

2-0 Introduzione

In tutti i casi ove si debbono studiare o descrivere dei fenomeni a carattere aleatorio, che diano cioè luogo a variabili od a funzioni aleatorie, l'analisi deterministica deve essere sostituita da un tipo di analisi a carattere statistico.

Nei fenomeni di tipo aleatorio non è possibile conoscere esattamente il valore che la variabile o la funzione assumeranno ad osservazioni successive, ma è solo possibile predire il valore più probabile.

Tale predizione è basata sulla assunzione della regolarità del fenomeno in esame. Questo viene osservato un numero grande (infinito) di volte ed, in base a tali osservazioni, è possibile ottenere la frequenza di ricorrenza dei vari valori possibili assunti dalla variabile o dalla funzione aleatoria. A maggior frequenza di ricorrenza di un dato valore nelle osservazioni dovrà corrispondere maggior probabilità, assunta la regolarità del fenomeno, che tale valore si ripeta in una osservazione successiva. Il concetto di probabilità viene quindi introdotto onde assegnare alla variabile od alla funzione aleatoria una misura quantitativa, con un numero compreso tra 0 ed 1, della frequenza di ricorrenza di un dato valore in un numero grande di eventi.

Si noti che la conoscenza delle probabilità relative ad un determinato fenomeno aleatorio può derivare sia dalla conoscenza a priori derivante dalle proprietà del fenomeno, sia da misurazioni empiriche su di un numero grande di esperienze.

Si esaminano di seguito alcune definizioni e proprietà delle variabili aleatorie. Si passerà poi a trattare le funzioni aleatorie o processi aleatori.

2-1 Le Variabili Aleatorie

Si consideri di operare una osservazione ϵ di un fenomeno aleatorio; sia X il risultato ad essa associabile, rappresentabile con un valore numerico. I valori $X=X(\epsilon)$ costituiscono una variabile aleatoria.

La X può essere una variabile continua, discontinua o mista.

Nel caso di variabile discontinua si definisce la probabilità che la X assuma il valore X_i come la frequenza relativa di tale valore, con il seguente simbolismo:

$$p_i = P [X=X_i] \quad , \quad \text{ove } 0 \leq p_i \leq 1$$

Se si considera l'insieme $[X_1, X_2, \dots, X_k]$ di possibili valori, la probabilità che X assuma uno di questi valori è data dalla somma delle probabilità di ogni valore, ovvero:

$$P[X=X_1], \text{ oppure } [X=X_2], \text{ oppure } \dots \dots [X=X_k] = p_1 + p_2 + p_3 + \dots + p_k \quad (2-2)$$

Se l'insieme sopra considerato racchiude tutti i possibili valori si ha la certezza che la X assuma uno di questi valori, ovvero:

$$\sum_{i=1}^k p_i = 1 \quad (2-3)$$

Nel caso di variabile aleatoria X continua, la probabilità che essa assuma un valore stabilito x è nulla. Si può invece parlare della probabilità che sia $x \leq X \leq x + dx$. Per questo si ricorre alla funzione densità di probabilità $p(x)$ mediante la quale la probabilità suddetta è esprimibile come:

$$p(x) \, dx = P [x \leq X \leq x + dx] \quad (2-4)$$

Si definisce poi la funzione di distribuzione della probabilità, o funzione di probabilità cumulativa, come:

$$F(x) = P [X \leq x] = \int_{-\infty}^x p(x) \, dx \quad (2-5)$$

Similmente alla (2-3), si ha qui:

$$\int_{-\infty}^{\infty} p(x) dx = 1 \quad (2-6)$$

Consideriamo ora alcuni esempi:

- Un sistema di controllo a relay ha una tensione di errore che assume i valori $-1, 0, +1$. Tali errori ricorrono con probabilità $1/4, 1/2$ ed $1/4$ rispettivamente. La funzione di distribuzione corrispondente a tale situazione avrà perciò l'andamento di Fig. 10, a) e la $p(x)$ è da questa ottenibile mediante derivazione della $F(x)$, ovvero:

$$p(x) = dF(x)/dx = 1/4 \delta(x+1) + 1/2 \delta(x) + 1/4 \delta(x-1)$$

La $p(x)$ è poi rappresentata in Fig. 10, b).

