Lecture 4

M-ary Hypothesis Testing

1. Review of Binary Hypothesis Testing

Thus far we have studied the Bayesian Decision:

$$\boxed{\delta_B(y) = \arg\min_i \sum_{j=0}^{M-1} C(i,j)\,\pi_j(y)} \qquad (4.1)$$

We remark that (4.1) is a very general result, of which we can derive two special cases:

(1) if $C_{11} = C_{00} = 0$; $C_{10} = C_{01} = 1$
(This is the uniform cost),
then the Bayesian decision rule becomes a likelihood ratio test (LRT)

$$\boxed{L(y) \underset{H_1}{\overset{H_0}{\lessgtr}} \frac{\pi_0}{\pi_1},} \qquad (4.2)$$

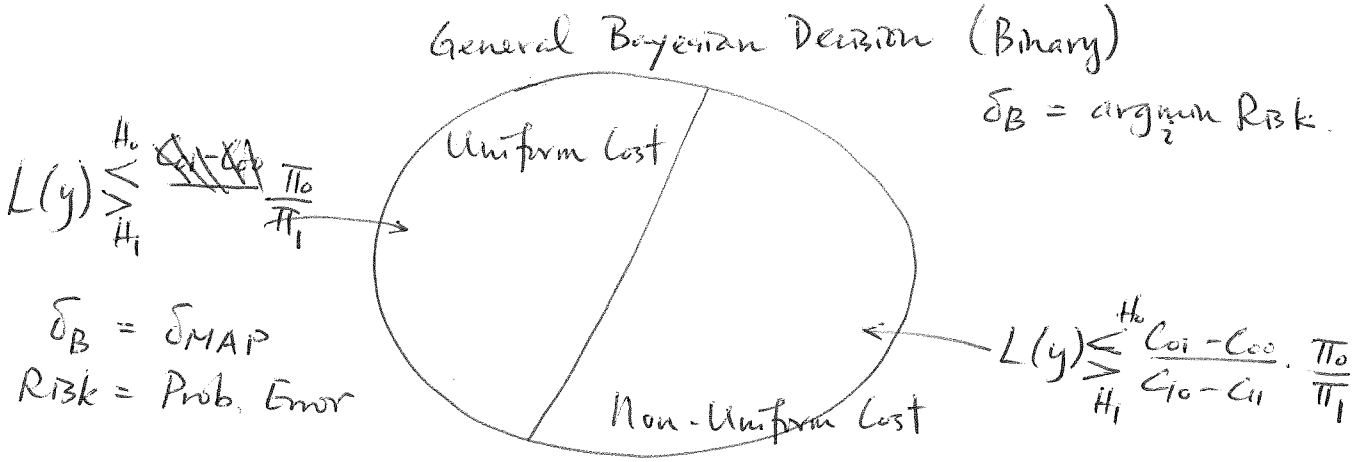
where $L(y) \overset{\text{def}}{=} \dfrac{f_1(y)}{f_0(y)}$.

(2) if, in addition to (1), we also have $\pi_0 = \pi_1$, then (4.2) reduces to $L(y) \underset{H_1}{\overset{H_0}{\lessgtr}} 1$.

As long as the cost is uniform, Bayesian Decision rule is also a Maximum a Posteriori rule. In this case, the probability of error = Bayesian risk:

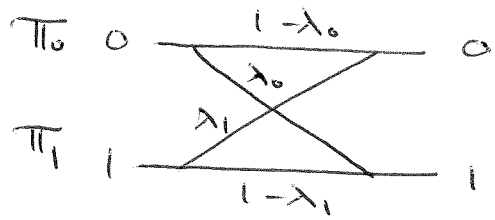
$$\mathbb{P}(\Theta \neq \delta(Y)) = \mathbb{E}_Y\left[\mathbb{E}_{\Theta|Y}[C(\delta(Y),\Theta) \mid Y=y]\right].$$

The relationship between these rules is shown below:



<u>Example</u>: Binary Symmetric Channel

Consider a communication channel as follow



Suppose that the communication channel is symmetric, i.e. $\lambda_0 = \lambda_1 = \lambda$, and assume a uniform cost, $\pi_0 = \pi_1$. Given an observation $y \in \{0, 1\}$, find the Bayesian decision rule.

