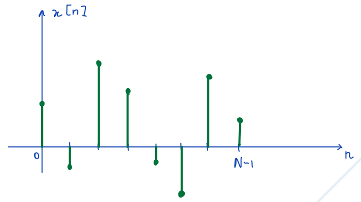
DETECTION THEORY

- Steven M. Kay, "FUNDAMENTALS OF STATISTICALLY SIGNAL PROCESSING: detection theory", Prentice-hall, 1998

GOAL OF DETECTION THEORY : To detect an event occurring in a signal

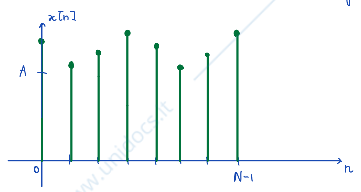
EX: Detect a change of mean in white Gaussian noise.
The idea is that we have a normal situation (normal mean that it's a nominal situation, that one that usually happe (may is given by the design specification).



→ this is the situation when the mean value is $\emptyset$.
We call this normal situation, the scenario H_0 which is called also: NULL HYPOTHESIS

$$\mu = \mathbb{E}[x[n] \mid H_0] = \emptyset$$

now suppose that there is an anomaly:



→ for this anomaly the mean value is changed $= A$
→ this anomaly we call H_1: ALTERNATIVE HYPOTHESIS

$$\mu : \mathbb{E}[x[n] \mid H_1] = A$$

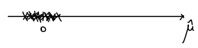
this is usually an anomalous situation which we have to detect as fast as soon possible
(APPLICATION EX ATOMIC CLOCKS)

⇒ How we should design a detector?

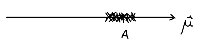
We use our intuition and we first compute the sample mean $\hat{\mu} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$

If I want to detect the change of the mean, then we have to understand that the mean $\emptyset$ and A don't exist at the same time!, they are separate ⇒ so we estimate the mean value and then we want to understand if the mean is $\emptyset$ or A, since we have just one num which is the sample mean over N samples ⇒ we have to set up BOH

If we are in the H_0 ipothesis ⇒ for $\hat{\mu}$ we have

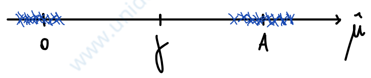


If we are in the H_1 hypothesis ⇒ for $\hat{\mu}$ we have



→ so we have to choose a THRESHOLD γ

and then we want to decide H_1 if $\hat{\mu}$ is large of γ



we decide H_1 if is $\hat{\mu} > \gamma$

NOTE:

Binary hypothesis testing : we decide between the two hypotheses H_0 and H_1.

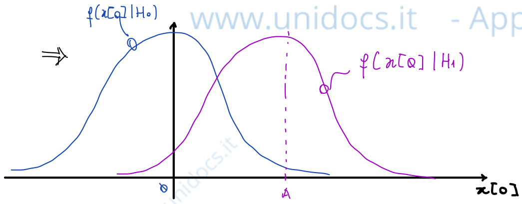
- How do we design a detector?
What are the fundamental quantities of detection theory?

To understand the fundamental quantities of detection theory, we consider the case of a Gaussian random variable $x[\emptyset]$, $N=1$.

$$\Rightarrow \quad x[\emptyset] \sim \begin{cases} \mathcal{N}(\emptyset, \sigma^2) : H_0 \\ \mathcal{N}(A, \sigma^2) : H_1 \end{cases}$$

→ so the variance doesn't change and only the mean value changes.

