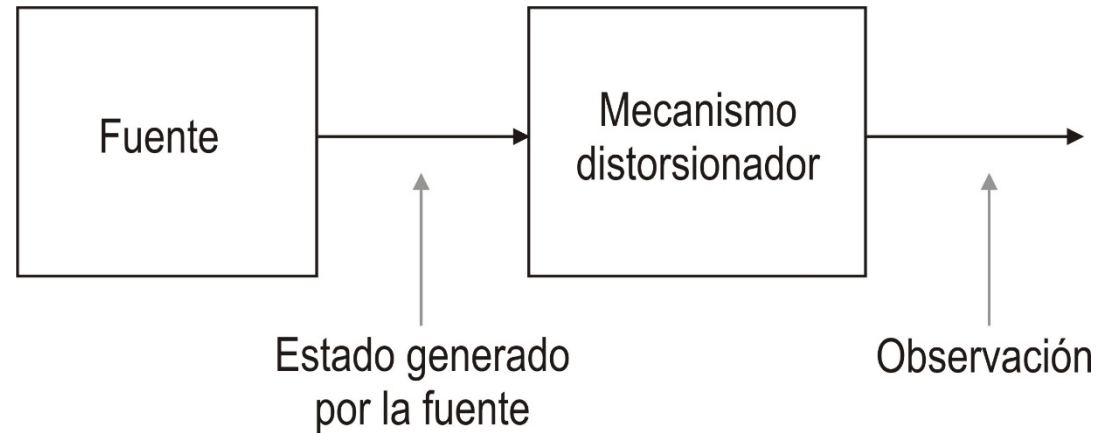


Teoría de la Decisión

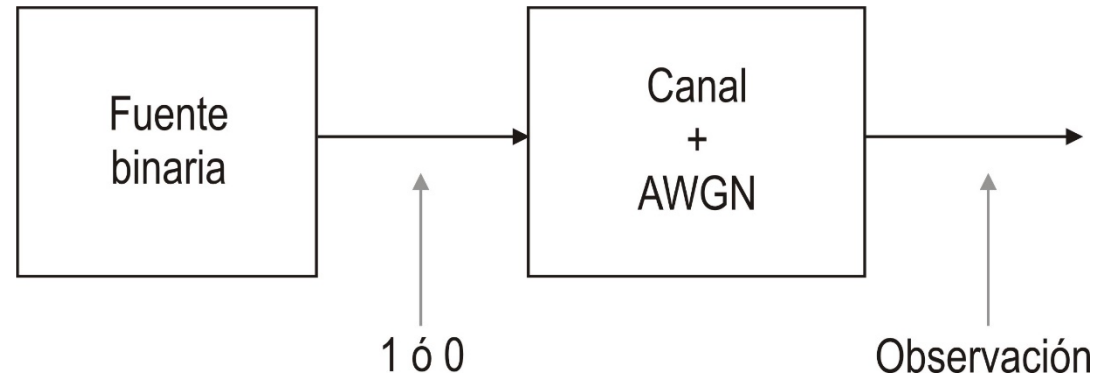
Decisión, criterios MAP y ML, mínima P_e