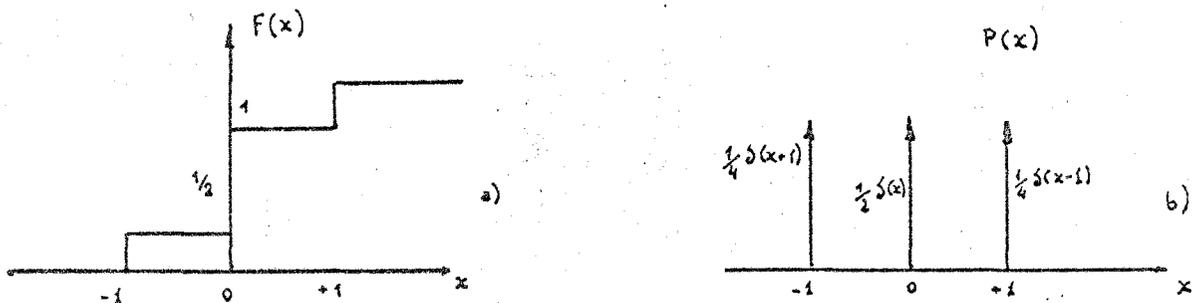


Fig. 10

- Si consideri la funzione $X(t)$ di Fig. 11, a), ove la pendenza dei tratti rettilinei è aleatoria. Si consideri la variabile aleatoria associabile alla misurazione dell'ampiezza della $X(t)$ a vari istanti di tempo. La probabilità che sia $x \leq X \leq x+dx$ è uguale alla probabilità che l'istante di misurazione cada in uno degli intervalli dt corrispondenti. Essendo poi ognuno di questi intervalli dt indipendente dal valore di x , ne consegue che ogni valore $0 \leq x \leq 1$ ha la stessa probabilità. Le Fig. 11, b) e c) rappresentano rispettivamente la funzione $F(x)$ e la sua derivata $p(x)$.

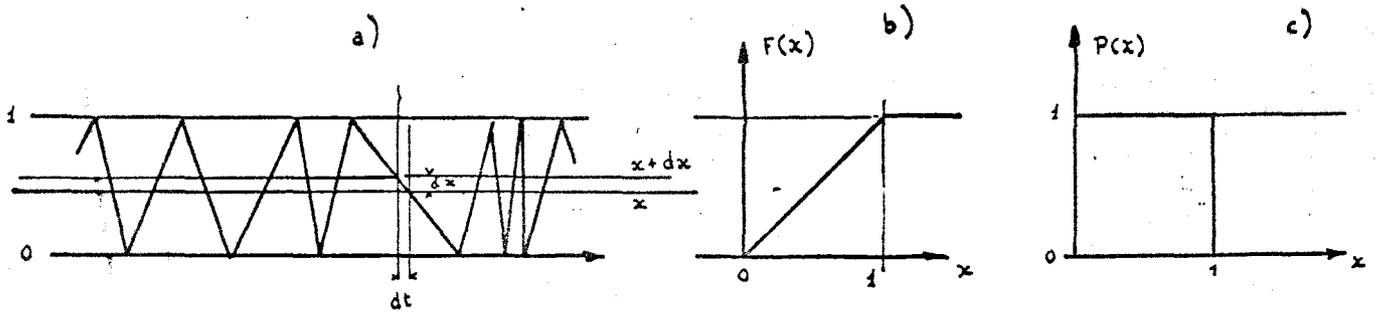


Fig. 11

- Consideriamo infine il caso di una sinusoide, a modulazione di fase aleatoria, $X(t) = \cos \Theta(t)$, ove, per ogni dato istante t , sono possibili con ugual probabilità tutti i valori di Θ tra 0 e 2π .

Se si considera ora la variabile aleatoria associabile alla misurazione dell'ampiezza a vari istanti t , questa sarà legata funzionalmente alla variabile aleatoria Θ . Precisamente la probabilità che la X cada nell'intervallo dx sarà uguale ^{al doppio} dalla probabilità che la corrispondente Θ cada nell'intervallo $d\Theta$. Quindi si avrà:

$$2 \cdot p_{\Theta}(\Theta) \cdot d\Theta = p_X(x) \cdot dx,$$

$$\text{dalla quale deriva: } p_X(x) = \left(\frac{d\Theta}{dx} \right)^2 = \frac{1}{\pi \sqrt{1-x^2}}$$

In Fig. 12 sono rappresentate la funzione $X(t)$, la funzione di distribuzione e la densità di probabilità.

