

Fundamentos en teoría de detección

A. M. Alvarez-Meza, Ph.D.

amalvarezme@unal.edu.co

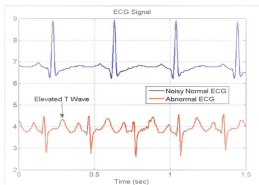
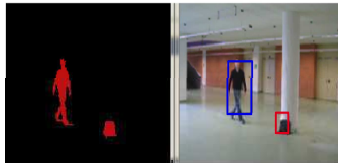
Departamento de ingeniería eléctrica, electrónica y computación
Universidad Nacional de Colombia-sede Manizales



- 1 Teoría de detección y análisis de datos
- 2 Conceptos básicos
- 3 FOB por Neyman-Pearson
- 4 FOB por riesgo Bayesiano

- 1 Teoría de detección y análisis de datos
- 2 Conceptos básicos
- 3 FOB por Neyman-Pearson
- 4 FOB por riesgo Bayesiano

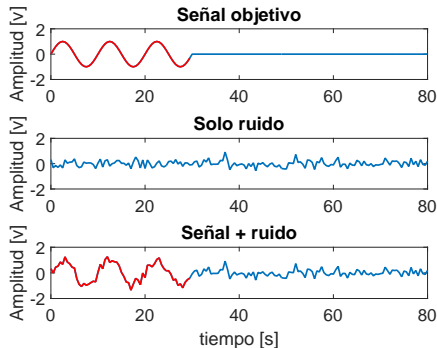
Teoría de detección y sus aplicaciones



- Comunicaciones
- Bioseñales
- Imágenes y video
- **Análisis de datos**

La teoría de detección es fundamental para identificar patrones discriminantes desde datos

Teoría de detección desde filtración óptima binaria (FOB)



- Detección de eventos de interés.
- Extraer información del evento.
- Desde estadística:
Test de hipótesis [Kay, 1998].
- Test de hipótesis binaria:
¿Solo ruido o Señal + ruido? = FOB.

¡La FOB puede entenderse como un problema de **detección**!

Campos afines

- Análisis de aleatoriedad [Liptser and Shiryaev, 2013].
- Teoría de estimación [Kay, 1993].
- Habilidades en programación.

Objetivos

- Comprender los principios básicos de FOB desde teoría de detección.
- Aplicar la FOB en problemas reales orientados al procesamiento de señales y al análisis de datos (clasificación → Aprendizaje de máquina → Analítica de datos).

- 1 Teoría de detección y análisis de datos
- 2 Conceptos básicos
- 3 FOB por Neyman-Pearson
- 4 FOB por riesgo Bayesiano

Detección de nivel DC: problema matemático

- Se dispone de una única señal de muestra $x[0] \in \mathbb{R}$.
- Se asume una amplitud de nivel DC $A \in \mathbb{R}$, distorsionada con ruido blanco Gaussiano (RBG) $w[0] \sim \mathcal{N}(0, \sigma^2)$,
 $\sigma^2 \in \mathbb{R}^+$: varianza del ruido.
- **Objetivo de la FOB:** decidir (*discriminar*) entre:
 - Hipótesis solo ruido $\mathcal{H}_0 : x[0] = w[0]$.
 - Hipótesis señal + ruido $\mathcal{H}_1 : x[0] = A + w[0]$.
- **Posible solución:** dado que el ruido tiene media cero:
 - Para $x[0] < A/2 \implies \mathcal{H}_0$.
 - Para $x[0] \geq A/2 \implies \mathcal{H}_1$.

Se espera detectar de forma correcta la mayoría de las veces

Detección de nivel DC: modelado de incertidumbre

- Repetir el experimento varias veces.
- Función de densidad de probabilidad (fdp) del ruido:

$$p(w[0]) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} w^2[0]\right)$$

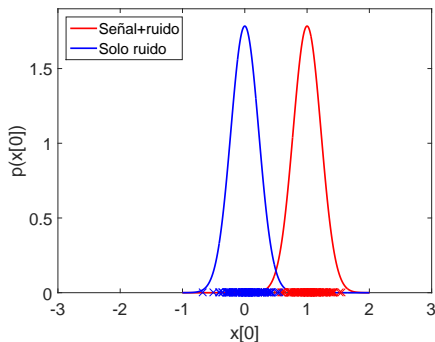
- El desempeño del detector depende de la diferencia entre fdps:

$$p(x[0]; \mathcal{H}_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} x^2[0]\right)$$

