Exercice 2

1.1 $p(x_i | C_k) = (p_{ik})^{x_i} (1-p_{ik})^{(1-x_i)}$

1.2
$$p(C_k | x) = \frac{p(x | C_k)\, p(C_k)}{p(x)}$$
$$= \frac{\prod_{i=1}^{n} p(x_i | C_k)\, p(C_k)}{p(x)}$$

$$\log p(C_k | x) = \log p(C_k) + \sum_{i=1}^{n} \log\left((p_{ik})^{x_i} (1-p_{ik})^{(1-x_i)}\right) - \log p(x)$$
$$= \log p(C_k) + \sum_{i=1}^{n} x_i \log \frac{p_{ik}}{1-p_{ik}} + \sum_{i=1}^{n} \log(1-p_{ik}) - \log p(x)$$

1.3

1.3.1
$$\log \frac{P(C_1 | x)}{P(C_2 | x)} = \log \frac{P(C_1)}{P(C_2)} + \sum_{i=1}^{n} x_i \log \frac{p_{i1}(1-p_{i2})}{(1-p_{i1})\, p_{i2}} + \sum_{i=1}^{n} \log\left(\frac{1-p_{i1}}{1-p_{i2}}\right)$$

1.3.2

règle de Bayes : décide C_1 si $\dfrac{P(C_1 | x)}{P(C_2 | x)} > 1$

i.e $\log \dfrac{P(C_1 | x)}{P(C_2 | x)} > 0$

Représentation fréquentielle

2-1

$$p(y_1 \dots y_q) = \prod_{i=1}^{q} p(y_i) = \prod_{i=1}^{q} p_{ik}^{y_i} (1-p_{ik})^{1-y_i}$$

$$= p_{ik}^{\sum_{i=1}^{q} y_i} \cdot (1-p_{ik})^{\sum_1^q 1-y_i}$$

$$= p_{ik}^{q_i} (1-p_{ik})^{q-q_i}$$

2-2 $C_q^{q_i}$ sequences.

2-3 $p(x_i = q_i | C_k) = C_q^{q_i} p_{ik}^{q_i} (1-p_{ik})^{q-q_i}$

2-4

2-4.1 $p(x_1^2 | C_k) = p(x_2 | x_1, C_k) \, p(x_1 | C_k)$

2-4.2 $p(x_2 | x_1 = q_1, C_k)$

$p_{1k} + p_{2k} = 1$ $\quad$ $p_{2k} = 1 - p_{1k}$

$q_1 + q_2 = q$ $\quad$ $q_2 = q - q_1$

Si on a fixé q_1 termes x_1, il reste à fixer q_2 termes qui sont forcément x_2.

$p(x_2 = q_2 | x_1 = q_1, C_k) = 1$

2-4.3 $p(x = (q_1, q_2) | C_k) = p(x_1 = q_1 | C_k) \cdot 1$

$$= C_q^{q_1} p_{1k}^{q_1} (1-p_{1k})^{q-q_1}$$

$$p(x=(q_1,q_2)\mid C_k) = \frac{q!}{q_1!\,q_2!}\, p_{1k}^{q_1}\, p_{2k}^{q_2}$$

2.5 $$\log \frac{P(C_1|x)}{P(C_2|x)} = \log \frac{p(x|C_1)}{p(x|C_2)} + \log \frac{p(C_1)}{p(C_2)}$$

$$\log \frac{P(C_1|x)}{P(C_2|x)} = \log \prod_{i=1}^{n} \left(\frac{p_{i1}}{p_{i2}}\right)^{q_i} + \log \frac{P(C_1)}{P(C_2)}$$

$$\log \frac{P(C_1|x)}{P(C_2|x)} = \sum_{i=1}^{n} \left(q_i \log \frac{p_{i1}}{p_{i2}}\right) + \log \frac{P(C_1)}{P(C_2)}$$

fonction discriminante linéaire.