

# Probabilistic Retrieval Models: Basic Idea

Query    Doc    Rel

**q**            **d**            **R**

q1            d1            1

q1            d2            1

q1            d3            0

q1            d4            0

q1            d5            1

...

q1            d1            0

q1            d2            1

q1            d3            0

q2            d3            1

q3            d1            1

q4            d2            1

q4            d3            0

$$f(q,d)=p(R=1 \mid d,q)=?$$

$$\frac{\text{count}(q, d, R = 1)}{\text{count}(q, d)}$$

$$P(R=1 \mid q1,d1) = ? \quad 1/2$$

$$P(R=1 \mid q1,d2) = ? \quad 2/2$$

$$P(R=1 \mid q1,d3) = ? \quad 0/2$$

What about unseen documents?

Unseen queries?