Decizie și Estimare în Prelucrarea Informației

Capitolul II. Elemente de teorie statistică a detecției

II.1 Introducere

Introducere

- Detecția semnalelor = a decide care semnal este prezent dintre două sau mai multe posibilități
  - inclusiv că nu există nici un semnal (este 0)
- Avem la dispoziție observații **cu zgomot**
  - semnalele sunt afectate de zgomot
  - zgomotul este aditiv (se adună la semnalul original)

Schema bloc a detecției semnalelor

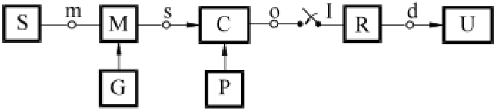


Figure 1: Signal detection model

- Conținut:
  - Sursa de informație: generează mesajele a_n cu probabilitățile $p(a_n)$
  - Generator: generează semnalele diferite $s_1(t), \ldots s_n(t)$
  - Modulator: transmite semnalul $s_n(t)$ la mesajul a_n
  - Canal: adaugă zgomot aleator
  - Eșantionare: ia eșantioane din semnalul $s_n(t)$
  - Receptor: **decide** ce mesaj a_n s-a fost recepționat
  - Utilizator: primește mesajele recuperate

Scenarii practice

- Transmisie de date
  - nivele constante de tensiune (de ex. $s_n(t) =$ constant 0 sau 5V)
  - modulație PSK (Phase Shift Keying): $s_n(t) =$ cosinus cu aceeași frecvență dar faze inițiale diferite
  - modulație FSK (Frequency Shift Keying): $s_n(t) =$ cosinus cu frecvențe diferite
  - modulație OFDM (Orthogonal Frequency Division Multiplexing): caz particular de FSK
- Radar
  - se emite un semnal; în cazul unui obstacol, semnalul se reflectă înapoi
  - receptorul așteaptă posibilele reflecții ale semnalului emis și decide
    - nu este prezentă o reflecție -> nici un obiect
    - semnalul reflectat este prezent -> obiect detectat

Generalizări

- Decizie între mai mult de două semnale
- Numărul de eșantioane (observații):
  - un singur eșantion
  - mai multe eșantioane
  - observarea întregului semnal continuu, pentru un timp T

II.2 Detecția semnalelor folosind 1 eșantion

Detecția unui semnal cu 1 eșantion

- Cel mai simplu caz: detecția unui semnal afectat de zgomot, folosind un singur eșantion
  - două mesaje a_0 și a_1
  - mesajele sunt modulate cu semnalele $s_0(t)$ și $s_1(t)$
    - pentru a_0: se transmite $s(t) = s_0(t)$
    - pentru a_1: se transmite $s(t) = s_1(t)$
  - peste semnal se suprapune zgomot aditiv, alb, $n(t)$
  - se recepționează un semnal cu zgomot, $r(t) = s(t) + n(t)$
  - eșantionarea preia un singur eșantion la timpul t_0, $r(t_0)$
  - decizie: pe baza $r(t_0)$, care semnal a fost cel transmis?

Ipoteze și decizii

- Există două **ipoteze**:
  - H_0: semnalul adevărat este $s(t) = s_0(t)$ (s-a transmis a_0)
  - H_1: semnalul adevărat este $s(t) = s_1(t)$ (s-a transmis a_1)
- Receptorul poate lua una din două **decizii**:
  - D_0: receptorul decide că semnalul corect este $s(t) = s_0(t)$
  - D_1: receptorul decide că semnalul corect este $s(t) = s_1(t)$

Rezultate posibile

- Există 4 situații posibile:
  1. **Rejecție corectă**: ipoteza corectă este H_0, decizia este D_0
     - Probabilitatea este $P_r = P(D_0 \cap H_0)$
  2. **Alarmă falsă** (detecție falsă): ipoteza corectă este H_0, decizia este D_1
     - Probabilitatea este $P_{af} = P(D_1 \cap H_0)$
  3. **Pierdere** (rejecție falsă): ipoteza corectă este H_1, decizia este D_0
     - Probabilitatea este $P_p = P(D_0 \cap H_1)$
  4. **Detecție corectă**: ipoteza corectă este H_1, decizia este D_1
     - Probabilitatea este $P_d = P(D_1 \cap H_1)$

Originea termenilor

- Terminologia are la origine aplicații radar (prima aplicație a teoriei detecției)
    - un semnal se emite de către sursă
    - semnal recepționat = o posibilă reflecție din partea unei ținte, puternic afectată de zgomot
    - H_0 = nu există un obiect, nu există semnal reflectat (doar zgomot)
    - H_1 = există un obiect, există un semnal reflectat
    - de aceea numele celor 4 scenarii sugerează “detecția unui obiect”

Zgomotul

- În general se consideră zgomot **aditiv**, **alb**, **staționar**
  - aditiv = zgomotul se adună ci semnalul
  - alb = două eșantioane distincte sunt necorelate
  - staționar = are aceleași proprietăți statistice la orice moment de timp
- Semnalul de zgomot $n(t)$ este necunoscut
  - este o realizare a unui proces aleator
  - se cunoaște doar distribuția sa, nu și valorile particulare

Eșantionul preluat la recepție

- La recepție se primește semnalul $r(t) = s(t) + n(t)$
  - $s(t)$ = semnalul original, fie $s_0(t)$, fie $s_1(t)$
  - $n(t)$ = semnalul de zgomot necunoscut
- Valoarea eșantionului luat la momentul t_0 este $r(t_0) = s(t_0) + n(t_0)$
  - $s(t_0)$ = fie $s_0(t_0)$, fie $s_1(t_0)$
  - $n(t_0)$ este un eșantion din semnalul de zgomot

Eșantionul preluat la recepție

- Eșantionul $n(t_0)$ este o **variabilă aleatoare**
  - fiind un eșantion de zgomot (un eșantion dintr-un proces aleator)
  - presupunem o v.a. continuă , adică intervalul valorilor posibile e continuă
- $r(t_0) = s(t_0) + n(t_0 =$ o constantă $+$ o variabilă aleatoare
  - este de asemenea o variabilă aleatoare
  - $s(t_0)$ este o constantă, egală fie cu $s_0(t_0)$, fie cu $s_1(t_0)$
- Care e distribuția lui $r(t_0)$?
  - o constantă $+$ o v.a. $=$ aceeași distribuție ca v.a., dar translată cu valoarea constantei

Funcții de plauzibilitate

- Fie distribuția zgomotului $w(x)$, cunoscută
  - aceasta este distribuția v.a. $n(t_0)$
- Distribuția lui $r(t_0) = s(t_0) + n(t_0) = w(x)$ translată cu $s(t_0)$
- În ipoteza H_0, distribuția eșantionului este $w(r|H_0) = w(x)$ translată cu $s_0(t_0)$
- În ipoteza H_1, distribuția eșantionului este $w(r|H_1) = w(x)$ translată cu $s_1(t_0)$
- Distribuțiile $w(r|H_0)$ și $w(r|H_1)$ se numesc **distribuții condiționate** sau **funcțiile de plauzibilitate**
  - "$|$" înseamnă "condiționat de", "dat fiind"
  - adică dat fiind una sau cealaltă dintre ipoteze
  - r reprezintă necunoscuta funcției

Criteriul plauzibilității maxime (Maximum Likelihood)

- Cum se decide care ipoteză este adevărată, pe baza eșantionului observat $r = r(t_0)$?
- **Criteriul plauzibilității maxime**: se alege ipoteza care este **cea mai plauzibilă** a fi generat eșantionul observat $r = r(t_0)$
  - se alege valoarea maximă dintre $w(r(t_0)|H_0)$ și $w(r(t_0)|H_1)$
  - în engleză: Maximum Likelihood (ML)
- Criteriul ML exprimat la un **raport de plauzibilitate**:

$$\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

  - criteriul este evaluat pentru eșantionul observat $r = r(t_0)$

Exemplu: zgomot gaussian

- Fie cazul în care zgomotul are distribuție normală
- La tablă:
  - schiță a celor două distribuții condiționate $w(r|H_0)$ și $w(r|H_1)$
  - discuție: ce decizie se ia pentru diferite valori ale lui r
  - discuție: care este pragul T pentru decizii

Zgomot cu distribuție normală (AWGN)

- Caz particular: zgomotul are distribuția normală $\mathcal{N}(0, \sigma^2)$
  - zgomot de tip AWGN
- Raportul de plauzibilitate este $\frac{w(r|H_1)}{w(r|H_0)} = \frac{e^{-\frac{(r-s_1(t_0))^2}{2\sigma^2}}}{e^{-\frac{(r-s_0(r_0))^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} 1$
- Pentru distribuția normală, e preferabil să aplicăm **logaritmul natural**
  - logaritmul este o funcție monoton crescătoare, deci nu schimbă rezultatul comparației
  - dacă $A < B$, atunci $\log(A) < \log(B)$
- Valoarea **log-likelihood** al unui observații = logaritmul plauzibilității (likelihood)
  - de obicei se folosește logaritmul natural, dar poate fi în orice bază