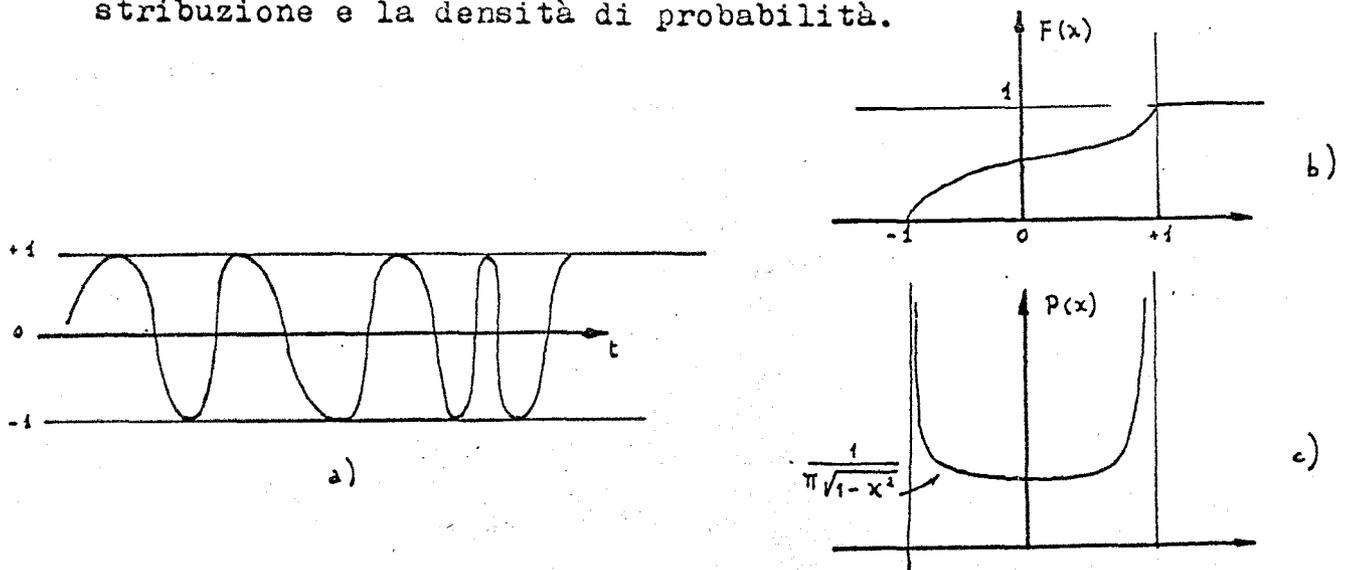


Fig. 12

2-2 Descrizione della variabile aleatoria.

Nel descrivere statisticamente la variabile aleatoria X bisogna tener presente che generalmente l'insieme dei valori di X che hanno una probabilità non trascurabile è piuttosto raggruppato. E' in tal caso sufficiente dare il valore centrale di tale raggruppamento ed un valore che dia una misura della dispersione attorno al valore centrale.

Per descrivere tali grandezze si ricorre all'operazione di media statistica, o speranza matematica. Questa consiste nell'operazione di media dei valori della variabile aleatoria, o di valori di funzione della variabile aleatoria, ponderati con la probabilità della variabile aleatoria in esame. Se f è una funzione certa, si può generalmente considerare la funzione di variabile aleatoria $f(X)$, corrispondente alla variabile aleatoria X , per la quale si definisce la media statistica, o speranza matematica:

$$(1) \quad E[f(X)] = \int_{-\infty}^{\infty} f(x) \cdot p(x) \cdot dx \quad (\text{per } x \text{ continua})$$
$$E[f(X)] = \sum_i f(x_i) \cdot p_i \quad (\text{per } x \text{ discontinua}) \quad (2-7)$$

In particolare nel caso in cui $f(X) = X^n$, con n intero, la $E[X^n]$ prende il nome di momento di ordine n della variabile aleatoria X , e si rappresenta col simbolo m_n .

Nel caso in cui $n=1$, si ha:

$$(1) \quad m_1 = E[X] = \int_{-\infty}^{\infty} x \cdot p(x) \cdot dx \quad (\text{per } x \text{ continua}) \quad (2-8)$$
$$m_1 = E[X] = \sum_i x_i \cdot p_i \quad (\text{per } x \text{ discontinua})$$

E' questo il valore statistico medio della variabile aleatoria, ovvero il valore più probabile o valore centrale. Se si rappresenta la probabilità come una massa distribuita sull'asse delle x , la $E[X]$ rappresenta l'ascissa del centro di gravità.

Si consideri ora la nuova variabile aleatoria centrata $X' = X - m_1$,

(1) Il simbolo E per la media statistica ha origine dalla parola inglese "expectation".-

./...

I momenti di tale variabile sono detti centrati; quello di ordine n viene rappresentato nel seguente modo:

$$\mu_n = E [X^n] = E [(X - m_1)^n] \quad (2-9)$$

Tra questi momenti centrati è interessante particolarmente quello di ordine due, ovvero:

$$\mu_2 = \sigma^2 = E [X^2] = E [(X - m_1)^2] \quad (2-10)$$

Questa grandezza prende il nome di varianza, e dà una valutazione energetica dell'entità della dispersione dei valori della variabile aleatoria attorno al valore centrale m_1 . La grandezza $\sigma = (\mu_2)^{1/2}$ prende il nome di scarto quadratico medio centrato, o deviazione standard.

