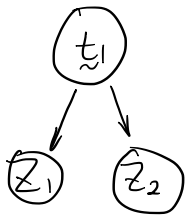
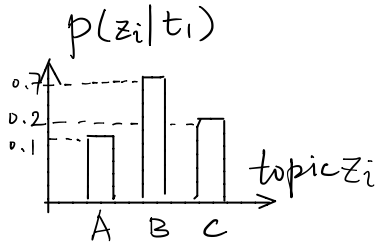
Probabilistic Latent semantic analysis (PLSA)

$$P(w,t) = P(t) \sum_z P(z|t) \cdot P(w|z)$$

w: word, t: document, z: topic

One document example:





Topic z_i for document 1

$z_i \sim \text{Categorical}(t_1)$

e.g., $z_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $z_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ ← A, ← B, ← C

$t_1 = (0.1, 0.7, 0.2)$

[For LDA, $t_1 \sim \text{Dirichlet}(\alpha)$]

$f_{z_i} \rightarrow w_i$

$w_i = \begin{bmatrix} w_i^{(1)} \\ \vdots \\ w_i^{(n)} \end{bmatrix} \in \mathbb{N}^n$ (n: # of vocabulary) $\sim \text{Multinomial}(f_{z_i})$, e.g., $f_{z_i} = (0.01, 0.01, 0.005, \cdots, 0.002)$

[For LDA, $f_{z_i} \sim \text{Dir}(\beta)$]

Dirichlet distribution:

Support : $x_1, \cdots, x_K \in (0,1)$, $\sum_{i=1}^{K} x_i = 1$

Parameters : $\alpha_1, \cdots, \alpha_K > 0$

PDF : $f(x_1, \cdots, x_K) \propto x_1^{\alpha_1 - 1} \cdot x_2^{\alpha_2 - 1} \cdot \cdots \cdot x_K^{\alpha_K - 1}$