- 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

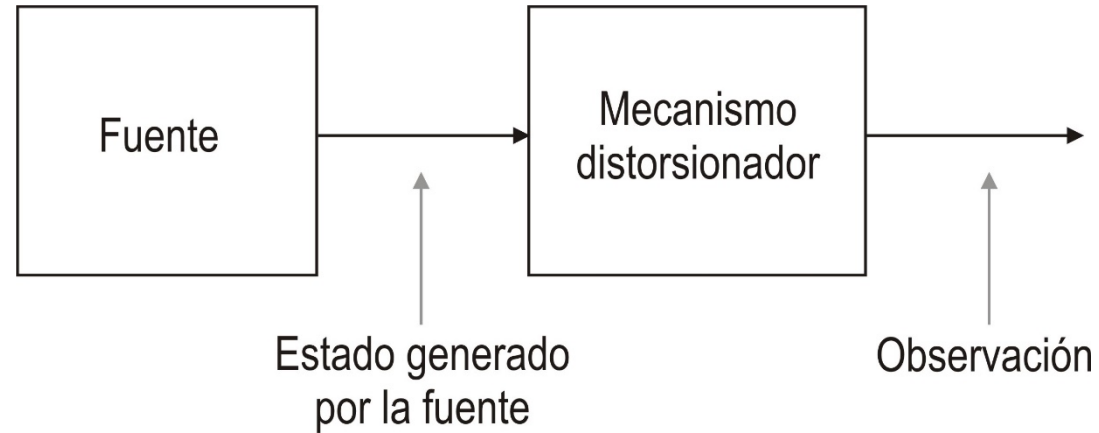


- Se trata de un problema con dos hipótesis (1 y 0)
- El mecanismo distorsionador consta del propio canal de comunicaciones y del ruido aditivo blanco gaussiano

-
- Cuando sólo existen dos hipótesis el problema se denomina **decisión** o **detección**
 - Cuando existen más de dos hipótesis se denomina **clasificación**

- La resolución de estos problemas se basa en técnicas estadísticas
- El mecanismo que produce la distorsión se modela como un proceso estocástico cuyas propiedades se conocen:
 - Media
 - Varianza
 - Función de densidad de probabilidad
- Conocidas las propiedades del proceso estocástico que produce la distorsión, y conocidas las hipótesis, se pueden conocer las propiedades del proceso resultante (observación) y determinar cuál es la hipótesis más probable dada la observación realizada

- Modelo típico:



- Un problema de decisión simple se divide en cuatro componentes importantes:
 - La fuente
 - Un mecanismo probabilístico de transición (*mecanismo distorsionador*)
 - El espacio de observaciones
 - La regla de decisión

- **La fuente** genera una salida de entre una serie de posibles elecciones (*hipótesis*). En el caso más sencillo, dos opciones (caso binario)
- En el caso de M hipótesis, se etiquetan como

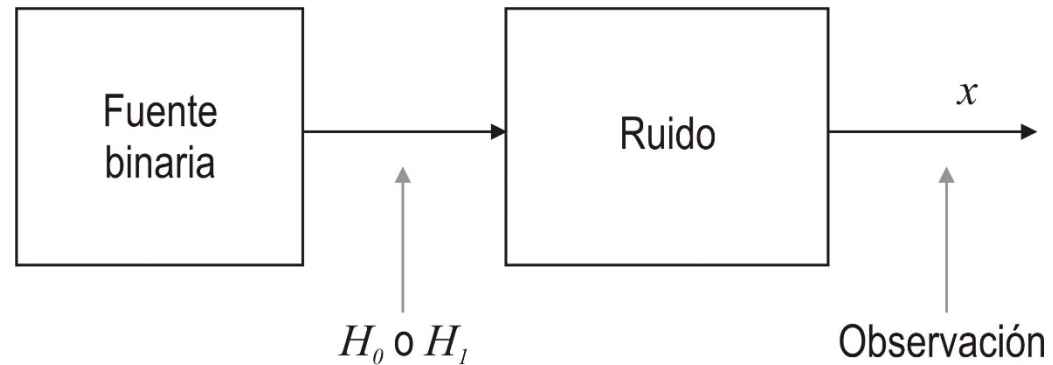
$$H_0, H_1, \dots, H_{M-1}$$

- Algunos mecanismos de fuente ejemplo:
 - Sistema de comunicaciones digitales (transmisor)
 - Envía “1” y “0” etiquetados como H_1 y H_0
 - Observación de un RADAR en una determinada dirección del espacio
 - H_1 representa la presencia de un blanco, y H_0 que no lo hay
 - Problema de diagnóstico médico a través del examen de un ECG
 - Presencia de un ataque al corazón (H_1) o no (H_0)
 - Clasificación de locutores mediante un registro de voz:
 - Alemán, Inglés o Español, y masculino o femenino: H_0 - H_5
 - Etc.