Raportul “log-likelihood” în cazul ML

- Aplicarea logaritmului natural la ambii termeni ai relației conduce la:

$$-(r - s_1(t_0))^2 + (r - s_0(t_0))^2 \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

- Care este echivalent cu:

$$|r - s_0(t_0)| \underset{H_0}{\overset{H_1}{\gtrless}} |r - s_1(t_0)|$$

- Notă: $|r - A| =$ distanța dintre r și A
  - $|r| =$ distanța de la r la 0
- Așadar, se alege distanța minimă dintre $r(t_0)$ și $s_1(t_0)$ sau $s_0(t_0)$

Criteriul ML pentru zgomot gaussian

- Criteriul ML **pentru zgomot gaussian**: ipoteza se alege pe baza **celei mai apropiate** valori dintre $s_0(t_0)$ și $s_1(t_0)$ față de eșantionul $r = r(t_0)$
  - principiul **cel mai apropiat vecin** ("*nearest neighbor*")
  - un principiu foarte general, întâlnit în multe alte scenarii
  - un receptor ce folosește ML se mai numește **receptor de distanță minimă** ("*minimum distance receiver*")

Etape pentru decizia pe baza ML

1. Se schițează cele două distribuții condiționate $w(r|H_0)$ și $w(r|H_1)$
2. Se determină care dintre cele două funcții este mai mare în dreptul valorii eșantionului observat $r = r(t_0)$

Etape pentru decizia pe baza ML, zgomot gaussian

- Doar dacă zgomotul este gaussian, identic pentru toate ipotezele:
  1. Se determină $s_0(t_0) =$ valoarea semnalului original, în absența zgomotului, în cazul ipotezei H_0
  2. Se determină $s_1(t_0) =$ valoarea semnalului original, în absența zgomotului, în cazul ipotezei H_1
  3. Se compară cu eșantionul observat $r(t_0)$, se alege cea mai apropiată valoare

Decizie pe bază de prag

- Alegerea valorii celei mai apropiate = identic cu compararea r cu un prag $T = \frac{s_0(t_0)+s_1(t_0)}{2}$
  - i.e. dacă cele doup valori sunt 0 și 5, decidem prin compararea lui r cu 2.5
- În general, pragul = punctul de intersecție al celor două distribuții condiționate

Exerciți u

- Un semnal poate avea două valori posibile, 0 sau 5. Semnalul este afectat de zgomot alb, gaussian, cu distribuția $\mathcal{N}\ (\mu = 0, \sigma^2 = 2)$. Receptorul ia un singur eșantion, cu valoarea $r = 2.25$
  1. Scrieți expresiile celor două distribuții condiționate, și reprezentați-le
  2. Ce decizie se ia pe baza criteriului plauzibilității maxime?
  3. Dar dacă semnalul 0 este afectat de zgomot gaussian $\mathcal{N}(0, 0.5)$, iar semnalul 5 de zgomot uniform $\mathcal{U}[-4, 4]$?
  4. Repetați b. și c. dacă valoarea 0 se înlocuiește cu -1

Regiuni de decizie

- **Regiuni de decizie** = intervalul de valori ale eșantionului r pentru care se ia o anumită decizie
- Regiunea de decizie R_0 = intervalul de valori ale lui r care conduc la decizia D_0
- Regiunea de decizie R_1 = intervalul de valori ale lui r care conduc la decizia D_1
- Regiunile de decizie acoperă întreg domeniul de valori ale lui r (toată axa reală)
- Exemplu: indicați regiunile de decizie la exercițiul anterior
  - $R_0 = [-\infty, 2.5]$
  - $R_1 = [2.5, \infty]$

Funcția de plauzibilitate

- Să notăm în mod generic ipotezele cu H_i, și semnalele $s_i(t)$, unde i este 0 sau 1
- Să considerăm distribuția condiționată $w(r|H_i)$
  - fie cea de le exemplul anterior:

$$w(r|H_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r - s_i(t_0))^2}{2\sigma^2}}$$

- Care este variabila necunoscută în această expresie?
  - nu r, din moment ce acesta ni se dă în problemă
  - i este necunoscuta

Terminologie: probabilitate și plauzibilitate

- În aceeași expresie matematică a funcției de distribuție:
  - dacă se cunosc parametrii statistici (de ex. μ, σ, H_i), și necunoscuta este valoarea însăși (de ex. r, x) atunci funcția reprezintă densitatea de **probabilitate**
  - dacă se cunoaște valoarea însăși (de ex. r, x), și necunoscuta o reprezintă un parametru statistic (de ex. μ, σ, i), atunci avem o **funcție de plauzibilitate**

Funcția de plauzibilitate

- În cazul detecției semnalelor, funcția $w(r|H_i) = f(i)$ este o funcție de plauzibilitate
  - necunoscuta este i
- Funcția este definită doar pentru $i = 0$ și $i = 1$
  - sau, în general, pentru $i =$ câte ipoteze are problema
- Criteriul ML = se alege i pentru care această funcție este maximă

$$Decizia \;\; D_i = \arg\max_i w(r|H_i)$$

  - Notație:
    - $\arg\max f(x)$ = argumentul x pentru care funcția $f(x)$ este maximă
    - $\max f(x)$ = valoarea maximă a funcției $f(x)$
    - a se vedea exemplul grafic la tablă
- Criteriul plauzibilității maxime înseamnă “se alege i care maximizează funcția de plauzibilitate $f(i) = w(r|H_i)$”

Generalizări

- Dacă zgomotul are altă distribuție?
  - Se schițează distribuțiile condiționate
  - Se evaluează pentru $r = r(t_0)$
  - Criteriul ML = se alege cea mai mare funcție $w(r|H_i)$ în punctul r dat
- Regiunile de decizie sunt date de punctețe de intersecție ale distribuțiilor condiționate
  - Pot fi mai multe intersectări, în general, deci mai multe praguri

Generalizări

- Dacă zgomotul are distribuție diferită în ipoteza H_0 față de ipoteza H_1?
- Similar:
  - Se schițează distribuțiile condiționate
  - Se evaluează pentru $r = r(t_0)$
  - Criteriul ML = se alege cea mai mare funcție $w(r|H_i)$ în punctul r dat

Generalizări

- Dacă cele două semnale $s_0(t)$ și $s_1(t)$ sunt constante / nu sunt constante?
- Nu contează forma semnalelor
  - Tot ce contează sunt valorile celor două semnale la momentul de eșantionare t_0:
    - $s_0(t_0)$
    - $s_1(t_0)$

Generalizări

- Dacă avem mai mult de 2 ipoteze?
- Se extinde raționamentul la n ipoteze
  - Avem n semnale posibile $s_0(t), \dots \; s_{n-1}(t)$
  - Avem n valori diferite $s_0(t_0), \dots \; s_{n-1}(t_0)$
  - Avem n distribuții condiționate $w(r|H_i)$
  - Pentru $r = r(t_0)$ dat, se alege valoarea maximă dintre cele n valori $w(r|H_i)$

Generalizări

- Dacă se iau mai multe eșantioane din semnale?
- Va fi tratat separat într-un subcapitol ulterior

Exercițiu

- Un semnal poate avea patru valori posibile: -6, -2, 2, 6. Fiecare valoare durează timp de o secundă. Semnalul este afectat de zgomot alb cu distribuție normală. Receptorul ia un singur eșantion pe secundă. Folosind criteriul plauzibilității maxime, decideți ce semnal s-a transmis, dacă receptorul primește eșantioanele următoare:

$$4,\ 6.6,\ -5.2,\ 1.1,\ 0.3,\ -1.5,\ 7,\ -7,\ 4.4$$

Probabilități condiționate

- Putem calcula **probabilitățile condiționate** ale celor 4 rezultate posibile
- Fie regiunile de decizie:
  - R_0: dacă $r \in R_0$, decizia este D_0
  - R_1: daca $r \in R_1$, decizia este D_1
- Probabilitatea condiționată a rejecției corecte
  - $=$ probabilitatea de a lua decizia D_0 când ipoteza este H_0
  - $=$ probabilitatea ca r să fie în R_0, calculată pe distribuția $w(r|H_0)$

$$P(D_0|H_0) = \int_{R_0} w(r|H_0)dx$$

- Probabilitatea condiționată a alarmei false
  - $=$ probabilitatea de a lua decizia D_1 când ipoteza este H_0
  - $=$ probabilitatea ca r să fie în R_1, calculată pe distribuția $w(r|H_0)$

$$P(D_1|H_0) = \int_{R_1} w(r|H_0)dx$$

Probabilități condiționate

- Probabilitatea condiționată de pierdere
  - $=$ probabilitatea de a lua decizia D_0 când ipoteza este H_1
  - $=$ probabilitatea ca r să fie în R_0, calculată pe distribuția $w(r|H_1)$

$$P(D_0|H_1) = \int_{R_0} w(r|H_1)dx$$

- Probabilitatea condiționată a detecției corecte
  - $=$ probabilitatea de a lua decizia D_1 când ipoteza este H_1
  - $=$ probabilitatea ca r să fie în R_1, calculată pe distribuția $w(r|H_1)$

$$P(D_1|H_1) = \int_{R_1} w(r|H_1)dx$$

Probabilități condiționate

- Relații între probabilitățile condiționate
  - suma rejecție corectă + alarmă falsă = 1
  - suma pierdere + detecție corectă = 1
  - De ce? Justificați.

