

Likelihood Function: $p(d)=?$

$$p(d) = p(\theta_1)p(d | \theta_1) + p(\theta_2)p(d | \theta_2)$$

$$= p(\theta_1) \prod_{i=1}^L p(x_i | \theta_1) + p(\theta_2) \prod_{i=1}^L p(x_i | \theta_2)$$

$d = x_1 x_2 \dots x_L$

the 0.000001

the 0.03

Topic
Choice

How can we generalize it to include k topics/clusters?

$d \leftarrow p(\theta_1 | z)$

we 0.01
food 0.003
...
text 0.000006

L