<u>Solution</u> Under the Bayesian decision framework, we know that $\delta_B(y)$ is to check

$$L(y) \underset{H_1}{\overset{H_0}{\lessgtr}} \frac{\pi_0}{\pi_1} = 1$$

Therefore, since

$$L(y) = \frac{f_1(y)}{f_0(y)} = \cancel{\frac{\mathbb{P}(H_1 \mid Y=y)}{\mathbb{P}(H_0 \mid Y=y)} = \begin{cases} \end{cases}}$$

$$= \frac{\mathbb{P}(Y=y \mid H_1)}{\mathbb{P}(Y=y \mid H_0)} = \begin{cases} \frac{\lambda}{1-\lambda}, & y=0, \\ \frac{1-\lambda}{\lambda}, & y=1, \end{cases}$$

We have:

if $y=0$, $\frac{\lambda}{1-\lambda} \underset{H_1}{\overset{H_0}{\lessgtr}} 1 \iff \lambda \underset{H_1}{\overset{H_0}{\lessgtr}} \frac{1}{2}$

So $\delta_B(0) = \begin{cases} 0, & \text{if } \lambda \leq \frac{1}{2} \\ 1, & \text{if } \lambda > \frac{1}{2} \end{cases}$

If $y=1$, $\frac{1-\lambda}{\lambda} \underset{H_1}{\overset{H_0}{\lessgtr}} 1 \iff \lambda \underset{H_1}{\overset{H_0}{\gtrless}} \frac{1}{2}$

So $\delta_B(1) = \begin{cases} 0, & \text{if } \lambda \geq \frac{1}{2} \\ 1, & \text{if } \lambda \leq \frac{1}{2} \end{cases}$

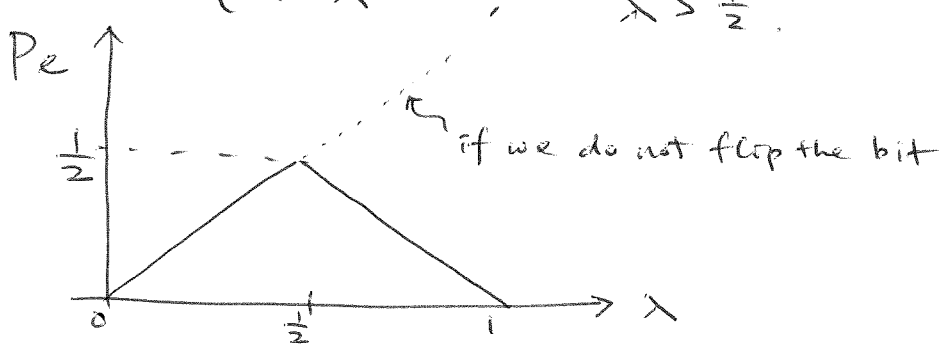
So in general, we have

$$\delta_B(y) = \begin{cases} y, & \text{if } \lambda \leq \frac{1}{2} \\ 1-y, & \text{if } \lambda > \frac{1}{2}. \end{cases} \qquad (4.3)$$

Intuitively, what (4.3) implies is that if $\lambda > \frac{1}{2}$, then we need to "flip" the bit. (Does it make sense physically?)

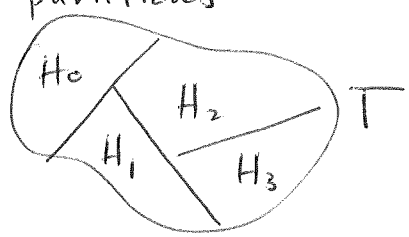
Finally, we can check the probability of error, which is

$$\begin{aligned} P_e &= \pi_0 \mathbb{P}(\delta(y) \neq 0 \mid H_0) + \pi_1 \mathbb{P}(\delta(y) \neq 1 \mid H_1) \\ &= \begin{cases} \pi_0 \lambda + \pi_1 \lambda, & \lambda \leq \frac{1}{2} \\ \pi_0 (1-\lambda) + \pi_1 (1-\lambda), & \lambda > \frac{1}{2} \end{cases} \\ &= \begin{cases} \lambda, & \lambda \leq \frac{1}{2} \\ 1-\lambda, & \lambda > \frac{1}{2}. \end{cases} \end{aligned}$$



2. M-ary Hypothesis Testing

As a generalization of binary hypothesis testing, the M-ary hypothesis testing partitions the observation space Γ into M partitions:



