pi}(2n)^{1/2}(2n)^{(2n)}e^{-(2n)}}{\sqrt{2\pi}(n+k)^{1/2}(n+k)^{(n+k)}e^{-(n+k)}\sqrt{2\pi}(n-k)^{1/2}(n-k)^{(n-k)}e^{-(n-k)}}$$

$$= \frac{(2n)^{1/2}(2n)^{(2n)}e^{-(2n)}}{\sqrt{2\pi}(n+k)^{1/2}(n+k)^{(n+k)}e^{-2n}(n-k)^{1/2}(n-k)^{(n-k)}e^{-(n-k)}}$$

3) Combine and cancel e terms:

$$= \frac{2^{2n+1/2}n^{2n+1/2}e^{-2n}}{\sqrt{2\pi}(n+k)^{1/2+n+k}(n-k)^{n-1/2-k}e^{-2n)}}$$

$$= \frac{2^{2n+1/2}n^{2n+1/2}}{\sqrt{2\pi}(n+k)^{1/2+n+k}(n-k)^{n-1/2-k}}$$

4) rearrange and simplify bold terms to $n^{1/2}$

$$= \frac{2^{2n+1/2}\mathbf{n^{2n+1/2}}}{\sqrt{2\pi}\mathbf{n^{1/2+n+k}}(1+\frac{k}{n})^{1/2+n+k}\mathbf{n^{1/2+n-k}}(1-\frac{k}{n})^{1/2+n-k}}$$

$$= \frac{2^{2n+1/2}}{\sqrt{2\pi}n^{1/2}(1+\frac{k}{n})^{1/2+n+k}(1-\frac{k}{n})^{1/2+n-k}}$$

5) simplify and rearrange

$$= \frac{2^{2n+1/2}}{\sqrt{2\pi}n^{1/2}(1-\frac{k^2}{n^2})^{1/2+n}(1-\frac{k}{n})^{-k}(1+\frac{k}{n})^k}$$

6) take the natural log of the bolded

$$= \frac{2^{2n+1/2}}{\sqrt{2\pi}n^{1/2}} * \frac{1}{\mathbf{(1-\frac{k^2}{n^2})^{1/2+n}(1-\frac{k}{n})^{-k}(1+\frac{k}{n})^k}}$$

$$= \frac{2^{2n+1/2}}{\sqrt{2\pi}n^{1/2}}e^{\ln\frac{1}{(1-\frac{k^2}{n^2})^{1/2+n}(1-\frac{k}{n})^{-k}(1+\frac{k}{n})^k}}$$

$$= \ln1 - (1/2+n)\ln(1-\frac{k^2}{n^2}) - k\ln(1-\frac{k}{n}) + k\ln(1+\frac{k}{n})$$

7) apply taylor series expansion $\ln(1+z) \simeq z$ as n gets large k goes to infinity, bold expressions cancel

$$= 0 - (1/2+n)\frac{k^2}{n^2} \mathbf{-k\ln(1-\frac{k}{n}) + k\ln(1+\frac{k}{n})}$$

$$= -(1/2+n)\frac{k^2}{n^2}$$

8) rearrange and $1/2n$ goes to 0 as n becomes very large, n cancels with n^2 in denominator

$$= -n(1/2n+1)\frac{k^2}{n^2}$$

$$= -\frac{k^2}{n}$$

9) take natural log and substitute back into original expression

$$= \frac{2^{2n+1/2}}{\sqrt{2\pi}n^{1/2}}e^{-\frac{k^2}{n}}$$

10) Given: $n = t/2$ & $k = x/2$. Substitute in for n and k and simplify

$$= \frac{2^{2(t/2)+1/2}}{\sqrt{2\pi}(t/2)^{1/2}} e^{-\frac{(x/2)^2}{(t/2)}}$$

11) cancel $\sqrt{2}$

$$= \frac{2^{2t+1}}{\sqrt{2\pi t}} e^{-\frac{(x)^2}{(2t)}}$$

12) replace numerator with above expression and assume $2^t I(t = 2n) = 1/2$ and cancel numerator

$$= \frac{\binom{2n}{n-k}}{2^t} I(t = 2n)$$

$$= \frac{\frac{2^{2t+1}}{\sqrt{2\pi t}} e^{-\frac{(x)^2}{(2t)}}}{2^t}(1/2)$$

13) Answer

$$Pr(x_t \equiv x) \simeq \frac{1}{(\sqrt{2\pi t})^{e-\frac{x^2}{2t}}}$$

Problem 7:

