How to Estimate p(w|d)

p(w|d)

Max. Likelihood Estimate

$$p_{ML}(w \mid d) = \frac{c(w,d)}{|d|}$$

Smoothed LM

$$p(w \mid d) > 0$$
$$even \ \ if \ \ c(w,d) = 0$$

w

Word