

W și Rao



O INTRODUCERE ÎN DETECȚIA SEMNALELOR

Problemele de detecție apar în: RADAR, telecomunicații, sonar, medicină, prelucrarea și recunoașterea imaginilor, seismologie, automatizări ș. a.

În toate aceste domenii trebuie să decidem dacă un eveniment este prezent sau nu și apoi să obținem mai multă informație despre el. Detectarea evenimentelor ține de "teoria detecției", cunoscută și sub denumirea de "testarea ipotezelor" sau "teoria deciziilor statistice"

În anumite cazuri trebuie să decidem între două ipoteze posibile, situație în care se vorbește despre ipoteze binare. Acesta este cazul detecției unei ținte prin RADAR. Alteori trebuie să decidem între mai multe ipoteze, situație în care se vorbește despre ipoteze multiple. Astfel de cazuri apar în telecomunicații, în clasificarea și recunoașterea formelor (pattern recognition) ș. a.

Trebuie să decidem între două sau mai multe ipoteze, având la dispoziție date observate, care, prin natura lor sunt afectate de zgomot. Datele observate fiind aleatoare, pentru detecție este necesară o abordare statistică. Admitem că, uneori, deciziile luate pot fi eronate. Urmărim ca probabilitatea deciziilor corecte să fie cât mai mare, iar rata erorilor cât mai mică.

Vom vedea că sunt posibile două clase de detecție, cea bazată pe abordarea Neyman-Pearson și cea bazată pe abordarea Bayes.

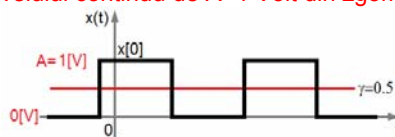
Pentru exemplificare, considerăm problema recepției unui semnal util, $s(t)$, din care se prelevează un eșantion, $x[0]$. Ca urmare a prezenței unui zgomot aditiv de putere cunoscută eșantionul poate fi

\mathcal{H}_0 : $x[0]=0+w[0]$; este prezent doar zgomotul

\mathcal{H}_1 : $x[0]=A+w[0]=1+w[0]$; este prezent semnalul util de 1 Volt, afectat de zgomot

$$w[0] \sim \mathcal{N}(0, \sigma^2)$$

Avem acces experimental doar la eșantionul $x[0]$. Cunoscând doar acest eșantion și caracteristicile statistice ale eșantionului de zgomot alb, gaussian, trebuie să spunem care dintre cele două ipoteze este adevărată. Se poate vorbi și despre detectarea nivelului continuu de $A=1$ Volt din zgomot



Deoarece zgomotul este de medie nulă, "bunul simț tehnic" ne spune că e bine să comparăm eșantionul $x[0]$ cu un prag $A/2=0.5$ Volt. Decidem că:

$x[0] \geq 0.5$ Volt, semnalul $A=1$ Volt este prezent sau \mathcal{H}_0 este ipoteza adevărată

$x[0] < 0.5$ Volt, semnalul $A=1$ Volt nu este prezent sau \mathcal{H}_1 este ipoteza adevărată

În cele două ipoteze, repartițiile eșantionului $x[0]$ sunt normale, cu aceeași dispersie, dar cu medii diferite, 0 și 1, așa cum rezultă din relațiile

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

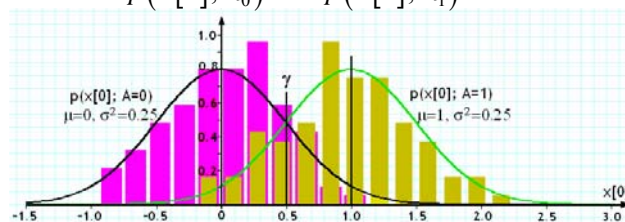
$$p(x[0]; \mathcal{H}_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x[0]-1)^2}{2\sigma^2}\right\}$$

A decide între cele două ipoteze, înseamnă să stabilim conform căreia dintre cele două repartiții a fost generat eșantionul $x[0]$. Putem privi decizia și altfel. Considerăm familia de densități de probabilitate, dependentă de parametrul A

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

Pentru $A=0$ obținem sau $A=1$, obținem, respectiv,

$$p(x[0]; \mathcal{H}_0) \text{ sau } p(x[0]; \mathcal{H}_1)$$



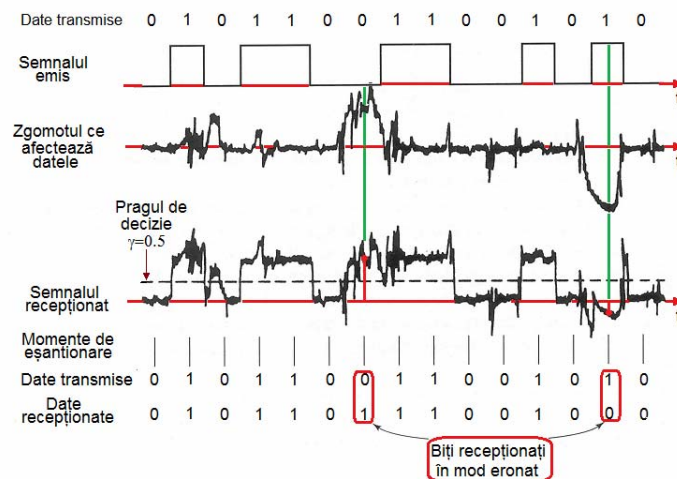
Problema de detectie devine una de testare a valorii parametrului A . Având dată observația $x[0]$, trebuie să testăm dacă $A=0$, sau $A=1$. Testul se descrie prin

$$\mathcal{H}_0 : A=0$$

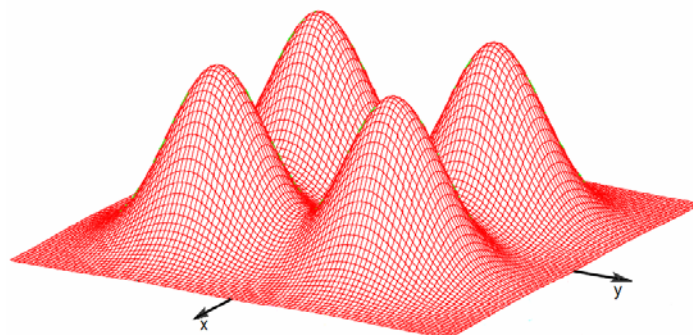
$$\mathcal{H}_1 : A=1$$

Am definit astfel un test al parametrului din densitatea de probabilitate (repartiție)