1) set up problem

- displacement = j - i
- P = # positive steps
- N = # negative steps
- P + N = t
- p - N = displacement

Find....

$$\binom{t}{P} = number\ of\ random\ walks$$

2) solve linear set of equations by addition

$$P + N = t$$

$$P - N = j - i$$

$$2P = t + j - i$$

3) divide both sides by 2

$$P = \frac{t + j - i}{2}$$

4) t choose p is equal to...

$$\binom{t}{\frac{t+j-i}{2}} = number\ of\ random\ walks$$

Problem 8:

1)

$$N_{fr} = N(1, 1, 2n - 2) - N(-1, 1, 2n - 2)$$

2) plug into $\binom{t}{\frac{t+j-i}{2}}$, i=first number in parens, j=second and t=third

$$= \frac{(2n - 2)}{\frac{2n-2+1-1}{2}} - \frac{(2n - 2)}{\frac{2n-2+1+1}{2}}$$

$$= \binom{2n - 2}{n - 1} - \binom{2n - 2}{n}$$

3) this equals...

$$= \frac{(2n - 2)!}{(n - 1)!(n - 1)!} - \frac{(2n - 2)!}{n!n!}$$

4) simplify...

$$= \frac{(2n - 2)!n!n!}{(n - 1)!(n - 1)!n!n!} - \frac{(2n - 2)!(n - 1)!(n - 1)!}{(n - 1)!(n - 1)!n!n!}$$

$$= \frac{(2n - 2)!n!n! - (2n - 2)!(n - 1)!(n - 1)!}{(n - 1)!(n - 1)!n!n!}$$

$$= \frac{(2n-2)!}{(n-1)!n!}$$

5) apply sterling's approximation:

$$= \frac{\sqrt{2\pi(2n-2)}(\frac{2n-2}{e})^{2n-2}}{\sqrt{2\pi(n)}(\frac{n}{e})^n \sqrt{2\pi(n-1)}(\frac{n-1}{e})^{n-1}}$$

6) simplify (cancel $\sqrt{2\pi}$ and $(n-1)^{1/2}$ and e terms)

$$= \frac{\sqrt{2\pi}2^{1/2}(n-1)^{1/2}(2n-1)^{(2n-1)}e^{-2n}}{\sqrt{2\pi}n^{1/2}n^n e^{-n}\sqrt{2\pi}(n-1)^{1/2}(n-1)^{(n-1)}e^{-(n-1)}}$$

$$= \frac{2^{1/2}(2n-1)^{(2n-1)}e}{\sqrt{2\pi}n^{1/2}n^n(n-1)^{(n-1)}}$$

7) cancel $(n-1)^{(n-1)}$ and combine terms

$$= \frac{2^{1/2}2^2 n^{(n-1)}(n-1)^{(n-1)}(n-1)^2 e}{\sqrt{2\pi}n^{1/2}n^n(n-1)^{(n-1)}}$$

$$= \frac{2^{1/2+2}n^{(n-1)}(n-1)^2 e}{\sqrt{2\pi}n^{1/2+n}}$$

8) simplify to...

$$= \frac{2^{2n-1}e}{\sqrt{2\pi}}(1-1/2)^n$$

9) as n gets large $(n-1/2)^n$ goes to 1 and we can cancel constants and set proportional

$$= \frac{2^{2n-1}e}{\sqrt{2\pi}}(n-1/2)^n$$

$$= \frac{(2n-2)!}{(n-1)!n!} \propto \frac{2^{2n-1}}{\sqrt{2}}(1)$$

As n goes to infinity, the numerator of the leading term becomes very large and the denominator becomes very small.