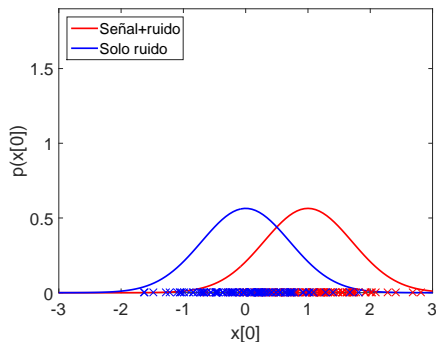
$$p(x[0]; \mathcal{H}_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x[0] - A)^2\right)$$

- La detección de nivel DC es un **test de parámetro** desde la familia de fdps: $p(x[0]; A) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (x[0] - A)^2\right)$.

Detección de nivel DC: modelado de incertidumbre



$$A = 1, \sigma^2 = 0.05$$



$$A = 1, \sigma^2 = 0.5$$

Mayor distancia entre fdps = Mayor relación señal ruido (RSR): A^2/σ^2
¡Mejor desempeño en FOB!

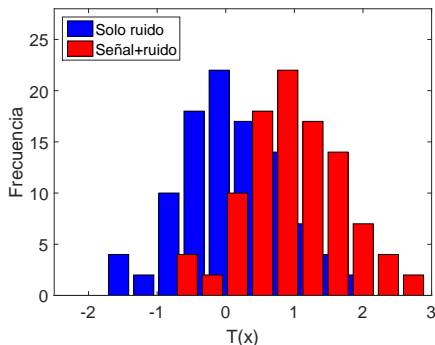
Consideraciones alrededor de la FOB

- En la práctica se trabaja con RSR bajas.
- Se debe **restringir** el valor de RSR requerido.
- El éxito de la FOB dependerá entonces de la **cantidad** de datos procesados.
- Detección nivel DC desde la señal $\{x[n] \in \mathbb{R}\}_{n=0}^{N-1}$:
 - Hipótesis solo ruido $\mathcal{H}_0 : x[n] = w[n]$
 - Hipótesis señal + ruido $\mathcal{H}_1 : x[n] = A + w[n]$donde $w[n] \sim \mathcal{N}(0, \sigma^2)$.

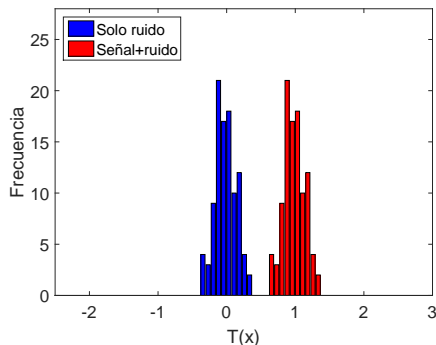
FOB desde estadístico $T(\mathbf{x})$

- Para aprovechar las N muestras $x[n]$ se detecta a partir de algún estadístico estimado desde los datos: $T : \mathbb{R}^N \rightarrow \mathbb{R}$.
 - Ej: Media muestral $T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$.
 - La FOB se resume entonces:
 - Para $T(\mathbf{x}) < \zeta \implies \mathcal{H}_0$.
 - Para $T(\mathbf{x}) \geq \zeta \implies \mathcal{H}_1$.
- $\zeta \in \mathbb{R}$: umbral.
- Desde teoría de estimación: a **mayor** número de **muestras**, mejor será el **desempeño** en FOB.

Estadístico $T(x)$ en FOB: detección nivel DC



$$A = 1, \sigma^2 = 0.5, N = 1, d^2 = 2$$



$$A = 1, \sigma^2 = 0.5, N = 20, d^2 = 40$$

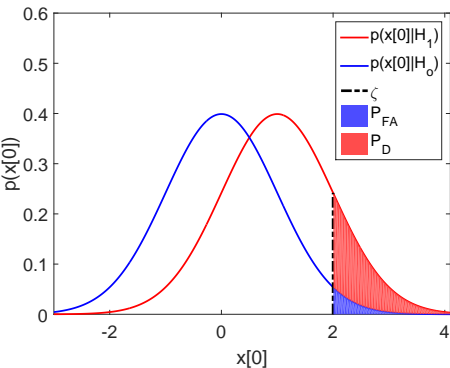
Experimento con 100 repeticiones

- **Coeficiente de desvío.** $d^2 = \frac{(\mathbb{E}\{T(x); \mathcal{H}_1\} - \mathbb{E}\{T(x); \mathcal{H}_0\})^2}{\text{var}\{T(x); \mathcal{H}_0\}}$
- En nivel DC $d^2 = NA^2/\sigma^2$; Si $N \rightarrow \infty$, $d \rightarrow \infty$