- **El mecanismo de transición** (o de distorsión) puede verse como un sistema que conoce cuál es la hipótesis cierta y, basado en este conocimiento, genera un punto en el **espacio de observación** según una cierta regla probabilística
- El **espacio de observación**, tal y como lo consideramos aquí, es de dimensión finita:
 - Las observaciones consisten en un conjunto de N números
 - La observación se puede representar como un punto en el espacio N -dimensional
 - Este problema se denomina **problema clásico de detección**
- **La regla de decisión** es el procedimiento mediante el cual se “determina” qué hipótesis es la correcta
 - Asigna cada punto del espacio de observación a una de las hipótesis
 - Es el objetivo del problema
- El diseño de la regla de decisión adecuada depende de cómo interactúan estos cuatro componentes del problema

Criterio de mínima probabilidad de error

- Al plantear la decisión a partir de la observación es necesario establecer un criterio de error adecuado
- Modelo binario típico:



- Dos posibles hipótesis: H_0 y H_1
- Por ejemplo, la decisión es óptima si la probabilidad de error es mínima
- Situaciones:
 - Decidir H_0 cuando se trataba de $H_1 \rightarrow$ *Error*
 - Decidir H_0 cuando se trataba de $H_0 \rightarrow$ *Acierto*
 - Decidir H_1 cuando se trataba de $H_1 \rightarrow$ *Acierto*
 - Decidir H_1 cuando se trataba de $H_0 \rightarrow$ *Error*

- Según el teorema de Probabilidad Total, la probabilidad de error viene dada por:

$$P_e = P(H_0)P(error|H_0) + P(H_1)P(error|H_1)$$

- $P(H_0)$ y $P(H_1)$ son conocidos: *Probabilidades a priori*
- La observación x puede ser:

$$x = h_0 + n$$

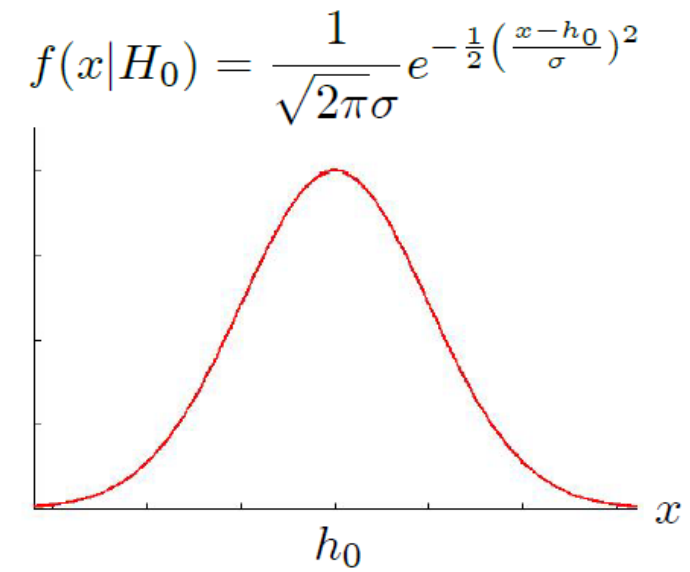
$$x = h_1 + n$$

donde n es el ruido que distorsiona la observación

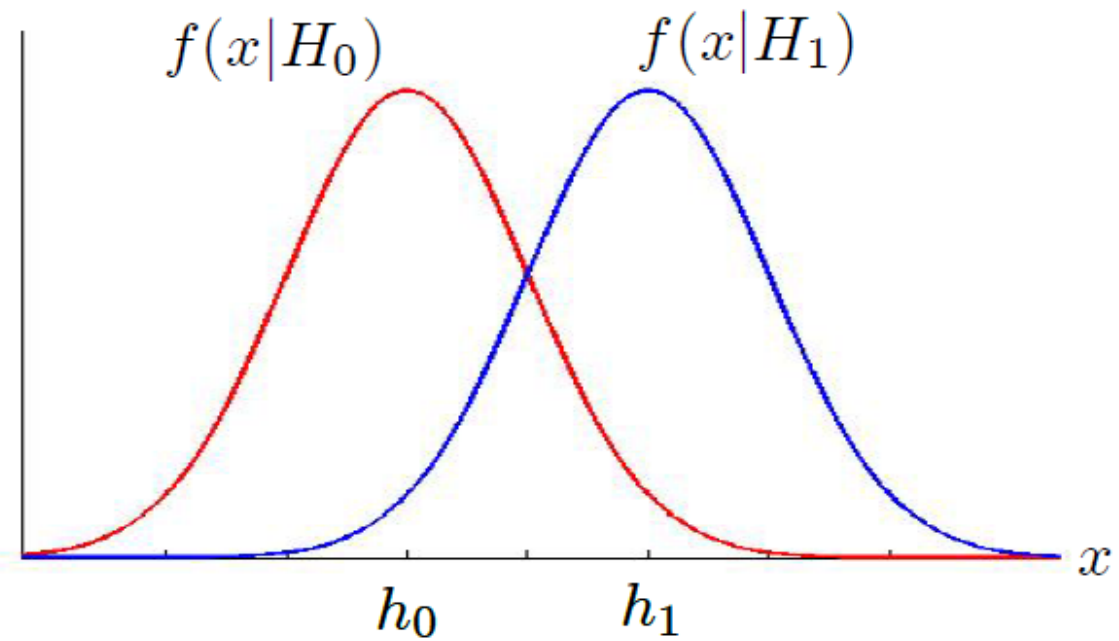
- De este modo, la función densidad de probabilidad (fdp) de la observación condicionada a la hipótesis H_0 dependerá de la fdp del ruido
- Por ejemplo, si el ruido es gaussiano de media 0 y varianza σ^2

$$f(x|H_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_0}{\sigma}\right)^2}$$

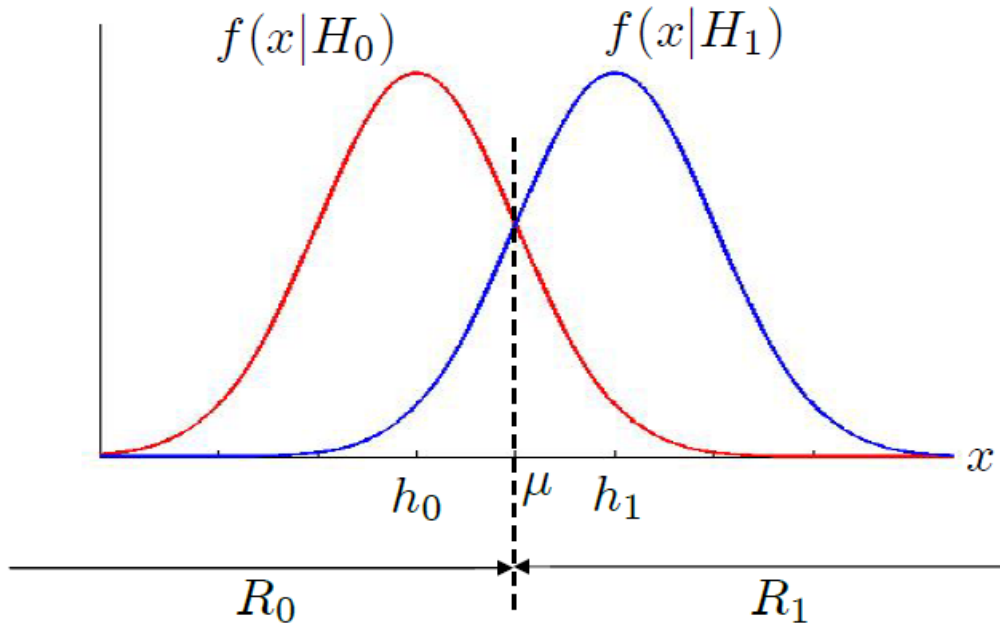
Criterio de mínima
probabilidad de error



Representando ambas fdp:



Criterio de mínima probabilidad de error

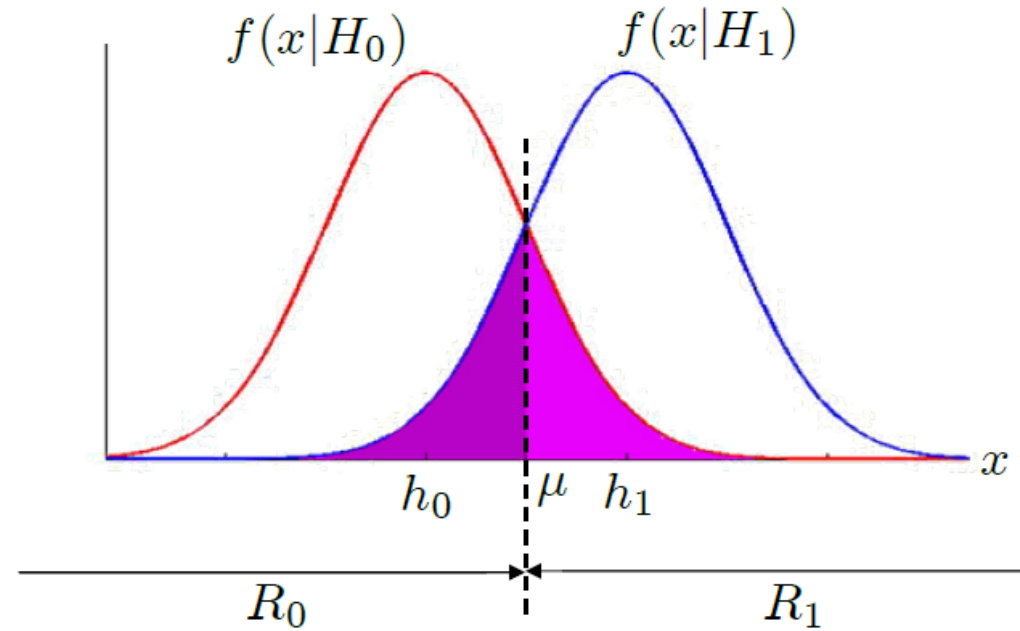


$$x \in R_0 \Leftrightarrow H_0$$

$$x \in R_1 \Leftrightarrow H_1$$

- La decisión se fundamenta en la hipótesis más probable según la observación obtenida.
- Se trata de minimizar la probabilidad de error

Criterio de mínima probabilidad de error



$$x \in R_0 \Leftrightarrow H_0$$

$$x \in R_1 \Leftrightarrow H_1$$

- La probabilidad de error para cada hipótesis es el área bajo su fdp en la región que “no le corresponde”. Por ejemplo:

$$P(\text{error}|H_0) = \int_{R_1} f(x|H_0) dx$$

$$P_e = P(H_0)P(error|H_0) + P(H_1)P(error|H_1)$$

$$P(error|H_0) = \int_{R_1} f(x|H_0)dx$$

$$P(error|H_1) = \int_{R_0} f(x|H_1)dx$$

Sumando y restando (la igualdad no cambia)

