Exercise 1a p. 94

$$L(y) = \left(\frac{\lambda_{1-y}}{1-\lambda_y}\right)^{1-2y}$$

Compare to (uniform costs and equal priors)

$$\tau = \frac{\pi_0 C_{10}}{\pi_1 C_{01}} = 1$$

$$\Rightarrow \quad L(y) \underset{H_0}{\overset{H_1}{\gtrless}} \tau = 1$$

Bayes test: $\delta_B(y) = \begin{cases} y & \text{if } 1-\lambda_1 \geq \lambda_0 \\ 1-y & \text{"} \ 1-\lambda_1 < \lambda_0 \end{cases}$

$\boxed{\text{For } \lambda_0 = \lambda_1 = \lambda}$, we obtain

$$\delta_B(y) = \begin{cases} y & \text{if } \lambda \leq 1/2 \\ 1-y & \text{"} \ \lambda > 1/2 \end{cases}$$

$\Rightarrow$ decision region

$$\Gamma_1 = \begin{cases} \{y=1\} & \text{if } \lambda \leq 1/2 \\ \{y=0\} & \text{if } \lambda > 1/2 \end{cases}$$

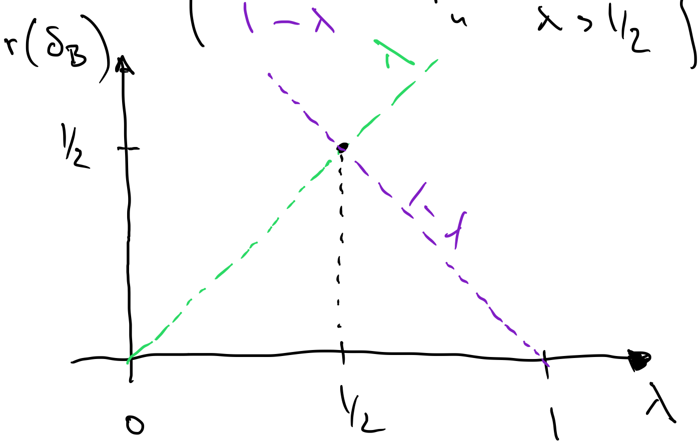
For equal priors, i.e. $\pi_0 = \pi_1 = 1/2$,

$\Rightarrow$ Bayes risk $r(\delta_B) = \pi_0 P(\Gamma_1 | H_0) + (1-\pi_0) P(\Gamma_0 | H_1)$

$$\Leftrightarrow r(\delta_B) = \begin{cases} \frac{1}{2}\Big[\underbrace{P(y=1 \mid H_0)}\ + \underbrace{P(y=0 \mid H_1)}\Big], & \lambda \le 1/2 \\ \text{given by channel model} & \\ \frac{1}{2}\Big[\overbrace{P(y=0 \mid H_0)}\ + \overbrace{P(y=1 \mid H_1)}\Big], & \lambda > 1/2 \end{cases}$$

$$= \begin{cases} \frac{1}{2}\left[\lambda + \lambda\right] & \text{for } \lambda \le 1/2 \\ \frac{1}{2}\left[1-\lambda + 1-\lambda\right] & \text{"}\ \lambda > 1/2 \end{cases}$$

$$= \left\{ \begin{array}{ll} \lambda & \text{for } \lambda \le 1/2 \\ 1-\lambda & \text{"}\ \lambda > 1/2 \end{array} \right\} = \min(\lambda, 1-\lambda)$$



Exercise 15 p.94

First find Bayes risk $V(\pi_0)$ as a function of π_0 for the symmetric case $\lambda_0 = \lambda_1 = \lambda$