- 1 Teoría de detección y análisis de datos
- 2 Conceptos básicos
- 3 FOB por Neyman-Pearson**
- 4 FOB por riesgo Bayesiano

- La FOB se resuelve desde decisión estadística como un **test de hipótesis** simple.
- Se asume que las fdps $p(x; \mathcal{H}_0)$ y $p(x; \mathcal{H}_1)$ son conocidas y se impone un **nivel de significancia**.
- **Neyman-Pearson** (NP) se puede entender como la solución clásica (frecuentista) en detección desde teoría de estimación.
- NP: **maximización** de la probabilidad de **detección** P_D sujeta a un valor **deseado** de probabilidad de **falsa alarma** P_{FA} .

Neyman-Pearson: falsa alarma y detección



Regla de detección

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_o)} > \zeta$$

$\zeta \in \mathbb{R}$ se estima a partir de un nivel de significancia $\alpha \in \mathbb{R}^+$:

$$P_{FA} = P(\mathcal{H}_1; \mathcal{H}_o) = P(x[0] > \zeta; \mathcal{H}_o)$$

$$P_{FA} = \int_{\{\mathbf{x}: L(\mathbf{x}) > \zeta\}} p(\mathbf{x}; \mathcal{H}_o) d\mathbf{x} = \alpha,$$

Se utilizan funciones **cumulativas complementarias**. Además:

$$P_D = P(\mathcal{H}_1; \mathcal{H}_1) = P(x[0] > \zeta; \mathcal{H}_1)$$

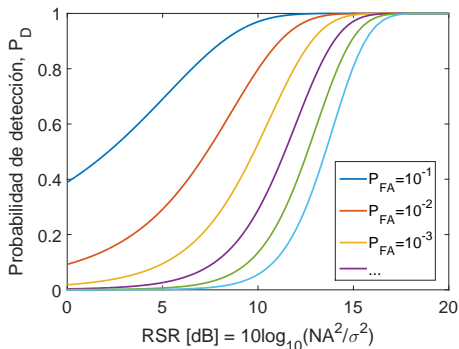
$$P_D = \int_{\{\mathbf{x}: L(\mathbf{x}) > \zeta\}} p(\mathbf{x}; \mathcal{H}_1) d\mathbf{x}$$

Neyman-Pearson: detección de nivel DC

- Para una señal $\mathbf{x} = \{x[n]\}_{n=0}^{N-1}$, las hipótesis se fijan como:
 - Hipótesis solo ruido $\mathcal{H}_0 : \mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$
 - Hipótesis señal + ruido $\mathcal{H}_1 : \mathbf{x} \sim \mathcal{N}(A\mathbf{1}, \sigma^2 \mathbf{I})$
- $L(\mathbf{x}) = \frac{-1}{2\sigma^2} \left(-2A \sum_{n=0}^{N-1} x[n] + NA^2 \right) > \log(\zeta)$
- $L(\mathbf{x}) = T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{\sigma^2}{NA} \log(\zeta) + \frac{A}{2} = \zeta'$
- La detección de nivel DC se basa en la estimación de A desde:

$$T(\mathbf{x}) \sim \begin{cases} \mathcal{N}(0, \sigma^2/N) & \mathcal{H}_0 \\ \mathcal{N}(A, \sigma^2/N) & \mathcal{H}_1 \end{cases}$$

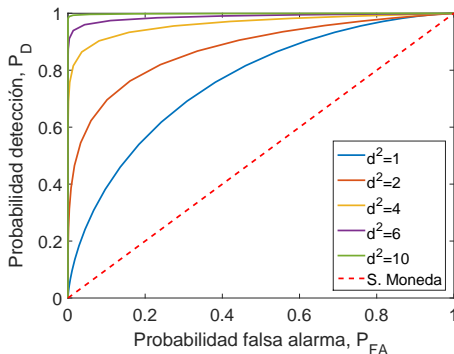
Neyman-Pearson: detección de nivel DC



$$P_{FA} \downarrow \text{RSR} \uparrow \implies P_D \uparrow$$
$$d^2 = \frac{N(\mu_o - \mu_1)^2}{\sigma^2}$$