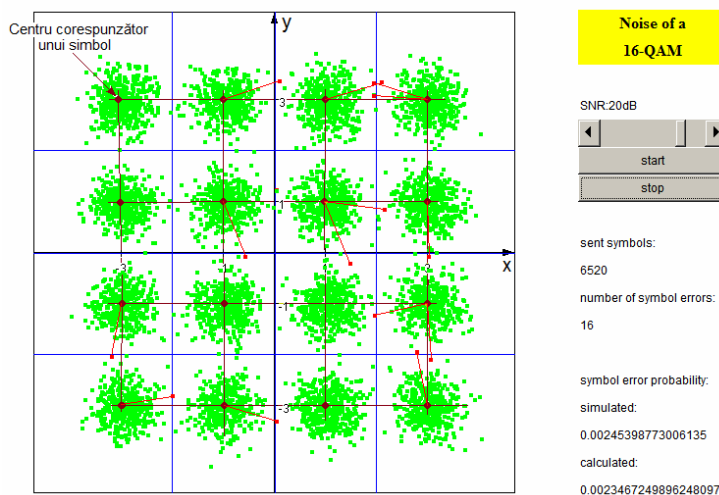
Pentru o mai bună înțelegere a detecției binare, exemplificăm cu un flux de date binare. la recepția căruia apar două erori. ca urmare a zgomotului



În unele cazuri, în comunicațiile de date, se transmit mai mult de două simboluri. Pot fi 4 sau mai multe. Reprezentarea simbolurilor se poate face, în funcție de tipul de modulație digitală, într-un plan. Simbolurile se definesc printr-un punct (x,y) din plan. Zgomotul va determina apariția unor suprafețe de densitate de probabilitate, așa cum se arată în figură. Datele recepționate sunt "proiectate" pe un sistem de două semnale ortogonale, ce constituie axele figurii. Proiecția se face prin corelare.



Un sistem cu PAM pe două axe ortogonale este cel din figură. Detecția se poate face fie prin proiecții, fie prin calculul distanței față de punctele ce reprezintă cele 16 simboluri. Zgomotul determină "împrăștierea" punctelor ce reprezintă semnalul recepționat (datele) în jurul punctelor de centru "de masă" corespunzătoare simbolurilor. Se remarcă o quasi simetrie circulară a dispunerii punctelor de date.



Abordarea "clasică" în detecție presupune că parametrii sunt neprecizați dar determiniști, în timp ce abordarea bayesiană consideră ca parametrii iau valori, dintre cele posibile, dar în mod aleator

Dacă putem asigna probabilități apriorice celor două ipoteze, vom recurge la cele două densități de repartiție condiționată

$$p(x[0]|\mathcal{H}_0) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x[0])^2}{2\sigma^2}\right\}; \quad \text{Trebuie cunoscută probabilitatea apriorică } P\{\mathcal{H}_0\}$$

$$p(x[0]|\mathcal{H}_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x[0]-A)^2}{2\sigma^2}\right\}; \quad \text{Trebuie cunoscută probabilitatea apriorică } P\{\mathcal{H}_1\}$$

Trebuie deci să facem o distincție netă între notațiile

$$p(x[0]; \mathcal{H}_0) \quad \text{și} \quad p(x[0]|\mathcal{H}_0)$$

Comportări asimptotice

În unele aplicații trebuie să detectăm prezența unor semnale "slabe", ce au un raport semnal/zgomot (SNR) de valoare redusă. De multe ori succesul în astfel de cazuri depinde de lungimea N a secvenței de semnal de date, $x[n]$, disponibile. Pentru a exemplifica comportamentul asimptotic în detecție, vom considera că dispunem nu de unul, ci de mai multe eșantioane de date

$$\{x[0], x[1], \dots, x[N-1]\}$$

În cele două ipoteze datele $x[n]$ sunt

$$w[n] \sim \mathcal{N}(0, \sigma^2) \quad n=0, 1, \dots, N-1$$

Zgomotul ce afectează datele, este un zgomot alb, gaussian, deci cu eșantioane IID

$$\mathcal{H}_0 : x[n]=w[n] \quad n=0, 1, \dots, N-1$$

$$\mathcal{H}_1 : x[n]=A+w[n] \quad n=0, 1, \dots, N-1$$

O abordare "bună" pare a fi medierea eșantioanelor, ceea ce reduce din dispersia inițială a datelor, și compararea statisticii rezultate, T , cu un prag, γ . Vom decide că \mathcal{H}_1 este

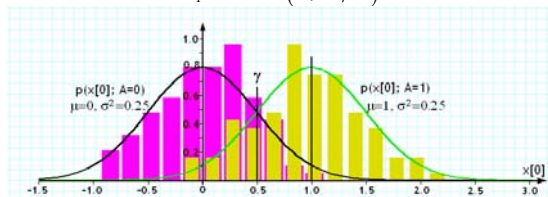
ipoteza adevărată, dacă

$$T = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \geq \gamma$$

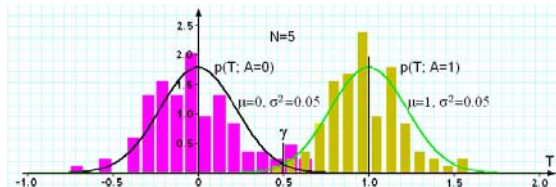
Se pare că e "bine" să luăm pentru pragul γ valoarea de $A/2$. Pentru $N=1$ și $\gamma=1/2$, regăsim un exemplu deja discutat de noi. Variabila aleatoare T are, în cele două ipoteze, repartițiile

