

Special Case of N-Gram LM: Unigram LM

- Generate text by generating each word INDEPENDENTLY
- $p(w_m | w_1, \dots, w_{m-n+1}, \dots, w_{m-1}) = p(w_m)$: History didn't matter!
- How to estimate a unigram LM?
 - Text data: d
 - Maximum Likelihood estimator:

$$p_{ML}(w | d) = \frac{c(w, d)}{\sum_{w' \in V} c(w', d)}$$

Count of word w in d

Total Counts of all words in d