CAP. 10 - ELEMENTI DI TEORIA STATISTICA DELLA DECISIONE (Bibl. N° 5)

10-1. Introduzione

Il problema di ricevere un messaggio da un punto ad un altro, o quello di misurare un parametro mediante osservazione di un segnale, si riduce in definitiva ad un problema di decisione. Generalmente si domanda di procedere nella decisione in modo da ridurre al minimo, in media, i costi che gli errori, dovuti alla presenza del rumore, comportano.

Il problema può impostarsi come segue :

Chiamamo x lo stato vero di un certo fenomeno, sotto osservazione. Siano $x_1, x_2, \dots x_n$, i possibili stati, di probabilità a priori $P(x_1)$, $P(x_2), \dots P(x_n)$.

Diciamo y il risultato dell'osservazione del fenomeno, ovvero lo stato ricevuto.

Generalmente si osserva $y \neq x$ a causa del rumore.

Siano $y_1, y_2 \dots y_n$ i possibili stati osservati. Per un dato x_i esiste la probabilità $P(y_j / x_i) = P_{ij}$ che si osservi lo stato y_j .

Si avrà cioè la seguente matrice di (condizionali) probabilità:

$$\begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_m \end{array} \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & \dots & y_m \end{array} \\ \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \vdots & \vdots & & \vdots \\ P_{n1} & P_{n2} & \dots & P_{nm} \end{bmatrix} \end{array} \qquad (10\text{-}1)$$

Indichiamo ora con a_K la decisione che x_K è lo stato del fenomeno. In realtà lo stato vero potrebbe essere x_i , nel qual caso a_K è in errore.

Indichiamo con C_{iK} il costo di decidere $x = x_K$, quando in realtà è $x = x_i$ lo stato del fenomeno.

Potremo rappresentare i vari costi nella seguente matrice :

$$\begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{array}{c} \begin{array}{cccc} a_1 & a_2 & \dots & a_m \end{array} \\ \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1m} \\ C_{21} & C_{22} & \dots & C_{2m} \\ \vdots & \vdots & & \vdots \\ C_{n1} & C_{n2} & & C_{nm} \end{bmatrix} \end{array} \qquad (10\text{-}2)$$

A questo punto il problema è genericamente impostato. Ovvero, per descrivere la situazione dal punto di vista statistico è necessario conoscere :

- probabilità a priori degli stati possibili del fenomeno,
- probabilità condizionali dei possibili stati osservati, ed
- i costi di errore.

.../...

Vediamo ora di procedere mediante un esempio.

Consideriamo un canale di trasmissione binario, e sia :

p_0 la probabilità che 0 sia ricevuto come 1 e
p_1 quella di ricevere 1 come 0.

Siano :

q_0 e q_1 le rispettive probabilità di ricezione corretta.

Siano :

$x_1 = 11$ ed $x_2 = 00$ i due possibili stati del messaggio.

Se la ricezione è corretta assumiamo che non vi sia costo di errore. In effetti potrebbe essere necessario in alcuni casi considerare un costo negativo, o premio. Diciamo C_f il costo del ricevere 11 quando lo stato vero è 00 , e C_m quello di ricevere 00 quando lo stato è 11 .

Chiamamo $P = P(x_1)$ e $Q = P(x_2)$.

La matrice delle probabilità di osservazione è la seguente :

$$\begin{array}{c} \\ x = \\ 11 \\ 00 \end{array} \begin{array}{c} \begin{array}{cccc} y = & 11 & 10 & 01 & 00 \end{array} \\ \begin{bmatrix} q_1^2 & q_1 p_1 & q_1 p_1 & p_1^2 \\ p_0^2 & q_0 p_0 & q_0 p_0 & q_0^2 \end{bmatrix} \end{array} \qquad (10\text{–}3)$$

E quella dei costi è :

$$\begin{array}{c} \\ x = \\ 11 \\ 00 \end{array} \begin{array}{c} \begin{array}{cc} a = & 11 \quad 00 \end{array} \\ \begin{bmatrix} 0 & C_m \\ C_f & 0 \end{bmatrix} \end{array} \qquad (10\text{–}4)$$

Detto ciò, il problema della decisione può essere enunciato come quello di trovare il criterio con il quale la decisione conseguente ad una determinata osservazione presenta il minor costo di errore. Tale criterio è detto <u>criterio della decisione di Bayes</u>.

A volte non è possibile seguire aderentemente tale procedura, o perché risulta troppo complessa oppure alcuni dati essenziali non sono noti. In tali casi vengono seguiti altri criteri, con conseguenti maggiori costi di errore.

.../...

10-2. La Rivelazione del segnale.

Vediamo ora il criterio della decisione nel caso della rivelazione. Si hanno qui 2 soli possibili stati del fenomeno, segnale presente e segnale assente. Tale situazione è detta di prova delle ipotesi (hypothesis testing). Rifacendoci all'esempio precedente del canale binario, si hanno i due seguenti possibili stati veri del segnale :

$$x_2 = 00 \quad = \text{segnale assente}$$

$$x_1 = 11 \quad = \text{segnale presente}$$

Si è visto che le seguenti sequenze possono essere ricevute :

$$y_1 = 11;\ y_2 = 10;\ y_3 = 01\ ;\ y_4 = 00$$

Le possibili decisioni sono, per ogni stato ricevuto y_j : segnale presente, a_1 , $(x = 11)$; segnale assente a_2 , $(x = 00)$.

Supponiamo di aver ricevuto y_j e di aver deciso per a_i . Il costo della decisione a_i , è, in media :

$$z(y_j, a_i) = C_{11} \cdot P(x_1/y_j) + C_{21} \cdot P(x_2/y_j) \qquad (10\text{-}4)$$

ove $z(y_j, a_i)$ è funzione di y_j .

Ora, dalla (2-14) del capitolo 2, e dalla relazione :

$$P(y_j) = \sum_x P(x) \cdot P(y_j/x) ,$$

si ricava il Teorema di Bayes :

$$P(x_i/y_j) = \frac{P(y_j/x_i) \cdot P(x_i)}{\sum_x P(x)\, P(y_j/x)} = \frac{P_{ij} \cdot P(x_i)}{P_{1j}\, P(x_1) + P_{2j} \cdot P(x_2)} \qquad (10\text{-}5)$$

Le probabilità condizionali nella (10-4) sono quindi ricavabili mediante la (10-5).