Mathematically, the M-ary decision rule is

$$\delta_B(y) = \arg\min_i \sum_{j=0}^{M-1} C(i,j)\,\pi_j(y)$$

$$= \arg\min_i \sum_{j=0}^{M-1} C(i,j) \frac{\pi_j f_j(y)}{\sum_{\ell=0}^{M-1} \pi_\ell f_\ell(y)}$$

$$= \arg\min_i \sum_{j=0}^{M-1} C(i,j)\,\pi_j f_j(y) \qquad \left(\because \sum_{\ell=0}^{M-1} \pi_\ell f_\ell(y) \text{ is indep of } i\right)$$

$$= \arg\min_i \sum_{j=0}^{M-1} C(i,j)\,\pi_j \frac{f_j(y)}{f_0(y)} \qquad \left(\because f_0(y) \text{ is indep of } i\right)$$

$$\boxed{\delta_B(y) \overset{(a)}{=} \arg\min_i \sum_{j=0}^{M-1} C(i,j)\,\pi_j L_j(y).} \qquad (4.4)$$

where in (a) we define $L_j(y) \overset{\text{def}}{=} \dfrac{f_j(y)}{f_0(y)}$.

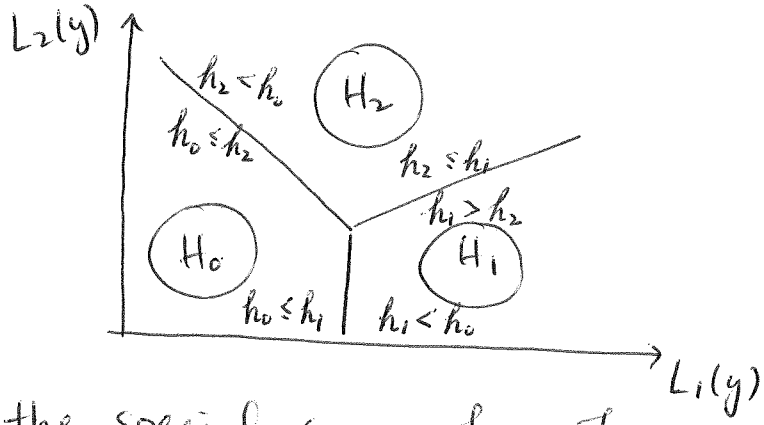
If we further define

$$h_i(y) = \sum_{j=0}^{M-1} C(i,j)\,\pi_j L_j(y), \qquad (4.5)$$

then (4.4) becomes

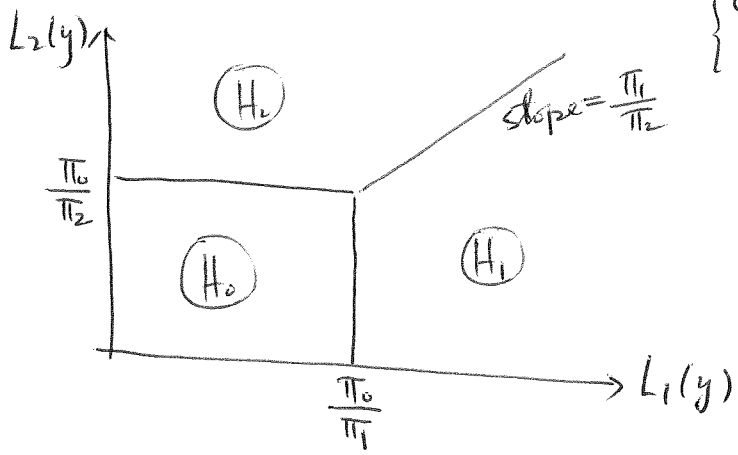
$$\delta_B(y) = \arg\min_i h_i(y).$$

Pictorially, (4.5) suggests that $(M=3)$



In the special case of <u>uniform cost</u>, the partition looks like follows

$$\begin{cases} C_{00} = C_{11} = C_{22} = 0 \\ C_{ij} = 1 \quad i \neq j \end{cases}$$



<u>Example</u>

Let $\mu_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\mu_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, and let $N \sim \mathcal{N}(0, \sigma^2 I_3)$.

Suppose the hypothesis is:

$$H_i : \text{"} Y = \mu_i + N \text{"}.$$

Assume uniform cost, assume $\pi_0 = \frac{1}{4}$, $\pi_1 = \frac{1}{4}$, $\pi_2 = \frac{1}{2}$.

Find the Bayes decision rule.

