

BM25/Okapi Approximation

(Robertson et al. 94)

- Idea: Approximate $p(R=1 | Q,D)$ with a simpler function that share similar properties
- Observations:
 - $\log O(R=1 | Q,D)$ is a sum of term weights W_i
 - $W_i = 0$, if $TF_i = 0$
 - W_i increases monotonically with TF_i
 - W_i has an asymptotic limit
- The simple function is

$$W_i = \frac{TF_i(k_1 + 1)}{k_1 + TF_i} \log \frac{p_i(1 - q_i)}{q_i(1 - p_i)}$$