Similmente alla (10-4), se a_2 è la decisione conseguente ad $y = y_j$, si ha il costo :

$$z(y_j/a_2) = C_{12} \cdot P(x_1/y_j) + C_{22} \cdot P(x_2/y_j) \qquad (10\text{-}6)$$

.../...

Il costo totale, o rischio, mediato su tutti i possibili valori di y, quando si sia adottato uno stabilito criterio di decisione, sarà minimo se si opera ogni volta la decisione, alla quale consegue il minimo rischio condizionale. Ad esempio, per ogni valore y_j, si dovrà decidere a_1, se $r(y_j, a_1) < r(y_j, a_2)$.

Dalle (10-4), (10-5) e (10-6) segue che si deve scegliere a_1 se è :

$$\lambda = \frac{P(y_j/x_1)}{P(y_j/x_2)} > \frac{P(x_2)}{P(x_1)} \quad \frac{C_{21} - C_{22}}{C_{12} - C_{11}} = \beta \qquad (10\text{-}7)$$

β è una costante che comprende la conoscenza della probabilità a priori e dei costi di errore.
λ è un rapporto di probabilità condizionali, ed è chiamato rapporto probabile (likelihood ratio).

La (10-7) dice che si deve scegliere a_1 se $\lambda > \beta$, ed a_2 se $\lambda < \beta$.

Tornando all'esempio del canale binario, si ha :

$$\beta = \frac{Q}{P} \quad \frac{C_f}{C_m} \qquad (10\text{-}8)$$

C_f è anche chiamato costo di falso allarme e C_m costo di mancato allarme.

Dalla (10-8), facendo uso della (10-3), si ha ad esempio che, se il messaggio ricevuto è $y = 00$ ed è $q_1^2/p_0^2 < \beta$, allora la decisione è a_2. Se viceversa è $q_1^2/p_0^2 > \beta$, allora la decisione è a_1.
Se indichiamo con $P(a_2/x_1)$ e $P(a_1/x_2)$ rispettivamente le probabilità di mancato allarme e di falso allarme, il costo medio di errore quando il segnale è veramente presente e si decide per a_2 è :

$$r_1 = C_m \cdot P(a_2/x_1) \qquad (10\text{-}9)$$

Similmente, quello quando il segnale è effettivamente assente e si decide per a_1, è :

$$r_2 = C_f \cdot P(a_1/x_2) \qquad (10\text{-}10)$$

r_1 ed r_2 sono i rischi condizionali.

.../...

Il rischio medio totale è :

$$R = r_1 \cdot P + r_2 \cdot Q \qquad (10\text{-}11)$$

10-3. Soglia di decisione ottima.

Proseguiamo ora con un'altro esempio di rivelazione ottima di un segnale che si conosce poter assumere solo i valori 1 volt (segnale presente), o 0 volt (segnale assente), in presenza di rumore additivo. Sia il rumore gaussiano centrato e di deviazione standard σ .

In questo caso i possibili stati y sono infiniti, essendo y una variabile continua. Si dovrà perciò introdurre una densità di probabilità condizionale.

Sia x_1 , di probabilità P, lo stato corrispondente ad 1 volt, ed x_2, di probabilità Q, quello di 0 volt .

Si ha, similmente all'esempio precedente :

$$\begin{matrix} & y \\ x_1 \\ x_2 \end{matrix}\begin{bmatrix} p(y/x_1) \\ p(y/x_2) \end{bmatrix} \quad , \quad \begin{matrix} & a_1 & a_2 \\ x_1 \\ x_2 \end{matrix}\begin{bmatrix} 0 & C_m \\ C_f & 0 \end{bmatrix} \qquad (10\text{-}12)$$

Le densità di probabilità condizionali sono rappresentate nella figura seguente :

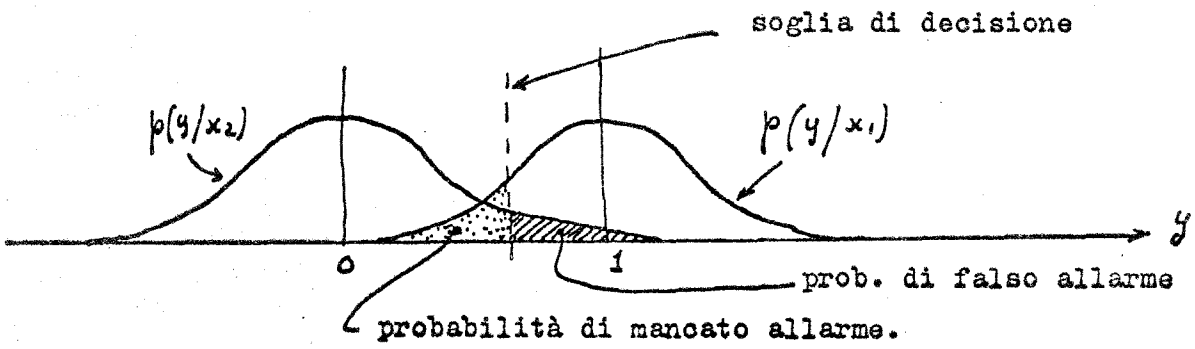


Fig. 104

.../...

I costi di errore medi, corrispondenti alle decisioni a_1 ed a_2 , per un arbitrario valore y sono :

$$r(y,a_1) = 0 \cdot P(x_1/y) + C_f \cdot P(x_2/y)$$
$$r(y,a_2) = C_m \cdot P(x_1/y) + 0 \cdot P(x_2/y) \qquad (10\text{-}13)$$

Entrambe tali espressioni sono funzioni di y .

Il rischio totale è il costo di errore mediato per tutti i valori di y che si ha per un dato criterio di decisione. Naturalmente tale rischio sarà minimizzato se, per ogni valore di y , si sceglie l' ipotesi che dà minor valore di $r(y,a_1)$, o di $r(y/a_2)$. Perciò il criterio di decisione divide il campo dei valori di y in due regioni, una nella quale la decisione a_1 è preferibile e l'altra nella quale lo è a_2.

