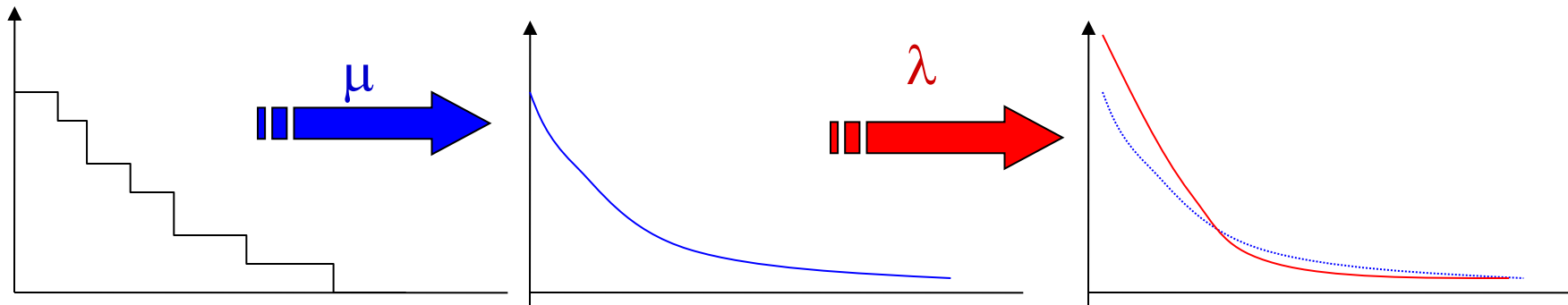# Two-stage Smoothing [Zhai & Lafferty 02]

**Stage-1**

-Explain unseen words
-Dirichlet prior(Bayesian)

**Stage-2**

-Explain noise in query
-2-component mixture



**Collection LM**

$$P(w|d) = (1-\lambda) \frac{c(w,d) + \mu p(w|C)}{|d| + \mu} + \lambda p(w|U)$$

**User background model**
**Can be approximated by p(w|C)**