Probabilități condiționate

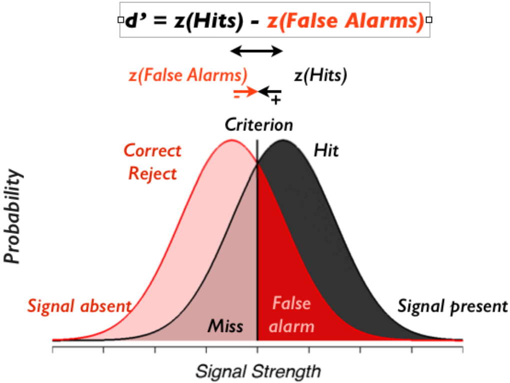


Figure 2: Probabilități condiționate

- Ignorați textul, contează zonele colorate
- [sursa: hhttp://gru.stanford.edu/doku.php/tutorials/sdt]*

Probabilitățile celor 4 rezultate

- Probabilitățile condiționate sunt calculate **dat fiind** una sau alta dintre ipoteze
- Nu includ și probabilitățile *ipotezelor înselor*
  - adică, $P(H_0) =$ probabilitatea de a avea ipoteza H_0
  - $P(H_1) =$ probabilitatea de a avea ipoteza H_1
- Pentru a le lua în calcul, se multiplică cu $P(H_0)$ sau $P(H_1)$
  - $P(H_0)$ și $P(H_1)$ se numesc probabilitățile **inițiale** (sau **a priori**) ale ipotezelor

Reamintire (TCI): regula lui Bayes

- Reamintire (TCI): regula lui Bayes

$$P(A \cap B) = P(B|A) \cdot P(A)$$

- Interpretare
  - Probabilitatea $P(A)$ este extrasă din $P(B|A)$
  - $P(B|A)$ nu mai conține nici o informație despre $P(A)$, șansele ca A chiar să aibă loc
  - Exemplu: P(gol | șut la poartă) $= \frac{1}{2}$. Câte goluri se înscriu?

Exercițiu

- Un semnal constant poate avea două valori posibile, 0 sau 5. Semnalul este afectat de zgomot gaussian $\mathcal{N}\left(\mu = 0, \sigma^2 = 2\right)$. Receptorul decide pe baza criteriului plauzibilității maxime, folosind un singur eșantion din semnal.
  1. Calculați probabilitatea condiționată a alarmei false
  2. Calculați probabilitatea condiționată de pierdere
  3. Dacă $P(H_0) = \frac{1}{3}$ și $P(H_1) = \frac{2}{3}$, calculați probabilitatea rejecției corecte și a detecției corecte (nu cele condiționate)

Dezavantaje ale criteriului plauzibilității maxime

- Criteriul ML compară distribuțiile **condiționate** ale eșantionului observat
  - condiționate de ipotezele H_0 sau H_1
- Condiționarea de ipotezele H_0 și H_1 ignoră probabilitatea celor două ipoteze H_0 și H_1
  - Decizia e aceeași indiferent dacă $P(H_0) = 99.99\%$ și $P(H_1) = 0.01\%$, sau invers
- Dacă $P(H_0) > P(H_1)$, am vrea să împingem pragul de decizie înspre H_1, și vice-versa
  - pentru că este mai probabil ca semnalul să fie $s_0(t)$
  - și de aceea vrem să "favorizăm"/"încurajăm" decizia D_0
- Avem nevoie de un criteriu mai general ...

Criteriul probabilității minime de eroare

- Se iau în calcul probabilitățile $P(H_0)$ și $P(H_1)$
- Se urmărește **minimizarea probabilității totale de eroare** P_e
  - erori = alarme false și ratări
- Trebuie să găsim regiunile de decizie R_0 și R_1

Probabilitatea de eroare

- Probabilitatea unei alarme false

$$\begin{aligned}P(D_1 \cap H_0) = & P(D_1|H_0) \cdot P(H_0) \\ = & \int_{R_1} w(r|H_0)dx \cdot P(H_0) \\ = & (1 - \int_{R_0} w(r|H_0)dx \cdot P(H_0)\end{aligned}$$

- Probabilitatea unei ratări

$$\begin{aligned}P(D_0 \cap H_1) = & P(D_0|H_1) \cdot P(H_1) \\ = & \int_{R_0} w(r|H_1)dx \cdot P(H_1)\end{aligned}$$

- Suma lor este

$$P_e = P(H_0) + \int_{-\infty}^{T} [w(r|H_1) \cdot P(H_1) - w(r|H_0) \cdot P(H_0)]dx$$

Probabilitatea de eroare minimă

- Urmărim minimizarea P_e, adică să minimizăm integrala
- Pentru a minimiza integrala, se alege R_0 astfel încât pentru toți $r \in R_0$, termenul din integrala este **negativ**
  - integrarea pe întregul interval în care o funcție este negativă conduce la valoarea minimă
- Așadar, când $w(r|H_1) \cdot P(H_1) - w(r|H_0) \cdot P(H_0) < 0$ avem $r \in R_0$, adică decizia D_0
- Invers, dacă $w(r|H_1) \cdot P(H_1) - w(r|H_0) \cdot P(H_0) > 0$ avem $r \in R_1$, adică decizia D_1
- Astfel

$$w(r|H_1) \cdot P(H_1) - w(r|H_0) \cdot P(H_0) \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

$$\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)}{P(H_1)}$$

Interpretare

- Similar cu criteriul plauzibilității maxime, dar depinde de probabilitățile celor două ipoteze (cazuri, simboluri)
  - Când una dintre ipoteze este mai probabilă decât cealaltă, pragul este împins în favoarea sa, înspre cealaltă ipoteză
- De asemenea bazat pe raportul de plauzibilitate, ca și primul criteriu

Criteriul probabilității minime de eroare - zgomot gaussian

- Presupunând că zgomotul este gaussian (normal), $\mathcal{N}(0, \sigma^2)$

$$w(r|H_1) = e^{-\frac{(r-A)^2}{2\sigma^2}}$$

$$w(r|H_0) = e^{-\frac{r^2}{2\sigma^2}}$$

- Se aplică logaritmul natural

$$-\frac{(r-A)^2}{2\sigma^2} + \frac{r^2}{2\sigma^2} \underset{H_0}{\overset{H_1}{\gtrless}} \ln\left(\frac{P(H_0)}{P(H_1)}\right)$$

- Echivalent

$$2rA - A^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \cdot \ln\left(\frac{P(H_0)}{P(H_1)}\right)$$

$$r \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{A^2 + 2\sigma^2 \cdot \ln\left(\frac{P(H_0)}{P(H_1)}\right)}{2A}}_{T}$$

Regiuni de decizie

- Se compară eșantionul tot cu un prag T, dar valoarea acestuia este împinsă înspre ipoteza mai puțin probabilă
  - T depinde de raportul $\frac{P(H_0)}{P(H_1)}$
- Regiuni de decizie
  - $R_0 = (-\infty, T]$
  - $R_1 = [T, \infty)$
  - pot fi diferite pentru alte tipuri de zgomot

Exerciții

- O sursă de informație furnizează două mesaje cu probabilitățile $p(a_0) = \frac{2}{3}$ și $p(a_1) = \frac{1}{3}$. Mesajele se transmit prin semnale constante cu valorile -5 (a_0) și 5 (a_1). Semnalele sunt afectate de zgomot alb cu distribuție gaussiană $\mathcal{N}(0, \sigma^2 = 1)$ Receptorul ia un singur eșantion cu valoarea r. Decizia se face prin compararea valorii r cu un prag T, astfel: dacă $r < T$ se decide că s-a transmis mesajul a_0, altfel se decide mesajul a_1.
  1. Să se găsească valoarea pragului T conform criteriul probabilității minime de eroare
  2. Dar dacă semnalul 5 este afectat de zgomot uniform $\mathcal{U}[-4, 4]$?
  3. Calculați probabilitatea unei alarme false și a unei ratări

Criteriul riscului (costului) minim

- Dacă ne afectează mai mult un anume tip de erori (de ex. alarme false) decât celelalte?
- Criteriul riscului (sau costului) minim: deciziile au un cost, se minimizează costul mediu
  - C_{ij} = costul deciziei D_i când ipoteza adevărată este H_j
  - C_{00} = costul unei rejecții corecte
  - C_{10} = costul unei alarme false
  - C_{01} = costul unei ratări
  - C_{11} = costul unei detecții corecte
- Definim **riscul** = costul mediu

$$R = C_{00}P(D_0 \cap H_0) + C_{10}P(D_1 \cap H_0) + C_{01}P(D_0 \cap H_1) + C_{11}P(D_1 \cap H_1)$$