Let's draw the two probability density functions for this situations:

⇒ 

(we are gonna use both x[0] for the random variable and for the value taken just for simplicity)

So we have: $f(x[0]|H_0) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2[0]}{2\sigma^2}}$

$$f(x[0]|H_1) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x[0]-A)^2}{2\sigma^2}}$$

simplest example: CAR ALARM ⟶ H_0: nobody is stealing the car

⟶ H_1: somebody is stealing the car

The detection rule is: we look at x[0] (which is the output of our sensor)
we decide H_1 when $x[0] > \gamma$

so this is our detection rule

Now let's consider the types of events that can occur:

⇒ The first event is DETECTION:

DETECTION:

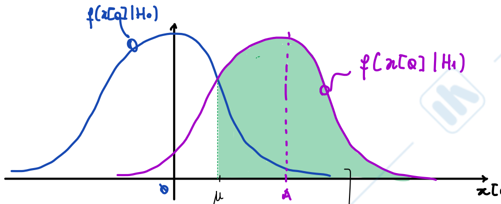
We have a detection when we decide H_1 and H_1 is true.

The probability of detection P_D is given by:

$$P_D = P(H_1 | H_1) = P(x[0] > \gamma | H_1) =$$

$$= \int_{\gamma}^{+\infty} f(x[0]|H_1)\, d x[0] =$$

$$= \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x[0]-A)^2}{2\sigma^2}}\, d x[0]$$



we are in the H_1 hypothesis and this is the probability the x[0] is larger then γ
↓
So this area gives the probability of DETECTION

What happen if the thing goes wrong and we have a false alarm?

FALSE ALARM [in statistic is called TYPE I ERROR]

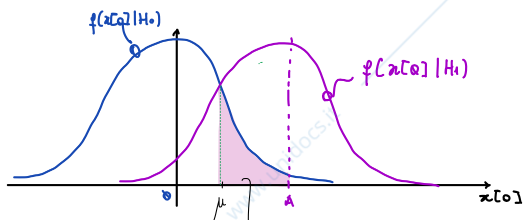
We have a false alarm when we decide H_1 and H_0 is true.

The probability of false alarm P_{FA} is given by:

$$P_{FA} = P(H_1 | H_0) = P[x[0] > \gamma | H_0] = \int_{\mu}^{+\infty} f(x[0]|H_0)\, d x[0] =$$

from a geometrical point of view is:

$$= \int_{\mu}^{+\infty} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2[0]}{2\sigma^2}}\, d x[0]$$



P_{FA} = this is the probability of false alarm since: we are in H_0 hypothesis (nobody is in the car) but the sensor given a value x' which is larger then the treshold because of the fluctuations.

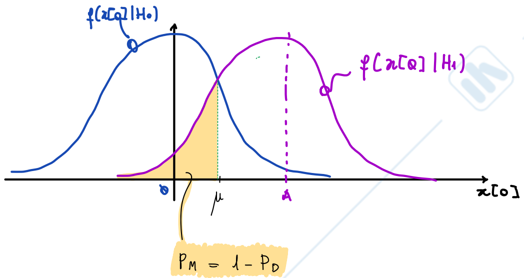
But there is another situation:

MISS [TYPE II ERROR]

We have a miss when we decide H_0 and H_1 is true.

(nobody steal the car (F) while are stealing the car)

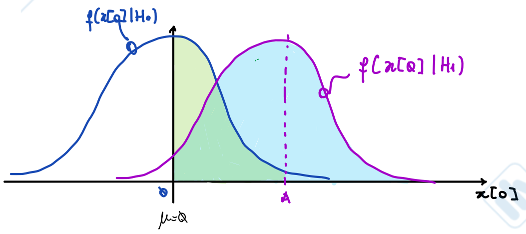
$$P_M = P(H_0 \mid H_1) = P(x[0] \le \gamma \mid H_1) = 1 - \underbrace{P(x[0] > \gamma \mid H_1)}_{P_D}$$



$$\Rightarrow \boxed{P_M = 1 - P_D}$$

Fundamental quantities of detection theory: P_D and P_{FA}

- Tradeoff between P_D and P_{FA}



(the rule is always: we decide H_1 when $x[0] > \gamma$)

Suppose we want increase P_D, so we reduce γ.

