Exercice 5.1 :

Let's consider the following prior: $p(\theta) = \sum_k p(z=k)\, p(\theta \mid z=k)$

the posterior after observing D is:

$$p(\theta \mid D) = \frac{p(\theta \mid D)}{p(D)}$$

the low of total probability:

$$p(\theta, D) = \sum_k p(\theta, D \mid z=k)\, p(z=k)$$

then,

$$\frac{p(\theta \mid D)}{p(D)} = \frac{\sum_k p(\theta, D \mid z=k)\, p(z=k)}{p(D)}$$

$$\frac{p(\theta \mid D)}{p(D)} = \frac{\sum_k p(z=k)\, p(D \mid z=k)\, p(\theta \mid D, z=k)}{p(D)}$$

since, $p(z=k \mid D) = \dfrac{p(D \mid z=k)\, p(z=k)}{p(D)}$ (bayes rule)

then,

$$p(\theta \mid D) = \sum_k p(z=k \mid D)\, p(\theta \mid D, z=k)$$

Exercise 5.2 :

a)

$\rho(\hat{y}=0 \mid x) = \lambda_{01}\, p(y=1 \mid x) = \lambda_{01}\, p_1$ (if we choose 0)

$\rho(\hat{y}=1 \mid x) = \lambda_{10}\, p(y=0 \mid x) = \lambda_{10}\, p_0$ (if we choose 1)

Therefore, if we pick the label 0 : $\rho(\hat{y}=0 \mid x) < \rho(\hat{y}=1 \mid x) \Rightarrow \lambda_{01} p_1 < \lambda_{10} p_0 = \lambda_{10}(1-p_1)$

$$\Rightarrow p_1 < \frac{\lambda_{10}}{\lambda_{01} + \lambda_{10}} \sim \theta$$

otherwise we pick the label 1. (the same process)

b) We want $0.1 = \dfrac{\lambda_{10}}{\lambda_{01} + \lambda_{10}} \Rightarrow 9\lambda_{10} = \lambda_{01} \Rightarrow$ any loss matrix satisfy this condition

Exercice 5.3 :

a)
The loss function is defined by

$$\ell(a|x) = \sum_{i=1}^{c} \lambda(\alpha_i | Y=j)\, P(y_i|x) = \sum_{j=1}^{c} \lambda(\alpha_i | Y=j)\, P_j$$

because $\lambda(\alpha_i | y=j) = \begin{cases} 0 & \text{if } i=j \in \{1,\dots,c\} \\ \lambda_r & \text{if } i = c+1 \\ \lambda_s & \text{otherwise} \end{cases}$

Then,

$$\ell(a|x) = \lambda_s (1 - P_a)$$

The minimum risk is when, $\lambda_s(1-P_a) < \lambda_r$

$$\Rightarrow \boxed{P_a \geq 1 - \frac{\lambda_r}{\lambda_s}}$$