- Criteriul riscului minim: **se minimizează riscul R**

Calcule

- Demonstrație la tablă
  - se folosește regula lui Bayes
- Concluzie: regula de decizie este

$$\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(C_{10} - C_{00})p(H_0)}{(C_{01} - C_{11})p(H_1)}$$

Interpretare

- Similar cu primele două criterii, bazat tot pe **raportul de plauzibilitate**
- Atât probabilitățile cât și costurile pot împinge pragul T într-o parte sau alta
- Caz particular: dacă $C_{10} - C_{00} = C_{01} - C_{11}$, se reduce la criteriul probabilității de eroare minime
  - de ex.: dacă $C_{00} = C_{11} = 0$ și $C_{10} = C_{01}$

În zgomot gaussian

- Dacă zgomotul este gaussian (normal), se aplică logaritmul natural, ca la celelalte criterii
- Se obține valoarea pragului T:

$$-(r-A)^2 + r^2 \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{2\sigma^2 \cdot \ln\left(\frac{(C_{10}-C_{00})p(H_0)}{(C_{01}-C_{11})p(H_1)}\right)}_{C}$$

$$r \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{A^2 + 2\sigma^2 \cdot \ln\left(\frac{(C_{10}-C_{00})p(H_0)}{(C_{01}-C_{11})p(H_1)}\right)}{2A}}_{T}$$

În zgomot gaussian

- În general, pentru un raport de plauzibilitate comparat cu K, $\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} K$, pragul este

$$T = \frac{A^2 + 2\sigma^2 \cdot \ln K}{2A}$$

Exemplu

- Exemplu la tablă: 0 / 5, zgomot alb $N(0, \sigma^2)$, un eșantion

Criteriul Neymar-Pearson

- Criteriul Neymar-Pearson: se maximizează probabilitatea de detecție ($P(D_1 \cap H_1)$) păstrând probabilitatea alarmei false sub o limită fixată ($P(D_1 \cap H_0) \leq \lambda$)
- Se deduce pragul T din constrângerea la limită $P(D_1 \cap H_0) = \lambda$

Exercițiu

- O sursă de informație produce două mesaje cu probabilitățile $p(a_0) = \frac{2}{3}$ și $p(a_1) = \frac{1}{3}$.
- Mesajele sunt codate ca semnale constante cu valorile -5 (a_0) și 5 (a_1).
- Semnalele sun afectate de zgomot alb cu distribuție *triunghiulară* în intervalul $[-5, 5]$.
- Receptorul ia un singur eșantion r.
- Decizia se ia prin compararea r cu un prag T: dacp $r < T$ se decide că mesajul este a_0, altfel este a_1.
    1. Găsiți pragul T conform criteriului Neymar-Pearson, pentru $P_{fa} \leq 10^{-2}$
    2. Care este probabilitatea de detecție corectă?

Două nivele de semnal nenule

- Dacă semnalul $s_0(t)$ nu este 0, ci are o altă valoare constantă $s_0(t) = B$?
- Distribuția zgomotului $w(r|H_0)$ va fi centrată pe B în loc de 0
- În rest, totul rămâne la fel
- Performanțele sunt determinate de diferența dintre cele două valori $(A - B)$
  - cazul $s_0 = 0$, $s_1 = A$ este identic cu cazul $s_0 = -\frac{A}{2}$, $s_1 = \frac{A}{2}$
- Valabil pentru toate criteriile de decizie

Semnale diferențiale sau unipolare

- Semnal unipolar: o valoare este 0, cealaltă este nenulă
  - $s_0 = 0$, $s_1 = A$
- Semnal diferențial: două valori nenule cu semne contrare, aceeași valoare absolută
  - $s_0 = -\frac{A}{2}$, $s_1 = \frac{A}{2}$
- Care metodă este mai bună?

Semnale diferențiale sau unipolare

- Cu aceeași diferență între nivele, performanțele deciziei sunt identice
- Dar puterea medie a semnalelor diferă
- Pentru semnale diferențiale: $P = \left(\pm\frac{A}{2}\right)^2 = \frac{A^2}{4}$
- Pentru semnale unipolare: $P = P(H_0) \cdot 0 + P(H_1)\,(A)^2 = \frac{A^2}{2}$
  - presupunând probabilități egale $P(H_0) = P(H_1) = \frac{1}{2}$
- Semnalul diferențial necesită putere la jumătate față de cel unipolar (mai bine)

Sumar: criterii de decizie

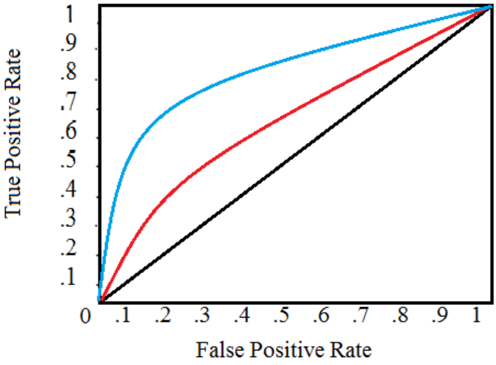
- Am văzut: decizie între două nivele constante, bazată pe 1 eșantion r
- Toate criteriile au la bază un test al raportului de plauzibilitate

$$\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Criterii diferite conduc la valori diferite pentru K (pragul de plauzibilitate)
- În funcție de distribuția zgomotului, axa reală este împărțită în regiuni
  - regiunea R_0: dacă r este aici, se decide D_0
  - regiunea R_1: dacă r este aici, se decide D_1
  - de ex. $R_0 = (-\infty, \frac{A+B}{2}]$, $R_1 = (\frac{A+B}{2}, \infty)$ (pentru crit. plauz. max)
- Pentru zgomot gaussian, pragul este $T = \frac{A^2+2\sigma^2 \cdot \ln K}{2A}$

Caracteristica de operare a receptorului (ROC)

- Performanța unui receptor este ilustrată cu un grafic numit **"Caracteristica de operare a receptorului"** (**"Receiver Operating Characteristic", ROC)**
- Reprezintă probabilitatea detecției corecte $P_d = P(D_1 \cap H_1)$ în funcție de probabilitatea alarmei false $P_{fa} = P(D_1 \cap H_0)$



Caracteristica de operare a receptorului (ROC)

- Există întotdeauna un **compromis** între P_d și P_{fa}
  - creșterea P_d implică și creșterea P_{fa}
  - pentru a fi siguri că nu ratăm nici un semnal (creșterea P_d), plătim prin creșterea probabilității de alarme false
- Criterii diferite = diferite praguri K = diferite puncte pe grafic = compromisuri diferite
- Cum să creștem performanțele unui receptor?
  - adică să creștem P_D menținând P_{fa} la aceeași valoare

Performanțele detecției în zgomot alb gaussian

- Considerăm probabilități egale $P(H_0) = P(H_1) = \frac{1}{2}$
- Deciziile se iau pe baza raportului de plauzibilitate

$$\frac{w(r|H_1)}{w(r|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Probabilitatea detecției corecte este

$$\begin{aligned}
P_d =& P(D_1|H_1)P(H_1) \\
=& P(H_1) \int_T^{\infty} w(r|H_1) \\
=& P(H_1)(F(\infty) - F(T)) \\
=& \frac{1}{4} \left(1 - erf\left(\frac{T - A}{\sqrt{2}\sigma}\right)\right) \\
=& Q\left(\frac{T - A}{\sqrt{2}\sigma}\right)
\end{aligned}$$

Performanțele detecției în zgomot alb gaussian

- Probabilitatea alarmei false este

$$\begin{aligned}
P_{fa} =& P(D_1|H_0)P(H_0)\\
=& P(H_0)\int_T^\infty w(r|H_0)\\
=& P(H_0)(F(\infty) - F(T))\\
=& \frac{1}{4}\left(1 - erf\left(\frac{T-0}{\sqrt{2}\sigma}\right)\right)\\
=& Q\left(\frac{T}{\sqrt{2}\sigma}\right)
\end{aligned}$$

- Rezultă că $\frac{T}{\sqrt{2}\sigma} = Q^{-1}\left(P_{fa}\right)$
- Înlocuind în P_d se obține

$$P_d = Q\left(\underbrace{Q^{-1}\left(P_{fa}\right)}_{constant} - \frac{A}{\sqrt{2}\sigma}\right)$$

Raportul semnal zgomot

- **Raportul semnal zgomot (SNR)** $= \frac{\text{puterea semnalului original}}{\text{puterea zgomotului}}$
- Puterea medie a unui semnal = valoarea pătratică medie $= \overline{X^2}$
  - Puterea semnalului original este $\frac{A^2}{2}$
  - Puterea zgomotului este $\overline{X^2} = \sigma^2$ (pentru valoare medie nulă $\mu = 0$)
- În cazul nostru, SNR $= \frac{A^2}{2\sigma^2}$

$$P_d = Q\left(\underbrace{Q^{-1}\left(P_{fa}\right)}_{constant} - \sqrt{SNR}\right)$$

