4.9

$$p(\boldsymbol{\phi}_n, C_k) = p(C_k)p(\boldsymbol{\phi}_n|C_k) = \pi_k p(\boldsymbol{\phi}_n|C_k)$$

The likelihood function is given by

$$p(\mathbf{t}|\pi_1, \pi_2, \ldots, \pi_K) = \prod_{n=1}^{N}\prod_{k=1}^{K}(\pi_k p(\boldsymbol{\phi}_n|C_k))^{t_{nk}}$$

The log-likelihood function is given by

$$p(\mathbf{t}|\pi_1, \pi_2, \ldots, \pi_K) = \sum_{n=1}^{N}\sum_{k=1}^{K}\ln\left((\pi_k p(\boldsymbol{\phi}_n|C_k))^{t_{nk}}\right)$$

$$= \sum_{n=1}^{N}\sum_{k=1}^{K} t_{nk}\ln(\pi_k) + t_{nk}\ln(p(\boldsymbol{\phi}_n|C_k))$$

We need to maximizing w.r.t K parameters : π_k for all $k = 1 \ldots K$.

But we also have a constraint : $\sum_{i=1}^{K}\pi_k = 1$.

If we use constrained optimization, we can define a Lagrangian function (using this reference):

$$L(\pi_1, \pi_2, \ldots, \pi_K, \lambda) = \sum_{n=1}^{N}\sum_{k=1}^{K}(t_{nk}\ln(\pi_k) + t_{nk}\ln(p(\boldsymbol{\phi}_n|C_k))) + \lambda\left(1 - \sum_{i=1}^{K}\pi_k\right)$$

For any particular π_k,

$$\frac{dL(\pi_1, \pi_2, \ldots, \pi_K, \lambda)}{d\pi_k} = \sum_{n=1}^{N}\left(\frac{t_{nk}}{\pi_k} + 0\right) - \lambda = 0$$

$$\implies \lambda = \sum_{n=1}^{N}\frac{t_{nk}}{\pi_k}$$

$$\implies \lambda = \frac{N_k}{\pi_k}$$

$$\implies \pi_k = \frac{N_k}{\lambda}$$

Using the constraint,

$$\sum_{i=1}^{K} \pi_i = 1$$

$$\implies \sum_{i=1}^{K} \frac{N_k}{\lambda} = 1$$

$$\implies \frac{N}{\lambda} = 1$$

$$\implies \lambda = N$$

Thus, we can see that

$$\pi_k = \frac{N_k}{\lambda} = \frac{N_k}{N}, \quad \forall\, k = \{1, 2, \ldots, K\}.$$