$$\mathcal{H}_0 : T \sim N(0, \sigma^2/N)$$

$$\mathcal{H}_1 : T \sim N(A, \sigma^2/N)$$

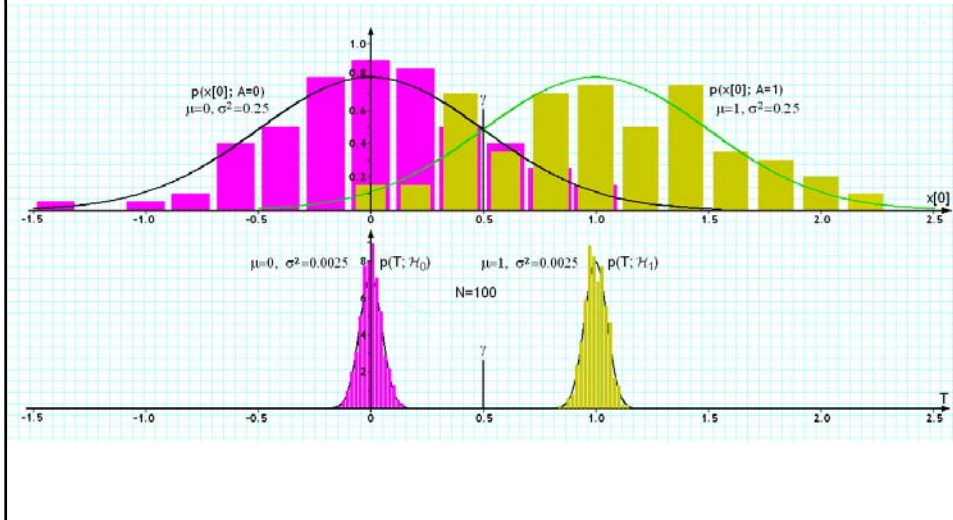


În figură se prezintă repartiția datelor în cele două ipoteze, la o dispersie a zgomotului de 0.25. După cum se observă, există multe cazuri în care, deși suntem în ipoteza \mathcal{H}_0 , există multe eșantioane de semnal ce depășesc pragul, ceea ce conduce la decizia, eronată, că ipoteza \mathcal{H}_1 este adevărată; multe date ce apar în ipoteza \mathcal{H}_1 se situează sub prag, ceea ce conduce la decizia, eronată, că ipoteza \mathcal{H}_0 este adevărată



Dacă decizia se ia pe baza statisticii T , calculată pentru doar $N=5$ eșantioane dispersia scade de la 0.25 (dispersia datelor $x[n]$), la valoarea 0.05 (dispersia statisticii T). După cum se vede din figură, scade numărul deciziilor incorecte (al erorilor de decizie). Crește, în schimb, timpul necesar luării unei decizii!

Dacă creștem și mai mult numărul eșantioanelor ce se mediază pentru a obține statistica T , numărul deciziilor incorecte scade. În figură se arată cazul datelor cu aceleași proprietăți statistice, în cele două ipoteze ca și în exemplele anterioare, dar pentru care se prelucrează $N=100$ eșantioane. Dispersia statisticii T scade la 0.0025 și, după cum se vede, practic nu mai apar erori. Abaterea standard a statisticii T este de doar 0.05 , așa că $3\sigma=0.15 < 0.5$, ceea ce înseamnă că erorile sunt extrem de rare.



Intuim că dacă mediile celor două repartiții diferă mult între ele, numărul detecțiilor corecte este mare. Dar "întinderea" curbelor de repartiție depinde de σ , abaterea standard, care este aceeași în ambele ipoteze. Pe măsură ce valoarea lui σ este mai redusă, scade numărul deciziilor incorecte. Raportul

$$d = \frac{\mu_1 - \mu_0}{\sigma}$$

numit și coeficient de deflexie, este un indicator, pentru cazul de față, al "detectabilității".

Uneori se folosește pătratul expresiei, dar i se spune tot coeficient de deflexie

$$d^2 = \frac{(E\{T; \mathcal{H}_1\} - E\{T; \mathcal{H}_0\})^2}{Disp\{T; \mathcal{H}_0\}}$$

Deoarece

$$Disp\{T; \mathcal{H}_1\} = Disp\{T; \mathcal{H}_0\} = \sigma^2 / N$$

și

$$E\{T; \mathcal{H}_1\} = A \quad \text{și} \quad E\{T; \mathcal{H}_0\} = 0$$

rezultă, pentru coeficientul de deflexie

$$d^2 = \frac{A^2}{\sigma^2 / N} = \frac{NA^2}{\sigma^2} = N \cdot SNR$$

Performanțele de detecție se pot îmbunătăți fie crescând valoarea SNR, fie crescând numărul de eșantioane prelucrate (evident, pot fi luate și ambele măsuri)

Ca o concluzie ce rezultă din acest exemplu, pare să fie adecvată o analiză a comportamentului asimptotic, atunci când N crește foarte mult. Comportarea asimptotică permite determinarea mai ușoară a unor detectoare și permite analiza performanțelor lor statistice. Ca să dăm un exemplu, dacă $w[n]$ sunt eșantioane IID dar negausiene, atunci nici statistica T nu va avea o repartiție gaussiană. Cu toate acestea, pentru N foarte mare, putem invoca teorema limită centrală pentru a utiliza, totuși, o repartiție gaussiană pentru T . Pentru a evalua performanțele detectorului, va fi necesar, într-un astfel de caz, să evaluăm doar primele două momente ale statisticii T .

Teorema limită centrală

Dacă variabilele aleatoare

$$\{x[0], x[1], \dots, x[N-1]\}$$

sunt statistic independente și identic distribuite (IID), cu media și dispersia

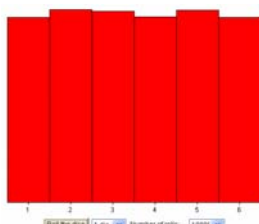
$$\mu \text{ și } \sigma^2$$

dar cu o repartiție oarecare, atunci variabila aleatoare y tinde spre o repartiție gaussiană, de medie nulă și dispersia unitară, adică

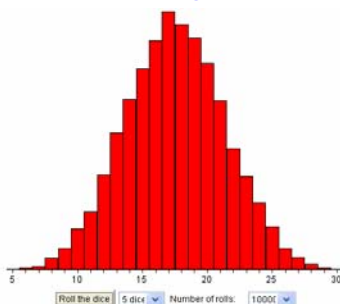
$$y = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \frac{x[n] - \mu}{\sigma} \xrightarrow{N \rightarrow \infty} \mathcal{N}(0, 1)$$

De fapt, teorema are o formă mai generală, în care se arată că suma unui număr foarte mare de variabile aleatoare chiar cu repartiții diferite, tinde spre o repartiție normală, cu condiția ca nici unul din termenii sumei să nu fie predominant

