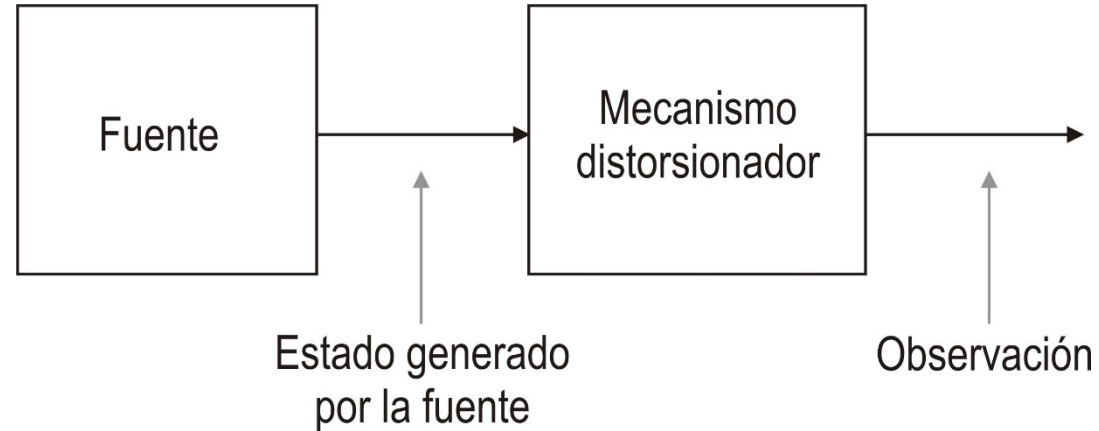


Teoría de la Decisión

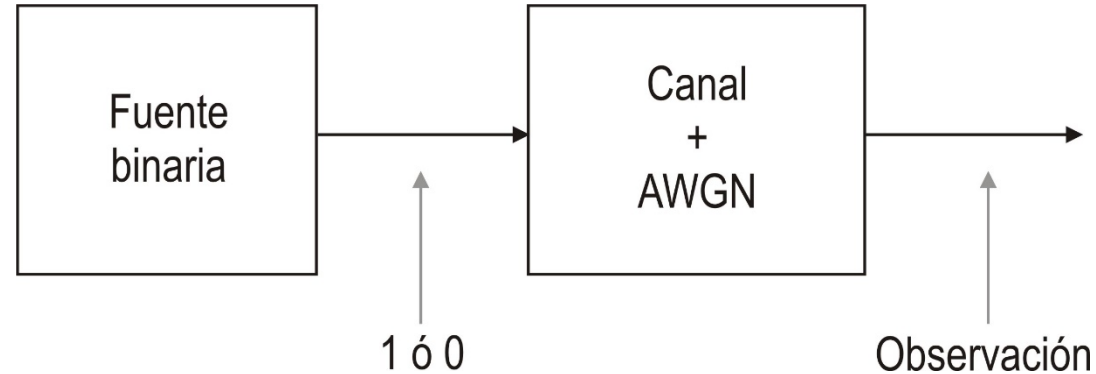
Clasificación

- Problema genérico



- La fuente puede generar un conjunto finito de estados $H_0, \dots, H_N$ llamados *hipótesis*
- **Objetivo:** determinar la hipótesis más probable a partir de la observación realizada
 - Basado en:
 - Conocimiento de las hipótesis
 - Conocimiento de las características del mecanismo distorsionador (ej.: canal, ruido, etc.)

