

# PoCS Assignment 3

*P. Alexander Burnham*

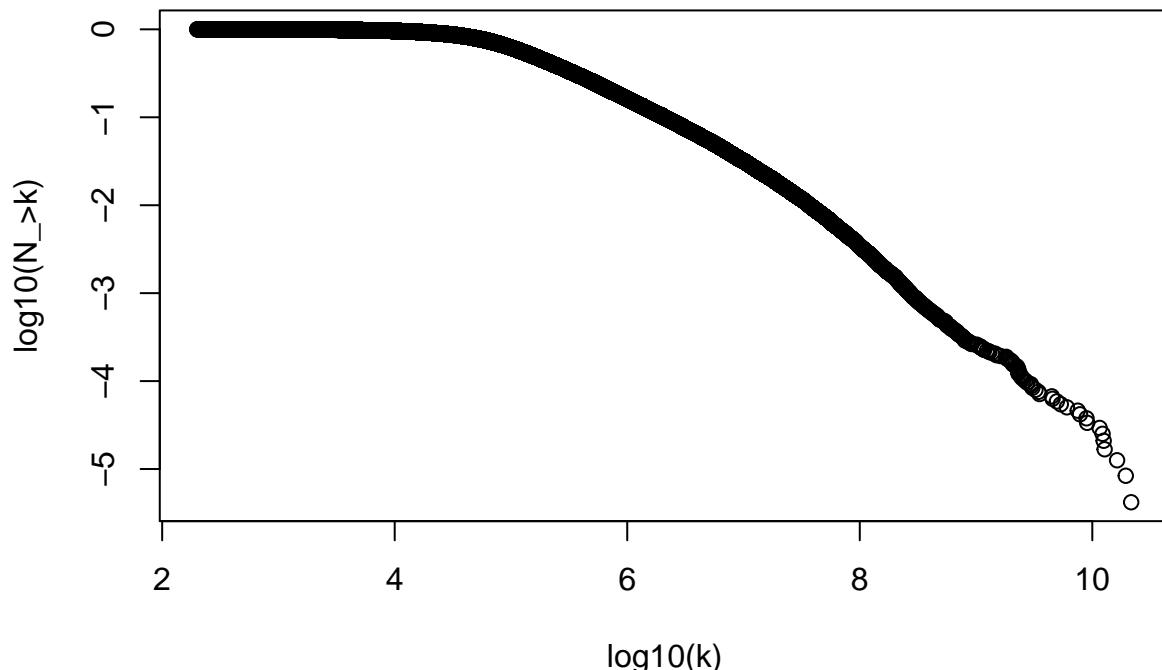
*September 18, 2018*

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

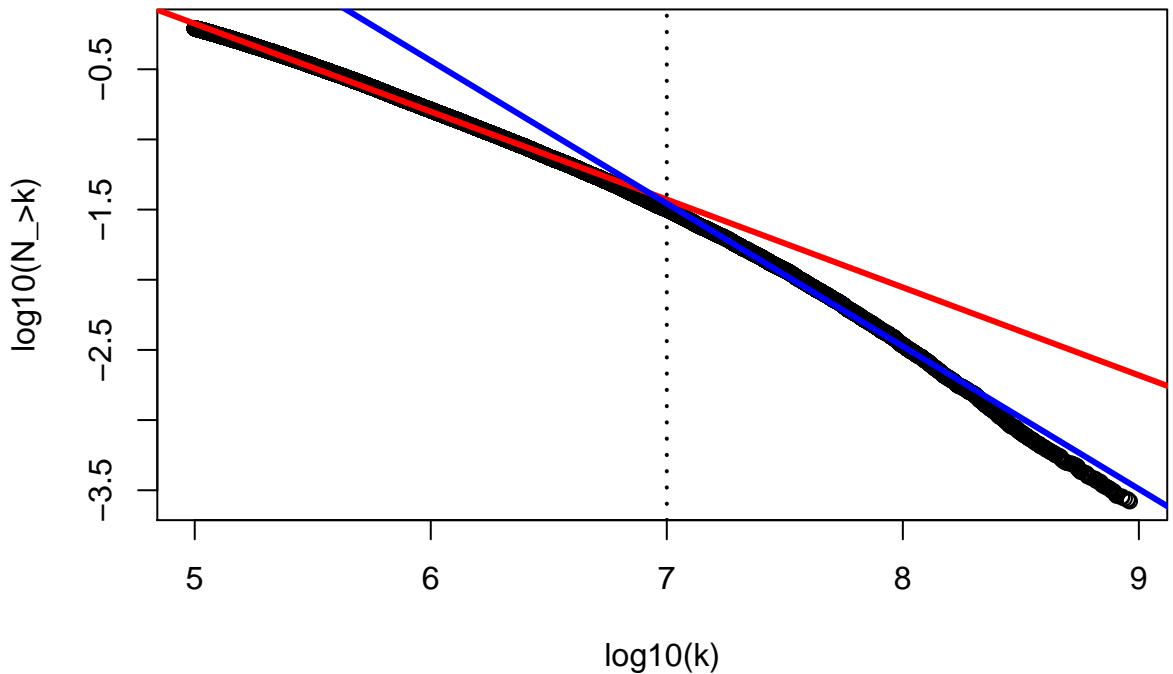
Worked with: NA

**Problem 1:**

**ccdf in log-log space**



## Trimmed CCDF – Two Scaling Regimes



**Problem 2:**

```

## 
## Call:
## lm(formula = dfTrimOne$logN_k ~ dfTrimOne$logk)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -0.069833 -0.010643  0.005324  0.012753  0.018645 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 2.957e+00 5.121e-04  5773 <2e-16 ***
## dfTrimOne$logk -6.264e-01 9.055e-05 -6917 <2e-16 ***
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.01654 on 140087 degrees of freedom