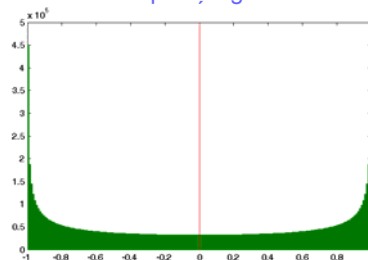
Pentru a ilustra valoarea teoremei, vom considera un experiment de aruncare a unui zar, de 10000 de ori. Se reprezintă frecvența de apariție a numerelor, rezultând o histogramă uniformă, foarte depărtată de una gaussiană, așa cum se vede din figură



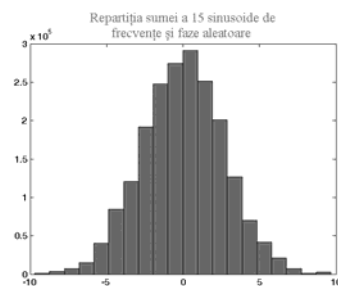
Dacă se aruncă, deodată 5 zaruri, se face suma numerelor generate în aruncare și se repetă experimentul de 10000 de ori, frecvența relativă de apariție a sumelor, de la 5 la 30, relevă o tendință de comportament gaussian, cum se și vede din figură



În figură se arată histograma unui semnal sinusoidal. Ea nu seamănă, nici pe departe, cu o repartiție gaussiană



Histograma unei sume de 15 sinusoides, nici una predominantă, având frecvențe ce nu sunt în rapoarte armonice, seamănă însă cu o repartiție gaussiană. Este o dovadă a aplicabilității teoremei limită centrală



TEORIA DECIZIEI STATISTICE

Ne vom ocupa de cazul în care cunoaștem, în întregime, densitățile de probabilitate (repartiție) în toate ipotezele posibile. Acesta este cel mai simplu caz. Prima abordare posibilă este una de tip "clasic" ce se bazează pe teorema Neyman-Pearson. Ea se aplică în RADAR, SONAR, căutări în baze de date, ș.a., adică în acele cazuri în care nu cunoaștem probabilitățile apriorice ale ipotezelor implicate



Jerzy Neyman 1894-1981



Egon Sharpe Pearson 1895-1980



Thomas Bayes, 1701-1761

În domeniul telecomunicațiilor, cel al recunoașterii formelor, și în oricare problemă de luare a deciziilor în care se cunosc probabilitățile apriorice ale ipotezelor, se recurge la o metodică de lucru bazată pe așa numita strategie bayesiană

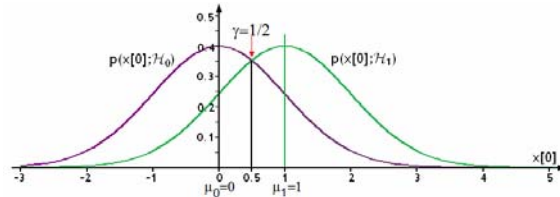
Teorema Neyman-Pearson

Începem cu un exemplu de testare a ipotezelor. Vom presupune că observăm o realizare a unei variabile aleatoare, care poate fi repartizată normal, $\mathcal{N}(0, 0)$, sau $\mathcal{N}(0, 1)$. Trebuie să decidem dacă $\mu=0$ sau $\mu=1$, pe baza unui singur eşanţion de date, $x[0]$. Fiecare dintre cele două valori ale mediei poate fi privită ca şi o ipoteză. Trebuie deci să decidem care dintre ipotezele

$$\mathcal{H}_0 : \mu=0$$

$$\mathcal{H}_1 : \mu=1$$

este adevărată, având la dispoziţie un singur eşanţion de date, $x[0]$. Ca terminologie, \mathcal{H}_0 se numeşte "ipoteza nulă" iar \mathcal{H}_1 se numeşte "ipoteza alternativă". Problema de decizie pe care o discutăm, se numeşte şi problema de "test al ipotezei binare".



Densităţile de probabilitate ale datelor, în cele două ipoteze, sunt arătate în figură. Dacă trebuie să decidem, pe baza unui singur eşanţion de date, este greu de spus care dintre cele două ipoteze l-a generat, sau cătei densităţi de repartiţie aparţine acel eşanţion.

Dacă pragul de decizie se ia $\gamma=1/2$, atunci decizia se ia conform cu regula:

Dacă: $x[0] \geq 1/2$ atunci $p(x[0]; \mathcal{H}_1) \geq p(x[0]; \mathcal{H}_0) \Rightarrow \mathcal{H}_1$ este ipoteza adevărată

Dacă: $x[0] < 1/2$ atunci $p(x[0]; \mathcal{H}_0) > p(x[0]; \mathcal{H}_1) \Rightarrow \mathcal{H}_0$ este ipoteza adevărată

Notăm

$P\{\mathcal{H}_i; \mathcal{H}_j\}$ probabilitatea de a decide că ipoteza \mathcal{H}_i este adevărată, atunci când, de fapt, ipoteza \mathcal{H}_j este cea adevărată

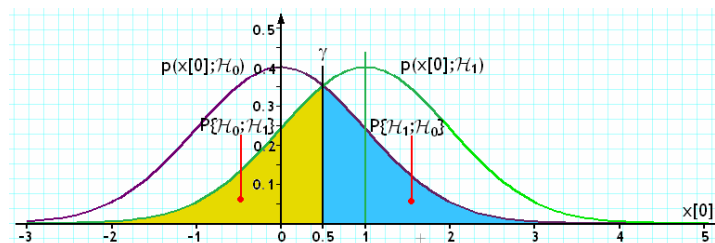
şi rezultă că putem avea două tipuri de erori

$$P\{\mathcal{H}_1; \mathcal{H}_0\} = P\{x[0] \geq 1/2; \mathcal{H}_0\} \text{ si } P\{\mathcal{H}_0; \mathcal{H}_1\} = P\{x[0] < 1/2; \mathcal{H}_1\}$$

Dacă declarăm

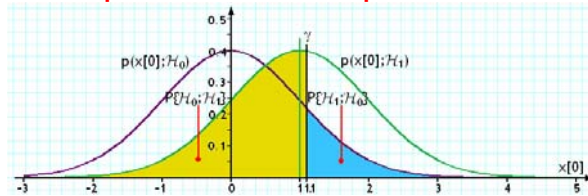
ipoteza \mathcal{H}_1 ca adevărată, atunci când ipoteza \mathcal{H}_0 este cea adevărată, comitem o "eroare de tipul I"

ipoteza \mathcal{H}_0 ca adevărată, atunci când ipoteza \mathcal{H}_1 este cea adevărată, comitem o "eroare de tipul II"



Probabilităţile celor două tipuri de erori sunt numeric egale cu ariile indicate în figură.