- Ejemplo: comunicaciones digitales



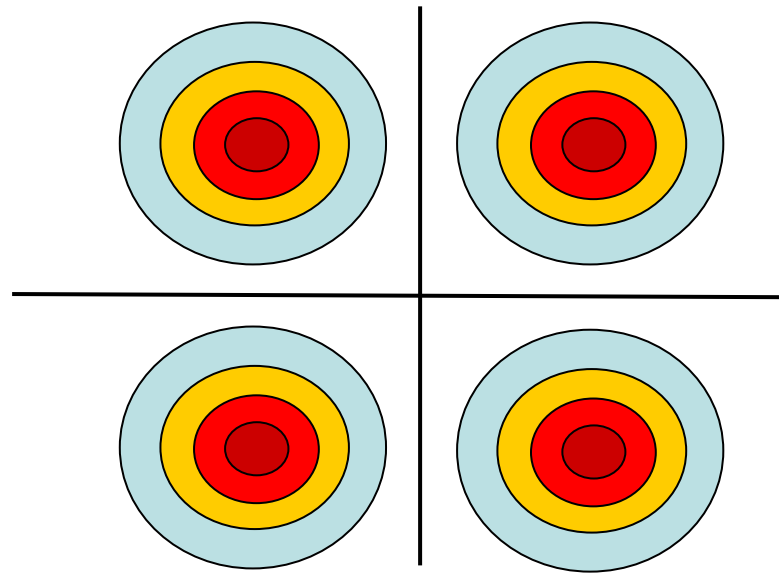
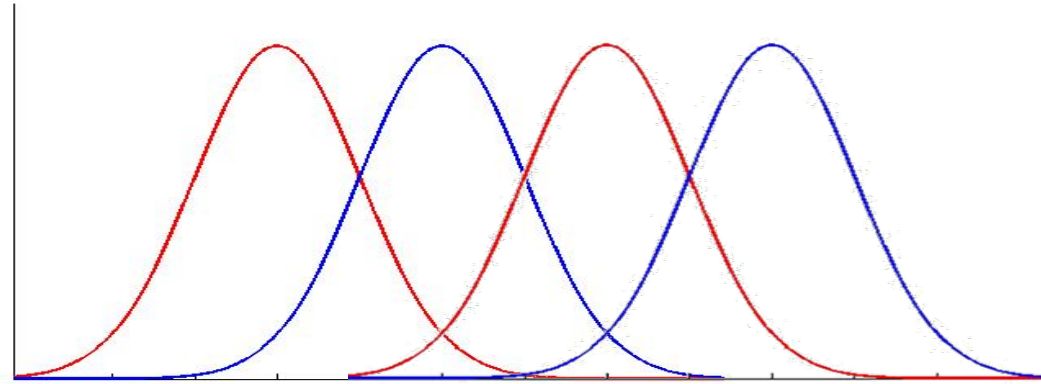
- Se trata de un problema con dos hipótesis (1 y 0)
- El mecanismo distorsionador consta del propio canal de comunicaciones y del ruido aditivo blanco gaussiano
- Cuando el número de hipótesis es dos, el problema se denomina **problema de decisión binario básico** (o simple)
- Cuando el número de hipótesis es mayor que dos, el problema de decisión se suele denominar **problema de clasificación**

Decisión entre M hipótesis

- El próximo caso de interés es aquel en el que debe escogerse entre M posibles hipótesis (no necesariamente 2)
- En el caso simple binario se consideran dos posibles salidas de la fuente, cada una de las cuales corresponde a una hipótesis
- En este nuevo caso se consideran M posibles salidas de la fuente entre las que hemos de tomar una decisión a partir de una observación
 - Es habitual que el problema con más de dos hipótesis se denomine **clasificación**
 - También es habitual que en clasificación las hipótesis se refieran como **clases**
- El fundamento del problema de clasificación es el mismo que el de decisión binaria, aunque, lógicamente, más general
- El espacio se divide en diferentes regiones asignadas a cada hipótesis en función de criterios de distancia o distancia ponderada

Decisión entre M hipótesis

- Ejemplos:

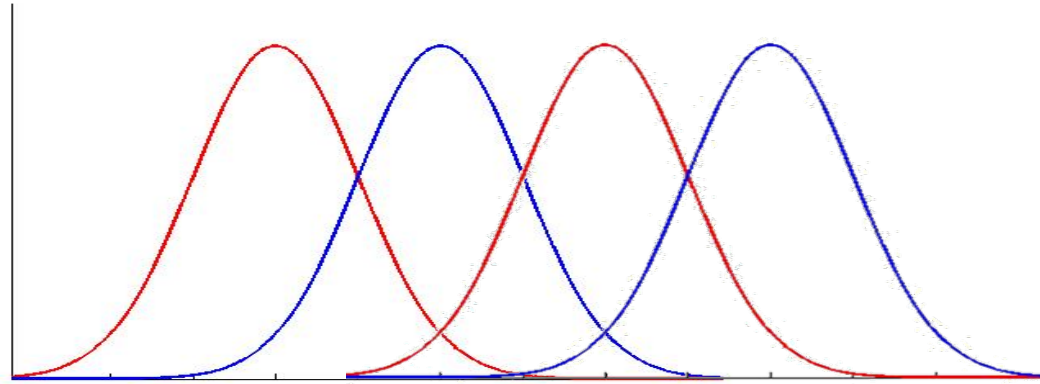


Decisión entre M hipótesis

- El problema de decisión entre M hipótesis puede plantearse según los mismos criterios que el binario:
 - Criterio de mínima probabilidad de error
Se maximiza la probabilidad de detección correcta, calculado a partir del teorema de probabilidad total
 - Decisión/clasificación Bayesiana
Se asigna un coste a cada decisión (M^2 posibles), se asume un conjunto de probabilidades a priori $P_0, P_1, \dots, P_{M-1}$, y se minimiza el riesgo
 - Generalización del criterio de Neyman-Pearson
Es posible, pero generalmente no se utiliza en la práctica

Decisión entre M
hipótesis
Criterio de mínima
probabilidad de error

- Partamos de un caso con M hipótesis equiprobables y una única observación x (caso unidimensional)
 - Ejemplo: comunicaciones digitales, 4 símbolos (-3, -1, 1 y 3)



- Teorema de probabilidad total: la probabilidad de error es

$$P_e = \sum_{i=1}^M P(H_i)P(error|H_i)$$

- Y la de clasificación correcta

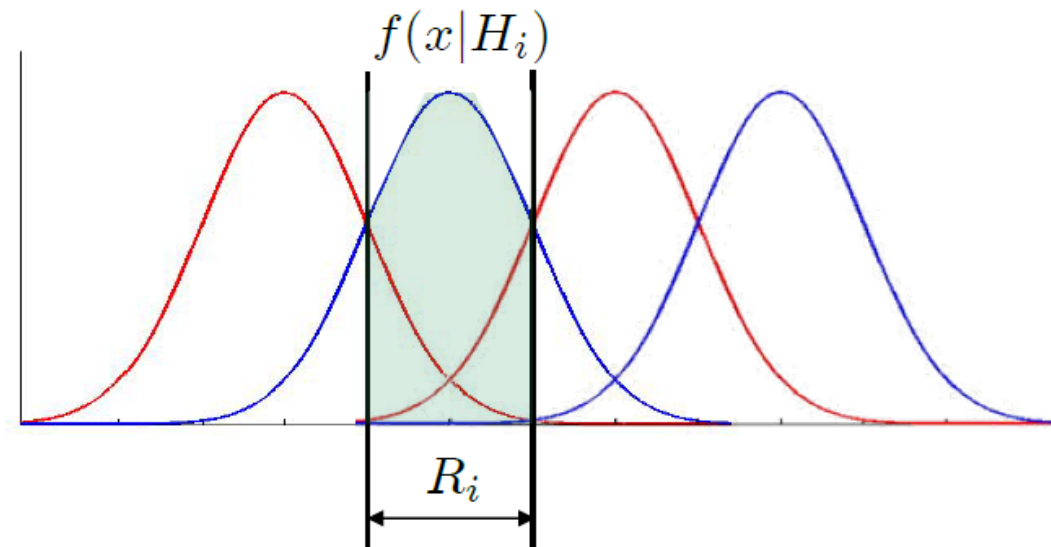
$$P_c = \sum_{i=1}^M P(H_i)P(acierto|H_i)$$

Decisión entre M
hipótesis
Criterio de mínima
probabilidad de error

- La minimización del error o la maximización del acierto conducen a la obtención de la región R_i :

$$R_i / f(x|H_i) > f(x|H_j) \quad \forall j \neq i$$

- Puede entenderse intuitivamente:
 - Comparación hipótesis a hipótesis por pares
 - Región de máxima densidad de probabilidad



$$R_i / f(x|H_i) > f(x|H_j) \quad \forall j \neq i$$

- Esta expresión realmente es un cociente de verosimilitud equivalente al de decisión ML (o caso equiprobable):

$$\frac{f(x|H_i)}{f(x|H_j)} \geq \frac{R_i}{R_j} \quad \forall j \neq i$$

- Si $R_{\bar{i}}$ representa todo el espacio distinto a la región R_i :