(first find test, define the risk r_B, where π_0 is a parameter):

$$L(y) = \left\{ \begin{array}{ll} \frac{\lambda}{1-\lambda}, & y=0 \\ \frac{1-\lambda}{\lambda}, & y=1 \end{array} \right\} = \left(\frac{\lambda}{1-\lambda}\right)^{1-2y} \gtrless \frac{\pi_0}{1-\pi_0}$$

$$\tau\Big|_{\substack{\text{uniform}\\\text{costs}}} = \frac{\pi_0}{\pi_1}\,\frac{\cancel{c_{10}}^{\,1} - \cancel{c_{00}}^{\,0}}{\cancel{c_{01}}^{\,1} - \cancel{c_{11}}^{\,0}} = \frac{\pi_0}{1-\pi_0}$$

$$\Rightarrow \Gamma_1 = \left\{ y \in \{0,1\} : \lambda^{1-2y}(1-\pi_0) \geq \pi_0 (1-\lambda)^{1-2y} \right\}$$

$$= \left\{ y = 0 \wedge \lambda(1-\pi_0) \geq \pi_0 (1-\lambda) \right\}$$

$$\cup \left\{ y = 1 \wedge \frac{1-\pi_0}{\lambda} \geq \frac{\pi_0}{1-\lambda} \right\}$$

$$= \underbrace{\left\{ y = 0 \wedge \pi_0 \leq \lambda \right\} \cup \left\{ y = 1 \wedge 1 - \pi_0 - \lambda \geq 0 \right\}}_{\Gamma_1(y,\pi_0,\lambda)}$$

$$\Rightarrow \Gamma_0 = \{0,1\} \setminus \Gamma_1(y,\pi_0,\lambda) = \Gamma_0(y,\pi_0,\lambda)$$

"Usual" approach:

$$L(y,\lambda) \underset{H_0}{\overset{H_1}{\gtrless}} f(\pi_0) = \frac{\pi_0}{1-\pi_0}$$

Problem: Γ_0, Γ_1 given as a function of y, π_0, λ

$\Rightarrow$ instead try another approach:

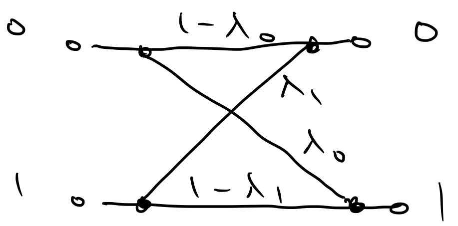
formulate $V(\pi_0)$ as a function of π_0 only

$\Rightarrow$ requires to express Γ_0 and Γ_1 as a function of π_0 first and then find $R_0(\delta_{\pi_0})$ and $R_1(\delta_{\pi_0})$ by an interpretation of the corresponding conditional PDFs.

Generalize the approach to the binary channel

Distinguish four cases of (Γ_0, Γ_1) which can be indeed expressed as a function of π_0 (namely via the dependence of the test threshold on π_0).

Consider generic binary channel



Bayes risk:

$$V(\pi_0) = r(\delta_{\pi_0}) = \pi_0 R_0(\delta_{\pi_0}) + (1-\pi_0) R_1(\delta_{\pi_0})$$

$$= \pi_0 \left[C_{00} P(\Gamma_0 | H_0) + C_{10} P(\Gamma_1 | H_0) \right] +$$

$$(1-\pi_0)\left[C_{01} P(\Gamma_0 | H_1) + C_{11} P(\Gamma_1 | H_1)\right]$$

Bayes test: $L(y) = \dfrac{p_1(y)}{p_0(y)} = \begin{cases} \dfrac{\lambda_1}{1-\lambda_0} & \text{for } y=0 \\ \dfrac{1-\lambda_1}{\lambda_0} & \text{"} \ y=0 \end{cases}$

to be compared to

$$\tau = \frac{\pi_0}{1-\pi_0} \frac{C_{10}-C_{00}}{C_{01}-C_{11}} .$$

In view of the fact that $L(y)$ depends on both λ_0 and λ_{11} and the threshold τ_B is already a function of π_0, we 'parametrize' the four possible cases for Γ_0 (and thus Γ_1)

case A : $\Gamma_0 = \Gamma = \{0,1\}$ ←

" B : $\Gamma_0 = \{0\}$

" C : $\Gamma_0 = \{1\}$

" D : $\Gamma_0 = \{\} = \phi$,

namely for the cases that can arise for the Bayes rule :

