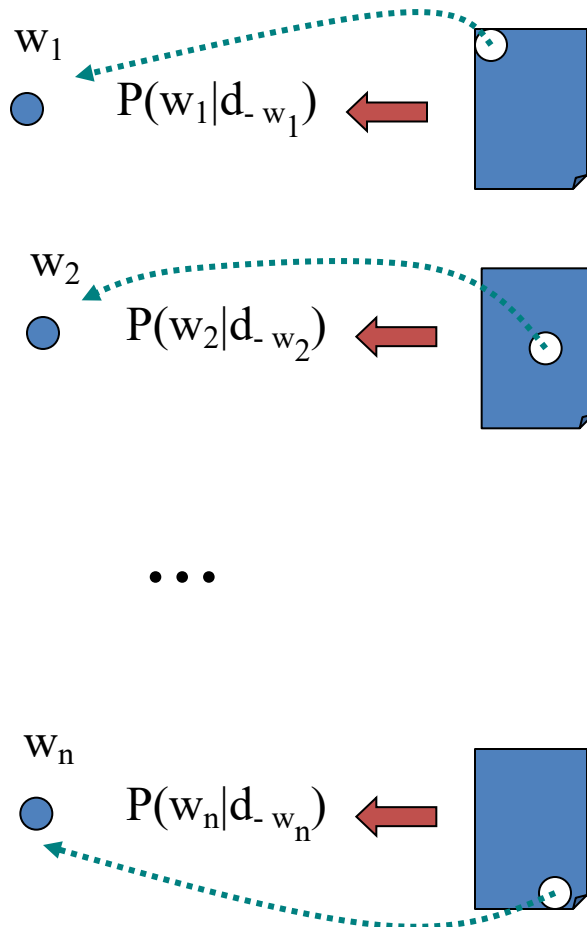


# Estimating $\mu$ using leave-one-out [Zhai & Lafferty 02]



**log-likelihood**

$$l_{-1}(\mu | C) = \sum_{i=1}^N \sum_{w \in V} c(w, d_i) \log \left( \frac{c(w, d_i) - 1 + \mu p(w | C)}{|d_i| - 1 + \mu} \right)$$

**Leave-one-out**

**Maximum Likelihood Estimator**

$$\hat{\mu} = \operatorname{argmax}_{\mu} l_{-1}(\mu | C)$$

**Newton's Method**