- $P_{FA} = Q(\zeta' / \sqrt{\sigma^2/N})$
- $P_D = Q\left((\zeta' - A) / \sqrt{\sigma^2/N}\right)$
- $\zeta' = \sqrt{\frac{\sigma^2}{N}} Q^{-1}(P_{FA})$
- Q : función de distribución cumulativa complementaria para $\mathcal{N}(0, 1)$.
- P_{FA} y P_D pueden aproximarse mediante simulación de **Monte Carlo**.

Neyman-Pearson: Curva ROC detección nivel DC

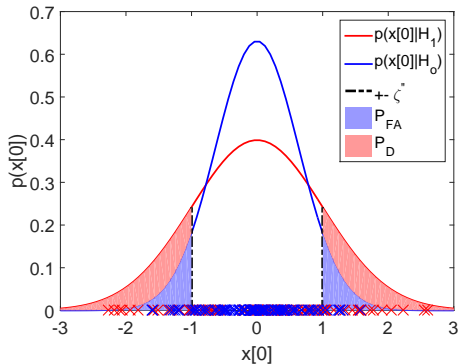


$$P_{FA} \downarrow RSR \uparrow \implies P_D \uparrow$$
$$d^2 = \frac{N(\mu_o - \mu_1)^2}{\sigma^2}$$

- $P_{FA} = 1 - \text{especificidad} = \text{FPR}$
- $P_D = \text{sensibilidad} = \text{TPR}$
- FPR y TPR comunmente usados en **reconocimiento de patrones!**
- ROC: Receiver Operating Characteristic (Característica Operativa del Receptor) - gráfica de sensibilidad vs especificidad en clasificador binario al variar umbral.

Neyman-Pearson:

¿Qué información extraer desde $T(\mathbf{x})$?



$$\begin{aligned}\mathcal{H}_0 : x[n] &\sim \mathcal{N}(0, \sigma_o^2) \\ \mathcal{H}_1 : x[n] &\sim \mathcal{N}(0, \sigma_1^2) \\ \sigma_o^2 &> \sigma_1^2\end{aligned}$$

$$\bullet L(\mathbf{x}) = T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]^2 > \zeta''$$

$$\bullet \zeta'' = \frac{\frac{2}{N} \log(\zeta) + \log\left(\frac{\sigma_o^2}{\sigma_1^2}\right)}{\frac{1}{\sigma_o^2} - \frac{1}{\sigma_1^2}}$$

- NP: estimación de $\theta \in \mathbb{R}$ desde una **fdp parametrizada** $p(\mathbf{x}; \theta)$.
- $T(\mathbf{x})$ codifica la **información relevante** en \mathbf{x} para FOB (clasificación binaria).

- 1 Teoría de detección y análisis de datos
- 2 Conceptos básicos
- 3 FOB por Neyman-Pearson
- 4 FOB por riesgo Bayesiano

FOB desde Bayes: Mínima probabilidad de error

- **Información a priori** = asignar probabilidades a \mathcal{H}_c .
- **Probabilidad de error** desde Bayes (FOB $c \in \{0, 1\}$):

$$P_e = P(\mathcal{H}_o|\mathcal{H}_1)P(\mathcal{H}_1) + P(\mathcal{H}_1|\mathcal{H}_o)P(\mathcal{H}_o)$$

- $P(\mathcal{H}_c|\mathcal{H}_{c'})$: probabilidad de decidir \mathcal{H}_c dado que $\mathcal{H}_{c'}$ es cierto (condicional).
- $P(\mathcal{H}_c)$: prior = **suposiciones** sobre ocurrencia de \mathcal{H}_c .
- **Regla de detección**

$$L(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_o)} > \frac{P(\mathcal{H}_o)}{P(\mathcal{H}_1)} = \zeta$$