→ Suppose $\gamma = 0$, so:

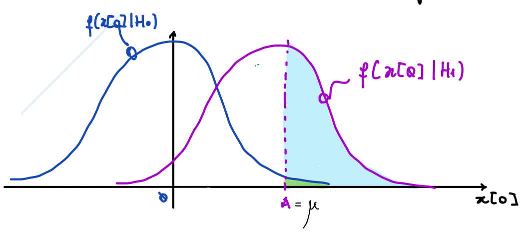
To increase P_D we reduce γ → the probability of detection is (almost 1)

the prob. of false alarm is which is a huge quantity.

⇓

P_{FA} increase! (the minimum variation → turn on the alarm)

Now let's take a look for the other hypothesis:



→ Let's increase the treshold: $\mu = A$

- the prob. of false alarm is (very small)
- the probability of detection is

To reduce P_{FA}, we increase γ ⇒ P_D decreases unfortunately!

For make things work, ⇒ we need more measurement → MORE ANSWER

But first we consider:

HOW DO WE DESIGN A DETECTOR?

➡ A common approach is to fix P_{FA}, and then design the detector by maximing P_D.

The design strategy is the NEYMAN - PEARSON APPROACH.

THE NEYMAN - PEARSON APPROACH:

The optimal detector which maximizes P_D for a given P_{FA} decides H_1 when:

$$L(\underline{x}) = \frac{f(\underline{x} \mid H_1)}{f(\underline{x} \mid H_0)} > \gamma_L$$

where γ_L must satisfy: $P_{FA} = \int f(\underline{x} \mid H_0)\, d\underline{x}$ w\ $\{\underline{x} : L(\underline{x}) > \gamma_L\}$ and $\underline{x} = \begin{bmatrix} x[0] \\ \vdots \\ x[N-1] \end{bmatrix}$

TERMINOLOGY

$L(\underline{x})$: LIKELIHOOD RATIO

$L(\underline{x}) > \gamma$: LIKELIHOOD RATIO TEST [LRT]

⇒ so we decide H_1 when the likelihood ratio is $> \gamma_L$. γ_L is fixed for the probab. of false alarm. → So we decide the probability of fal. alarm and we get γ so that we get this P_{FA}. Then we decide H_1, namely we decide when this anomaly is true.

THE DETECTOR IS THE INEQUALITY

. Now we want to apply that in some example (1 - sample ÷ N - sample)

EX

We consider one sample of white Gaussian noise:

$$x[0] \sim \begin{cases} N(0,1) & : H_0 \\ N(1,1) & : H_1 \end{cases} \quad \rightarrow [N=1,\ A=1,\ \sigma^2=1]$$

It is $\underline{x} = x[0]$, and we decide H_1 when: $L(\underline{x}) = L(x[0]) = \dfrac{f(x[0] \mid H_1)}{f(x[0] \mid H_0)} > \gamma_L$

where $f(x[0] \mid H_0) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2[0]}{2}}$

$$f(x[0] \mid H_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x[0] - \overset{1}{\overset{\shortparallel}{A}})^2}{2}}$$

now we substitute this $f(x[0]\mid i)$ into the likelihood:

$$L = \frac{\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(x[0]-1)^2}{2}}}{\frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2[0]}{2}}} = e^{-\frac{(x[0]-1)^2 - x^2[0]}{2}} =$$

$$= e^{-\frac{\cancel{x^2[0]} + 1 - 2x[0] - \cancel{x^2[0]}}{2}} =$$

$$= \boxed{e^{x[0]-\frac{1}{2}} = L} \rightarrow \text{LIKELIHOOD RATIO}$$

We decide H_1 when $L > \gamma_L \Rightarrow e^{x[0]-\frac{1}{2}} > \gamma_L$

$$\Rightarrow x[0] - \frac{1}{2} > \log(\gamma_L)$$

$$\Rightarrow x[0] > \underbrace{\log(\gamma_L) + \frac{1}{2}}_{\substack{\text{THIS IS ANOTHER TREASURE} \\ = \gamma}}$$