La (10-5) è ora della seguente forma :

$$P(x_i/y) = \frac{P(x_i) \cdot p(y/x_i)\,dy}{\sum_x P(x_1) \cdot p(y/x_i)\,dy} \qquad (10\text{-}14)$$

Dalla (10-13) e dalla (10-14) si ricava che è $r(y,a_1) < r(y,a_2)$ ovvero si dovrà decidere per a_1 , se è :

$$\lambda(y) = \frac{p(y/x_1)}{p(y/x_2)} > \frac{P(x_2)}{P(x_1)} \cdot \frac{C_{21} - C_{22}}{C_{12} - C_{11}} = \frac{Q \cdot C_f}{P \cdot C_m} = \beta \qquad (10\text{-}15)$$

Nel caso in esame, con rumore gaussiano, si ha :

$$p(y/x_2) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-y^2/2\sigma^2} \quad ; \quad p(y/x_1) = \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot e^{-(y-1)^2/2\sigma^2} \qquad (10\text{-}16)$$

dalla quale si ricava :

$$\lambda(y) = e^{(2y-1)/2\sigma^2} \qquad (10\text{-}17)$$

.../...

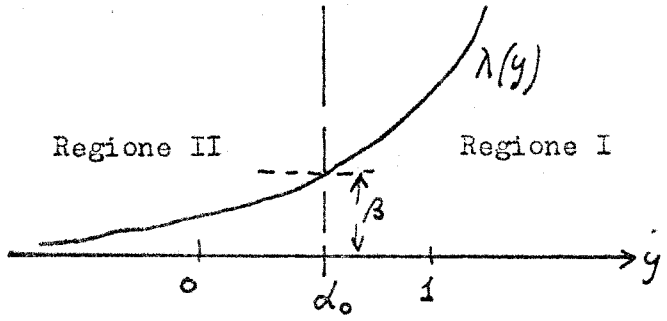


Fig. 105

La (10-17) è rappresentata in Fig. 105 . Il valore di y in corrispondenza di $\lambda(y) = \beta$ è dato da :

$$\alpha_0 = 1/2 + \sigma^2 \ln \beta \qquad (10\text{-}18)$$

ed è detto <u>valore di soglia ottimo</u> per un dato valore di β .

Tale valore divide il campo dei valori di y in due regioni :

1) <u>regione I</u> : ove si decide segnale presente (a_1), $\lambda(y) > \beta$.
2) <u>regione II</u> : ove si decide segnale assente (a_2), $\lambda(y) < \beta$.

Pensiamo ora di stabilire un generico valore di soglia α , indipendentemente per un momento da β , e di suddividere il campo di y nelle due suddette regioni e di decidere come sopra.

Possiamo allora definire le seguenti probabilità, <u>funzioni del valore scelto per α</u> . Si ha, osservando la Fig. 105 :

- probabilità di rivelazione (hit probability) :

$$P(H) = P[y > \alpha / x_1] = P[a_1 / x_1] = \int_{y>\alpha} p(y/x_1)\, dy \qquad (10\text{-}19)$$

è la probabilità di decidere "segnale presente" quando è veramente presente.

- probabilità di falso allarme :

$$P(FA) = P[y > \alpha / x_2] = P[a_1 / x_2] = \int_{y>\alpha} p(y/x_2)\, dy \qquad (10\text{-}20)$$

è la probabilità di decidere "segnale presente" quando è assente.

-probabilità di mancato allarme :

$$P(M) = P[y < \alpha / x_1] = P[a_2 / x_1] = \int_{y<\alpha} p(y/x_1)\, dy \qquad (10\text{-}21)$$

<u>Il rischio totale medio</u> relativo alla soglia α è dato da :

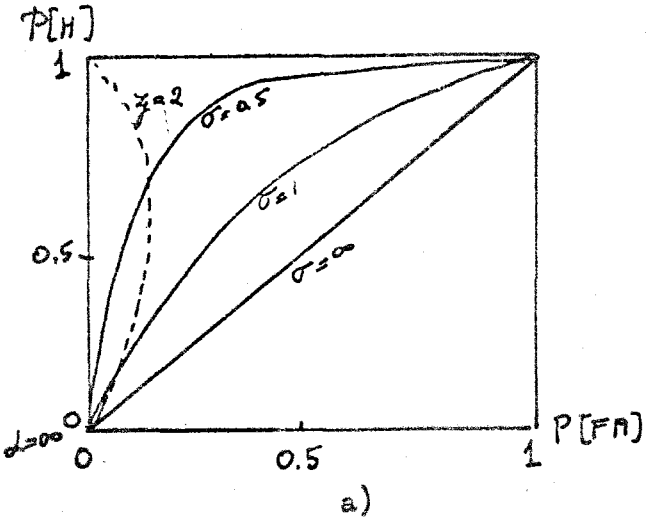
$$R = P \cdot P[M] \cdot C_m + Q \cdot P[FA] \cdot C_f \qquad (10\text{-}22)$$

ove $P[M] \cdot C_m = z_1$ e $P[FA] \cdot C_f = z_2$ (10-23)

.../...

Se ci si chiede come scegliere α onde minimizzare il valore di R, noti Q, P, C_m e C_f, la risposta proviene dal criterio di decidere generalmente per ogni valore di y in modo da avere il minor costo di errore ; ovvero si dovrà scegliere $\alpha = \alpha_0$.

Consideriamo ora le grandezze $P[H]$ e $P[FA]$, funzioni di α , Per ogni assegnato valore di σ è possibile tracciare le due grandezze in funzione di α ; si ottiene in tal modo la <u>caratteristica di funzionamento del ricevitore</u>, riportata nel diagramma di Fig. 106-a).



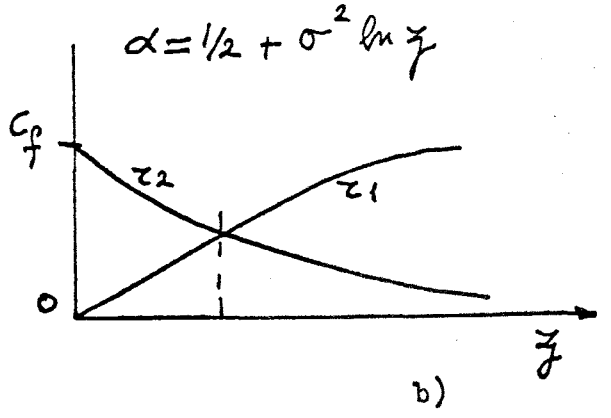


Fig. 106

Dalle (10-19) e (10-20) si vede che la sola grandezza dalla quale $P(H)$ e $P(FA)$ dipendono è α , per un dato σ . Nel caso ad esempio in cui β non sia noto, il valore di α può essere scelto in base alla caratteristica del ricevitore, vale a dire per una soddisfacente coppia di valori $P(H)$ e $P(FA)$.
Questa soluzione non sarà l'ottimo dal punto di vista statistico.