case	$L(0)$	$L(1)$	Bayes rule	risk in form of $r(\delta_{\pi_0}(y)) = \text{const.} + \pi_0 \alpha$
A	$L(0)<\tau_B$	$L(1)<\tau_B$	$\delta_{\pi_0}(y)=0$	$\Gamma_0=\{0,1\}=\Gamma$, $\Gamma_1=\phi$ $P(\Gamma_0\mid H_0)=P(\Gamma\mid H_0)=1$ $P(\Gamma_0\mid H_1)=P(\Gamma\mid H_1)=1$ $P(\Gamma_1\mid H_0)=P(\Gamma_1\mid H_1)=P(\phi)=0$ $\Rightarrow r(\delta_{\pi_0}(y))=C_{00}\pi_0+(1-\pi_0)C_{01}$ $=C_{01}+\pi_0\underbrace{(C_{00}-C_{01})}_{\alpha}$ $=V_A(\pi_0)$
B	$L(0)<\tau_B$	$L(1)\geq\tau_B$	$\delta_{\pi_0}(y)=y$	$\Gamma_0=\{0\}$, $\Gamma_1=\{1\}$ $P(\Gamma_0\mid H_0)=P\{\{0\}\mid H_0\}=1-\lambda_0$ $P(\Gamma_1\mid H_0)=P\{\{1\}\mid H_0\}=\lambda_0$ $P(\Gamma_0\mid H_1)=\lambda_1$ $P(\Gamma_1\mid H_1)=1-\lambda_1$ $\Rightarrow r(\delta_{\pi_0}(y))=$ $\pi_0[C_{00}(1-\lambda_0)+C_{10}\lambda_0]$ $+(1-\pi_0)[C_{01}\lambda+C_{11}(1-\lambda_1)]$ $=C_{01}\lambda_1+C_{11}(1-\lambda_1)$

				$+\pi_0\left[(c_{00}-c_{11})+\lambda_0(c_{10}-c_{00}) + \lambda_1(c_{11}-c_{01})\right]$ $= V_B(\pi_0)$
C	$L(0) \geq \tau_B$	$L(1) < \tau_B$	$\delta_{\pi_0}(y) = 1-y$	$\Gamma_0 = \{1\}$, $\Gamma_1 = \{0\}$ $P(\Gamma_0 \mid H_0) = \lambda_0$, $Pr\{\Gamma_1 \mid H_0\} = 1-\lambda_0$ $P(\Gamma_0 \mid H_1) = 1-\lambda_1$, $Pr\{\Gamma_1 \mid H_1\} = \lambda_1$ $\Rightarrow r(\delta_{\pi_0}(y)) =$ $\pi_0\left[c_{00}\lambda_0 + c_{10}(1-\lambda_0)\right]$ $+(1-\pi_0)\left[c_{01}(1-\lambda_1) + c_{11}\lambda_1\right]$ $= c_{11}\lambda_1 + c_{01}(1-\lambda_1)$ $+\pi_0\left[(c_{10}-c_{01})+\lambda_0(c_{00}-c_{10}) + \lambda_1(c_{01}-c_{11})\right]$ $= V_C(\pi_0)$
D	$L(0) \geq \tau_B$	$L(1) \geq \tau_B$	$\delta_{\pi_0}(y) = 1$	$\Gamma_0 = \phi$, $\Gamma_1 = \Gamma$ $\Rightarrow P(\Gamma_0 \mid H_0) = P(\Gamma_0 \mid H_1) = 0$ $P(\Gamma_1 \mid H_0) = P(\Gamma_1 \mid H_1) = 1$ $\Rightarrow r(\delta_{\pi_0}(y)) = \pi_0 c_{10} + (1-\pi_0)c_{11}$ $= c_{11} + \pi_0(c_{10} - c_{11})$ $= V_D(\pi_0)$

