ΘΕΜΑ 4: Εκτίμηση παραμέτρων και ταξινόμηση.

① Δείγματα $\{x_1^{y_i}, x_2^{y_i}, \dots, x_n^{y_i}\}$:

Ορίζω $\underline{x_k^{y_i} \triangleq x_k}$

Αφού τα χαρακτηριστικά ακολουθούν Bernoulli

$$P(x_k | p^{y_i}) = (p^{y_i})^{x_k} (1 - p^{y_i})^{1 - x_k}$$

Θα υπολογίσουμε τον $p^{y_i}$ με τον εκτιμητή μέγιστης πιθανοφάνειας ⊛

$$L_X(p^{y_i}) = P(x | p^{y_i}) = \prod_{k=1}^{n} P(x_k | p^{y_i})$$

$$\Rightarrow l_X(p^{y_i}) = \log L_X(\theta) = \sum_{k=1}^{n} \log P(x_k | p^{y_i})$$

$$\Rightarrow \boxed{l_X(p^{y_i}) = \sum_{k=1}^{n} x_k \log p^{y_i} + (1 - x_k) \log(1 - p^{y_i})}$$

Το $p^{y_i}$ που ζητάω είναι $p^{y_i}_{MLE} = \underset{p^{y_i}}{\arg\max}\, P(x | p^{y_i})$

$$= \underset{p^{y_i}}{\arg\max}\, l_X(p^{y_i})$$

Πρέπει $\frac{\partial}{\partial p^{y_i}}\left( l_x(p^{y_i}) \right) = 0$

$$\Rightarrow \frac{\partial}{\partial p^{y_i}} \sum_{k=1}^{n} X_k \log p^{y_i} + (1-X_k)\log(1-p^{y_i}) = 0$$

$$\Rightarrow \sum_{k=1}^{n} \left( \frac{X_k}{p^{y_i}} - \frac{(1-X_k)}{1-p^{y_i}} \right) = 0$$

$$\Rightarrow \sum_{k=1}^{n} \left( \frac{X_k}{p^{y_i}} + \frac{X_k}{1-p^{y_i}} - \frac{1}{1-p^{y_i}} \right) = 0$$

$$\Rightarrow \frac{1}{p^{y_i}} \sum_{k=1}^{n} X_k + \frac{1}{1-p^{y_i}} \sum_{k=1}^{n} X_k - \frac{n}{1-p^{y_i}} = 0$$

$$\Rightarrow \left( \frac{1}{p^{y_i}} + \frac{1}{1-p^{y_i}} \right) \sum_{k=1}^{n} X_k = \frac{n}{1-p^{y_i}}$$

$$\Rightarrow \frac{1}{p^{y_i}\cancel{(1-p^{y_i})}} \sum_{k=1}^{n} X_k = \frac{n}{\cancel{1-p^{y_i}}}$$

$$\Rightarrow \boxed{p^{y_i} = \frac{1}{n} \cdot \sum_{k=1}^{n} X_k}$$