

Gibbs sampling in LDA: Illustration

i	w_i	d_i	iteration	
			1	2
1	MATHEMATICS	1	2	?
2	KNOWLEDGE	1	2	
3	RESEARCH	1	1	
4	WORK	1	2	
5	MATHEMATICS	1	1	
6	RESEARCH	1	2	
7	WORK	1	2	
8	SCIENTIFIC	1	1	
9	MATHEMATICS	1	2	
10	WORK	1	1	
11	SCIENTIFIC	2	1	
12	KNOWLEDGE	2	1	
·	·	·	·	
·	·	·	·	
·	·	·	·	
50	JOY	5	2	

What's the most likely topic for w_i in d_i ?

How likely would d_i choose topic j ?

How likely would topic j generate word w_i ?

$$P(z_i = j | \mathbf{z}_{-i}, \mathbf{w}) \propto \frac{n_{-i,j}^{(w_i)} + \beta}{n_{-i,j}^{(\cdot)} + W\beta} \frac{n_{-i,j}^{(d_i)} + \alpha}{n_{-i,\cdot}^{(d_i)} + T\alpha}$$

Slide source: Tom Griffiths' presentation at <https://cocosci.berkeley.edu/tom/talks/compling.ppt>