Paper gives word counts, word type counts

884,647 words,
14,376 only once
31,534 word types

Q: If we saw another 884,647 words, how many new words would there be?

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

\[ = 1 \]

\[ \Delta(t) = 11,430 = \Delta(1) \]

\[ \text{Var}(\Delta(1)) \text{ by assuming the } n_x \text{ are independent, Poisson} \]

\[ \text{Var} \{ \Delta(1) \} \approx \sum_{i=1}^{8} n_i t^i = 31,534 \]

\[ \text{std} = 178 \]