Dacă, așa cum se arată în figură, creștem valoarea pragului de decizie γ , scade probabilitatea de a comite erori (de decizie) de tipul I, dar crește în schimb probabilitatea de a comite erori de tipul II. **Este imposibil să reducem, simultan, probabilitățile de producere a ambelor tipuri de erori!!**



Adoptând o terminologie din radar vorbim despre

$P_{FA} = P\{\mathcal{H}_1; \mathcal{H}_0\}$, probabilitatea "alarmei false" (false alarm probability)

$P_M = P\{\mathcal{H}_0; \mathcal{H}_1\}$, probabilitatea de "pierdere a țintei" (miss probability)

$P_D = P\{\mathcal{H}_1; \mathcal{H}_1\} = 1 - P_M$, probabilitatea de "detectare corectă a țintei"

O abordare posibilă pentru determinarea unor detectoare optimale, conform abordării Neyman-Person, (NP) constă în a menține probabilitatea alarmei false (probabilitatea de ripostă neprovocată) la o valoare impusă α , evident mică, și a minimiza apoi tipul II de eroare, ceea ce înseamnă maximizarea probabilității de detecție (corectă) a țintei.

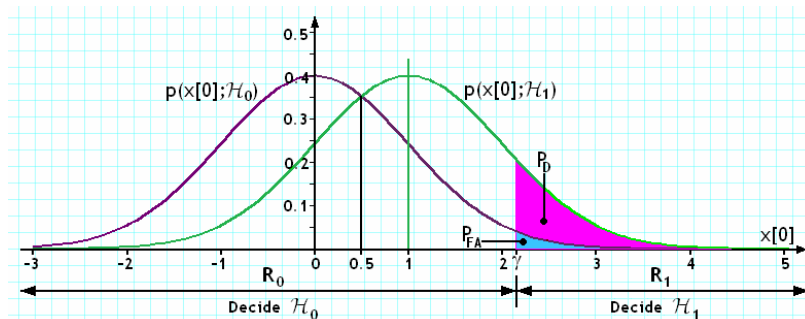
Impunem

$$P_{FA} = P\{\mathcal{H}_1; \mathcal{H}_0\} = \alpha$$

și maximizăm

$$P_D = P\{\mathcal{H}_1; \mathcal{H}_1\} = 1 - P_M$$

Probabilitățile alarmei false și de detecție corectă sunt numeric egale cu ariile arătate în figură. Se observă că dacă se crește valoarea pragului de decizie scade probabilitatea alarmei false dar, în același timp, scade și probabilitatea detecției corecte. Sunt indicate și regiunile de decizie, în care dacă intră eșantionul de semnal, $x[0]$, luăm decizia arătată.



Revenind la exemplu, căutăm valoarea pragului, γ , care satisface condițiile expuse.

Avem

$$P_{FA} = P\{\mathcal{H}_1; \mathcal{H}_0\} = p(x[0] \geq \gamma; \mathcal{H}_0) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt = Q(\gamma)$$

în care funcția $Q(x)$ are relația de definiție și proprietățile

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt; \quad Q(0) = 0.5; \quad Q(-x) = 1 - Q(x)$$

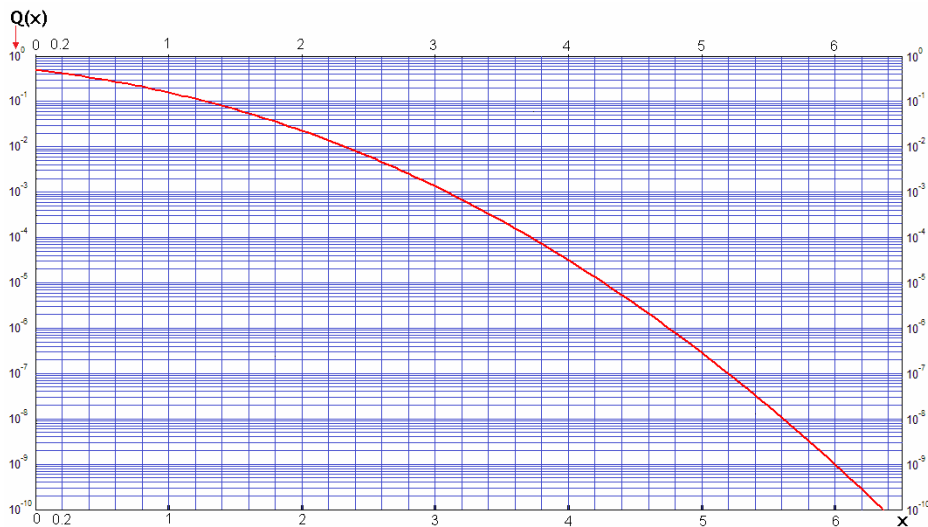
Se mai utilizează funcția lui Laplace, $\Phi(x)$ și funcția complementară de eroare, $\text{erfc}(x)$

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp(-t^2/2) dt; \quad \text{erfc}(x) = \int_x^{\infty} \frac{2}{\sqrt{\pi}} \exp(-t^2) dt$$