Se ora si sviluppa l'espressione di σ^2 si ottiene:

$$\begin{aligned} \sigma^2 &= \int (x - E[x])^2 \cdot p(x) \cdot dx = \int x^2 \cdot p(x) dx + \int E[x]^2 \cdot p(x) dx - \\ &- 2 \int x \cdot E[x] \cdot p(x) dx = E[x^2] + E[x]^2 \cdot 1 - 2 \cdot E[x]^2 = \\ &= E[x^2] - E[x]^2 \end{aligned}$$

$$\text{ovvero: } \sigma^2 = \mu_2 = m_2 - (m_1)^2 \quad (2-11)$$

E' interessante notare, considerando il significato energetico della (2-11), l'analogia che intercorre con la distribuzione della potenza di un segnale elettrico. Si può fare infatti il seguente raffronto:

la componente della potenza σ^2 , prodotta dalla componente alternativa del segnale, è uguale alla potenza totale (m_2), meno la potenza prodotta dalla componente continua, ovvero il valor medio al quadrato (m_1^2). La σ rappresenta quindi il valore efficace.

2-3 Insieme di due variabili aleatorie.

Si ha spesso a dover considerare insiemi di più variabili aleatorie; quello di due variabili è il più frequente.

Si considera qui l'insieme di due variabili aleatorie discontinue X ed Y, nel qual caso si possono introdurre le seguenti probabilità:

- Probabilità congiunta, ovvero la probabilità che le due variabili assumano in uno stesso evento due valori stabiliti: X_i ed Y_j :

$$P[X_i, Y_j] = P[X = X_i, Y = Y_j] \quad (2-12)$$

- Probabilità condizionale, ovvero la probabilità che una delle variabili assuma un valore prestabilito quando l'altra variabile assume un dato valore:

$$P[X_i/Y_j] = P[X = X_i \text{ quando } Y = Y_j] \quad (2-13)$$

tra la probabilità congiunta e quella condizionale sussiste la relazione:

$$P[X_i, Y_j] = P[X_i] \cdot P[Y_j/X_i] = P[Y_j] \cdot P[X_i/Y_j] \quad (2-14)$$

La probabilità poi che almeno uno dei due eventi X_i ed Y_j si verifichi è data da:

$$P[X_i] + P[Y_j] - P[X_i, Y_j] \quad (2-15)$$

Se i due eventi sono poi mutualmente esclusivi è allora:

$$P[X_i, Y_j] = 0 \quad (2-16)$$

In tal caso la (2-15) si ridurrà dalla somma $P[X_i] + P[Y_j]$

Se i due eventi sono tra loro indipendenti è allora:

$$P[X_i, Y_j] = P[X_i] \cdot P[Y_j]$$

$$P[X_i/Y_j] = P[X_i]; \quad P[Y_j/X_i] = P[Y_j] \quad (2-17)$$

Se le due variabili aleatorie X ed Y sono continue si può introdurre, similmente al caso unidimensionale, la densità di probabilità bidimensionale $\rho(x, y)$:

$$\rho(x, y) = \frac{1}{dx \cdot dy} \cdot P[x \leq X \leq x+dx, y \leq Y \leq y+dy] \quad (2-18)$$

La $\rho(x, y)$ può anch'essa considerarsi ottenuta mediante derivazione di una funzione di distribuzione bidimensionale; si ha anche qui, similmente alla (2-6):

$$\iint_{-\infty}^{\infty} \rho(x, y) \cdot dx \cdot dy = 1 \quad (2-19)$$

Se si conosce la $\rho(x, y)$ può interessare risalire da questa alla densità di probabilità di una delle variabili, ad esempio $p(x)$.

Questa prende in tal caso il nome di densità di probabilità a priori in quanto necessita a priori di tutte le indicazioni sulla variabile Y per poter essere ottenuta tramite la $\rho(x, y)$. Consideriamo infatti, nel piano dei valori di X ed Y , il punto $M(X, y)$, ove $x \leq X \leq x+dx$.