Un criterio meno arbitrario può stabilirsi quando siano noti i costi C_f e C_m e non le probabilità a priori P e Q.

In tal caso è possibile tracciare le funzioni di rischio τ_1 ed τ_2 in funzione di vari possibili valori di z , secondo la (10-23) , ove con z si è rappresentata la funzione β quando la si considera come variabile indipendente (Fig. 106-b)).

Il rischio medio totale $R = P \cdot \tau_1 + Q \cdot \tau_2$ non è mai superiore al più grande dei due rischi condizionali τ_1 ed τ_2 .

La decisione più sicura è di scegliere z , e quindi α , in modo da minimizzare il più grande rischio condizionale. Tale situazione corrisponde a fare $\tau_1 = \tau_2$. Questo criterio è detto di <u>minimo-massimo</u>.

.../...

10-4. Rappresentazione multidimensionale.

Consideriamo ora il caso più interessante di un fenomeno multidimensionale, ovvero il segnale è un punto nello spazio multidimensionale. Sia ad esempio il caso di un segnale da essere rivelato contenuto nella banda W ed avente durata T. Tale segnale può essere rappresentato mediante un vettore ad n dimensioni, con $n = 2\,WT$; ovvero :

$$\vec{X} = x_1, x_2 \cdots\cdots x_{2TW} \qquad (10\text{-}24)$$

Supponiamo che il rumore sia anch'esso rappresentabile nello stesso spazio mediante un vettore $\vec{N}$.

L'osservazione $\vec{Y}$ consiste, o di segnale più rumore $\vec{X}+\vec{N}$, oppure di solo rumore $\vec{N}$. Indichiamo con S ed N rispettivamente le condizioni segnale presente e segnale assente.

Si possono ancora definire i seguenti gruppi di grandezze :

- probabilità a priori: $P(S)=P$; $P(N)=Q$

- matrice dei costi :

$$\begin{array}{c} \\ S \\ N \end{array}\begin{array}{c} \begin{array}{cc} a_1 & a_2 \end{array} \\ \begin{bmatrix} 0 & C_m \\ C_f & 0 \end{bmatrix} \end{array} \qquad (10\text{-}25)$$

- informazione sul rumore :

$$\begin{array}{c} \\ S \\ N \end{array}\begin{array}{c} \begin{array}{c} \vec{Y} \end{array} \\ \begin{bmatrix} p(\vec{Y}/S) \\ p(\vec{Y}/N) \end{bmatrix} \end{array} \qquad (10\text{-}26)$$

ove è :

$$p(\vec{Y}/S) = p(y_1, y_2 \cdots\cdots y_{2TW}/S) \qquad \text{e}$$

$$p(\vec{Y}/N) = p(y_1, y_2 \cdots\cdots y_{2TW}/N) \qquad (10\text{-}27)$$

Si può pensare alle probabilità condizionali (10-27) come a delle "nuvole" nello spazio ad n dimensioni (Fig. 107).

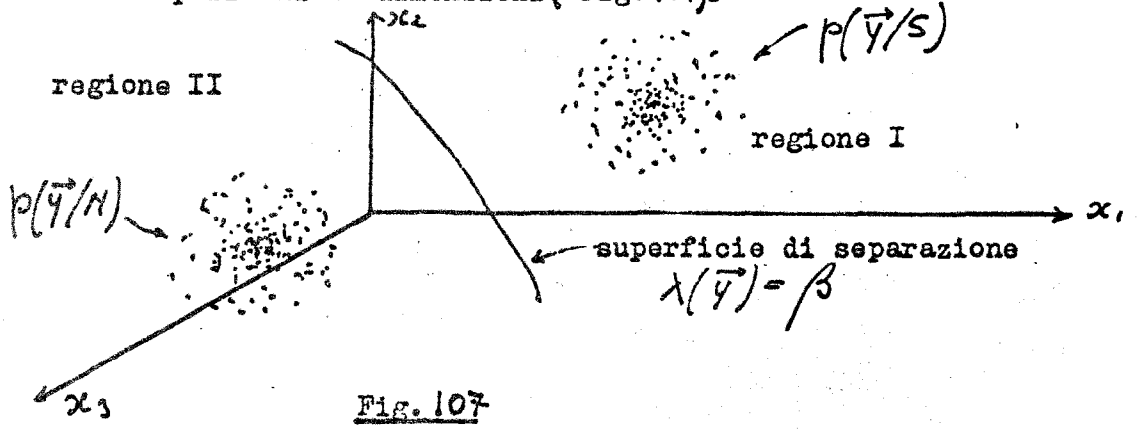


Fig. 107

.../...

Ancora possiamo pensare di voler dividere lo spazio del segnale in due regioni :

- <u>regione I</u> : ove si decide segnale presente se $\vec{y}$ cade in tale regione (decisione a_1).

- <u>regione II</u> : si decide segnale assente (a_2) se $\vec{y}$ cade in tale regione.

La divisione va fatta in modo da minimizzare il rischio medio. Si ha ancora che tale condizione corrisponde a :

$$\lambda(\vec{y}) = \frac{p(\vec{y}/S)}{p(\vec{y}/N)} = \beta \qquad (10\text{-}28)$$

Si può anche qui definire, similmente a quanto visto precedentemente :

$$P[H] = P[a_1/S] = \int_I p(\vec{y}/S)\, d\vec{y} \qquad (10\text{-}29)$$

$$P[FA] = P[a_1/N] = \int_I p(\vec{y}/N)\, d\vec{y} \qquad (10\text{-}30)$$

$$P[M] = P[a_2/S] = \int_{\bar{I}} p(\vec{y}/S)\, d\vec{y} \qquad (10\text{-}31)$$

Nel caso in cui il segnale sia noto ed esprimibile con i campioni s_i , ed il rumore descrivibile come rumore gaussiano bianco, di densità spettrale N_o , si ha :

$$p(\vec{y}/N) = \left(\frac{1}{2\pi N_o W}\right)^{TW} \cdot e^{-\sum_1^{2WT} y_i^2 / 2N_o W} \qquad (10\text{-}32)$$

$$p(\vec{y}/S) = \left(\frac{1}{2\pi N_o W}\right)^{TW} \cdot e^{-\sum_1^{2WT} (y_i - s_i)^2 / 2N_o W} \qquad (10\text{-}33)$$

$$\lambda(\vec{y}) = e^{-\sum s_i^2 / 2N_o W} \cdot e^{-\sum y_i s_i / N_o W} \qquad (10\text{-}34)$$

Col metodo della teoria della decisione si chiede al ricevitore di <u>calcolare</u> $\lambda(\vec{y})$ per il particolare segnale $\vec{y}$ ricevuto e di confrontare tale valore con il valore di β . Questo è supposto noto a priori ed è funzione delle probabilità a priori e dei costi.