- Pentru P_{fa} de valoare fixă, P_d crește odată cu SNR
  - Q este o funcție monoton descrescătoare

Performanța depinde de SNR

- Performanța receptorului crește odată cu creșterea SNR
  - SNR mare: performanță bună
  - SNR mic: performanță slabă

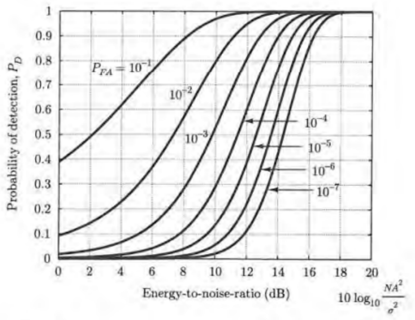


Figure 4: Performanțele detecției depind de SNR

[sursa: Fundamentals of Statistical Signal Processing, Steven Kay]

Decizii între ipoteze statistice

- Teoria statistică a detecției este utilă și în alte contexte în afară de detecția unor semnale propriu-zise
  - oriunde avem de ales între două ipoteze
- Decizia se face între două distribuții de probabilitate
  - indiferent ce semnificație au cele două distribuții
- În cazul detecției unui semnal constant, se alege între două distribuții care **diferă doar prin valoarea medie**, în general
  - o distribuție are valoarea medie 0, cealaltă A
- Dar se poate face decizie între distribuții care diferă prin alt parametru
  - valoarea medie, sau
  - varianța, or
  - forma distribuției, etc

Decizii între ipoteze statistice

- Exemplu: Un eșantion cu valoarea $r = 2.5$ poate proveni dintr-o distribuție $\mathcal{N}(0, \sigma^2 = 1)$ (ipoteza H_0) sau dintr-o alta $\mathcal{N}(0, \sigma^2 = 2)$ (ipoteza H_1). Care ipoteză se consideră adevărată?
  - Ceea ce diferă este varianța, nu valoarea medie
- Se pot folosi exact aceleași criterii
  - Se desenează cele două distribuții
  - Se calculează plauzibilitățile $w(r|H_0)$ și $w(r|H_1)$ for r
  - Se decide pe baza raportului de plauzibilitate, conform unui criteriu

II.3 Detecția unui semnal constant cu mai multe eșantioane

Eșantioane multiple dintr-un semnal constant

- Presupunem că avem mai multe eșantioane, nu doar unul
- Eșantioanele formează **vectorul eșantioanelor**

$$\mathbf{r} = [r_1, r_2, ...r_N]$$

- În ambele ipoteze, semnalul recepționat este un **proces aleator**
  - H_0: proces aleator cu valoarea medie 0
  - H_1: proces aleator cu valoarea medie A
- Dacă zgomotul este staționar și ergodic, semnalul recepționat este și el staționar și ergodic (semnalul = o constantă + zgomotul)
- Valorile vectorului **r** sunt descrise de **distribuția de ordin** N a procesului aleator, $w_N(\mathbf{r}) = w_N(r_1, r_2, ...r_N)$
- Dacă zgomotul este alb, momentele de timp când se iau eșantioanele nu contează

Plauzibilitatea vectorului de eșantioane

- Se aplică **aceleași criterii** bazate pe raportul de plauzibilitate în cazul unui singur eșantion

$$\frac{w_N(\mathbf{r}|H_0)}{w_N(\mathbf{r}|H_1)} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Observații
  - $\mathbf{r}$ este un vector; prin el se consideră plauzibilitatea tuturor eșantioanelor
  - ipotezele H_0 și H_1 sunt aceleași ca în cazul cu 1 eșantion
  - $w_N(\mathbf{r}|H_0) =$ plauzibilitatea vectorului $\mathbf{r}$ în ipoteza H_0
  - $w_N(\mathbf{r}|H_1) =$ plauzibilitatea vectorului $\mathbf{r}$ în ipoteza H_1
  - valoarea lui K este dată de criteriul de decizie utilizat
- Interpretare: se alege ipoteza cea mai plauzibilă de a fi generat datele observate
  - identic ca la 1 eșantion, doar că acum datele = mai multe eșantioane

Descompunere pe fiecare eșantion

- Presupunând că zgomotul este alb, eșantioanele r_i sunt **realizări independente ale aceleiași distribuții**
- În acest caz, distribuția totală $w_N(\mathbf{r}|H_j)$ se poate descompune ca un produs

$$w_N(\mathbf{r}|H_j) = w(r_1|H_j) \cdot w(r_2|H_j) \cdot ... \cdot w(r_N|H_j)$$

- Termenii $w(r_i|H_j)$ sunt plauzibilitățile fiecărui eșantion în parte
    - de ex. plauzibilitatea obținerii vectorului $[5.1, 4.7, 4.9]$ = plauzibilitatea obținerii lui 5.1 $\times$ plauzibilitatea obținerii lui 4.7 $\times$ plauzibilitatea obținerii lui 4.9

Descompunere pe fiecare eșantion

- Prin urmare, criteriile bazate pe raportul de plauzibilitate devin

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{w(r_1|H_1)}{w(r_1|H_0)} \cdot \frac{w(r_2|H_1)}{w(r_2|H_0)} ... \frac{w(r_N|H_1)}{w(r_N|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Raportul de plauzibilitate al unui vector de eșantioane = produsul rapoartelor plauzibilitate ale fiecărui eșantion

Caz particulae: AWGN

- AWGN = "Additive White Gaussian Noise" = Zgomot alb, gaussian, aditiv
- În ipoteza H_1: $w(r_i|H_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r_i-A)^2}{2\sigma^2}}$
- În ipoteza H_0: $w(r_i|H_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{r_i^2}{2\sigma^2}}$
- Raportul de plauzibilitate al vectorului $\mathbf{r}$

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum(r_i-A)^2}{2\sigma^2}}}{e^{-\frac{\sum(r_i)^2}{2\sigma^2}}}$$

- Se pot găsi trei interpretări ale raportului de plauzibilitate

Interpretarea 1: media eșantioanelor

- Interpretarea 1: media eșantioanelor

$$\begin{aligned}
\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} &= \frac{e^{-\frac{\sum (r_i - A)^2}{2\sigma^2}}}{e^{-\frac{\sum (r_i)^2}{2\sigma^2}}} \\
&= e^{-\frac{\sum (r_i - A)^2 - \sum (r_i)^2}{2\sigma^2}} \\
&= e^{-\frac{\sum (r_i^2 - 2 r_i A + A^2) - \sum (r_i)^2}{2\sigma^2}} \\
&= e^{-\frac{\sum (-2 r_i A + A^2)}{2\sigma^2}} \\
&= e^{-\frac{-2A \sum (r_i) + N A^2}{2\sigma^2}} \\
&= e^{-\frac{-2A \frac{\sum (r_i)}{N} + A^2}{2 \frac{\sigma^2}{N}}}
\end{aligned}$$

Media a N variabile aleatoare normale

- Fie U_r = media aritmetică a eșantioanelor r_i

$$U_r = \frac{1}{N} \sum r_i$$

- Care este distribuția sa?
- Fie suma $S_r = \sum r_i$ a celor N eșantioane r_i
  - Din cap.I: suma unor v.a. normale cu distribuția $\mathcal{N}(\mu, \sigma^2)$ este:
  - cu distribuție normală $\mathcal{N}(\mu_S, \sigma_S^2)$, unde:
  - valoarea medie: $\mu_S = N \cdot \mu$
  - varianța: $\sigma_S^2 = N \cdot \sigma^2$
- Așadar $U_r = \frac{1}{N} S_r$, din proprietățile mediei se obține:
  - U_r are distribuție normală, cu:
  - valoarea medie $= \frac{1}{N} \mu_S = \frac{1}{N} N \mu = \mu$
  - varianța $= \left(\frac{1}{N}\right)^2 \sigma_S^2 = \left(\frac{1}{N}\right)^2 N \sigma_S^2 = \frac{1}{N} \sigma^2$

Media a N variabile aleatoare normale

- Media a N realizări ale unei distribuții normale are tot o distribuție normală, cu
  - aceeași valoare medie
  - varianța de N ori mai mică
- Dacă N este foarte mare, media aritmetică este un **estimator** foarte bun pentru valoarea medie a distribuției
  - distribuția sa devine foarte "îngustă" în jurul valorii medii

Interpretarea 1: media eșantioanelor

$$
\begin{aligned}
\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} &= e^{-\frac{-2AU_r+A^2}{2\frac{\sigma^2}{N}}} \\
&= \frac{e^{-\frac{U_r^2-2AU_r+A^2}{2\frac{\sigma^2}{N}}}}{e^{-\frac{U_r^2}{2\frac{\sigma^2}{N}}}} \\
&= \frac{e^{-\frac{(U_r-A)^2}{2\frac{\sigma^2}{N}}}}{e^{-\frac{U_r^2}{2\frac{\sigma^2}{N}}}} \\
&= \frac{w(U_r|H_1)}{w(U_r|H_0)}
\end{aligned}
$$