E' questo un generico punto compreso tra le due rette passanti per i punti x ed $x+dx$ e parallele all'asse y ; la probabilità della sua esistenza è data da:

$$p(x) \cdot dx = P[x \leq X \leq x+dx]$$

Tale probabilità è anche esprimibile tramite la $\rho(x, y)$, integrando solo rispetto ad y , ovvero:

$$p(x) \cdot dx = dx \cdot \int_{-\infty}^{\infty} \rho(x, y) \cdot dy, \quad \text{da cui si ottiene:}$$

$$p(x) = \int_{-\infty}^{\infty} \rho(x, y) \cdot dy \quad (2-20)$$

ed analogamente per la variabile Y :

$$q(y) = \int_{-\infty}^{\infty} \rho(x, y) \cdot dx \quad (2-21)$$

Vediamo ora come la (2-14) possa essere espressa nel caso di variabili continue. Si può scrivere:

$$P[x \leq X \leq x+dx, y \leq Y \leq y+dy] = P[y \leq Y \leq y+dy] \cdot P[x \leq X \leq x+dx / y \leq Y \leq y+dy] \quad (2-22)$$

Si noti ora come il primo membro della (2-22) sia uguale a $\rho(x, y) \cdot dx \cdot dy$, mentre nel secondo membro il primo termine è esprimibile con la probabilità a priori $q(y) \cdot dy$. Introduciamo ora la densità di probabilità condizionale $p_y(x)$, condizionata al valore $y=Y$; in tal modo il secondo termine del prodotto al secondo membro della (2-22) potrà esprimersi, nel caso in cui $dy \rightarrow 0$, con $p_y(x) \cdot dx$. Pertanto la (2-22) diviene:

$$\rho(x, y) \cdot dx \cdot dy = q(y) \cdot dy \cdot p_y(x) \cdot dx \quad (2-23)$$

Da questa si ottiene:

$$p_y(x) = \frac{\rho(x, y)}{q(y)} \quad (2-24)$$

e similmente:

$$q_x(y) = \frac{\rho(x, y)}{p(x)} \quad (2-25)$$

Ancora, nel caso in cui X ed Y siano indipendenti:

$$p_y(x) = p(x); \quad q_x(y) = q(y); \quad \rho(x, y) = p(x) \cdot q(y) \quad (2-26)$$

Ovvero, nel caso di indipendenza delle due variabili, la densità bidimensionale è uguale al prodotto delle due densità a priori.

2-4 Alcune proprietà della distribuzione di probabilità bidimensionale - coefficiente di correlazione -

Similmente al caso unidimensionale è anche qui possibile operare la media statistica di una funzione delle due variabili aleatorie X ed Y . Si introduca pertanto la funzione certa $f(x, y)$, e si consideri la variabile aleatoria ottenuta quando $x=X$ ed $y=Y$.

Si ha allora:

$$E[f(X,Y)] = \iint_{-\infty}^{\infty} f(x,y) \cdot \rho(x,y) \cdot dx \cdot dy \quad (2-27)$$

Si può ora definire il generico momento centrato di ordine j, k :

$$\mu_{j,k} = E \left[(X-m_{1X})^j \cdot (Y-m_{1Y})^k \right] = \iint_{-\infty}^{\infty} (x-m_{1X})^j \cdot (y-m_{1Y})^k \cdot \rho(x,y) \cdot dx \cdot dy \quad (2-28)$$

ove m_{1X} ed m_{1Y} sono i valori statistici medi di X e di Y.

Il più interessante tra i momenti è la covarianza μ_{11} :

$$\mu_{11} = E \left[(X-m_{1X}) \cdot (Y-m_{1Y}) \right] = E[X' \cdot Y'] \quad (2-29)$$

E' facilmente dimostrabile che:

$$\mu_{11} = E[X' \cdot Y'] = E[X \cdot Y] - E[X] \cdot E[Y] \quad (2-30)$$

Si definisce ora il coefficiente di correlazione tra X' ed Y', C:

$$C = \frac{E[X' \cdot Y']}{\sigma_x \cdot \sigma_y} = \frac{\mu_{11}}{\sigma_x \cdot \sigma_y} \quad (2-31)$$

Il coefficiente C dà una misura del legame esistente tra le due variabili aleatorie. Precisamente, nel caso di indipendenza statistica tra le due variabili si ha $E[X \cdot Y] = E[X] \cdot E[Y]$, e dalla (2-30) $\mu_{11} = 0$; quindi $C = 0$.

Se poi X ed Y sono legate tra loro da una relazione lineare si ha $C = 1$. Infatti, considerando X ed Y centrate e posto che sia $Y = aX$, si ha:

$$E[X \cdot Y] = E[X \cdot (aX)] = E[aX^2] = a^2 \cdot \sigma_x^2 = a \sigma_x^2;$$

$$E[X^2] = \sigma_x^2; \quad E[Y^2] = E[a^2 \cdot X^2] = a^2 \sigma_x^2;$$

Da queste relazioni, sostituendo nella (2-31), si ha : $C = \frac{a \sigma_x^2}{a \sigma_x^2} = 1$.