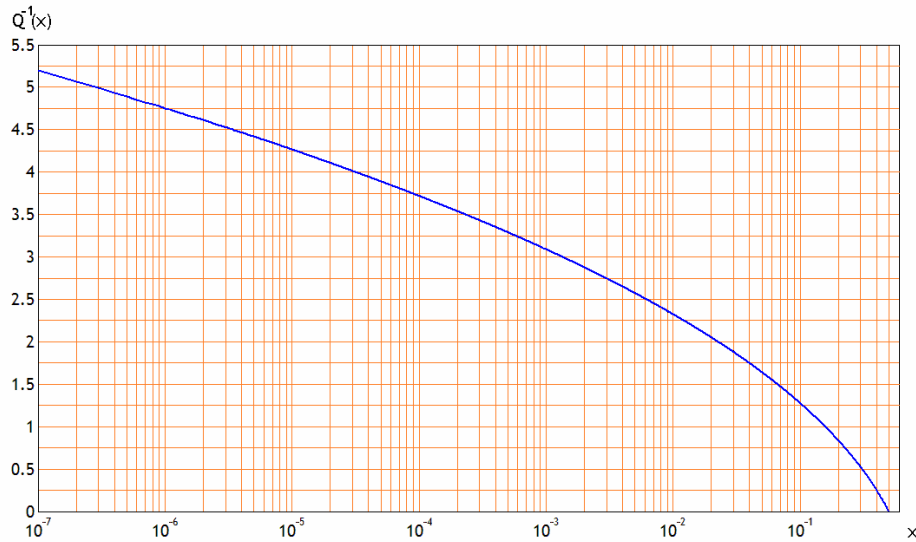
Se pot verifica relațiile de legătură

$$Q(x) = 1 - \Phi(x); \quad Q(x) = \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right); \quad \text{erfc}(x) = 2Q(\sqrt{2}x)$$

În figură se prezintă funcția $Q(x)$. Se vede că este o funcție strict descrescătoare și, prin urmare, este inversabilă. În MATLAB funcția $Q(x)$ se calculează cu `qfunc(x)`



Inversa funcției Q este arătată în figură. Ea se poate calcula în MATLAB cu `qfuncinv(x)`



Reluăm exemplul discutat. Am înpus o probabilitate a alarmei false de valoare α , prin

$$P_{FA} = Q(\gamma) = \alpha$$

ecuație din care se determină pragul testului, γ

$$\gamma = Q^{-1}(\alpha)$$

Dacă, spre exemplu, considerăm că este admisibilă o probabilitate a alarmei false de $\alpha=0.001$ rezultă pentru pragul testului valoarea $\gamma=qfuncinv(0.001)=3.0902$. Pentru a determina probabilitatea de detecție corectă, avem

$$\begin{aligned} P_D &= P\{\mathcal{H}_1; \mathcal{H}_1\} = P\{x[0] \geq 3.09; \mathcal{H}_1\} = \int_{3.09}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(t-1)^2}{2}\right) dt \\ &= \int_{2.09}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du = Q(2.09) \cong 1.83 \cdot 10^{-2} \end{aligned}$$

Nu prea detectăm corect prezența semnalului dacă probabilitatea alarmei false este de 0.001. Valoarea probabilității de detecție se mai numește și "puterea testului".

Ne putem pune întrebarea dacă valoarea de mai sus pentru probabilitatea de detecție este tot ce se poate obține?

Dacă detectorul dispune de N date observate,

$$\{x[0], x[1], \dots, x[N-1]\}$$

situația se poate îmbunătăți radical

Mulțimea vectorilor \mathbf{x} de date, N dimensional, se separă în două submulțimi (regiuni), corespunzătoare deciziilor ce trebuie să fie luate. Se definesc mulțimile

$$\mathbb{R}_1 = \{\mathbf{x} : \text{se decide } \mathcal{H}_1, \text{ adică se rejectează } \mathcal{H}_0\} \text{ (se numeste "regiune critică")}$$

$$\mathbb{R}_0 = \{\mathbf{x} : \text{se decide } \mathcal{H}_0, \text{ adică se rejectează } \mathcal{H}_1\}$$

$$\mathbb{R}_1 \cup \mathbb{R}_0 = \mathbb{R}_N \text{ si } \mathbb{R}_1 \cap \mathbb{R}_0 = \emptyset$$

Într-un caz general constrângerea privind probabilitatea alarmei false se exprimă cu

$$P_{FA} = \int_{\mathbb{R}_1} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha$$

Valoarea α se numește și "nivelul de semnificație al testului"

Există mai multe grupuri de vectori de date, \mathbf{x} , adică mai multe regiuni critice pentru care relația de constrângere este satisfăcută. Dintre toate acestea trebuie aleasă cea regiune critică, care maximizează probabilitatea de detecție corectă

$$P_D = \int_{\mathbb{R}_1} p(\mathbf{x}; \mathcal{H}_1) d\mathbf{x}$$

Teorema Neyman-Pearson (NP) ne arată cum trebuie aleasă regiunea critică, dacă se cunosc densitățile de probabilitate ale vectorilor de date \mathbf{x} în cele două ipoteze

$$p(\mathbf{x}; \mathcal{H}_0) \text{ si } p(\mathbf{x}; \mathcal{H}_1)$$

și se impune nivelul de semnificație al testului, α

Teorema Neyman-Pearson. Pentru a maximiza probabilitatea de detecție, la o probabilitate impusă a alarmei false, α , se decide \mathcal{H}_1 dacă

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma; \quad \mathbf{x} \text{ fixat, dar oarecare}$$

Pragul testului, γ , este soluția ecuației

$$P_{FA} = \int_I p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha; \quad \text{unde } I = \{\mathbf{x} : L(\mathbf{x}) > \gamma\}$$

$L(\mathbf{x})$ se numește raport de plauzibilitate și este raportul a două plauzibilități, deoarece se consideră că vectorul \mathbf{x} este fixat, existând o realizare a datelor, pentru care se efectuează testul.

