Clasificación

Clasificación binaria

- **Planteamiento del problema**
  - **Hipótesis**: H_0 y H_1. Complementarias
  - **Vector de observaciones**: $\mathbf{x} = [\, x[0]\; x[1]\; \cdots\; x[N-1]\,]^T \in \mathbb{R}^N$
    - Si la hipótesis correcta es H_0, $\mathbf{x}$ es una realización del vector aleatorio $\mathbf{X}|H_0$.
    - Si la hipótesis correcta es H_1, $\mathbf{x}$ es una realización del vector aleatorio $\mathbf{X}|H_1$
  - **Modelo:**
    - $\mathbf{X}|H_0 \quad \sim \quad f(\mathbf{x}|H_0)$
    - $\mathbf{X}|H_1 \quad \sim \quad f(\mathbf{x}|H_1)$
  - **Probabilidades a priori**: $P(H_0)$ y $P(H_1)$.
  - **Probabilidades a posteriori**: $P(H_0|\mathbf{x})$ y $P(H_1|\mathbf{x})$.

- **Regla de clasificación MAP y LRT**

$$P(\mathbf{x}|H_1) \underset{H_0}{\overset{H_1}{\gtrless}} P(\mathbf{x}|H_0) \quad \Rightarrow \quad \frac{f(\mathbf{x}|H_1)}{f(\mathbf{x}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)} = \gamma_0 \tag{1}$$

Para otros criterios de clasificación cambia γ_0, pero el LRT siempre es la regla óptima:

$$LR(\mathbf{x}) = \frac{f(\mathbf{x}|H_1)}{f(\mathbf{x}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma_0 \quad \Rightarrow \quad T(\mathbf{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{2}$$

  - $T(\mathbf{x})$: estadístico
  - γ_0: umbral de decisión (su valor depende del criterio de optimalidad)

- **Casos posibles:**
  - Aciertos: $H_0|H_0 \iff T(\mathbf{x}|H_0) < \gamma, \qquad H_1|H_1 \iff T(\mathbf{x}|H_1) > \gamma$
  - Errores: $H_1|H_0 \iff T(\mathbf{x}|H_0) > \gamma\ , \qquad H_0|H_1 \iff T(\mathbf{x}|H_1) < \gamma$

Jesús Pérez Arriaga

- **Prestaciones del clasificador:**
  - **Probabilidad de falsa alarma:** $P_{FA}(\gamma) = P(H_1|H_0) = P(T(\mathbf{X}|H_0) > \gamma)$
  - **Probabilidad de detección:** $P_D(\gamma) = P(H_1|H_1) = P(T(\mathbf{X}|H_1) > \gamma)$
  - **Curva ROC:** $P_{FA}(\gamma)$ vs. $P_D(\gamma)$ para distintos valores de γ.
  - **Probabilidad de error:**

$$\begin{aligned} P_e(\gamma) &= P(H_1|H_0)\ P(H_0) + P(H_0|H_1)\ P(H_1) \\ &= P_{FA}(\gamma)\ P(H_0) + (1 - P_D(\gamma))\ P(H_1) \end{aligned}$$

  - **Coste Bayesiano:**

$$\begin{aligned} C(\gamma) &= C_{0,0}\ P(H_0|H_0)\ P(H_0) + C_{0,1}\ P(H_0|H_1)\ P(H_1) + \\ & \quad C_{1,0}\ P(H_1|H_0)\ P(H_0) + C_{1,1}\ P(H_1|H_1)\ P(H_1) \\ &= C_{0,0}\ (1 - P_{FA}(\gamma))\ P(H_0) + C_{0,1}\ (([1 - P_D(\gamma))\ P(H_1) + \\ & \quad C_{1,0}\ P_{FA}(\gamma)\ P(H_0) + C_{1,1}\ P_D(\gamma)\ P(H_1), \end{aligned}$$

    donde $C_{1,0} > C_{0,0}$ y $C_{0,1} > C_{1,1}$.
    - Coincide con la probabilidad de error si $C_{0,1} = C_{1,0} = 1$ y $C_{0,0} = C_{1,1} = 0$.

- **Clasificadores óptimos según el criterio de clasificación**

  Todos se basan en el LRT. Cada criterio de clasificación tiene asociado un valor de γ_0.

  - **Clasificador de mínima probabilidad de error:**

$$\gamma_0 = \frac{P(H_0)}{P(H_1)}$$

    - Es el clasificador MAP.
    - Se reduce al clasificador ML cuando $P(H_0) = P(H_1) = 1/2$

  - **Clasificador de mínimo coste Bayesiano**

$$\gamma_0 = \frac{P(H_0)\ (C_{1,0} - C_{0,0})}{P(H_1)\ (C_{0,1} - C_{1,1})}$$

    - El clasificador MAP es un caso particular donde $C_{0,0} = C_{1,1} = 0$ y $C_{0,1} = C_{1,0}$.

  - **Clasificador Neyman-Pearson**

$$P_{FA}(\gamma) = \alpha \Rightarrow \gamma$$

    - Se conoce como clasificador/criterio CFAR (Constant False Alarm Rate).
    - No requiere conocer probabilidades a priori.
    - Teorema de Neyman-Pearson: Maximiza la P_D para la P_{FA} prefijada.

Clasificación M-aria

- **Planteamiento del problema**

  - **Hipótesis**: $H_0, H_1, \ldots, H_{M-1}$. Complementarias

  - **Vector de observaciones**: $\mathbf{x} = [\, x[0]\; x[1]\; \cdots\; x[N-1]\,]^T \in \mathbb{R}^N$

    Si la hipótesis correcta es H_m, $\mathbf{x}$ es una realización del vector aleatorio $\mathbf{X}|H_m$.

  - **Modelo:** $\mathbf{X}|H_m \quad \sim \quad f(\mathbf{x}|H_m)$

  - **Probabilidades a priori y a posteriori**: $P(H_m)$, $P(H_m|\mathbf{x})$.

- **Regla de clasificación MAP**

$$P(H_m|\mathbf{x}) \overset{H_m}{\gtrless} P(H_i|\mathbf{x}), \quad \forall i \;\; \Rightarrow \;\; g_m(\mathbf{x}) \overset{H_m}{\gtrless} g_i(\mathbf{x}), \quad \forall i$$

  - Funciones discriminantes: $g_m(\mathbf{x}) = f(\mathbf{x}|H_m)\; P(H_m)$.

- **Probabilidad de error**

$$P_e = \sum_{i=0}^{M-1} \sum_{j \neq i}^{M-1} P(H_j|H_i) P(H_i) = 1 - \sum_{m=0}^{M-1} P(H_m|H_m) P(H_m)$$

- **Clasificador de mínimo coste Bayesiano**

  - **Coste Bayesiano:** $C = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} C_{j,i}\; P(H_j|H_i) P(H_i)$

    Coincide con P_e cuando $C_{i,j} = 1, \forall j \neq i$ y $C_{i,i} = 0, \forall i$.

  - **Clasificador:**

$$\underset{m=0,\ldots,M-1}{\text{argmax}} \quad g_m(\mathbf{x}) = -\sum_{i=0}^{M-1} C_{m,i} P(H_i|\mathbf{x}) = -\sum_{i=0}^{M-1} C_{m,i} f(\mathbf{x}|H_i) P(H_i)$$

    El clasificador MAP es un caso particular donde $C_{m,i} = 1, \forall m \neq i$ y $C_{i,i} = 0, \forall i$.

- **Matriz de confusión**