▷ 在GDA中,我們需要學習五個參數($\mu_0, \sigma_0, \mu_1, \sigma_1, \phi$),但上課時沒有特別來展開$\sigma_0, \sigma_1$來告訴我們 MLE 中的$\sigma_0, \sigma_1$,我們來重新推一次,$\sigma_0, \sigma_1$,從GDA之中。

→ For $\{(x^k, y^k)\}_{k=1}^{M}$, $y^k \in \{0,1\}$

Define the set of index: $I_0 = \{k : y^k = 0\}$; $I_1 = \{k : y^k = 1\}$, where we all know that the number of samples: $m_0 = |I_0|$; $m_1 = |I_1|$, $M = m_0 + m_1$

$f(x|y=0) \sim \mathcal{N}(\mu_0, \sigma_0^2)$; $f(x|y=1) \sim \mathcal{N}(\mu_1, \sigma_1^2)$, Moreover, ① $P(y=1) = \phi$; $P(y=0) = 1-\phi$

$$\mathcal{L}(\mu_0, \mu_1, \sigma_0^2, \sigma_1^2, \phi) = \prod_{k \in I_0} [(1-\phi)\mathcal{N}(x^k | \mu_0, \sigma_0^2)] \prod_{k \in I_1} [\phi \mathcal{N}(x^k | \mu_1, \sigma_1^2)]$$

$$\longrightarrow \ell = m_1 \ln\phi + m_0 \ln(1-\phi) - \frac{m_0}{2}\ln\sigma_0^2 - \frac{1}{2\sigma_0^2}\sum_{k \in I_0}(x^k - \mu_0)^2 - \frac{m_1}{2}\ln\sigma_1^2 - \frac{1}{2\sigma_1^2}\sum_{k \in I_1}(x^k - \mu_1)^2$$

Next step, we derivative to μ_0, μ_1: $\frac{\partial \ell}{\partial \mu_0} = \frac{1}{\sigma_0^2}\sum_{k \in I_0}(x^k - \mu_0) = 0$, ① $\mu_0 = \frac{1}{m_0}\sum_{k \in I_0} x^k$

② $\mu_1 = \frac{1}{m_1}\sum_{k \in I_1} x^k$

★ We derivative to σ_0^2, σ_1^2:

① $$\frac{\partial \ell}{\partial \sigma_0^2} = \frac{-m_0}{2\sigma_0^2} + \frac{1}{2(\sigma_0^2)^2}\sum_{k \in I_0}(x^k - \mu_0)^2 = 0$$

→ $-m_0\sigma_0^2 + \sum_{k \in I_0}(x^k - \mu_0)^2 = 0$, then, we have $\sigma_0^2 = \frac{1}{m_0}\sum_{k \in I_0}(x^k - \mu_0)^2 \Rightarrow \sigma_0 = \sqrt{\frac{1}{m_0}\sum_{k \in I_0}(x^k - \mu_0)^2}$

② $\sigma_1^2 = \frac{1}{m_1}\sum_{k \in I_1}(x^k - \mu_1)^2 \Rightarrow \sigma_1 = \sqrt{\frac{1}{m_1}\sum_{k \in I_1}(x^k - \mu_1)^2}$

▨