- Si $P(\mathcal{H}_o) = P(\mathcal{H}_1)$, se escoge la hipótesis con **máxima verosimilitud** (MV) condicional $p(\mathbf{x}|\mathcal{H}_c)$.
- En MV $L(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_o)} > 1$
- Para detección nivel DC:
 - En NP: $L(\mathbf{x}) = T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{\sigma^2}{NA} \log(\zeta) + \frac{A}{2} = \zeta'$;
$$P_{FA} = Q\left(\sqrt{\frac{N}{\sigma^2}}(\zeta' - A)\right)$$
 - En MV se escoge \mathcal{H}_1 si: $L(\mathbf{x}) = T(\mathbf{x}) = \frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{A}{2}$;
$$P_e = Q\left(\sqrt{\frac{NA^2}{4\sigma^2}}\right)$$

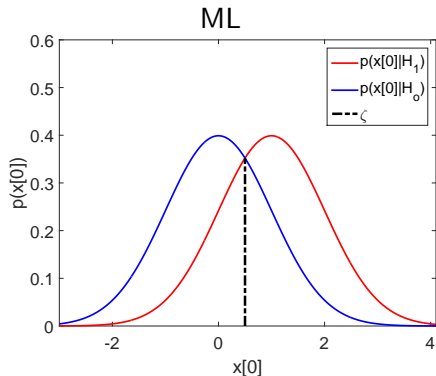
- Desde teorema de Bayes:

$$p(\mathbf{x}, \mathcal{H}_c) = p(\mathcal{H}_c|\mathbf{x})p(\mathbf{x}) = p(\mathbf{x}|\mathcal{H}_c)p(\mathcal{H}_c)$$

- Se escoge la hipótesis con **máximo posterior** (MAP) $p(\mathcal{H}_c|\mathbf{x})$.
- En MAP se escoge \mathcal{H}_1 si: $L(\mathbf{x}) = p(\mathcal{H}_1|\mathbf{x}) > p(\mathcal{H}_o|\mathbf{x})$
- MAP con respecto a las verosimilitudes:

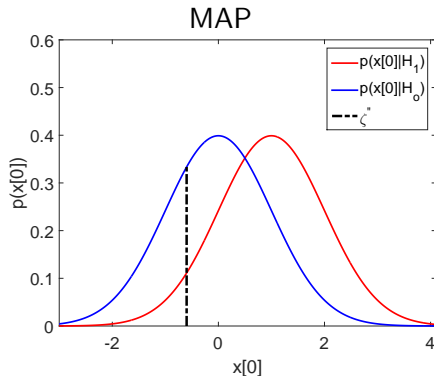
$$L(\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{H}_1)}{p(\mathbf{x}|\mathcal{H}_o)} > \frac{P(\mathcal{H}_o)}{P(\mathcal{H}_1)} = \zeta$$

Detección de nivel DC: ML vs MAP



$$A = 1, \sigma^2 = 1, N = 1 \quad P(\mathcal{H}_0) = P(\mathcal{H}_1)$$

$$T(\mathbf{x}) > \zeta = A/2$$



$$A = 1, \sigma^2 = 1, N = 1$$

$$P(\mathcal{H}_0) = 1/2 \neq P(\mathcal{H}_1) = 3/4$$

$$T(\mathbf{x}) > \zeta'' = \frac{\sigma^2}{NA} \log \left(\frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} \right) + A/2$$

Riesgo Bayesiano generalizado: múltiples hipótesis

- Múltiples hipótesis = **clasificación multi-clase**.
- Incluye **penalización** (costo) $\xi_{cc'} \in \mathbb{R}^+$, con $c, c' \in \{1, 2, \dots, C\}$ y $\xi_{cc} = 0$.
- **Riesgo de Bayes generalizado:**

$$\mathcal{R} = \sum_{c=1}^C \sum_{c'=1}^C \xi_{cc'} P(\mathcal{H}_c | \mathcal{H}_{c'}) P(\mathcal{H}_{c'})$$

- De FOB a múltiples hipótesis (clases) = clasificación = **reconocimiento de patrones**

- La **FOB** es la base de los métodos de **teoría de detección**.
- Neyman-Person y Bayes (ML y MAP) facilitan la **inclusión** de **restricciones** (FA o priors) de la aplicación.
- **Reconocimiento de patrones** = test múltiples hipótesis.



Kay, S. M. (1993).

Fundamentals of statistical signal processing, volume i: estimation theory.



Kay, S. M. (1998).

Fundamentals of statistical signal processing, vol. ii: Detection theory.

Signal Processing. Upper Saddle River, NJ: Prentice Hall.



Liptser, R. S. and Shiryaev, A. N. (2013).

Statistics of random Processes: I. general Theory, volume 5.

Springer Science & Business Media.