Normalmente è $0 \leq C \leq 1$. Questa affermazione può essere provata fa-

cendo uso della disuguaglianza di Schwarz. Sia infatti λ un numero reale e siano X ed Y due variabili aleatorie. Per ogni valore di λ si ha:

$$E[(\lambda X + Y)^2] = \lambda^2 \cdot E[X^2] + 2\lambda E[X \cdot Y] + E[Y^2] \geq 0 \quad (2-32)$$

ove la condizione ≥ 0 deriva dal fatto che la variabile $(\lambda X + Y)^2$ è sempre positiva. Se si considera ora λ come variabile e gli altri termini della disuguaglianza (2-32) come coefficienti, si può scrivere:

$$\lambda^2 \cdot a + 2 \cdot \lambda \cdot b + c \geq 0 \quad (2-33)$$

La (2-33) è verificata quando $\Delta = b^2 - ac \leq 0$, da cui risulta:

$$|E[X \cdot Y]|^2 \leq E[X^2] \cdot E[Y^2] \quad (2-34)$$

Da questa relazione e dalla (2-31) deriva che $0 \leq C \leq 1$.

Si è quindi visto come, date due variabili aleatorie X ed Y, nel caso di indipendenza statistica si abbia $C = 0$, e nel caso di dipendenza lineare $C = 1$.

Inversamente, se $C = 1$ si può dimostrare che tra X ed Y vi è dipendenza lineare; mentre la condizione $C = 0$ non significa necessariamente l'indipendenza statistica tra X ed Y; esse saranno però in tal caso certamente linearmente indipendenti.

Consideriamo infatti il seguente esempio, ove è $X = \cos \theta$ ed $Y = \sin \theta$, con θ uniformemente distribuito tra 0 e 2π . In questo caso X ed Y non sono linearmente dipendenti e non sono statisticamente indipendenti.

La loro dipendenza è infatti espressa dalla relazione $Y^2 = 1 - X^2$.

Si ha però: $E[X \cdot Y] = 0$ e $C = 0$.

Si vedrà in seguito come la condizione $C = 0$ assuma anche il significato di indipendenza statistica nel caso di variabili aleatorie gaussiane.

2-5 Convoluzione di probabilità

Si considerino due variabili aleatorie discontinue indipendenti, X ed Y, e si formi la nuova variabile aleatoria $Z = X+Y$. Si ponga il problema di determinare la probabilità $P [Z = Z_k]$ del valore $Z_k = X_i + Y_j$, note le probabilità dei valori X_i ed Y_j . Si noti che vi sono in genere molte coppie di valori X_i, Y_j che hanno come somma il valore Z_k e per ognuna di esse la probabilità di Z_k può scriversi, essendo X ed Y indipendenti:

$$P [X_i] \cdot P [Y_j] = P [X_i] \cdot P [Z_k - X_i] \quad (2-35)$$

La probabilità totale del valore Z_k si otterrà sommando tra loro tutte le probabilità parziali (2-35) corrispondenti a diversi valori della coppia X_i, Y_j , ovvero $X_i, Z_k - X_i$.

Si ottiene perciò

$$P [Z_k] = \sum_{X_i} P [X_i] \cdot P [Z_k - X_i] \quad (2-36)$$

Un simbolismo più efficace è quello in cui si sostituisca al simbolo P dei simboli diversi, a seconda della variabile cui ci si riferisce. In tal modo la (2-36) può scriversi:

$$R(Z) = \sum_X P [X] \cdot Q [Z - X] \quad (2-37)$$

La R è detta probabilità di convoluzione delle probabilità P e Q, e si può, anche qui come nel Cap. 1, usare il simbolismo: $R = P \otimes Q$.

Nel caso in cui le variabili X ed Y siano continue, la espressione (2-37) diviene un integrale di convoluzione, simile a quello introdotto nel Cap. 1.

Si ha in tal caso:

$$r(z) = p \otimes q = \int_{-\infty}^{+\infty} p(x) \cdot q(z - x) dx \quad (2-38)$$

ove ora $p(x)$, $q(y)$ e $r(z)$ rappresentano delle densità di probabilità. Quindi la densità di probabilità della somma di due o più variabili aleatorie indipendenti è data dalla convoluzione delle loro singole densità di probabilità.

