Gibbs Sampling for LDA (1 iteration)

Input

Documents:

- D1: apple, banana, apple
- D2: bus, train, banana

Vocabulary (V): apple, banana, bus, train
K (Topics): 2
We'll keep:

- Dirichlet priors: $\alpha = 1$ (for θ), $\beta = 1$ (for ϕ)
- Initial random topic assignment:
  - D1: apple (T0), banana (T1), apple (T0)
  - D2: bus (T1), train (T1), banana (T0)

Step 1: Count Tables (Before Sampling)

Word–Topic Counts (n_{wt}):

Word	T0	T1
apple	2	0
banana	1	1
bus	0	1
train	0	1

Document–Topic Counts (n_{dt}):

Document	T0	T1
D1	2	1
D2	1	2

Now we resample the topic for each word in turn.

Sampling Each Token

We'll update each word token's topic, one by one, using the Gibbs sampling formula:

Formula Probability a word token is assigned to topic t is:

$$P(t) \propto \frac{n_{wt}^{-i} + \beta}{n_t^{-i} + V\beta} \cdot \frac{n_{dt}^{-i} + \alpha}{n_d^{-i} + K\alpha}$$

Where:

- n_{wt}^{-i}: count of word w assigned to topic t, excluding current token
- n_t^{-i}: total words assigned to topic t, excluding current token
- n_{dt}^{-i}: number of words in document d assigned to topic t, excluding current token
- n_d^{-i}: total words in document d, excluding current token

Let's do one token in full detail, then summarize the rest.

Token 1: apple in D1 (currently T0)

Exclude current token from counts

$$n_{apple,T0}^{-i} = 1$$

$$n_{T0}^{-i} = 4 - 1 = 3$$
$$n_{D1,T0}^{-i} = 1$$
$$n_{D1}^{-i} = 3 - 1 = 2$$

Compute Probabilities:
P(T0):

$$P(T0) \propto \frac{1+1}{3+4\cdot 1} \cdot \frac{1+1}{2+2\cdot 1} = \frac{2}{7} \cdot \frac{2}{4} = \frac{4}{28} = 0.143$$

P(T1):

$$P(T1) \propto \frac{0+1}{2+4\cdot 1} \cdot \frac{1+1}{2+2\cdot 1} = \frac{1}{6} \cdot \frac{2}{4} = \frac{2}{24} \approx 0.083$$

Normalize: Total = 0.143 + 0.083 = 0.226

$$P(T0) \approx \frac{0.143}{0.226} \approx 0.63, \quad P(T1) \approx \frac{0.083}{0.226} \approx 0.37$$

$\implies$ Stays with Topic 0 (likely).
We can now repeat similar math for each of the 5 remaining tokens.