Estimation of Unigram LM

Unigram LM p(w|θ)=?   **Estimation**   **Text Mining Paper d**

Total #words=**100**

10/100 → **text ?**

5/100 → **mining ?**

3/100 → **association ?**

3/100 → **database ?**

…

1/100 → **query ?**

…
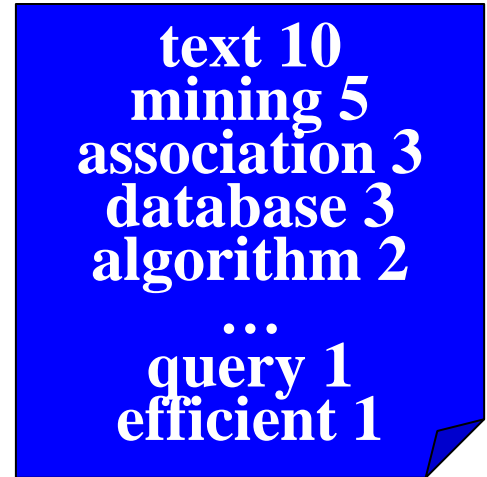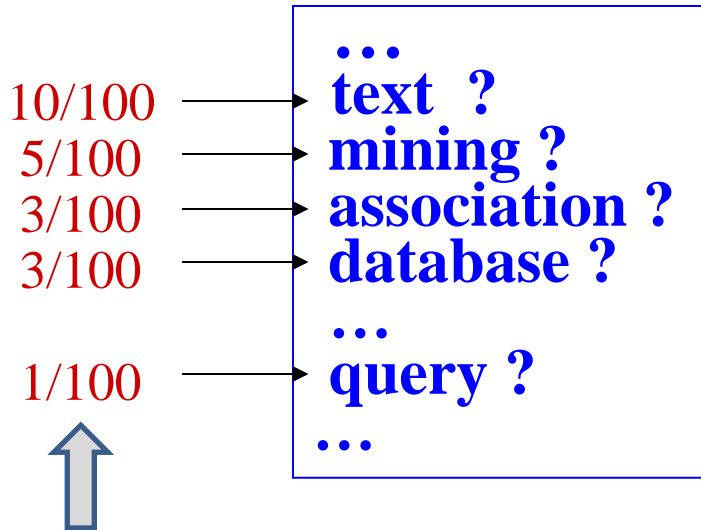
Maximum Likelihood Estimate

text 10
mining 5
association 3
database 3
algorithm 2
…
query 1
efficient 1

Is this our best estimate?
How do we define "best"?

6