We decide H_1 when $x[0] > \gamma$

when γ must satisfy
$$P_{FA} = \int_{\gamma}^{+\infty} f(x[0]\mid H_0)\, dx[0] = \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2[0]}{2}}\, dx[0]$$

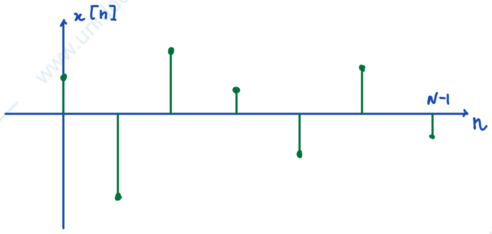
so we fix this quantity P_{FA} ($= 10^{-3}$) and then we get the γ for which this integral $\int_{\gamma}^{+\infty} f(x[0]\mid H_0)\,dx[0]$ is true (must return

Consequently,
$$P_D = \int_{\gamma}^{+\infty} f(x[0]\mid H_1)\, dx[0] = \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{(x[0]-1)^2}{2}}\, dx[0]$$

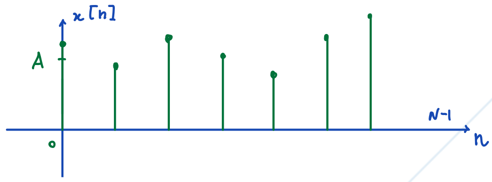
for example, if $\underbrace{P_{FA} = 10^{-3}}_{\text{we fixed}}$, then $\overset{\text{we compute}}{\gamma = 3}$ and $P_D \simeq 0.023 = 2.3\%$

EXAMPLE

Detect a change of mean in N samples of White Gaussian noise:



H_0 hypothesis



H_1 hypothesis

$\mathbb{E}[x[n] \mid H_0] = 0$

$x[n] = \xi[n]$

$\xi_n \sim \mathcal{N}(0, \sigma^2)$

ξ_n and ξ_m are statystically indipendent

$\mathbb{E}[x[n] \mid H_1] = A$

$x[n] = \xi[n] + A$

so it's the same WGN w\ different mean value

So we apply the Neyman - Pearson theorem, therefore we need to compute the likelihood ratio L.

The likelihood ratio is given by:

$$L(\underline{x}) = \frac{f(\underline{x} \mid H_1)}{f(\underline{x} \mid H_0)} \qquad \text{where} \quad \underline{x} = \begin{bmatrix} x[0] \\ \vdots \\ x[N-1] \end{bmatrix}$$

In the H_0 hypothesis, it is

the probab. density of WGN which we are going to consider is: we have N samples ⇒ we have N Gaussian rand. var. and each of the is $\xi_n \sim \mathcal{N}(0, \sigma^2)$ → In this case it's easy find the joint probability: when we have N r.v. and are statistically indipendent the joint prob. density is the product of 1-dimensional prob. density function.

$$f(x[0] \mid H_0) = f(x[0] \mid H_0) \cdots f(x[N-1] \mid H_0) =$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2[0]}{2\sigma^2}} \cdots \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2[N-1]}{2\sigma^2}} =$$

$$= \frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{x^2[0]}{2\sigma^2} \cdots - \frac{x^2[N-1]}{2\sigma^2}} =$$

$$= \boxed{\frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} x^2[n]} = f(\underline{x} \mid H_0)}$$

Similarly,

$$\boxed{f(\underline{x} \mid H_1) = \frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2}}$$