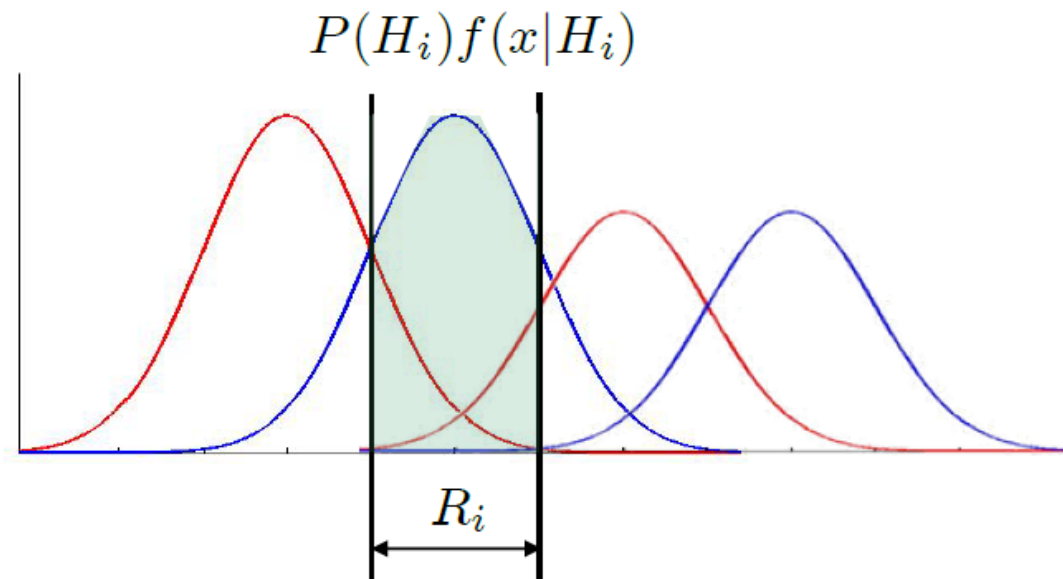
$$R_{\bar{i}} = \sum_{j=1, j \neq i}^M R_j$$

de algún modo se está realizando una decisión binaria simple con dos nuevas hipótesis, $\tilde{H}_i$ y $\tilde{H}_{noi}$, que ocupan respectivamente regiones del espacio R_i y $R_{\bar{i}}$

Decisión entre M
hipótesis
Criterio de mínima
probabilidad de error

- Si se tratase de hipótesis no equiprobables:

$$\frac{f(x|H_i)}{f(x|H_j)} \geq \frac{R_i}{R_j} \frac{P(H_j)}{P(H_i)} \quad \forall j \neq i$$



- Al igual que en decisión binaria, se ponderan las funciones densidad de probabilidad con las probabilidades a priori de cada clase

Decisión entre M
hipótesis
Criterio de mínima
probabilidad de error

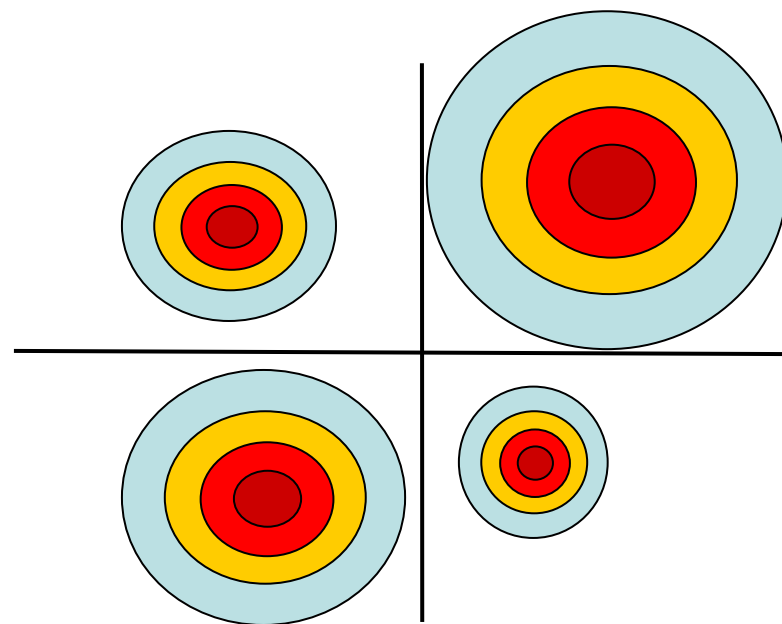
- Caso N -dimensional

$$\frac{f(\vec{x}|H_i)}{f(\vec{x}|H_j)} \geq \frac{R_i}{R_j} \frac{P(H_j)}{P(H_i)} \quad \forall j \neq i$$

con

$$\vec{x} = [x_1, x_2, \dots, x_N]$$

$$f(\vec{x}|H_i) = \prod_{n=1}^N f(x_n|H_i)$$



Decisión entre M
hipótesis
Criterio de mínima
probabilidad de error
Caso gaussiano