$$+P(H_0) \int_{R_0} f(x|H_0)dx - P(H_0) \int_{R_0} f(x|H_0)dx$$

La minimización de P_e conduce a

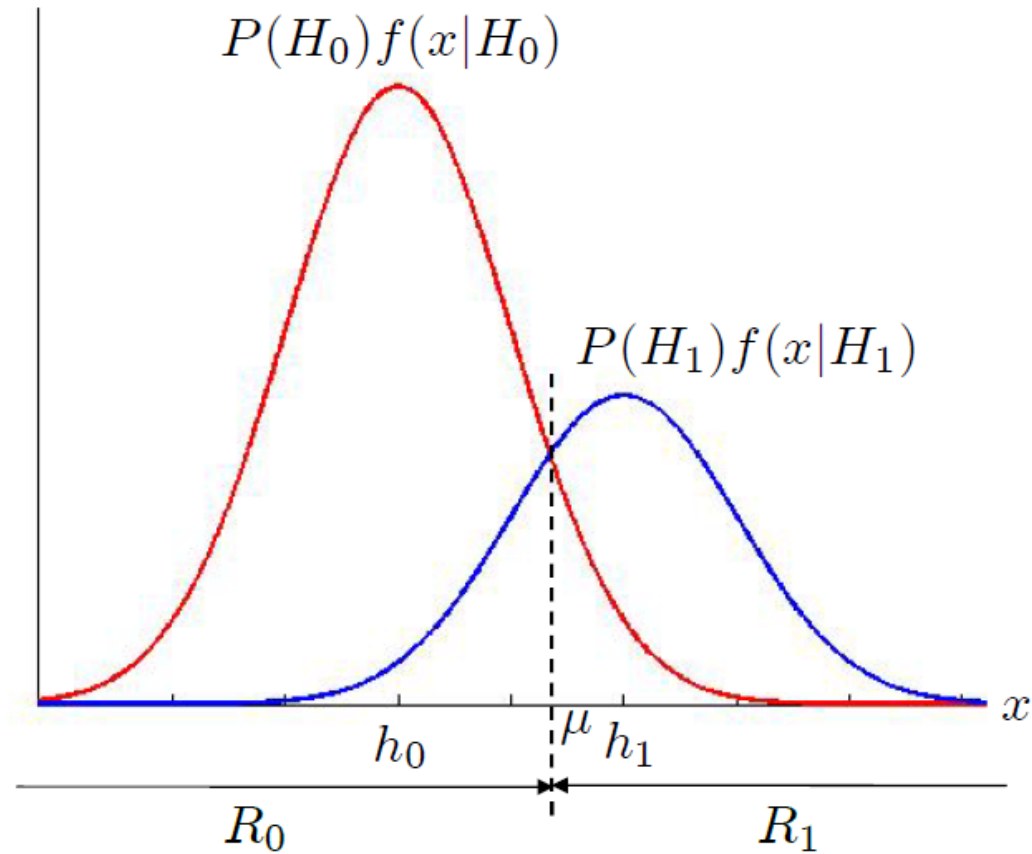
$$\frac{f(x|H_1)}{f(x|H_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)}{P(H_1)}$$

Cociente de verosimilitud

CRITERIO DE DECISIÓN MAP
(Maximum a posteriori)

Criterio de mínima probabilidad de error

- Gráficamente:



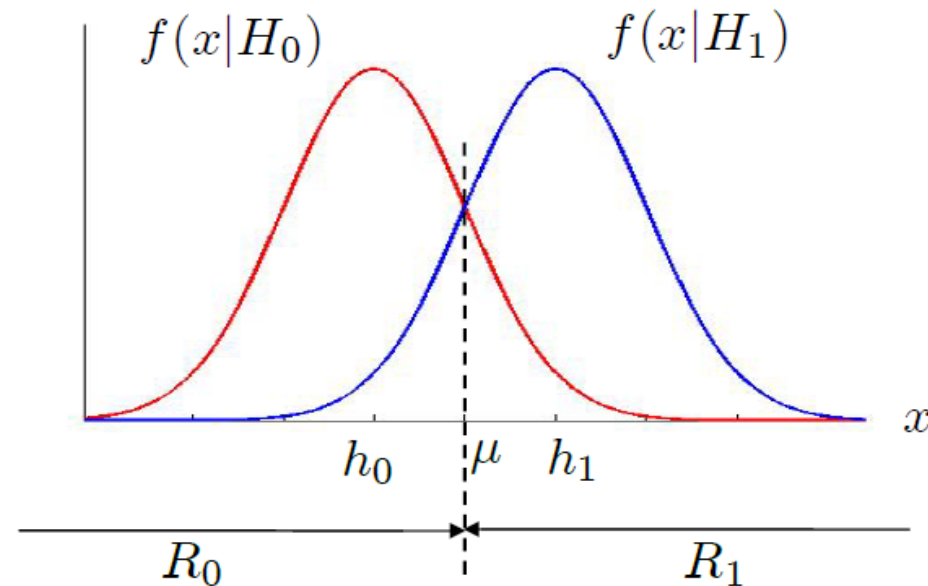
- El umbral se desplaza hacia la hipótesis menos probable reduciendo la región de decisión que le corresponde
- Mediante el Teorema de Bayes puede comprobarse que este criterio de decisión es equivalente a seleccionar la hipótesis H_i que tenga mayor probabilidad *a posteriori* $P(H_i|x)$

