P.M.F

$$p(x) = \mu^{x}(1-\mu)^{1-x} \quad x \in \{0,1\}$$

$$\mathbb{E}[x] = \sum_{x\in 0,1} x \cdot p(x) = \mu \cdot (1-\mu)^{1-1} = \mu.$$

$$\mathrm{Var}[x] = \mathbb{E}[(x-\mu)^2]$$

$$= \mathbb{E}[x^2] - \underbrace{(\mathbb{E}(x))^2}_{\mu^2}$$

$$= \sum_{x\in 0,1} x^2 \cdot p(x)$$

$$= \mu - \mu^2 = \mu(1-\mu)$$

$x_1 \cdots x_N$ binary

$\mu(1-\mu)$

$$x = x_1 + \cdots + x_N.$$

$$\mathbb{E}[x] = \underbrace{\mathbb{E}[x_1]}_{\mu} + \cdots + \underbrace{\mathbb{E}[x_N]}_{\mu} = N \cdot \mu.$$

$$\mathrm{Var}[x] = \mathrm{Var}[x_1] + \cdots + \mathrm{Var}[x_N] = N \cdot \mu(1-\mu)$$

$x_1 \cdots x_N \sim \mathrm{Cat}.\ \mu.$

$$p(D|\mu) = \prod_{n=1}^{N} p(x_n|\mu) = \prod_{n=1}^{N}\prod_{k=1}^{K} p(x_{nk}|\mu)$$

$$= \prod_{n=1}^{N}\prod_{k=1}^{K} \mu_k^{x_{nk}} = \prod_{k=1}^{K}\prod_{n=1}^{N} \mu_k^{x_{nk}}.$$

$$= \prod_{k=1}^{K} \mu_k^{\sum_{n=1}^{N} x_{nk}} \quad m_k \quad = \prod_{k=1}^{K} \mu_k^{m_k}$$

$$m_k = \sum_{n=1}^{N} x_{nk}$$

$$\mathcal{L} = \ln p(D|\mu) = \sum_{k=1}^{K} m_k \ln \mu_k$$

$$\text{s.t.} \quad \sum_{k=1}^{K} \mu_k = 1$$

$$\mathcal{L} = \sum_{k=1}^{K} m_k \ln \mu_k + \lambda\left(\sum_{k=1}^{K} \mu_k - 1\right)$$

$$\mu_1 \cdots \mu_K$$

$$\frac{\partial \mathcal{L}}{\partial \mu_k} = \frac{m_k}{\mu_k} + \lambda = 0$$

$$\frac{m_k}{\mu_k^*} = -\lambda \rightarrow$$

$$\mu_k = -\frac{m_k}{\lambda}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 + \sum_{k=1}^{K} \mu_k - 1 \rightarrow$$

$$0 + \sum_{k=1}^{K} -\frac{m_k}{\lambda} - 1 = 0$$

$$\sum_{k=1}^{K} m_k = N$$

$$-\frac{\sum_{k=1}^{K} m_k}{\lambda} - 1 = 0 \quad -\frac{N}{\lambda} - 1 = 0$$

$$\lambda = -N$$

$$\frac{m_k}{\mu_k^*} = -\lambda = N$$

$$\mu_k^* = \frac{m_k}{N}$$