In order to determine the function $V(\pi_0)$ for $0 \leq \pi_0 \leq 1$, we can express the conditions in the first two columns of the above table as a function of π_0

$\Rightarrow$ convert conditions of $L(0) \overset{(\geq)}{<} \tau_B$ and $L(1) \overset{(\geq)}{<} \tau_B$ to conditions on π_0:

in the derivation below

derivation

① $L(0) \geq \tau_B \Leftrightarrow \frac{\lambda_1}{1-\lambda_0} \geq \tau_B = \frac{\pi_0}{1-\pi_0} \frac{c_{10}-c_{00}}{c_{01}-c_{11}}$

$\Leftrightarrow \lambda_1(1-\pi_0)(c_{01}-c_{11}) \geq \pi_0(1-\lambda_0)(c_{10}-c_{00})$

$$\Leftrightarrow \quad \pi_0 \leq \underbrace{\frac{\lambda_1(c_{01}-c_{11})}{(1-\lambda_0)(c_{10}-c_{00})+\lambda_1(c_{01}-c_{11})}}_{\pi_{0,1}}$$

② $L(1) \geq \tau_B \Leftrightarrow \frac{1-\lambda_1}{\lambda_0} \geq \tau_B = \frac{\pi_0}{1-\pi_0} \frac{c_{10}-c_{00}}{c_{01}-c_{11}}$

$\Leftrightarrow (1-\lambda_1)(1-\pi_0)(c_{01}-c_{11}) \geq \lambda_0\pi_0(c_{10}-c_{00})$

$$\Leftrightarrow \quad \pi_0 \leq \underbrace{\frac{(1-\lambda_1)(c_{01}-c_{11})}{(1-\lambda_1)(c_{01}-c_{11})+\lambda_0(c_{10}-c_{00})}}_{\pi_{0,2}}$$

'Connect' the four cases with the resulting intervals for π_0 by intersecting the corresponding intervals:

$\boxed{\text{case A}}$ $\quad L(0) < \tau_B \wedge L(1) \leq \tau_B$

$$\Leftrightarrow \quad \pi_0 > \pi_{0,1} \ \wedge \ \pi_0 > \pi_{0,2}$$

$$\Leftrightarrow \quad \pi_0 > \max(\pi_{0,1}, \pi_{0,2}) = \overline{\pi}_0$$

$$\Rightarrow \quad V(\pi_0) = \text{risk for case A and } \boxed{\overline{\pi}_0 < \pi_0 \leq 1}$$

$$= V_A(\pi_0) = c_{01} + \pi_0 (c_{00} - c_{01})$$

$$\boxed{\text{case B}} \qquad L(0) < \tau_B \ \wedge \ L(1) \geq \tau_B$$

$$\Leftrightarrow \quad \pi_0 > \pi_{0,1} \ \wedge \ \pi_0 \leq \pi_{0,2}$$

$$\Leftrightarrow \quad \boxed{\pi_{0,1} < \pi_0 \leq \pi_{0,2}}$$

$$\Rightarrow \quad V(\pi_0) = V_B(\pi_0)$$

$$\boxed{\text{case C}} \qquad L(0) \geq \tau_B \ \wedge \ L(1) < \tau_B$$

mutually exclusive !

$$\Leftrightarrow \quad \pi_0 \leq \pi_{0,1} \ \wedge \ \pi_0 > \pi_{0,2}$$

$$\Leftrightarrow \quad \boxed{\pi_{0,2} < \pi_0 \leq \pi_{0,1}}$$

$$\Rightarrow \quad V(\pi_0) = V_C(\pi_0)$$

$$\boxed{\text{case D}} \qquad L(0) \geq \tau_B \ \wedge \ L(1) \geq \tau_B$$

$$\Leftrightarrow \quad \pi_0 \leq \pi_{0,1} \ \wedge \ \pi_0 \leq \pi_{0,2}$$

$$\Leftrightarrow \quad \pi_0 \leq \underbrace{\min(\pi_{0,1}, \pi_{0,2})}_{\underline{\pi}_0} \ \Rightarrow \ \boxed{0 \leq \pi_0 \leq \underline{\pi}_0}$$

$$\Rightarrow \quad V(\pi_0) = V_D(\pi_0) = c_{11} + \pi_0 (c_{10} - c_{11})$$

<u>How to deal with cases B and C?</u>

Only <u>one</u> case is relevant, which is given by case B if $\pi_{0,1} < \pi_{0,2}$

