

Estimation of Probabilities (Depending on the Data)

$$p(X_{w1} = 1) = \frac{\text{count}(w1)}{N}$$

$$p(X_{w2} = 1) = \frac{\text{count}(w2)}{N}$$

$$p(X_{w1} = 1, X_{w2} = 1) = \frac{\text{count}(w1, w2)}{N}$$

	W1	W2	
Segment_1	1	0	Only W1 occurred
Segment_2	1	1	Both occurred
Segment_3	1	1	Both occurred
Segment_4	0	0	Neither occurred
...			
Segment_N	0	1	Only W2 occurred

Count(w1) = total number segments that contain W1

Count(w2) = total number segments that contain W2

Count(w1, w2) = total number segments that contain both W1 and W2