

1.- Basics - Zipf's Law

a) Write a program to check Zipf's first law ($f = K/r$) on a real corpus: Count word frequencies, sort them by rank, and plot the curve.

b) Compute the proportionality constant (K) between rank and frequency for each word. Compute its average and deviation. Discuss the results. Are they consistent with Zipf's Law?

NOTE: Use the text files `corpus/en.txt` and `corpus/es.txt`