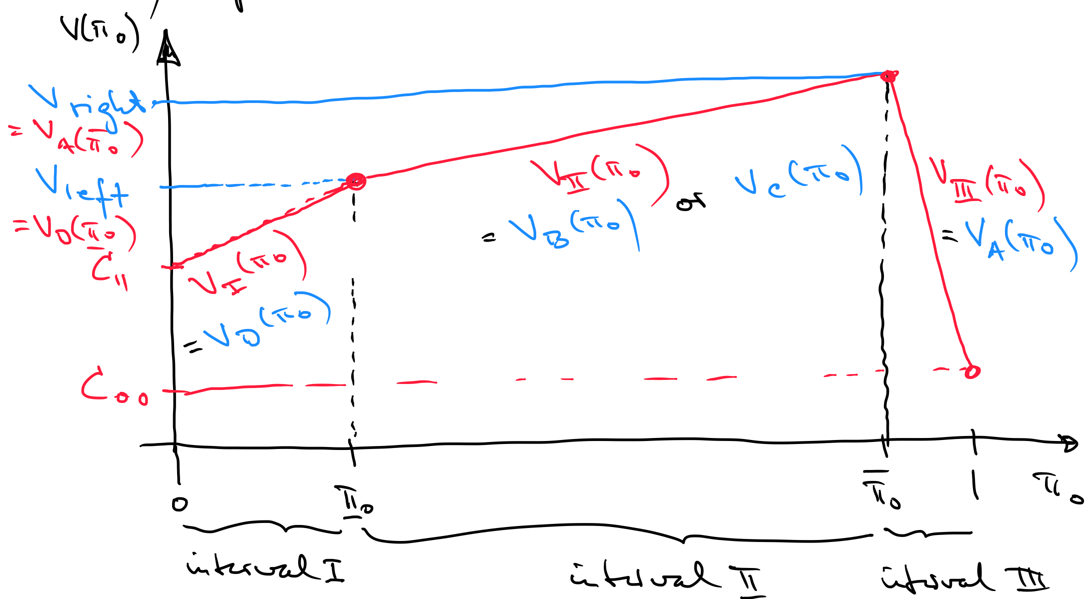
case C if $\pi_{0,2} < \pi_{0,1}$

In any case, the interval for π_0 is given by

$$\min(\pi_{0,1}, \pi_{0,2}) = \underline{\pi}_0 < \pi_0 < \overline{\pi}_0 = \max(\pi_{0,1}, \pi_{0,2}).$$

Clearly, one could check which case applies and choose the minimum Bayes risk as either $V_B(\pi_0)$ (if $\pi_{0,1} < \pi_{0,2}$) or $V_C(\pi_0)$ (if $\pi_{0,2} < \pi_{0,1}$).

This, in view of the continuity of $V(\pi_0)$ is not necessary since we know that $V(\pi_0)$ is piecewise linear and we only have THREE (not four) cases to distinguish for $0 \le \pi_0 \le 1$. We only thus connect $V(\underline{\pi}_0)$ in a linear way with $V(\overline{\pi}_0)$ to tell the shape of $V(\pi_0)$ for $\underline{\pi}_0 \le \pi_0 \le \overline{\pi}_0$.