.../...

Se risulta $\lambda(\vec{y}) > \beta$, allora il ricevitore decide A_1.

Se è $\lambda(\vec{y}) < \beta$, allora il ricevitore decide A_2.

Vedremo nel seguito con alcuni esempi come tale calcolatore possa essere realizzato praticamente.

10-5. Rivelazione ottima di un segnale determinato.

Consideriamo di applicare ora la (10-34) al caso in cui il segnale è una funzione determinata contenuta nella banda W, e di durata T. Sia inoltre 2 WT$\gg$ 1. La (10-34) può essere modificata grazie alla (4-36). Precisamente si ha :

$$\sum y_i \cdot s_i = 2 \cdot W \cdot \int_{-\infty}^{+\infty} y(t) \cdot s(t) \cdot dt \qquad (10\text{-}35)$$

E = energia del segnale = $\frac{1}{2W} \sum s_i^2 = \int_{-\infty}^{+\infty} s(t)^2 \, dt$ (10-36)

Essendo poi i campioni s_i nulli per t al difuori di - T , 0, si può scrivere la (10-34) nella seguente forma :

$$\lambda(\vec{y}) = e^{-E/N_0} \cdot e^{\frac{2}{N_0} \cdot \int_{-T}^{0} y(t) \cdot s(t) \, dt} > \beta \qquad (10\text{-}37)$$

ove :

$$\int_{-\infty}^{+\infty} y(t) \cdot s(t) \cdot dt \cong \int_{-T}^{0} y(t) \cdot s(t) \cdot dt \quad , \quad \text{se è } 2WT \gg 1.$$

La (10-37) equivale a scrivere :

$$u = \frac{2}{N_0} \int_{-T}^{0} y(t) \cdot s(t) \, dt > \alpha \qquad (10\text{-}38)$$

ove :

$$\alpha = \frac{E}{N_0} + \ln \beta \qquad (10\text{-}39)$$

Quindi il ricevitore dovrà calcolare u , per una data funzione ricevuta y(t), e confrontare poi u con la costante α .

Questo può essere ottenuto con un correlatore incrociato, od un filtro adattato, la cui risposta impulsiva sia h(t) = s(-t), e duri per il Tempo T.(Fig.108).

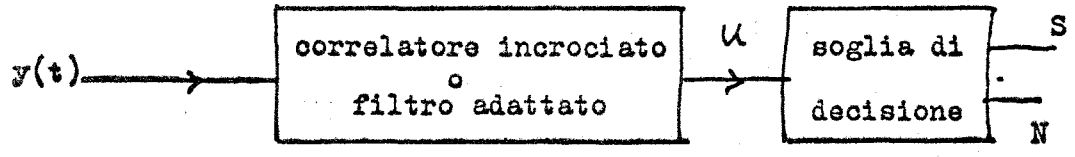


Fig. 108

.../...

Vediamo poi le proprietà statistiche di u . Si può scrivere :

$$u = \frac{1}{N_0 \cdot W} \cdot \Sigma y_i \cdot s_i \qquad (10\text{-}40)$$

Se non c'è segnale, allora $y_i = n_i$; ed essendo s_i delle costanti ed n_i delle variabili aleatorie gaussiane, si avrà che la distribuzione di probabilità di u , quando c'è rumore (N), $p(u/N)$, è gaussiana. Questa ha valore medio :

$$\overline{u} = \frac{1}{N_0 W} \cdot \Sigma s_i \cdot \overline{n_i} = 0 \qquad (10\text{-}41)$$

e varianza :

$$\sigma_u^2 = \overline{u^2} = \left(\frac{1}{N_0 W}\right)^2 \cdot \overline{\left(\Sigma s_i n_i\right)^2} =$$

$$= \left(\frac{1}{N_0 W}\right)^2 \cdot \Sigma s_i^2 \cdot \overline{n_i^2} = \frac{1}{N_0 W} \Sigma s_i^2 = \frac{2E}{N_0} \qquad (10\text{-}42)$$

Se vi è anche segnale, $y_i = s_i + n_i$, e $p(u/S)$ è ancora gaussiana.
Il valore medio è ora :

$$\overline{u} = \frac{1}{N_0 W} \cdot \Sigma \left(s_i^2 + s_i \cdot \overline{n_i}\right) = \frac{2E}{N_0} \qquad (10\text{-}43)$$

e la varianza è ancora data da $2E/N_0$ - $p(u/N)$ e $p(u/S)$ hanno quindi l'andamento di Fig. 109 .

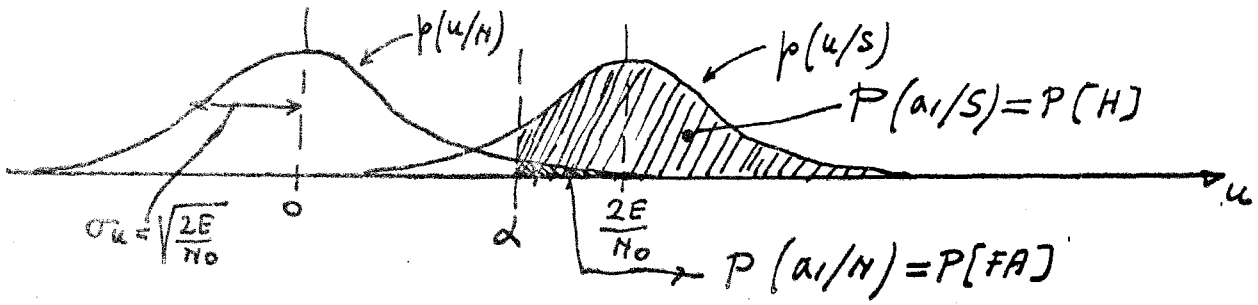


Fig. 109

.../...

per $P[M]$ e $P[FA]$:

Si ha qui ancora, dalle definizioni (10-29) e (10-30)

$$P[M] = \int_{u<\alpha} p(u/S)\,du \quad ; \quad P[FA] = \int_{u>\alpha} p(u/N)\,du \qquad (10\text{-}44)$$

La caratteristica di funzionamento del ricevitore è ora tracciabile in funzione di α, ove ogni curva ha $2E/N_0$ come parametro. Si noti che, dalla (7-53), $2E/N_0$ rappresenta pure il rapporto segnale-disturbo all'uscita del ricevitore.