- Raportul de plauzibilitate a N eșantioane gaussiene = raportul de plauzibilitate al **mediei eșantioanelor**

Interpretarea 1: media eșantioanelor

- Raportul de plauzibilitate a N eșantioane gaussiene = raportul de plauzibilitate al **mediei eșantioanelor**
  - media are o varianță mai mică, $\frac{1}{N}\sigma^2$, deci este mai precisă
  - e ca și cum distribuția zgomotului devine de N ori mai îngustă (datorită medierii)
- Detecția unui semnal constant cu N eșantioane este similaru cu detecția cu un singur eșantion, doar că
  - se folosește valoarea medie a eșantioanelor r_i
  - distribuția sa este de N ori mai îngustă (varianța e de N ori mai mică)
- Când N crește, probabilitatea erorilor scade => performanțe îmbunătățite

Exerciții

Exercițiu:

- Un semnal poate avea două valori, 0 (ipoteza H_0) sau 6 (ipoteza H_1). Semnalul este afectat de AWGN $\mathcal{N}(0, \sigma^2 = 1)$. Receptorul ia 5 eșantioane cu valorile $\{1.1, 4.4, 3.7, 4.1, 3.8\}$.
  1. Ce decizie se ia conform criteriului plauzibilității maxime?
  2. Ce decizie se ia conform criteriului probabilității minime de eroare. dacă $P(H_0) = 2/3$ și $P(H_1) = 1/3$?

Interpretarea 2: geometric

- Folositoare în special pentru criteriul plauzibilității maxime
- Raportul de plauzibilitate pentru vectorul **r**

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum (r_i - A)^2}{2\sigma^2}}}{e^{-\frac{\sum (r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- La criteriul plauzibilității maxime se compară cu 1

$$\frac{e^{-\frac{\sum (r_i - A)^2}{2\sigma^2}}}{e^{-\frac{\sum (r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

$$e^{-\frac{\sum (r_i - A)^2}{2\sigma^2} + \frac{\sum (r_i)^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

$$-\sum (r_i - A)^2 + \sum (r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

$$\sum (r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \sum (r_i - A)^2$$

Interpretarea 2: geometric

- $\sqrt{\sum(r_i)^2}$ este distanța geometrică (Euclidiană) între punctul $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctul $\mathbf{0} = [0, 0, ...0]$
- $\sqrt{\sum(r_i - A)^2}$ este distanța geometrică (Euclidiană) între punctul $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctul $\mathbf{A} = [A, A, ...A]$
- criteriul plauzibilității maxime alege **vectorul (punctul) semnalului cel mai apropiat** de vectorul (punctul) recepționat, într-un spațiu N-dimensional
  - receptorul se mai numește "receptor de distanță minimă"
  - aceeași interpretare ca în cazul 1-D
- Întrebare: care este interpretarea geometrică pentru celelalte criterii?

Exerciții

Exerciții:

- Un semnal poate avea două valori, 0 (ipoteza H_0) sau 6 (ipoteza H_1). Semnalul este afectat de AWGN $\mathcal{N}(0, \sigma^2 = 1)$. Receptorul ia două eșantioane cu valorile $\{1.1, 4.4\}$.
  1. Care este decizia conform criteriului plauzibilității maxime? Utilizați interpretarea geometrică.

Interpretarea 3: valoarea corelației

- Raportul de plauzibilitate al vectorului $\mathbf{r}$

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum(r_i-A)^2}{2\sigma^2}}}{e^{-\frac{\sum(r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$e^{-\frac{\sum(r_i-A)^2}{2\sigma^2}+\frac{\sum(r_i)^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$-\sum(r_i - A)^2 + \sum(r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$2\sum r_i A - NA^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$\frac{1}{N}\sum r_i A \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{A^2}{2} + \frac{1}{N}\sigma^2 \ln K}_{L=const}$$

Interpretarea 3: valoarea corelației

- **Valoarea de corelație** (sau “corelația”) a două semnale x and y este

$$< x, y >= \frac{1}{N} \sum x[n]y[n]$$

- Este valoarea funcției de corelație în 0

$$< x, y >= R_{xy}[0] = \overline{x[n]y[n+0]}$$

- Pentru semnale continue

$$< x, y >= \frac{1}{T} \int_{T/2}^{T/2} x(t)y(t)dt$$

- $\frac{1}{N} \sum r_i A =< \mathbf{r}, \mathbf{A} >$ este corelația vectorului recepționat $\mathbf{r} = [r_1, r_2, ...r_N]$ cu vectorul **țintă** $\mathbf{A} = [A, A, ...A]$

Interpretarea 3: valoarea corelației

- Dacă valoarea de corelație a vectorului recepționat cu vectorul țintă $\mathbf{A} = [A, A, ...A]$ este mai mare decât un prag L, se decide că semnalul este detectat.
  - altfel, semnalul este rejectat
- Decizia este **similară cu detecția semnalului cu singur eșantion**, unde valoarea eșantionului este $< \mathbf{r}, \mathbf{A} >$

Corelația ca măsura a similarității semnalelor

- În domeniul prelucrărilor de semnal, corelația este o formă de a măsura **similaritatea** a două semnale
- Interpretare: verificăm dacă vectorul recepționat este suficient de similar cu semnalul constant A
  - Da: (corelație mare) $=>$ semnalul este detectat
  - Nu: (corelație mică) $=>$ nu este detectat

Generalizare: două valori nenule

- Generalizare: două valori nenule, B și A
  - în zgomot Gaussian
- Interpretarea 1: media eșantioanelor
  - se folosește tot media eșantioanelor, cele două distribuții sunt centrate pe B și A
- Interpretarea 2: geometric (crit. plauzib. maxime)
  - se alege minimul distanței dintre $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctele $\mathbf{B} = [B, B, ...]$ și $\mathbf{A} = [A, A, ...]$
- Interpretarea 3: corelația
  - se calculează $< \mathbf{r}, \mathbf{B} >$ and $< \mathbf{r}, \mathbf{A} >$, corelația lui $\mathbf{r}$ cu $\mathbf{B} = [B, B, ...]$ și cu $\mathbf{A} = [A, A, ...]$.
  - pe slide-ul următor

Detecția a două valori nenule folosind corelația

$$e^{-\frac{\sum(r_i - A)^2}{2\sigma^2} + \frac{\sum(r_i - B)^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$-\sum(r_i - A)^2 + \sum(r_i - B)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$2\sum r_i A - NA^2 - 2\sum r_i B + NB^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$\frac{1}{N}\sum r_i A - \frac{A^2}{2} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{N}\sum r_i B - \frac{B^2}{2} + \frac{1}{N}\sigma^2 \ln K$$

Detecția a două valori nenule folosind corelația

- Pentru criteriul plauzibilității maxime ($K = 1$):

$$< \mathbf{r}, \mathbf{A} > - \frac{< \mathbf{A}, \mathbf{A} >}{2} \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, \mathbf{B} > - \frac{< \mathbf{B}, \mathbf{B} >}{2}$$

- Dacă valorile sunt opuse, $B = -A$, se alege cea mai similară cu **r**:
  - corelația este o măsură a similarității

$$< \mathbf{r}, \mathbf{A} > \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, -\mathbf{A} >$$

- Alte criterii: termen adițional $\frac{1}{N}\sigma^2 \ln K$

Exerciţiu:

- Un semnal poate avea două valori, -4 (ipoteza H_0) sau 5 (ipoteza H_1). Semnalul este afectat de zgomot alb Gaussian $\mathcal{N}(0, \sigma^2 = 1)$. Receptorul ia trei eşantioane cu valorile $\{1.1, 4.4, 2.2\}$.
    1. Care este decizia, conform criteriului plauzibilităţii maxime? Folosiţi toate cele trei interpretări.

II.4 Detecția semnal oarecare cu mai multe eșantioane

Eșantioane multiple dintr-un semnal oarecare

- Dorim detecția unui semnal **oarecare (ne-constant)** $s(t)$
- Cele N eșantioane se iau la momentele de timp $\mathbf{t} = [t_1, t_2, ...t_N]$ și formează **vectorul eșantioanelor**

$$\mathbf{r} = [r_1, r_2, ...r_N]$$

- Ce diferă față de cazul unui semnal constant?