From the figure, we obtain

$$V_{\underline{\pi}}(\pi_0) = V_{left} + \frac{V_{right} - V_{left}}{\overline{\pi}_0 - \underline{\pi}_0}(\pi_0 - \underline{\pi}_0)$$

In order to circumvent the general case of a minimax rule (depending on e.g. whether V_{left} or V_{right} is larger), we consider now the special case in the exercise, namely

$$\boxed{\lambda_0 = \lambda_1 = \lambda} \quad \text{and} \quad \boxed{\begin{matrix} C_{11} = C_{00} = 0 \\ C_{10} = C_{01} = 1 \end{matrix}}$$

$$\Leftrightarrow \boxed{C_{ij} = 1 - \delta_{i-j}}$$

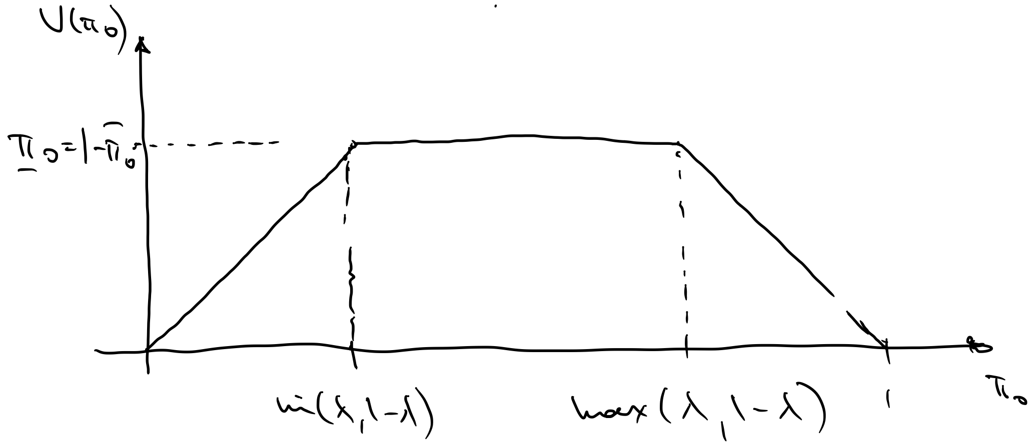
For these values

$$\Rightarrow \pi_{0,1} = \frac{\lambda}{1-\lambda+\lambda} = \lambda \quad ; \quad \pi_{0,2} = \frac{1-\lambda}{1-\lambda+\lambda} = 1-\lambda$$

$\Rightarrow$ so that

$$\underline{\pi}_0 = \min(\pi_{0,1}, \pi_{0,2}) = \min(\lambda, 1-\lambda)$$
$$\overline{\pi}_0 = \max(\pi_{0,1}, \pi_{0,2}) = \max(\lambda, 1-\lambda)$$

Furthermore, $V_D(\pi_0) = \pi_0$ and $V_A(\pi_0) = 1 - \pi_0$, so that $V_{left} = V_{right}$

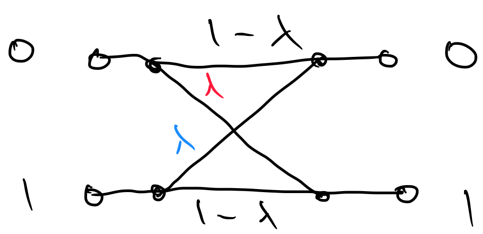


If we assume $\lambda < 1/2$ (or $\lambda > 1/2$), then $\pi_{0,1} < \pi_{0,2}$ (or $\pi_{0,2} < \pi_{0,1}$) and case B (or case C, resp.) provides a minimax test, since any π_0 with $\underline{\pi}_0 < \pi_0 < \bar{\pi}_0$ represents a least favorable prior π_L. Thus, $\lambda < 1/2$ (or $\lambda > 1/2$), we obtain

$$\delta_{\pi_L} = y \qquad \left(\text{or } \delta_{\pi_L} = 1 - y\right).$$

A randomization in this case is not required (only if $\lambda = 1/2$).

To see that this choice is indeed providing a risk being independent of π_0, consider again the BSC shown:



$$r(\delta_{\pi_L}) = \pi_0 P(T_1 \mid H_0) + (1-\pi_0) P(T_0 \mid H_1)$$

$$= \pi_0 \lambda + (1-\pi_0)\lambda = \lambda \neq f(\pi_0)$$