Criterio de mínima probabilidad de error

- Empleando el criterio de decisión MAP se considera información *a priori* de la fuente de datos (probabilidades de las hipótesis)
- Cuando dicha información no es accesible se emplea el criterio de decisión **ML** (Maximum Likelihood, máxima verosimilitud).

$$P(H_0) = P(H_1)$$

$$\frac{f(x|H_1)}{f(x|H_0)} \underset{H_0}{\overset{H_1}{\geq}} 1$$



Criterio de mínima
probabilidad de error
Caso unidimensional
gaussiano

- Partiendo del criterio de decisión MAP particularizado para ruido gaussiano de media 0 y varianza σ^2

$$f(x|H_0) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_0}{\sigma}\right)^2}$$

$$f(x|H_1) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_1}{\sigma}\right)^2}$$

$$\frac{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_1}{\sigma}\right)^2}}{\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_0}{\sigma}\right)^2}} \underset{H_0}{\overset{H_1}{\geq}} \frac{P(H_0)}{P(H_1)}$$

- Desarrollando se llega a

$$\underbrace{x \underset{H_0}{\overset{H_1}{\geq}} \frac{h_0 + h_1}{2} + \frac{\sigma^2}{h_1 - h_0} \ln \left(\frac{P(H_0)}{P(H_1)} \right)}_{\text{Umbral } \mu}$$

Observación (estadístico suficiente)

- En el caso ML:

$$P(H_0) = P(H_1)$$

$$x \underset{H_0}{\overset{H_1}{\gtrless}} \frac{h_0 + h_1}{2}$$

- Al no tener en cuenta información *a priori*, el umbral se fija en el punto medio entre las dos hipótesis
- Una vez determinado el umbral (MAP o ML) nos interesa determinar la **probabilidad de error**

$$P_e = P(H_0)P(error|H_0) + P(H_1)P(error|H_1)$$

$$P_e = P(H_0) \int_{\mu}^{\infty} f(x|H_0)dx + P(H_1) \int_{-\infty}^{\mu} f(x|H_1)dx$$

Criterio de mínima
probabilidad de error
Caso unidimensional
gaussiano

$$P_e = P(H_0) \int_{\mu}^{\infty} f(x|H_0)dx + P(H_1) \int_{-\infty}^{\mu} f(x|H_1)dx$$

Suponiendo, por ejemplo, $P(H_0)=P(H_1)=0.5$

$$P_e = 2 \cdot \frac{1}{2} \int_{\mu}^{\infty} f(x|H_0)dx$$

En el caso gaussiano:

$$P_e = \int_{\mu}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-h_0}{\sigma}\right)^2} dx$$

Con un cambio de variable:

$$z = \frac{x - h_0}{\sigma}; \quad dz = \frac{dx}{\sigma}$$

$$P_e = \int_{\frac{h_0+h_1}{2}-h_0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = \int_{d/2\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz \triangleq Q\left(\frac{d}{2\sigma}\right)$$

$$d = h_1 - h_0$$

Definición



Criterio de mínima
probabilidad de error
Caso unidimensional
gaussiano

$$P_e = Q\left(\frac{d}{2\sigma}\right)$$

donde

d es la distancia entre símbolos (hipótesis)

σ es la desviación típica del ruido

- La función Q se resuelve mediante tablas o numéricamente
- Se trata de una función decreciente, de manera que
 - La P_e disminuye si aumenta d
 - La P_e disminuye si disminuye σ
- Ejemplo: Diseño en comunicaciones digitales
 - Se fija la probabilidad de error especificada en el problema
 - Se estudia el tipo de ruido (σ)
 - Se determina por medio de las tablas de la función Q la distancia entre símbolos necesaria

Teoría de la Decisión

Decisión, criterios MAP y ML, mínima P_e