Ipoteze

- În fiecare ipoteză, semnalul este un proces aleator
  - H_0: proces aleator cu medie 0
  - H_1: proces aleator cu media $s(t)$
- Eșantionul r_i, de la momentul t_i, poate fi:
  - $0 +$ zgomot, în ipoteza H_0
  - $s(t_i) +$ zgomot, în ipoteza H_1
- Întregul vector al eșantioanelor $\mathbf{r}$ poate fi
  - $0 +$ zgomot, , în ipoteza H_0
  - $s(t) +$ zgomot, în ipoteza H_1, pentru $t =$ timpii de eșantionare t_i
- Distribuția vectorului $\mathbf{r}$ este descrisă de o funcție $w_N(\mathbf{r})$

Plauzibilitatea vectorului eșantioanelor

- Se folosesc **aceleași criterii** bazate pe raportul de plauzibilitate ca la semnale constante:

$$\frac{w_N(\mathbf{r}|H_0)}{w_N(\mathbf{r}|H_1)} \overset{H_1}{\underset{H_0}{\gtrless}} K$$

- Diferența este că semnalele "adevărate" sunt acum
  - $[0, 0, \dots\ 0]$ în ipoteza H_0
  - $[s(t_1), s(t_2), ...s(t_N)]$ în ipoteza H_1

Descompunere

- Distribuția vectorială $w_N(\mathbf{r}|H_j)$ se poate descompune ca un produs

$$w_N(\mathbf{r}|H_j) = w(r_1|H_j) \cdot w(r_2|H_j) \cdot ... \cdot w(r_N|H_j)$$

- Toate criteriile de decizie bazate pe raportul de plauzibilitate se pot scrie ca

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{w(r_1|H_1)}{w(r_1|H_0)} \cdot \frac{w(r_2|H_1)}{w(r_2|H_0)} ... \frac{w(r_N|H_1)}{w(r_N|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Raportul de plauzibilitate al unui singur eșantion r_i se calculează folosind cele două valori posibile ale semnalului, 0 și $s(t_i)$
  - la semnale constante, valorile erau 0 și A întotdeauna
  - acum sunt 0 și $s(t_i)$, în funcție de momentele de eșantionare t_i
  - momentele de eșantionare t_i trebuie alese astfel încât să maximizeze performanțele detecției

Caz particular: zgomot alb Gaussian ("AWGN")

- AWGN = "Additive White Gaussian Noise"
- In hypothesis H_1: $w(r_i|H_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(r_i - s(t_i))^2}{2\sigma^2}}$
- In hypothesis H_0: $w(r_i|H_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{r_i^2}{2\sigma^2}}$
- Raportul de plauzibilitate al vectorului $\mathbf{r}$

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum (r_i - s(t_i))^2}{2\sigma^2}}}{e^{-\frac{\sum (r_i)^2}{2\sigma^2}}}$$

- Sunt posibile două interpretări

Interpretarea 1: valoarea medie

- Interpretarea 1: valoarea medie
- Nu mai este valabilă, întrucât valorile $s(t_i)$ nu mai sunt identice

Interpretarea 2: geometric

- Folositoare mai ales în cazul criteriului plauzibilității maxime
- Raportul de plauzibilitate pentru vectorul $\mathbf{r}$

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum(r_i - s(t_i))^2}{2\sigma^2}}}{e^{-\frac{\sum(r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

- Criteriul plauzibilității maxime: $K = 1$

$$\frac{e^{-\frac{\sum(r_i - s(t_i))^2}{2\sigma^2}}}{e^{-\frac{\sum(r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

$$e^{-\frac{\sum(r_i - s(t_i))^2}{2\sigma^2} + \frac{\sum(r_i)^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} 1$$

$$-\sum(r_i - s(t_i))^2 + \sum(r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 0$$

$$\sum(r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \sum(r_i - s(t_i))^2$$

Interpretarea 2: geometric

- $\sqrt{\sum(r_i)^2}$ este distanța geometrică (Euclidiană) între punctul $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctul $\mathbf{0} = [0, 0, ...0]$
- $\sqrt{\sum(r_i - s(t_i))^2}$ este distanța geometrică (Euclidiană) între punctul $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctul $\mathbf{s(t)} = [s(t_1), s(t_2), ...s(t_N)]$
- Criteriul plauz. maxime alege **semnalul cel mai apropiat** de cel recepționat, într-un spațiu N-dimensional
  - se mai numește "receptor de distanță minimă"
  - aceeași interpretare ca în cazul 1-D
- Întrebare: interpretarea geometrică pentru celelalte criterii?

Exercițiu

Exercițiu:

- Fie detecția unui semnal $s(t) = 3\sin(2\pi f t)$ care poate fi prezent (ipoteza H_1) sau absent (ipoteza H_0). Semnalul este afectat de zgomot alb Gaussian $\mathcal{N}(0, \sigma^2 = 1)$. Receptorul ia două eșantioane.
  1. Care sunt cele mai bune momente de eșantionare t_1 și t_2 pentru a maximiza performanțele detecției?
  2. Receptorul ia două eșantioane $\{1.1, 4.4\}$, la momentele de timp $t_1 = \frac{0.125}{f}$ și $t_2 = \frac{0.625}{f}$. Care este decizia, conform criteriului plauz. maxime? Utilizați interpretarea geometrică.
  3. Dar dacă receptorul ia un al treilea eșantion la momentul $t_3 = \frac{0.5}{f}$. Se poate îmbunătăți detecția?

Interpretarea 3: valoarea corelației

- Raportul de plauzibilitate pentru vectorul $\mathbf{r}$

$$\frac{w_N(\mathbf{r}|H_1)}{w_N(\mathbf{r}|H_0)} = \frac{e^{-\frac{\sum(r_i - s(t_i))^2}{2\sigma^2}}}{e^{-\frac{\sum(r_i)^2}{2\sigma^2}}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$e^{-\frac{\sum(r_i - s(t_i))^2}{2\sigma^2} + \frac{\sum(r_i)^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$-\sum(r_i - s(t_i))^2 + \sum(r_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$2\sum r_i s(t_i) - \sum s(t_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$\frac{1}{N}\sum r_i s(t_i) \underset{H_0}{\overset{H_1}{\gtrless}} \underbrace{\frac{1}{2}\frac{\sum s(t_i)^2}{N} + \frac{1}{N}\sigma^2 \ln K}_{L = const}$$

Interpretarea 3: valoarea corelației

- $\frac{1}{N}\sum r_i s(t_i) =< \mathbf{r}, \mathbf{s(t_i)} >$ reprezintă valoarea corelației (sau "corelația") eșantioanelor recepționate $\mathbf{r} = [r_1, r_2, ...r_N]$ cu eșantioanele **țintă** $\mathbf{s(t_i)} = [s(t_1), s(t_2), ...s(t_N)]$
- Dacă corelația eșantioanelor recepționate **r** cu eșantioanele **țintă** $\mathbf{s(t_i)}$ este mai mare decât un prag L, se decide că semnalul este prezent.
    - Altfel, se decide că semnalul este absent
    - Corelația este o măsură a **similarității** a două semnale

Generalizare: două semnale oarecare

- Generalizare: se decide între **două semnale diferite** $s_0(t)$ și $s_1(t)$
  - în zgomot Gaussian
- Interpretarea 2: geometric
  - se alege distanța Euclidiană minimă dintre $\mathbf{r} = [r_1, r_2, ...r_N]$ și punctele $\mathbf{s_0(t)} = [s_0(t_1), s_0(t_2), ...]$ și $\mathbf{s_1(t)} = [s_1(t_1), s_1(t_2), ...]$
- Interpretarea 3: valoarea corelației
  - se calculează corelația $\mathbf{r}$ cu $\mathbf{s_0(t)} = [s_0(t_1), s_0(t_2), ...]$ și $\mathbf{s_1(t)} = [s_1(t_1), s_1(t_2), ...]$, $< \mathbf{r}, \mathbf{s_0} >$ and $< \mathbf{r}, \mathbf{s_1} >$.
  - pe slide-ul următor

Detecție între două semnale diferite folosind corelația

$$e^{-\frac{\sum(r_i - s_1(t_i))^2}{2\sigma^2} + \frac{\sum(r_i - s_0(t_i))^2}{2\sigma^2}} \underset{H_0}{\overset{H_1}{\gtrless}} K$$

$$-\sum(r_i - s_1(t_i))^2 + \sum(r_i - s_0(t_i))^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$2\sum r_i s_1(t_i) - \sum s_1(t_i)^2 - 2\sum r_i s_0(t_i) + \sum s_0(t_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} 2\sigma^2 \ln K$$

$$\frac{1}{N}\sum r_i s_1(t_i) - \sum s_1(t_i)^2 \underset{H_0}{\overset{H_1}{\gtrless}} \frac{1}{N}\sum r_i s_0(t_i) - \sum s_0(t_i)^2 + \frac{1}{N}\sigma^2 \ln K$$

Detecție între două semnale diferite folosind corelația

- Criteriul plauz. maxime ($K = 1$):

$$< \mathbf{r}, \mathbf{s_1} > - \frac{< \mathbf{s_1}, \mathbf{s_1} >}{2} \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, \mathbf{s_0} > - \frac{< \mathbf{s_0}, \mathbf{s_0} >}{2}$$

