Desenvolveu-se também abordagens probabilísticas para segmentação textual, por exemplo, o método proposto por [**?**] encontra a segmentação por meio de um modelo estatístico. Dado um texto representado por um conjunto de palavras $W = \{w_1, w_2, \ldots, w_n\}$ e um conjunto de segmentos $S = \{s_1, s_2, \ldots, s_m\}$ que segmenta W, a probabilidade da segmentação S é dada por:

$$P(S|W) = \frac{P(W|S)P(S)}{P(W)} \tag{1}$$

Com isso, é possível encontrar a sequência de segmentos mais provável $\hat{S} = argmax_S\ P(W|S)P(S)$. Nesse trabalho assume-se que os segmentos são estaticamente independentes entre si e as palavras nos segmentos são independentes dado o segmento que as contém. Essa simplificação permite decompor o termo $P(W|S)$ em um produtório de ocorrência de das palavras dado um segmento.

$$P(W|S) = \prod_{i=1}^{m} \prod_{j=1}^{n} P(w_j^i|S_i) \tag{2}$$

Onde $P(w_j^i|S_i)$ é a probabilidade da j-ésima palavra ocorrer no segmento S_i. Seja $f_i(w_j)$ a frequência da j-ésima palavra no i-ésimo segmento, n_i é o número de palavras em S_i e k é o número de palavas diferentes em W. Calcula-se:

$$P(w_j^i|S_i) = \frac{f_i(w_j) + 1}{n_i + k} \tag{3}$$

A suposição de independência entre segmentos e as palavras neles contidas, são é verificada no mundo real. Para segmentos muito pequenos a estimativa das probabilidades das palavras pode ser afetada, além disso, o modelo não leva em conta a importância relativa das palavras [**?**].

$$NCorte = \frac{corte(A, B)}{vol(A} + \frac{corte(A, B)}{vol(B)} \tag{4}$$