## Multiple R-squared:  0.9971, Adjusted R-squared:  0.9971 
## F-statistic: 4.785e+07 on 1 and 140087 DF,  p-value: < 2.2e-16

##                  2.5 %     97.5 %
## dfTrimOne$logk -0.6265445 -0.6261895 

## 
## Call:
## lm(formula = dfTrimTwo$logN_k ~ dfTrimTwo$logk)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -0.143495 -0.022795  0.007437  0.025831  0.038113 
## 
```

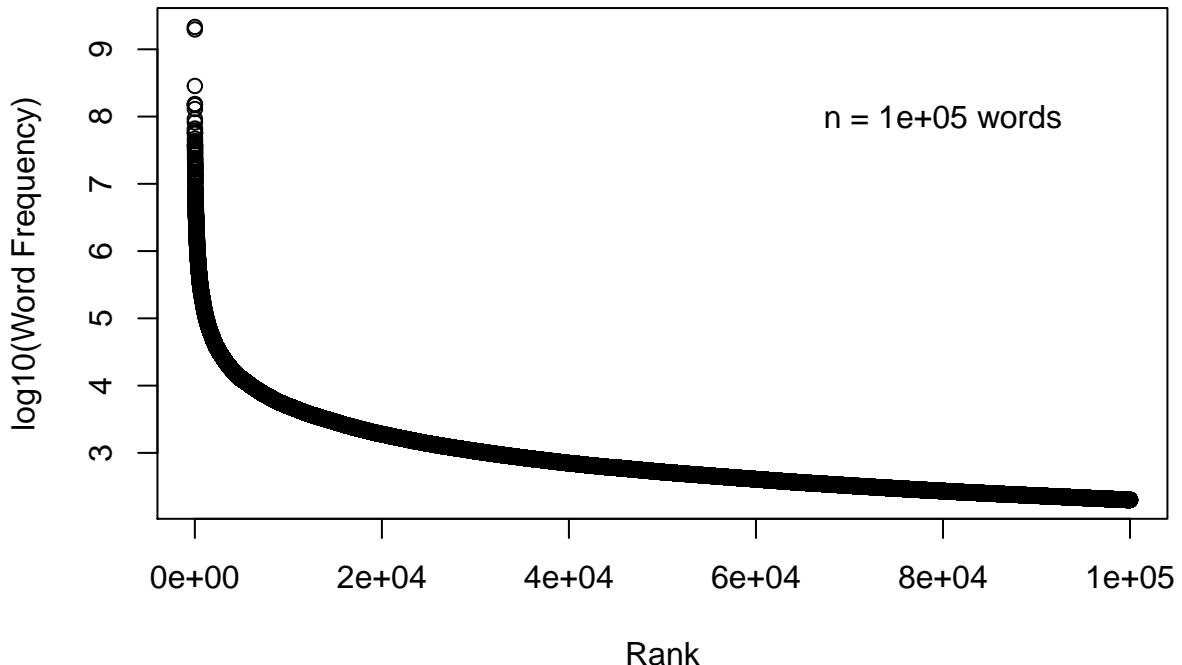
```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.6655094  0.0069849   811.1 <2e-16 ***
## dfTrimTwo$logk -1.0174302  0.0009363 -1086.6 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03105 on 7523 degrees of freedom
## Multiple R-squared:  0.9937, Adjusted R-squared:  0.9937
## F-statistic: 1.181e+06 on 1 and 7523 DF, p-value: < 2.2e-16
##                  2.5 %    97.5 %
## dfTrimTwo$logk -1.019266 -1.015595

```

Problem 3:

### Word Frequency by Rank



Problem 4:

### Word Frequency by Rank Trimmed (log-log)

