

Assume $p(y) = \text{prior}$
 $p(t|\phi) = \phi^t (1-\phi)^{1-t}$ (prior is Bernoulli)

Prove $\phi = \frac{1}{N} \sum_{n=1}^N \mathbb{I}[t^{(n)}=1]$

$$\ell(\phi, \mu_0, \mu_1, \Sigma) = -\ln \prod_{n=1}^N p(x^{(n)} | t^{(n)}, \mu_0, \mu_1, \Sigma) p(t^{(n)} | \phi) \quad (\text{slide 17})$$

(choose arbitrary class $k \in \{0, 1\}$, preserving generality, Assume $\Sigma_1 = \Sigma_2$
 (classes share same covar. matrix)

$$= -\ln \prod_{n=1}^N p(x^{(n)} | t^{(n)}, \mu_k, \Sigma) p(t^{(n)} | \phi) \quad (\text{Eqn from slide 6})$$

$$= -\ln \prod_{n=1}^N \left(\frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_k|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \right) \cdot \phi^t (1-\phi)^{1-t}$$

$$= -\left[\sum_{n=1}^N \ln \left(\frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma_k|^{\frac{1}{2}}} \right) + \ln \left(e^{-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k)} \right) + t \ln \phi + (1-t) \ln(1-\phi) \right]$$

$$= \sum_{n=1}^N \left\{ \frac{d}{2} \ln(2\pi) + \frac{1}{2} \ln |\Sigma_k| + (x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k) + t \ln \phi + (1-t) \ln(1-\phi) \right\} \quad \text{Eqn (1)}$$

Now equate derivative to 0

$$\frac{d}{d\phi} \ell(\phi, \mu_k, \Sigma_k) = 0 = \sum_{n=1}^N 0 + 0 + 0 + \frac{d}{d\phi} t \ln \phi + \frac{d}{d\phi} (1-t) \ln(1-\phi)$$

$$= \sum_{n=1}^N \frac{t^{(n)}}{\phi} + \frac{d}{d\phi} (1-t^{(n)}) + 0 \cdot \ln(1-\phi) + (1-t^{(n)}) \cdot \frac{1}{(1-\phi)} (-1) \quad (\text{chain rule})$$

$$= \sum_{n=1}^N \frac{t^{(n)}}{\phi} - \frac{(1-t^{(n)})}{(1-\phi)} = \sum_{n=1}^N \frac{t^{(n)}(1-\phi) - \phi(1-t^{(n)})}{\phi(1-\phi)} = 0$$

$$0 = \sum_{n=1}^N \frac{t^{(n)} - t\phi - \phi + t\phi}{\phi(1-\phi)} = \sum_{n=1}^N (t^{(n)} - \phi) = \sum_{n=1}^N t^{(n)} - N\phi$$

$$\Rightarrow \phi = \frac{1}{N} \sum_{n=1}^N t^{(n)}$$

$$\Rightarrow \phi = \frac{1}{N} \sum_{n=1}^N \mathbb{I}[t^{(n)}=1] \quad (\text{since prior is Bernoulli, } t^{(n)} \in \{0, 1\})$$

$$\therefore \phi = \frac{1}{N} \sum_{n=1}^N \mathbb{I}[t^{(n)}=1]$$