Substituting, the likelihood ratio becomes:

$$L(\underline{x}) = \frac{f(\underline{x} \mid H_1)}{f(\underline{x} \mid H_0)} = \frac{\frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2}}{\frac{1}{(2\pi)^{N/2}\sigma^N} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1} x^2[n]}} =$$

$$\Rightarrow \quad = e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2 + \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}x^2[n]} =$$

$$= e^{-\frac{1}{2\sigma^2}\left[\sum_{n=0}^{N-1}(x[n]-A)^2 - \sum_{n=0}^{N-1}x^2[n]\right]} =$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left[(x[n]-A)^2 - x^2[n]\right]} =$$

$$= e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}\left[\cancel{x^2[n]} + A^2 - 2A\,x[n] - \cancel{x^2[n]}\right]} =$$

$$= e^{-\frac{1}{2\sigma^2}\left[\underbrace{\sum_{n=0}^{N-1}A^2}_{NA^2} - 2A\sum_{n=0}^{N-1}x[n]\right]} =$$

$$= e^{-\frac{1}{2\sigma^2}\left[NA^2 - 2A\sum_{n=0}^{N-1}x[n]\right]}$$

$$\boxed{L(\underline{x}) = e^{\frac{A}{\sigma^2}\sum_{n=0}^{N-1}x[n] - \frac{NA^2}{2\sigma^2}}} \qquad \text{LIKELIHOOD RATIO}$$

According to the Neyman - Pearson theorem, we decide H_1 when $L(\underline{x}) > \gamma_L$.

Substituting,

$$L(\underline{x}) = e^{\frac{A}{\sigma^2}\sum_{n=0}^{N-1}x[n] - \frac{NA^2}{2\sigma^2}} > \gamma_L$$

$$\Rightarrow \frac{A}{\sigma^2}\sum_{n=0}^{N-1}x[n] - \frac{NA^2}{2\sigma^2} > \log\gamma_L$$

$$\frac{A}{\sigma^2}\sum_{n=0}^{N-1}x[n] > \log(\gamma_L) + \frac{NA^2}{2\sigma^2}$$

$$\sum_{n=0}^{N-1}x[n] > \frac{\sigma^2}{A}\log(\gamma_L) + \frac{NA}{2}$$

$$\underbrace{\frac{1}{N}\sum_{n=0}^{N-1}x[n]}_{\hat{\mu}} > \underbrace{\frac{\sigma^2}{AN}\log(\gamma_L) + \frac{A}{2}}_{\gamma}$$

We decide H_1 when $\hat{\mu} > \gamma$, where $\hat{\mu} = \frac{1}{N}\sum_{n=0}^{N-1}x[n]$

But right now we know the statistic of $\hat{\mu}$, of the sample mean estimator for WGN.

So we characterize the statics of $\hat{\mu}$.

$\hat{\mu}$ is Gaussian (linear operation on the Gaussian random variables $x[n]$)

$|$

↓ De Moivre

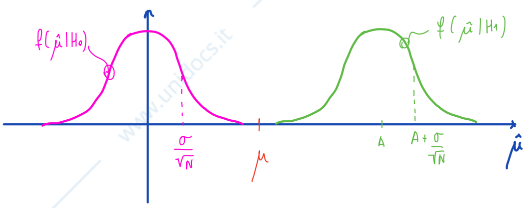
in H_0 hypothesis, $\mathbb{E}[\hat{\mu} | H_0] = 0$ and the variance for sample mean is: $\text{var}\{\hat{\mu}\} = \frac{\sigma^2}{N}$

$$\Rightarrow \left.\begin{aligned} H_0 : \ & \mathbb{E}[\hat{\mu} | H_0] = 0 \\ & \text{var}\{\hat{\mu}\} = \frac{\sigma^2}{N} \end{aligned}\right\} \Rightarrow \hat{\mu} \sim \mathcal{N}\left(0, \frac{\sigma^2}{N}\right)$$

in H_1 hypothesis:

$$\left.\begin{aligned} H_1 : \ & \mathbb{E}[\hat{\mu} | H_1] = A \\ & \text{var}\{\hat{\mu}\} = \frac{\sigma^2}{N} \end{aligned}\right\} \Rightarrow \hat{\mu} \sim \mathcal{N}\left(A, \frac{\sigma^2}{N}\right)$$