- Dacă semnalele au aceeași energie: $\sum s_1(t_i)^2 = \sum s_0(t_i)^2$, atunci $< \mathbf{s_1}, \mathbf{s_1} > = < \mathbf{s_0}, \mathbf{s_0} >$, și alegem semnalul **cel mai asemănător cu r**:
  - corelația este o măsură a similarității a două semnale

$$< \mathbf{r}, \mathbf{s_1} > \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, \mathbf{s_0} >$$

- Exemple:
  - Modulație BPSK: $s_1 = A\cos(2\pi f t)$, $s_0 = -A\cos(2\pi f t)$
  - Modulație 4-PSK: $s_{n=0,1,2,3} = A\cos(2\pi f t + n\frac{\pi}{4})$

Detecție pe baza corelației

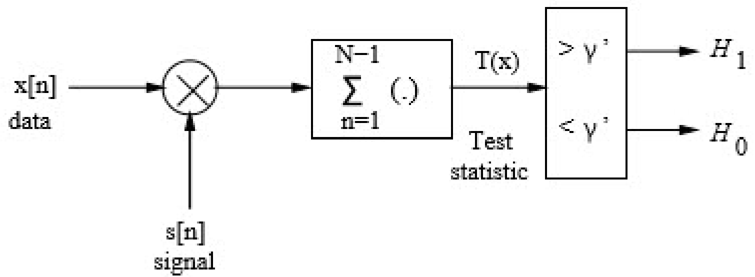


Figure 5: Detecția unui semnal folosind un corelator

[sursa: http://nptel.ac.in/courses/117103018/43]

Detecția a doua semnale

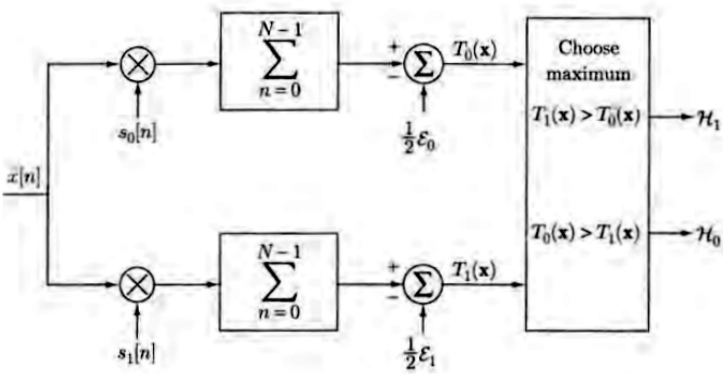


Figure 6: Decizie între două semnale diferite

[sursa: Fundamentals of Statistical Signal Processing, Steven Kay]

Filtru adaptat

- Cum se calculează corelația a două semnale $r[n]$ și $s[n]$ de lungime N?

$$< r, s >= \frac{1}{N} \sum r_i s(t_i)$$

- Fie $h[n]$ semnalul $h[n]$ **oglindit**
  - începând tot de la momentul 0, semnal cauzal

$$h[n] = s[N-1-n]$$

- Convoluția lui $r[n]$ cu $h[n]$ este

$$y[n] = \sum_k r[k]h[n-k] = \sum_k r[k]h[N-1-n+k]$$

- Rezultatul convoluției la finalul semnalului de intrare, $y[N-1]$ $(n = N-1)$, este chiar corelația
  - până la un factor de scalare $\frac{1}{N}$

$$y[N] = \sum_k r[k]s[k]$$

Filtru adaptat

- Pentru detecția unui semnal $s[n]$ se poate folosi un **filtru a cărui răspuns la impuls = oglindirea lui** $s[n]$, luându-se eșantionul de la finalul semnalului de intrare
  - se obține valoarea corelației

$$h[n] = s[N-1-n]$$

- **Filtru adaptat** = un filtru proiectat să aibă răspunsul la impuls egal cu oglindirea semnalului care se dorește a fi detectat (eng. "matched filter")
  - filtrul este *adaptat* semnalului dorit

Filtru adaptat

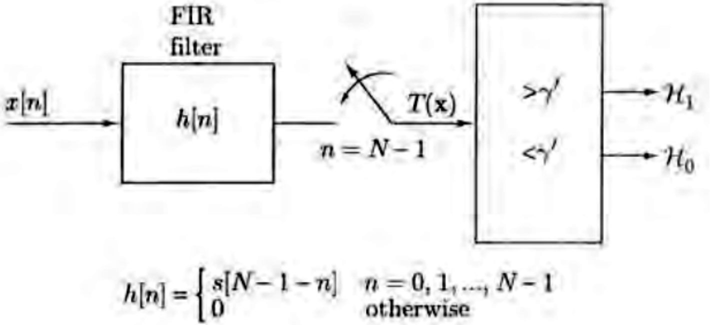


Figure 7: Detecție folosind un filtru adaptat

[sursa: Fundamentals of Statistical Signal Processing, Steven Kay]

II.5 Detecția unui semnal oarecare cu observare continuă

Observarea continuă a unui semnal oarecare

- Observare continuă = fără eșantionare, se folosește **întreg semnalul continuu**
  - similar cazului cu N eșantioane, dar cu $N \to \infty$
- Semnalul recepționat este $r(t)$
- Semnalul țintă este $s(t)$
- Presupunem doar zgomot Gaussian
- Cum are loc detecția?

Detecția semnalelor continue

- Se extinde cazul precedent cu N eșantioane la cazul unui semnal continuu, $N \to \infty$
- Interpretarea 1: media eșantioanelor
  - Nu mai este valabilă, întrucât $s(t)$ nu este constant

Interpretarea 2: geometric

- Interpretarea 2: geometric
- Fiecare semnal $r(t)$, $s(t)$ sau 0 reprezintă un punct într-un spațiu Euclidian infinit dimensional
- Distanța între două semnale este:

$$d(r,s) = \sqrt{\int (r(t) - s(t))^2 \, dt}$$

- Similar cu cazul N dimensional, dar cu integrală în loc de sumă
- Criteriul plauzibilității maxime:
  - se calculează distanța $d(r,s)$ între $r(t)$ și $s(t)$
  - se calculează distanța $d(r,0)$ între $r(t)$ și 0
  - se alege valoarea minimă

Interpretarea 3: corelația

- Corelația a două semnale continue $r(t)$ și $s(t)$ de lungime T

$$< \mathbf{r}, \mathbf{s} >= \frac{1}{T} \int_0^T r(t) \cdot s(t) dt$$

- Dacă corelația semnalului recepționat cu semnalul căutat $\mathbf{s(t_i)}$ este mai mare decât un prag L, se decide că semnalul este detectat.
  - Altfel, se decide că semnalul este absent
  - Corelația este o măsură a similarității a două semnale

Generalizări

- Detecția între **două semnale** $s_0(t)$ și $s_1(t)$
  - în zgomot Gaussian
- Interpretarea 2: geometric
  - se alege distanța Euclidiană minimă între punctul $\mathbf{r(t)}$ și punctele $\mathbf{s_0(t)}$ și $\mathbf{s_1(t)}$
    - folosind distanța dintre semnale definită mai sus
- Interpretarea 3: corelația
  - se calculează corelația lui $\mathbf{r(t)}$ cu $\mathbf{s_0(t)}$ și cu $\mathbf{s_1(t)}$.

Detecția între două semnale folosind corelația

- Criteriul plauz. maxime ($K = 1$):

$$< \mathbf{r}, \mathbf{s_1} > - \frac{< \mathbf{s_1}, \mathbf{s_1} >}{2} \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, \mathbf{s_0} > - \frac{< \mathbf{s_0}, \mathbf{s_0} >}{2}$$

- Dacă cele două semnale au energii egale: $\int s_1(t)^2 dt = \int s_0(t)^2 dt$, atunci $< \mathbf{s_1}, \mathbf{s_1} > = < \mathbf{s_0}, \mathbf{s_0} >$, așadar se alege **semnalul cel mai asemănător cu** $r(t)$:
  - Corelația este o măsură a similarității a două semnale

$$< \mathbf{r}, \mathbf{s_1} > \underset{H_0}{\overset{H_1}{\gtrless}} < \mathbf{r}, \mathbf{s_0} >$$

- Exemple
  - Modulația BPSK: $s_1 = A\cos(2\pi ft)$, $s_0 = -A\cos(2\pi ft)$
  - Modulația 4-PSK: $s_{n=0,1,2,3} = A\cos(2\pi ft + n\frac{\pi}{4})$

Filtru adaptat

- Corelația a două semnale se poate calcula cu un **filtru adaptat**
- **Filtru adaptat** = filtru proiectat să aibă răspunsul la impuls egal cu **oglindirea** semnalului căutat
  - filtrul este *adaptat* semnalului căutat
  - filtru continuu, cu răspuns la impuls continuu
- Pentru detecția unui semnal $s(t)$ se poate folosi un filtru adaptat, luând eșantionul de la ieșire în momentul final al semnalului de intrare
  - se obține chiar valoarea corelației