<u>Solution</u> By (4.4) and (4.5), we know that

$$\delta_B(y) = \underset{0 \le i \le 2}{\operatorname{argmin}} \; C_{i0}\pi_0 + C_{i1}\pi_1 L_1(y) + C_{i2}\pi_2 L_2(y),$$

where we can also define

$$h_0(y) = \cancel{C_{00}\pi_0}^{\,0} + C_{01}\pi_1 L_1(y) + C_{02}\pi_2 L_2(y)$$
$$h_1(y) = C_{10}\pi_0 + \cancel{C_{11}\pi_1 L_1(y)}^{\,0} + C_{12}\pi_2 L_2(y)$$
$$h_2(y) = C_{20}\pi_0 + C_{21}\pi_1 L_1(y) + \cancel{C_{22}\pi_2 L_2(y)}^{\,0}$$

With some calculations we can show that

$$h_0(y) = \frac{1}{4}L_1(y) + \frac{1}{2}L_2(y)$$
$$h_1(y) = \frac{1}{4} + \frac{1}{2}L_2(y)$$
$$h_2(y) = \frac{1}{4} + \frac{1}{4}L_1(y).$$

The M-ary hypothesis testing says we should claim H_0 if $h_0 \le h_1$ and $h_0 \le h_2$. So we must have

$$h_0 \le h_1 \Rightarrow \frac{1}{4}L_1(y) + \frac{1}{2}L_2(y) \le \frac{1}{4} + \frac{1}{2}L_2(y)$$
$$\Rightarrow \frac{1}{4}L_1(y) - \frac{1}{4} \le 0 \Rightarrow \boxed{L_1(y) \le 1}$$

$$h_0 \le h_2 \Rightarrow \frac{1}{4}L_1(y) + \frac{1}{2}L_2(y) \le \frac{1}{4} + \frac{1}{4}L_1(y)$$
$$\Rightarrow \frac{1}{2}L_2(y) - \frac{1}{4} \le 0 \Rightarrow \boxed{L_2(y) \le \frac{1}{2}}$$

Now, let's compute $L_1(y)$ and $L_2(y)$. By definition, we know that

$$f_j(y) \overset{\text{def}}{=} f_{Y|H_j}(y) = \frac{1}{\sqrt{(2\pi)^3\sigma^6}} \exp\left(-\frac{\|y-\mu_j\|^2}{2\sigma^2}\right).$$

(Remark: Multivariate Gaussian is

$$\frac{1}{(2\pi)^n |\Sigma|^{1/2}} \exp\left(-(x-\mu)^T \Sigma^{-1}(x-\mu)\right).$$

$$\Sigma = \sigma^2\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$)

Therefore, the likelihood ratio is

$$L_1(y) = \frac{f_1(y)}{f_0(y)} = \exp\left(\frac{-1}{2\sigma^2}\left(\|y-\underline{\mu}_1\|^2 - \|y-\underline{\mu}_0\|^2\right)\right)$$

$$= \exp\left(\frac{-1}{2\sigma^2}\left(\left\|\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\|^2 - \left\|\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\|^2\right)\right)$$

$$= \exp\left(\frac{-1}{2\sigma^2}\left((y_0^2 + (y_1-1)^2 + y_2^2) - ((y_0-1)^2 + y_1^2 + y_2^2)\right)\right)$$

$$= \exp\left(\frac{-1}{2\sigma^2}(2y_0 - 2y_1)\right) \cancel{= \exp\left(\frac{-1}{2\sigma^2}\right.}$$

$$= \exp\left(\frac{y_1 - y_0}{\sigma^2}\right).$$

Similarly,

$$L_2(y) = \exp\left(\frac{y_2 - y_0}{\sigma^2}\right).$$

Therefore, $L_1(y) \leq 1$ and $L_2(y) \leq \frac{1}{2}$ implies

$$\exp\left(\frac{y_1 - y_0}{\sigma^2}\right) \leq 1 \quad \text{and} \quad \exp\left(\frac{y_2 - y_0}{\sigma^2}\right) \leq \frac{1}{2}$$

$$\Rightarrow \quad \boxed{y_1 \leq y_0} \quad \text{and} \quad \boxed{y_2 \leq y_0 - 2\sigma^2 \ln 2}$$

And you can do the rest, following the same argument.

<u>Remark</u>: For any vector $\underline{y} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}$, the Euclidean norm (or the l_2-norm) is defined as

$$\|y\|_2^2 = y_0^2 + y_1^2 + y_2^2.$$

In general $\|y\|_2^2 = \sum_{j=0}^{M-1} y_j^2$ for any M-vector.