

11

Paper gives word counts, word type counts

884647 words,

14376 only once

31534 word types

Q: if we saw another 884647 words,  
how many new words would there  
be?

A:  $t = \frac{\# \text{ of words in new}}{\# \text{ in ref}}$

$= 1$

$\Delta(t) = 11430 = \Delta(1)$

$\text{Var}(\Delta(i))$  by assuming the  $n_x$   
are indep, poisson

$\text{Var}\{\Delta(i)\} \approx \sum_{i=1}^{\infty} n_i t^i = 31534$

std = 178