2-6 La Funzione Caratteristica

Si consideri la variabile aleatoria continua X , e $p(x)$ sia la sua densità di probabilità. Si consideri poi la variabile reale u ($-\infty < u < +\infty$) e si operi la trasformazione di Fourier della funzione $p(x)$, dal dominio della variabile aleatoria x a quello della variabile reale u . La funzione così ottenuta è detta "funzione caratteristica" e si scrive:

$$\varphi(u) = \int_{-\infty}^{+\infty} e^{jux} p(x) dx \quad (2-39)$$

Si noti che la $\varphi(u)$ può anche considerarsi come media statistica della funzione e^{juX} . Si ha infatti ancora:

$$E [e^{juX}] = \int_{-\infty}^{+\infty} e^{juX} \cdot p(x) \cdot dx = \varphi(u)$$

La $\varphi(u)$ gode della proprietà $\varphi(0) = 1$, e $|\varphi(u)| \leq 1$, che derivano dalle proprietà della $p(x)$.

Interessante è la relazione qui seguente, che lega la $\varphi(u)$ ai momenti della variabile aleatoria. Se si sviluppa in serie la e^{juX} , si ottiene

$$e^{juX} = 1 + \frac{j u X}{1!} + \dots + \frac{j^m \cdot u^m \cdot X^m}{m!} + \dots \quad (2-40)$$

Se ora si opera la media statistica, membro a membro, si ottiene

$$\varphi(u) = 1 + j \frac{u}{1!} E[X] + \dots + \frac{j^m u^m}{m!} E[X^m] + \dots \quad (2-41)$$

Da tale relazione consegue che, noti tutti i momenti della variabile X , si conosce univocamente la funzione $\varphi(u)$ e quindi la $p(x)$, sua trasformata. Viceversa è possibile ricavare i momenti dalla funzione $\varphi(u)$. Precisamente, ammessa la convergenza della serie (2-41) per $|u| \leq u_0$ (con $u_0 \neq 0$), si ha:

$$E[X^k] = (j)^{-k} \cdot \left(\frac{d^k \varphi(u)}{du^k} \right)_{u=0} \quad (2-42)$$

Un'altra funzione derivata dalla $\varphi(u)$ è la funzione:

$$\psi(u) = \log_e \varphi(u) \quad (2-43)$$

Spesso è sufficiente arrestare la serie (2-41) al secondo termine. Si ha in tal caso:

$$\varphi(u) = 1 + juE[X] - \frac{E[X]^2 + \sigma^2}{2} u^2 \quad (2-44)$$

per la funzione $\psi(u)$:

$$\psi(u) = juE[X] - \frac{\sigma^2}{2} u^2 \quad (2-45)$$

Un'interessante applicazione della funzione caratteristica si ha nel caso di addizione di variabile aleatorie statisticamente indipendenti. Siano queste n variabili X_1, X_2, \dots, X_n , e siano $\varphi_1, \varphi_2, \dots, \varphi_n$ le rispettive funzioni caratteristiche. Si formi la variabile:

$$Y = X_1 + X_2 + \dots + X_n$$

La funzione caratteristica di Y è data da

$$\begin{aligned} \varphi(u) &= E[e^{juY}] = E[e^{ju(X_1 + X_2 + \dots + X_n)}] = \\ &= E[e^{juX_1}] \cdot E[e^{juX_2}] \cdot \dots \cdot E[e^{juX_n}] = \\ &= \varphi_1(u) \varphi_2(u) \cdot \dots \cdot \varphi_n(u) \end{aligned} \quad (2-46)$$

E per la funzione $\psi(u)$:

$$\psi(u) = \psi_1(u) + \psi_2(u) + \dots + \psi_n(u) \quad (2-47)$$

Sostituendo poi in quest'ultima lo sviluppo (2-45), si ottiene:

$$\sigma_y^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \dots + \sigma_{X_n}^2 \quad (2-48)$$

Ci riserviamo di giustificare più avanti l'uso dello sviluppo approssimato (2-45) nella (2-48).

E' quest'ultima una assai importante relazione, soprattutto quando le si assegni un significato energetico.

Si supponga a tale scopo che le variabili $X_1, X_2 \dots X_n$ siano attestate campionando, ad intervalli di tempo ΔT , n segnali aleatori prodotti da altrettanti generatori di rumore. La (2-48) mostra allora che la componente alternativa della potenza spettante alla somma di tutti gli n treni di campioni è uguale alla somma delle potenze di ognuno dei singoli treni.

Si consideri infine due variabili aleatorie indipendenti e continue X_1 ed X_2 , e si formi la variabile $Y = X_1 + X_2$. La $\varphi(u)$ della Y è data da:

$$\begin{aligned} \varphi(u) &= \varphi_1(u) \cdot \varphi_2(u) = E \left[e^{juX_1} \right] \cdot E \left[e^{juX_2} \right] = \\ &= \int e^{jux_1} \cdot p(x_1) dx_1 \cdot \int e^{jux_2} \cdot q(x_2) \cdot dx_2 = \\ &= \iint p(x_1) \cdot q(y-x_1) e^{juy} dx_1 \cdot dy \end{aligned} \quad (2-49)$$

Dalla (2-49) si vede, data la definizione di $\varphi(u)$, che questa è la trasformata della densità $\int p(x_1) q(y-x_1) dx_1$.

