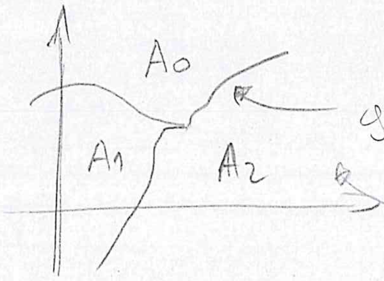
(7) Osnovne metode za klasifikaciju — (7.1) Bayesov klasifikator

↳ $Y \in S = \{0, 1, \dots, K-1\}$ za $K \geq 2$, $X \in \mathbb{R}^p$

⇒ tražimo $\hat{f} = \hat{f}(T) : \mathbb{R}^p \to S$

Svaka $\hat{f}$ dijeli $\mathbb{R}^p$ na $A_k := \hat{f}^{-1}(\{k\})$, $k \in S$

($p=2$)



granice odluke

[dakle, za $x \in A_2$, $\hat{f}(x) = 2$]

Svaka f-ja gubitka $L : S \times S \to \mathbb{R}$ je zapravo $K \times K$ matrica

[$L(a,b)$ = trošak ako stvarnu klasu a predvidimo s b]

↳ $L(f) := \mathbb{E}\left[L(Y, f(X))\right]$, te $\forall x \in \mathbb{R}^p$

$$L_x(f) := \mathbb{E}\left[L(Y, f(x)) \mid X = x\right]$$

$$= \sum_{k \in S} L(k, f(x)) \boxed{\underbrace{\mathbb{P}(Y = k \mid X = x)}_{=: p_k(x)}} \qquad (7.1)$$

↳ Bayesov klasifikator je $f^* : \mathbb{R}^p \to S$ dan s

$$f^*(x) = \underset{c \in S}{\operatorname{argmin}} \sum_{k \in S} L(k, c)\, p_k(x), \qquad (7.2)$$

te očito vrijedi $\forall f : \mathbb{R}^p \to S$, $\forall x \in \mathbb{R}^p$,

$$L_x(f^*) \leq L_x(f), \quad \boxed{L(f^*) \leq L(f)}$$

"Bayesov rizik"

[Naravno, f^* ne znamo jer ne znamo $p_k(x)$]

Pz Ako je $L(a,b) = \begin{cases} 1, & a=b \\ 0, & \text{inače} \end{cases}$ ("0-1 gubitak"),

vrijedi

$$f^*(x) = \underset{h\in S}{\mathrm{argmax}}\ p_h(x), \quad x\in\mathbb{R}^p \qquad (7.3)$$

[Naravno, f^* iz (7.2) ne znamo, jer ne znamo $\underline{\underline{p_h(x)}}$!]

↳ većina metoda prvo procjenjuje $p_h(x) = \mathbb{P}(Y=h \mid X=x)$, $x\in\mathbb{R}^p$ sa $\hat{p}_h(x)$ te imitira f^*:

$$\hat{f}(x) := \underset{c\in S}{\mathrm{argmax}} \sum_{h\in S} L(h,c)\,\boxed{\hat{p}_h(x)}, \quad x\in\mathbb{R}^p \qquad (7.4)$$

$$\left(0\text{-}1 \text{ gub.} \Rightarrow \hat{f}(x) = \underset{c}{\mathrm{argmax}}\ \hat{p}_c(x)\right) \qquad (7.5)$$