- Caso N -dimensional gaussiano con hipótesis equiprobables
- Partiremos del caso bidimensional

$$\frac{f(\vec{x}|H_i)}{f(\vec{x}|H_j)} \underset{R_j}{\overset{R_i}{\geq}} \frac{P(H_j)}{P(H_i)} \quad \forall j \neq i$$

$$P(H_1) = P(H_2) \dots = P(H_i) = P(H_j) = \dots$$

$$\left(\frac{1}{\sqrt{2\pi}\sigma} \right)^2 e^{-\frac{1}{2} \left[\left(\frac{x_1 - h_i}{\sigma} \right)^2 + \left(\frac{x_2 - h_i}{\sigma} \right)^2 \right]} \underset{R_j}{\overset{R_i}{>}} \left(\frac{1}{\sqrt{2\pi}\sigma} \right)^2 e^{-\frac{1}{2} \left[\left(\frac{x_1 - h_j}{\sigma} \right)^2 + \left(\frac{x_2 - h_j}{\sigma} \right)^2 \right]}$$

- Donde el vector de observación es

$$\vec{x} = [x_1, x_2]$$

- Desarrollando se obtiene que en la región R_i :

$$(x_1 - h_i)^2 + (x_2 - h_i)^2 < (x_1 - h_j)^2 + (x_2 - h_j)^2, \quad \forall j \neq i$$

que de nuevo es una medida de distancia (mínima distancia) donde se escoge la hipótesis i -ésima de coordenadas $[h_i, h_i]$

- Generalizando para el caso N -dimensional se obtiene que en R_i

$$\sum_{n=1}^N (x_n - s_i^{(n)})^2 < \sum_{n=1}^N (x_n - s_j^{(n)})^2, \quad \forall j \neq i$$

donde las coordenadas de cada hipótesis son

$$\vec{H}_i = [s_i^{(1)}, s_i^{(2)}, \dots, s_i^{(N)}]$$

y el vector de observaciones es

$$\vec{x} = [x_1, x_2, \dots, x_N]$$

¿Qué sucede si las hipótesis no son equiprobables en el caso general?

Decisión Bayesiana entre M hipótesis

- En el caso de dos hipótesis teníamos:
 - Se asume que se puede definir el **coste** de cada decisión:

C_{00} Coste de decidir H_0 cuando ha sucedido H_0 (acierto)

C_{10} Coste de decidir H_1 cuando ha sucedido H_0 (error)

C_{01} Coste de decidir H_0 cuando ha sucedido H_1 (error)

C_{11} Coste de decidir H_1 cuando ha sucedido H_1 (acierto)

- El coste total (riesgo) viene dado por

$$J = C_{00}P(H_0)P(H_0|H_0) + C_{10}P(H_0)P(H_1|H_0) \\ + C_{01}P(H_1)P(H_0|H_1) + C_{11}P(H_1)P(H_1|H_1)$$

- En el caso de M hipótesis generalizamos los costes:
 - El coste de cada decisión es C_{ij}
 - Primer subíndice: se decide la hipótesis i -ésima
 - Segundo subíndice: la hipótesis j -ésima es cierta
 - Se denota con R_i la región del espacio de observación en la que se decide la hipótesis H_i

- El coste (riesgo) total queda entonces como

$$J = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_j C_{ij} \int_{R_i} f(\vec{x}|H_j) dx$$

donde P_j es la probabilidad a priori de la hipótesis H_j

- Por sencillez estudiaremos el caso $M=3$ 
- Desarrollando encontramos que M hipótesis conducen a un espacio de decisión de un máximo de $M-1$ dimensiones

Bayesiano Ndimensional [Van Trees] Detection, Estimation, and Modulation Theory, Part I.pdf

Trabajo opcional:

Estudio del documento y presentación de un informe en Teams explicando el desarrollo y su significado.

Limite de entrega: 20 de abril (23:59h).

Teoría de la Decisión

Clasificación