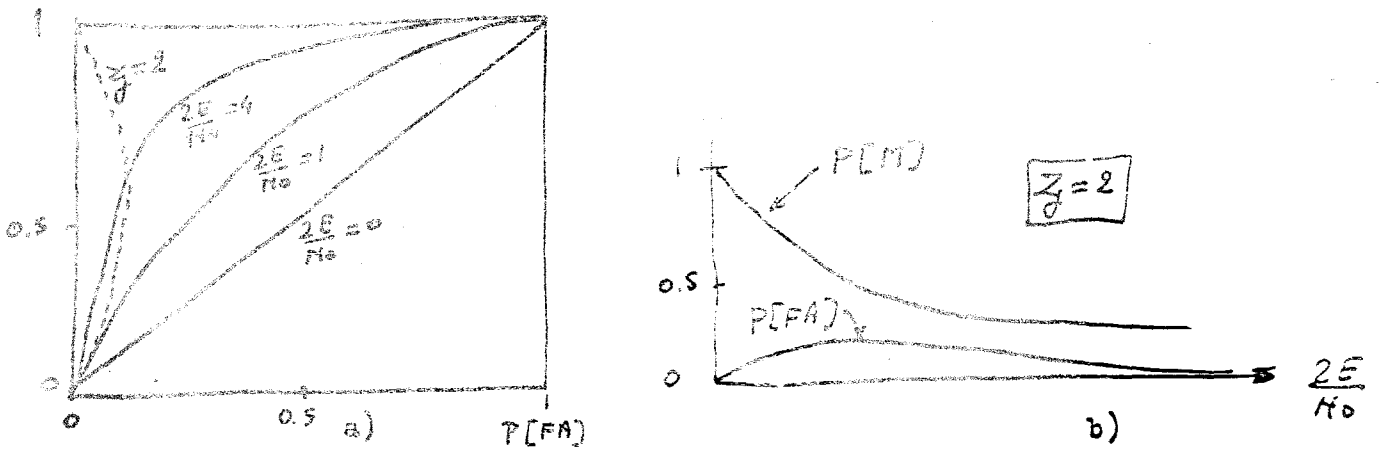


Fig. 110

La curva tratteggiata in Fig. 110-a), rappresenta l'andamento della $P(FA)$ e $P(M)$ per z_f costante, al variare di $2E/N_0$, ove con z_f si è indicata la β quando è considerata come variabile indipendente. Ovvero:

$$\alpha = \frac{E}{N_0} + \ln z_f$$

In Fig. 110-b), sono riportate le probabilità P(M), di mancato allarme, e P(FA), di falso allarme, per $z_f = 2$, ed in funzione del rapporto segnale-disturbo.

Quando non sia possibile assegnare le probabilità a priori ed i costi, è possibile seguire il metodo di Neyman-Person. Tale metodo consiste nell'assegnare un accettabile valore a P(FA), e poi decidere in modo da massimizzare P(M). In tal caso il luogo dei punti di funzionamento al variare di $2E/N_0$ è una linea verticale in Fig. 110-a).

Il caso menzionato precedentemente di minimo-massimo, nel quale siano noti c_f e c_m, e si operi in modo da minimizzare il più grande dei rischi condizionali r_1 ed r_2, dà luogo nella Fig. 110-a) ad una linea retta passante per il punto (0,1) e di pendenza $-c_f/c_m$.

.../...

10-6. Rivelazione di un segnale aleatorio di spettro di potenza assegnato.

Sia il segnale un membro di un processo aleatorio. Consideriamo segnale e rumore nella banda W e siano rispettivamente S_0 ed N_0 le densità spettrali, che supponiamo dapprima costanti. Sia il rumore bianco gaussiano additivo statisticamente indipendente dal segnale. Si ha, similmente alla (10-32), nel caso di solo rumore :

$$p(\vec{y}/N) = \left(\frac{1}{2\pi N_0 W}\right)^{TW} \cdot e^{-\Sigma y_i^2 / 2N_0 W} \qquad (10\text{-}45)$$

Nel caso in cui anche il segnale è presente si ha, grazie alle indipendenza tra segnale e rumore :

$$p(\vec{y}/S) = \left[\frac{1}{2\pi(S_0+N_0)\cdot W}\right]^{TW} \cdot e^{-\Sigma y_i^2 / 2(S_0+N_0)\cdot W} \qquad (10\text{-}46)$$

Si ha quindi la condizione per decidere segnale presente :

$$\lambda(\vec{y}) = \left(\frac{N_0}{S_0+N_0}\right)^{TW} \cdot \exp\left[\frac{1}{2N_0 W} \cdot \frac{S_0}{S_0+N_0} \Sigma y_i^2\right] > \beta \qquad (10\text{-}47)$$

Calcolare tale espressione equivale, similmente a quanto visto nel paragrafo 10-5, a calcolare la funzione u e confrontarla con la soglia α .

Ovvero :

$$u = \int_T y^2(t)\cdot dt > \alpha \qquad (10\text{-}48)$$

ove :

$$\alpha = N_0\left(1+\frac{N_0}{S_0}\right)\cdot \ln\left[\beta\left(1+\frac{S_0}{N_0}\right)^{TW}\right] \qquad (10\text{-}49)$$

Dalla (10-48) si vede quindi che il calcolatore dovrà calcolare l'integrale di $y^2(t)$ nel tempo T, ovvero l'energia del segnale totale ricevuto contenuta in T, e confrontare tale valore con la soglia α . Le distribuzioni $p(u/N)$ e $p(u/S)$ sono rappresentate in Fig. 111 .

