

Worked Example: Computing $P(\text{cat} \mid \text{the})$

Step 1: Count in Corpus

"...the cat sat...the dog ran...the cat jumped..."

Step 2: Count Bigrams

count("the")

= 3

count("the cat")

= 2

Step 3: Apply MLE Formula

$$\begin{aligned} P(\text{cat} \mid \text{the}) &= \text{count(the, cat)} / \text{count(the)} \\ &= 2 / 3 = 0.667 \end{aligned}$$

Result: 66.7% probability

"the" followed by "cat" in 2 out of 3 cases