

How to Estimate $p(w | d)$?

- Simplest solution: Maximum Likelihood Estimator
 - $P(w | d)$ = relative frequency of word w in d
 - What if a word doesn't appear in the text? $P(w | d)=0$
- In general, what probability should we give a word that has not been observed? Smoothing!