Let's draw right now:



if we consider μ here, things going to work well

for a given P_{FA}, the threshold γ must satisfy:

$$P_{FA} = \int_\gamma^{+\infty} f(\hat{\mu} | H_0)\, dz = \int_\gamma^{+\infty} \frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma} e^{-\frac{z^2 N}{2\sigma^2}}\, dz$$

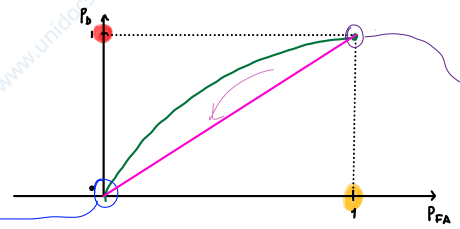
consequently, the probability of detection P_D is given by:

$$P_D = \int_\gamma^{+\infty} f(\hat{\mu} | H_1)\, dz = \int_\gamma^{+\infty} \frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma} e^{-\frac{(z-A)^2 N}{2\sigma^2}}\, dz$$

now that we have a detector, how we can visualize the probability of False alarm, detection respect to γ?
(If I change γ, the performance is changing, and also if we change the size of mean (= jump))
↓
We can visualize it by using the **RECEIVER OPERATING CHARACTERIC [ROC]**

• RECEIVER OPERATING CHARACTERIC [ROC]

The ROC is used to characterize the performances of a detector.



we know that every point is obtained by changing γ.

→ the γ value of this point is (considering $P_D = P_{FA} = 1$) ⇒ the threshould value is $\gamma = -\infty$ (if was $+\infty$ and we have a (frequency) jump, we never detect it, becouse it will never be longer then the threshould, so the P_D shoulcd be

the $\gamma = -\infty$ becouse if the threshould is $-\infty$, when we get a value (we estimate the value of sample mean $\hat{\mu}$) and it will be always longer then the threshould, so the $P_D = 1$.

But when there is no change of mean ($= 0$), any flactuation will be also large then the threshould → always detect false alarm and the $P_{FA} = 1$

$\gamma = +\infty$: becouse: if the threshould is $+\infty$, suppose that the mean value is 0, we are in H_0 hyp., it doesn't metter how large the flactuation are, there never will be larger then $+\infty$, → therefore we will never get false alarm → $P_{FA} = 0$, but if there is an anomaly / jump ⇒ it doesn't metter the size of jump, this will be always smaller then $\gamma = +\infty$ ⇒ so we will never detect it ⇒ $P_D = 0$

• so we have all this value by increasing γ.

● in this case $P_{FA} = 0$ and always detect $P_D = 1$ → IDEAL DETECTOR

● this is almost an ideal detector, we have just to reverse the decision → Is still an ideal detector, if we reverse the decision rule!

↓ EX

non azzecco mai se CLOUDLY or SUNLY → REVERSE DETECTOR

The worst detector is $50 \div 50$: (random detector)

this straight line is the random detector: which is the worst detector

The worst detector is the random detector: we flip a coin and we decide H_1 when we get "HEAD".

Suppose that $p(\text{head}) = p$, when $0 \leq p \leq 1$.

$$\Rightarrow \text{the } P_{FA} = P(H_1 \mid H_0) = P(\text{head} \mid H_0) = p(\text{head}) = p$$

is the prob. of detect an anomaly when it doesn't exists

$$P_D = P(H_1 \mid H_0) = P(\text{head} \mid H_1) = P(\text{head}) = p$$

$$\Rightarrow \boxed{P_{FA} = P_D = p}$$

If we increase the size of jump A ⇒ in which direction we go? In the ideal one or random?
In the ideal one since is easier to detect a large jump instead of thiny jump.
If we have more sample it will be more easier to detect the anomaly.