Pentru demonstrație aplicăm metoda multiplicatorilor lui Lagrange, pentru a maximiza probabilitatea de detecție, cu constrângerea impusă de valoarea admisă pentru probabilitatea alarmei false. Se construiește lagrangeanul F

$$F = P_D + \lambda(P_{FA} - \alpha) = \int_{\mathbb{R}_1} p(\mathbf{x}; \mathcal{H}_1) d\mathbf{x} + \lambda \left[\int_{\mathbb{R}_1} p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} - \alpha \right]$$

sau

$$F = \int_{\mathbb{R}_1} [p(\mathbf{x}; \mathcal{H}_1) + \lambda p(\mathbf{x}; \mathcal{H}_0)] d\mathbf{x} - \lambda \alpha$$

Pentru a maximiza F , vom include în regiunea critică doar acei vectori \mathbf{x} pentru care integrandul este pozitiv; integranzii pozitivi cresc valoarea integralei, în timp ce integranzii negativi scad valoarea integralei. Prin urmare:

$$\mathbf{x} \in \mathbb{R}_1 \iff p(\mathbf{x}; \mathcal{H}_1) + \lambda p(\mathbf{x}; \mathcal{H}_0) > 0$$

Deoarece densitățile de probabilitate sunt pozitive inegalitatea de mai sus se pune sub forma

$$\mathbf{x} \in \mathbb{R}_1 \iff \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > -\lambda$$

sau, cu notația $\gamma = -\lambda > 0$

$$\mathbf{x} \in \mathbb{R}_1 \iff \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma$$

Pragul testului satisface ecuația de constrângere deja amintită

$$P_{FA} = \int_I p(\mathbf{x}; \mathcal{H}_0) d\mathbf{x} = \alpha; \quad \text{unde } I = \{\mathbf{x} : L(\mathbf{x}) > \gamma\}$$

Vom relua exemplul introductiv. Raportul de plauzibilitate cerut de teorema NP se compară cu pragul testului, γ și

$$\text{decidem } \mathcal{H}_1 \text{ dacă: } L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} = \frac{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x[0]-1)^2}{2}\right\}}{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(x[0])^2}{2}\right\}} > \gamma;$$

După simplificări, condiția în care decidem că ipoteza \mathcal{H}_1 este adevărată, devine

$$\exp\{x[0] - 1/2\} > \gamma$$

sau, după logaritmare (care este posibilă, deoarece $\gamma > 0$) și rearanjare

$$x[0] > \ln \gamma + 1/2 = \gamma'$$

Pragul testului se determină din ecuația de constrângere

$$P_{FA} = P\{\exp\{x[0] - 1/2\} > \gamma; \mathcal{H}_0\} = \alpha = 0.001$$

sau din echivalența ei

$$P_{FA} = P\{x[0] > \gamma'; \mathcal{H}_0\} = \alpha = 0.001$$

Repartiția eșantionului $x[0]$ în ipoteza nulă este

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

asa că ecuația de constrângere a valorii probabilității alarmei false devine

$$P_{FA} = \int_{\gamma'}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-(x[0])^2/2\right] dx[0] = Q(\gamma') = 0.001$$

Se determină valoarea pragului, $\gamma' = 3.09$. Probabilitatea detecției corecte este cea anterior determinată, de 0.0183

În cele ce urmează vom utiliza frecvent relația

dacă: $x \sim \mathcal{N}(\mu, \sigma^2)$, atunci: $P\{x > \gamma\} = Q\left(\frac{\gamma - \mu}{\sigma}\right)$

Demonstrarea ei este un simplu exercițiu. Expresia densității de probabilitate a unei variabile aleatoare normale este

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

Se calculează probabilitatea ca variabila normală să depășească pragul γ' cu relația

$$P\{x > \gamma'\} = \int_{\gamma'}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \stackrel{u=(x-\mu)/\sigma}{=} \int_{(\gamma'-\mu)/\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} du = Q\left(\frac{\gamma'-\mu}{\sigma}\right)$$

Vom aborda un exemplu în care analizăm detecția unei componente continue, din zgomot alb, gaussian. Definim testul de ipoteze sub forma

$$\mathcal{H}_0: x[n] = w[n] \quad n = 0, 1, \dots, N-1$$

$$\mathcal{H}_1: x[n] = A + w[n] \quad n = 0, 1, \dots, N-1 \text{ și } A > 0$$

$$\mathcal{H}_0 \text{ și } \mathcal{H}_1: w[n] \sim \mathcal{N}(0, \sigma^2) \quad \mathbf{C}_w = \sigma^2 \mathbf{I}_u$$

Vectorul de date este compus din cele N eșantioane de semnal măsurabil

$$\mathbf{x} = [x[0] \ x[1] \ \dots \ x[N-1]]^T$$

Cele două ipoteze sunt caracterizate de medii diferite ale datelor

$$\mathcal{H}_0: \boldsymbol{\mu} = [0 \ 0 \ \dots \ 0]^T = \mathbf{0}$$

$$\mathcal{H}_1: \boldsymbol{\mu} = [A \ A \ \dots \ A]^T = A\mathbf{1}$$

Matricea de covarianță a zgomotului și deci și a datelor, în acest exemplu, este aceeași, în ambele ipoteze. Rezultă că repartițiile vectorilor de date, în cele două ipoteze sunt

$$\mathcal{H}_0: \mathbf{x} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_u)$$

$$\mathcal{H}_1: \mathbf{x} \sim \mathcal{N}(A\mathbf{1}, \sigma^2 \mathbf{I}_u)$$

Testul de ipoteze este echivalent cu testul de parametri

$$\mathcal{H}_0: \boldsymbol{\mu} = \mathbf{0}$$

$$\mathcal{H}_1: \boldsymbol{\mu} = A\mathbf{1}$$

Detectorul va decide că H_1 este ipoteza adevărată (semnalul util e prezent) dacă raportul de plauzibilitate depășește pragul γ , adică

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} = \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2\right\}}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n])^2\right\}} > \gamma$$

După simplificare rezultă

$$\exp\left\{-\frac{1}{2\sigma^2} \left[-2A \sum_{n=0}^{N-1} x[n] + \sum_{n=0}^{N-1} A^2\right]\right\} > \gamma$$

Se logaritmează

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

și se obține

$$\frac{A}{\sigma^2} \sum_{n=0}^{N-1} x[n] > \ln \gamma + \frac{NA^2}{2\sigma^2} \quad \text{se împarte cu } \frac{NA}{\sigma^2} > 0$$

Obținem, în final,

$$\bar{x} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] > \frac{\sigma^2}{NA} \ln \gamma + \frac{A}{2} = \gamma'$$

În acest exemplu, detectorul NP compară media eșantion cu un prag. Se știe deja că media eșantion este o estimare a componentei continue, A în cazul de față. Dacă media eșantion are o valoare pozitivă, suficient de mare, probabil că este prezentă A>0.

