

Given all the parameters, infer the distribution a word is from...

Is “**text**” more likely from θ_d or θ_B ?

From θ_d ($Z=0$)?

$p(\theta_d)p(\text{“text”}|\theta_d)$

From θ_B ($Z=1$)?

$p(\theta_B)p(\text{“text”}|\theta_B)$

$P(w|\theta_d)$

text 0.04
mining 0.035
association 0.03
clustering 0.005
...
the 0.000001

θ_d

$p(w|\theta_B)$

the 0.03
a 0.02
is 0.015
we 0.01
food 0.003
...
text 0.000006

θ_B

$p(\theta_d)+p(\theta_B)=1$

$P(\theta_d)=0.5$

$P(\theta_B)=0.5$

Topic
Choice

$p(z = 0 | w = \text{“text”}) =$

$$\frac{p(\theta_d)p(\text{“text”}|\theta_d)}{p(\theta_d)p(\text{“text”}|\theta_d) + p(\theta_B)p(\text{“text”}|\theta_B)}$$