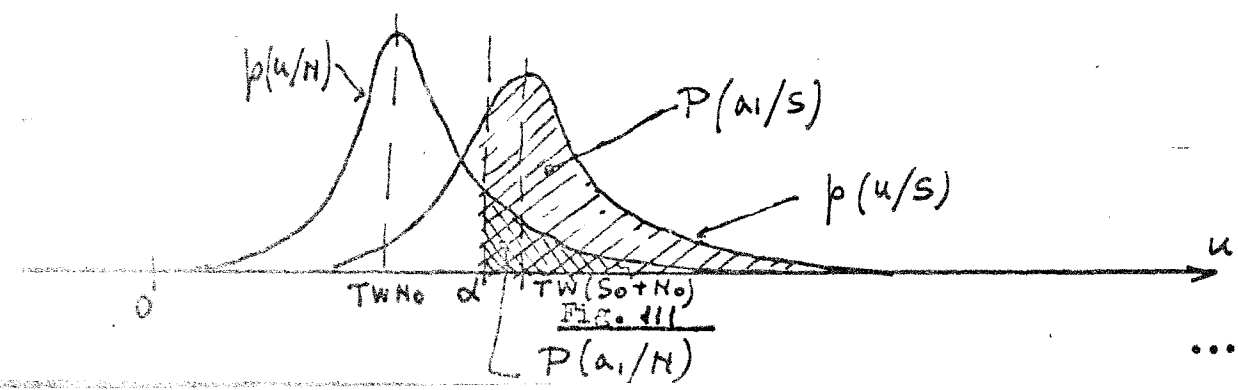


Fig. 111

.../...

Le distribuzioni di figura sono del tipo "chi-quadrato".
Dalla Fig. 111 è poi possibile calcolare, per ogni α, P(H) e P(FA) e tracciare la caratteristica di funzionamento del ricevitore.

Vediamo ora il caso in cui il segnale e rumore hanno densità spettrali arbitrarie, rispettivamente $W'_S(f)$ e $W'_N(f)$.

Supponiamo di poter sviluppare il processo aleatorio y(t), nell'intervallo T, in serie di Fourier. Si ha :

$$y(t) = \sum_{m=1}^{WT} a_m \cdot \cos m \cdot \omega_1 t + b_m \cdot \sin m \omega_1 t \qquad (10\text{-}50)$$

ove a_m e b_m sono variabili aleatorie ed $\omega_1 = 2\pi/T$.
E' dimostrabile che le variabili a_m e b_m sono approssimativamente indipendenti, e presentano varianza pari a $2 \cdot W'_N(m f_1) \cdot 1/T$ quando c'è solo rumore, è $2 \cdot \frac{1}{T}\left[W'_S(m f_1) + W'_N(m f_1)\right]$ quando il segnale è anche presente. Si ha quindi, descrivendo $\vec{y}$ mediante i coefficienti a_m e b_m :

$$p\left(\vec{y}/N\right) \cong K_N \cdot e^{-\frac{I}{4} \cdot \sum \frac{a_m^2 + b_m^2}{W'_N(m f_1)}} \qquad (10\text{-}51)$$

$$p\left(\vec{y}/S\right) \cong K_S \cdot e^{-\frac{I}{4} \cdot \sum \frac{a_m^2 + b_m^2}{W'_S(m f_1) + W'_N(m f_1)}} \qquad (10\text{-}52)$$

$$\lambda\left(\vec{y}\right) \cong K \cdot e^{\frac{I}{4} \cdot \sum \frac{(a_m^2 + b_m^2) \cdot W'_S(m f_1)}{\left[W'_S(m f_1) + W'_N(m f_1) \cdot W'_N(m f_1)\right]}} \qquad (10\text{-}53)$$

Il calcolatore dovrà quindi calcolare l'energia di un nuovo segnale $z(t)$ e confrontare tale misura con la soglia α . Il nuovo segnale è quello ottenuto moltiplicando i coefficienti di Fourier a_m e b_m per il fattore :

$$\sqrt{\frac{W'_S(m f_1)}{W'_N(m f_1) \cdot \left[W'_S(m f_1) + W'_N(m f_1)\right]}} \qquad (10\text{-}54)$$

Ciò può essere ottenuto mediante un filtro lineare di funzione di trasferta H(f), tale che sia :

$$|H(f)|^2 = \frac{W'_S(f)}{W'_N(f) \cdot \left[W'_S(f) + W'_N(f)\right]} \qquad (10\text{-}55)$$

.../...

Lo schema del ricevitore è pertanto il seguente :

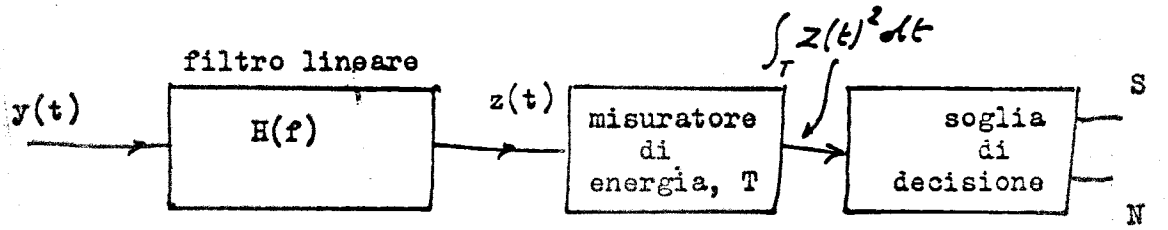


Fig. 112

Si noti che la (10-55) coincide con la (7-25) per piccoli valori del rapporto segnale-disturbo.

10-7. Stima di parametri.

Abbiamo precedentemente trattato il caso della prova delle ipotesi, con particolare riferimento al caso della rivelazione del segnale, vero di sole due ipotesi, segnale presente e segnale assente.

Consideriamo ora il caso del rilievo di parametri. Sia x, vero stato del fenomeno, funzione di uno o più parametri $a_1, a_2 \dots a_n$. Ad esempio nel caso di un solo parametro, a potrebbe trattarsi della distanza di bersaglio Sonar.

Lo stato osservato y è somma di rumore ed eventuale segnale ovvero :

$$y(t) = x(t) + n(t)$$

ove $n(t)$ è il rumore addittivo.

Dato lo stato $y(t)$, il problema è quello di trovare il criterio che porti alla stima $a'(y)$ del parametro a, con i minori costi di errore.

In questo caso sia gli stati veri del fenomeno, che quelli osservati, sono infiniti.

Il problema è ancora descritto mediante :

- le probabilità condizionali di osservazione $p(\vec{y}/a)$, ovvero $p(\vec{y}/a_1, a_2 \dots a_n)$

  queste descrivono le caratteristiche del rumore.

- informazione a priori riguardante la distribuzione dei parametri :

  $p(a)$, ovvero $p(a_1, a_2 \dots a_n)$

- la funzione dei costi $C(a', a)$, funzione di a ed a'.

.../...

La funzione dei costi dipende dal tipo di problema in esame. Spesso la sua scelta può essere determinata dalla necessaria semplificazione della trattazione matematica.

L'errore è definito dalla differenza $a' - a$.

