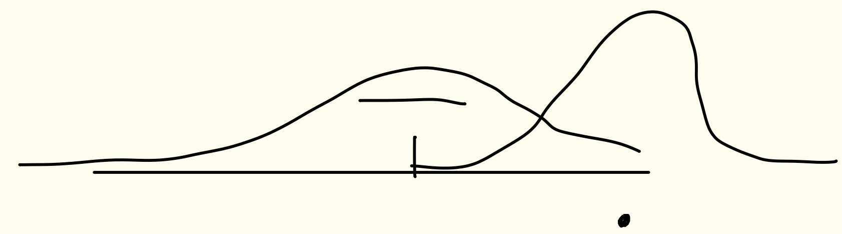
#B241/RPZ/SEM

$$p(x|k) = N(\mu, \sigma) \qquad T = \{x_1, \dots, x_n\} \qquad p(T|\theta)$$



MLE: $$\hat{\theta} = \underset{\theta}{\arg\max} \underbrace{\prod_i p(x_i|\theta)}_{L(\theta)} =$$

$$= \underset{}{\arg\max} \underbrace{\sum_i \log(p(x_i|\theta))}_{l(\theta)}$$

$$\frac{dl}{d\theta} = \dots = 0$$

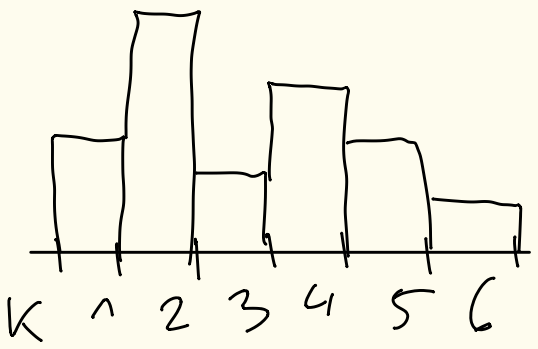
$$\theta = ?$$

MAP: $$\hat{\theta} = \underset{\theta}{\arg\max}\, p(\theta|T) = \text{—''—}\ \frac{p(T|\theta)\boxed{p(\theta)}}{\cancel{p(T)}}$$

$$p(x|T) = p(x|\hat{\theta})$$

BAYES: $$p(x|T) = \int_\theta p(x|\theta) \cdot p(\theta|T)\, d\theta =$$

$$= \int_\theta p(x|\theta) \cdot \frac{p(T|\theta)p(\theta)}{\int_\theta p(T|\theta)p(\theta)d\theta}\, d\theta$$



T

$N_1, N_2, \dots N_K$

$$L(\theta) = \prod_{K=1}^{K} p_k^{N_K} \qquad \ell(\theta) = \sum_K N_k \log(p_k)$$

$$\hat{\lambda}, \hat{\theta} = \underset{\theta, \lambda}{\operatorname{argmax}} \underbrace{\ell(\theta) + \lambda\left(\sum_K (p_k) - 1\right)}_{h}$$

$$\frac{\partial h}{\partial p_k} = \frac{N_k}{p_k} + \lambda = 0 \quad \Rightarrow \quad p_k = -\frac{N_k}{\lambda}$$

$$\frac{\partial h}{\partial \lambda} = \sum (p_k) - 1 = 0$$

$$\sum_K -\frac{N_k}{\lambda} = 1$$

$$\sum_K (-N_k) = \lambda \longrightarrow p_k = \frac{-N_k}{\sum_K (-N_k)} = \frac{N_k}{\sum_K N_k}$$

Show that:

$$\hat{p}_k = \frac{N_k}{\sum_{m=1}^{K} N_m}$$

$$L(\theta) = \prod_{k=1}^{K} p_k^{N_k} \xrightarrow{\log} \ell(\theta) = \sum_{k}^{K} N_k \log(p_k)$$

$$\hat{\lambda}, \hat{\theta} = \underset{\theta,\lambda}{\operatorname{argmax}} \underbrace{\ell(\theta) + \lambda\left(\sum_{k}^{K}(p_k) - 1\right)}_{\text{Lagrangian function } h}$$

$$\frac{\partial h}{\partial p_k} = \frac{N_k}{p_k} + \lambda = 0 \Longrightarrow \hat{p}_k = -\frac{N_k}{\lambda}$$

$$\frac{\partial h}{\partial \lambda} = \sum (p_k) - 1 = 0$$

$$\sum_{k}^{K} -\frac{N_k}{\lambda} = 1 \Rightarrow \sum_{k}^{K}(-N_k) = \lambda$$

thus:

$$\hat{p}_k = \frac{N_k}{\sum_{k=1}^{K}(-N_k)} = \frac{N_k}{\sum_{m=1}^{K} N_m}$$

□