Ajustând pragul γ' se controlează probabilitățile alarmei false și de detecție corectă. Știm că statistica

$$T(\mathbf{x}) = \bar{x}$$

are o repartiție gaussiană. Drept urmare, vom determina mediile și dispersiile ei, în cele două ipoteze. Avem

$$E\{T(\mathbf{x}); \mathcal{H}_0\} = E\left\{\frac{1}{N} \sum_{n=0}^{N-1} w[n]\right\} = \frac{1}{N} \sum_{n=0}^{N-1} E\{w[n]\} = 0$$

$$Disp\{T(\mathbf{x}); \mathcal{H}_0\} = Disp\left\{\frac{1}{N} \sum_{n=0}^{N-1} w[n]\right\} = \frac{1}{N^2} \sum_{n=0}^{N-1} Disp\{w[n]\} = \frac{\sigma^2}{N}$$

$$E\{T(\mathbf{x}); \mathcal{H}_1\} = E\left\{\frac{1}{N} \sum_{n=0}^{N-1} (A + w[n])\right\} = \frac{1}{N} \sum_{n=0}^{N-1} E\{A + w[n]\} = A$$

$$Disp\{T(\mathbf{x}); \mathcal{H}_1\} = Disp\{T(\mathbf{x}); \mathcal{H}_0\} = Disp\left\{\frac{1}{N} \sum_{n=0}^{N-1} w[n]\right\} = \frac{\sigma^2}{N}$$

În consecință, repartiția statisticii T, în cele două ipoteze, este

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

Aplicând relația roșie putem determina probabilitățile alarmei false și detecției corecte

$$P_{FA} = P\{T(\mathbf{x}) > \gamma'; \mathcal{H}_0\} = Q\left(\frac{\gamma' - 0}{\sqrt{\sigma^2/N}}\right) = \alpha$$

$$P_D = P\{T(\mathbf{x}) > \gamma'; \mathcal{H}_1\} = Q\left(\frac{\gamma' - A}{\sqrt{\sigma^2/N}}\right)$$

Din prima relație se determină valoarea pragului, cu expresia

$$\gamma' = \sqrt{\frac{\sigma^2}{N}} Q^{-1}(P_{FA})$$

Dacă substituim expresia pragului în relația de calcul a probabilității de detecție rezultă

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}}\right) = Q\left(Q^{-1}(P_{FA}) - \sqrt{ENR}\right); \quad ENR = \frac{NA^2}{\sigma^2}$$

Cu ENR am notat raportul dintre energia semnalului și puterea zgomotului (energy to noise ratio). În literatura de specialitate, de cele mai multe ori, acest raport este numit SNR

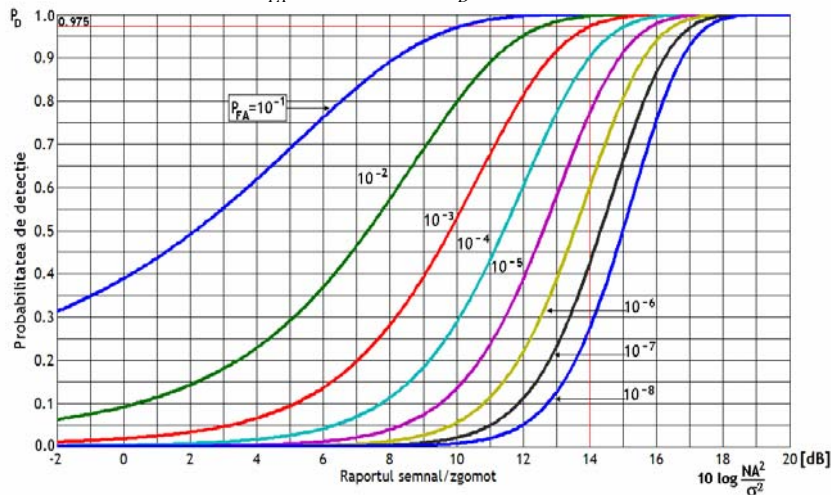
În figură sunt trasate curbele dependenței probabilității de detecție corectă, ca funcție de ENR, exprimat în dB, având ca și parametru de curbă probabilitatea alarmei false

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{ENR}\right)$$

Admițând $\alpha=0.001$, la o valoare ENR de 14 dB (un raport de aproximativ 25.12) se obține o probabilitate de detecție corectă de 0.975, adică 97.5%.

Pentru ENR=20dB (raport de 100)

$$P_{FA} = \alpha = 10^{-10} \Rightarrow P_D = 0.9999$$



Exemplul analizat este o ilustrare a unei probleme mai generale de testare a ipotezelor, numită și problema repartițiilor gaussiene cu valori medii diferite. În esență, se calculează valoarea unei statistici, T, și se compară cu un prag, γ' . Dacă T depășește pragul, se declară adevărată ipoteza H_1 , altfel, H_0 . În ambele ipoteze avem

$$T \sim \begin{cases} \mathcal{N}(\mu_0, \sigma^2) & \text{în ipoteza } \mathcal{H}_0 \\ \mathcal{N}(\mu_1, \sigma^2) & \text{în ipoteza } \mathcal{H}_1 \end{cases}; \mu_1 > \mu_0$$

În exemplul discutat, statistica T a fost media eșantion. În acest caz, mai general, probabilitatea alarmei false este

$$P_{FA} = P\{T > \gamma'; \mathcal{H}_0\} = Q\left(\frac{\gamma' - \mu_0}{\sigma}\right) = \alpha$$

ecuație din care rezultă valoarea pragului testului în această abordare

$$\frac{\gamma' - \mu_0}{\sigma} = Q^{-1}(P_{FA}) \Rightarrow \gamma' = \mu_0 + \sigma Q^{-1}(P_{FA})$$

Expresia valorii pragului testului se substituie în relația probabilității de detecție

$$P_D = P\{T > \gamma'; \mathcal{H}_1\} = Q\left(\frac{\gamma' - \mu_1}{\sigma}\right) = Q\left(Q^{-1}(P_{FA}) - \overbrace{\frac{\mu_1 - \mu_0}{\sigma}}^{>0}\right)$$

Definim coeficientul de deflexie cu relația

$$d^2 = \frac{(\mu_1 - \mu_0)^2}{\sigma^2}$$

care înlocuim în expresia probabilității de detecție dă

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{d^2}\right)$$

$Q(x)$ fiind o funcție descrescătoare, probabilitatea de detecție crește, odată cu creșterea coeficientului de deflexie.