Alcuni tipi di funzioni $C(a',a)$ sono le seguenti :

- $C(a',a) = |a'-a|$, costo proporzionale all'errore
- $C(a',a) = (a'-a)^2$, errore quadratico, di facile trattazione matematica
- $C(a',a) = K$, se $|a'-a| > \varepsilon$; $= 0$ altrove , errore limitato in $a-\varepsilon, a+\varepsilon$.

Supponiamo ora di dover stimare un particolare valore di a mediante l'osservazione $y(t)$, secondo un non ben definito procedimento.

Se operiamo ripetute osservazioni, statisticamente indipendenti tra loro, il costo di errore medio è dato da :

$$r(a) = \int C(a',a) \cdot p(\vec{y}/a) \cdot d\vec{y} \qquad (10\text{-}56)$$

che è chiamato rischio condizionale.

Ulteriormente si può mediare $r(a)$ per tutti i possibili valori di a , ottenendo il rischio totale medio :

$$R = \int r(a) \cdot p(a) \cdot da = \iint C(a',a) \cdot p(a) \cdot p(\vec{y}/a) \cdot da \cdot d\vec{y} =$$

$$= \iint C(a',a) \cdot p(a/\vec{y}) \cdot p(\vec{y}) \cdot da \cdot d\vec{y} \qquad (10\text{-}57)$$

Il metodo di decisione di Bayes consiste nello scegliere per ogni osservazione $\vec{y}$ quel valore di a' , tale che è mediamente minimizzato il rischio R.

Supponiamo ora che sia $C(a',a) = (a'-a)^2$. Si ha allora

$$R = \int \left\{ \int [a'(\vec{y}) - a]^2 \cdot p(a/\vec{y}) \cdot da \right\} \cdot p(\vec{y}) \cdot d\vec{y} \qquad (10\text{-}58)$$

E' poi , per un dato valore di $\vec{y}$:

$$\int [a'(\vec{y}) - a]^2 \cdot p(a/\vec{y}) \cdot da = \overline{(a-a')^2} = \sigma_a^2 + \bar{a}^2 - 2\bar{a}a' + a'^2 = \qquad (10\text{-}59)$$

$$= \sigma_a^2 + (\bar{a} - a')^2$$

.../...

La (10-59) è minima quando $\bar{a} = \hat{a}$. Quindi la (10-58) è minimizzata per ogni valore di $\vec{y}$ quando si sceglie :

$$a'(\vec{y}) = \bar{a} = \int a \cdot p(a/\vec{y}) \cdot da \qquad (10\text{-}60)$$

Più generalmente, se è $C(a',a) = f(a) \cdot (a'-a)^2$, allora si ottiene la condizione di minimo rischio o <u>criterio di Bayes</u> :

$$a'(\vec{y}) = \frac{\int a \cdot f(a) \cdot p(a/\vec{y}) \cdot da}{\int f(a) \cdot p(a/\vec{y}) \cdot da} \qquad (10\text{-}61)$$

Le (10-60) e (10-61) sono valide qualora sia nota la distribuzione a priori di a .

Se tale distribuzione non è nota si può sempre ricorrere al criterio di minimizzare il massimo dei rischi condizionali $r(a)$. Vi è poi un Teorema, in teoria della decisione, secondo il quale, se si può trovare una funzione di densità a priori $p(a)$ per la quale il rischio condizionale $r(a)$ è indipendente da a , questa è la più sfavorevole distribuzione di a , ed il corrispondente procedimento di decisione di Bayes soddisfa alla condizione di minimo-massimo.

Spesso si può ricorrere, quando vi siano notevoli difficoltà per il metodo di decisione ottimo di Bayes (o per quello di minimo-massimo), alla massimizzazione della funzione di densità $p(\vec{y}/a)$, chiamata <u>funzione probabile</u> (likelihood function) . Tale procedimento offre un metodo di stima, non il migliore. Questo può però essere confrontato con gli altri in termini del rischio medio. Spesso, come si può vedere dal teorema di Bayes, il massimo di $p(\vec{y}/a)$ capita per valore di a vicino a quello che massimizza $p(a/\vec{y})$, la funzione di densità a posteriori. In tal caso i due criteri di stima sono quasi equivalenti.

Facciamo ora un esempio.

Consideriamo il caso in cui lo stato vero a è rappresentat da una variabile avente distribuzione gaussiana centrata , di deviazione standard σ_a . Sia y lo stato misurato, perturbato da rumore gaussiano, di deviazione standard σ_n .

La funzione probabile è data da :

$$p(y/a) = \frac{1}{\sqrt{2\pi}\,\sigma_n} \cdot e^{-(y-a)^2/2\sigma_n^2} \qquad (10\text{-}62)$$

La densità di probabilità a posteriori è ricavata come segue

$$p(a,y) = \frac{p(a) \cdot p(y/a)}{p(y)} = \frac{1}{2\pi\sigma_a\sigma_n \cdot p(y)} e^{-a^2/2\sigma_a^2} \cdot e^{-(y-a)^2/2\sigma_n^2} \qquad (10\text{-}63)$$

.../...

Essendo qui poi $p(y)$ una costante (y è infatti noto), si può scrivere :

$$p(a/y) = \text{cost} \times e^{-(a - c\cdot y)^2 / 2\sigma_m^2 \cdot c}$$

$$\text{ove} \quad c = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_m^2} \qquad (10\text{-}64)$$

Tale distribuzione è rappresentata in Fig.

Assumiamo una funzione di errore quadratica.

Il metodo di stima di Bayes dà allora :

$$a'_{Bayes} = c \cdot y$$

ove $c\,y$ è il valor medio della distribuzione (10-64). Il corrispondente rischio è dato da :

$$R = \int r(a) \cdot p(a) \cdot da = c \cdot \sigma_m^2$$

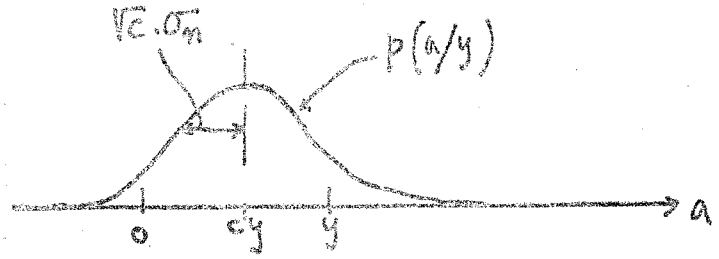


Fig. 113