Del resto allo stesso risultato si poteva arrivare ricordando la proprietà delle trasformate di Fourier: $f_1 \otimes f_2 \rightleftharpoons F_1 \cdot F_2$. Quindi, date due variabili indipendenti aleatorie, la densità di probabilità della loro somma è data dall'integrale di convoluzione delle loro densità di probabilità, e la trasformata, ovvero la funzione caratteristica è data dal prodotto delle singole funzioni caratteristiche.

2-7 La Distribuzione Normale - o Legge di Gauss ad una dimensione.

La variabile aleatoria continua X è gaussiana (o Laplaciana) quando la sua densità di probabilità è del tipo:

$$p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-m_1)^2}{2\sigma^2}} \quad (2-50)$$

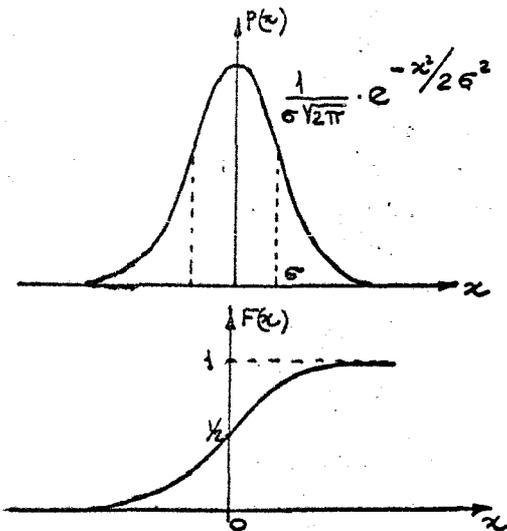
ove σ è la deviazione standard ed m_1 il valore statistico medio. Se X è una variabile centrata la (2-50) diventa:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \quad (2-51)$$

Tale densità di probabilità e la corrispondente funzione di distribuzione sono rappresentate in Fig. 13. La distribuzione dei valori della variabile aleatoria X che sia rappresentabile mediante la (2-50) è detta distribuzione normale o gaussiana.

Si calcoli ora la $\varphi(u)$ corrispondente alla densità (2-51) si ha:

$$\varphi(u) = E[e^{juX}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{juX} \cdot e^{-x^2/2\sigma^2} dx \quad (2-52)$$



Poniamo ora nella (2-52):

$$z = \frac{x}{\sigma\sqrt{2}} - ju \frac{\sigma}{\sqrt{2}} \quad (2-53)$$

da cui risulta:

$$\varphi(u) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-z^2} \cdot e^{-\frac{\sigma^2 u^2}{2}} \cdot dz = e^{-\frac{\sigma^2 u^2}{2}} \quad (2-54)$$

Fig. 13

La $\varphi(u)$ è quindi una funzione di andamento simile alla $p(x)$; cioè la trasformata di una funzione gaussiana è ancora una funzione di tipo gaussiano.

Per la $\psi(u)$ si ha:

$$\psi(u) = -\frac{\sigma^2}{2} u^2 \quad (2-55)$$

Ciò, la distribuzione normale o gaussiana è caratterizzata dal fatto che la $\psi(u)$ è una funzione di secondo grado in u . Inoltre tale funzione, e quindi la distribuzione, è completamente determinata quando si conosce il momento centrato del secondo ordine σ^2 .

Altre proprietà della distribuzione normale sono date dalle seguenti formule, che legano i momenti alla σ^2 .

$$E [X^2] = \sigma^2 \quad (X \text{ centrata})$$

$$E [X^{2k+1}] = 0.$$

$$E [X^{2k}] = \frac{(2k)!}{2^k \cdot (k)!} \sigma^{2k}$$

$$E [X^{2k+1}] = \frac{(k)!}{\sqrt{\pi}} \cdot (2 \sigma^2)^{\frac{2k+1}{2}} \quad (2-56)$$

2-8. Il Teorema del Limite Centrale

La distribuzione normale o gaussiana trova enorme riscontro nei fenomeni fisici; questo per il fatto che la somma di un numero grande di grandezze aleatorie indipendenti soddisfa quasi sempre alla distribuzione normale. E' questo il teorema del limite centrale, del quale si dà qui una prova più intuitiva che rigorosa. Si è accennato nel paragrafo 2-6 alla somma di variabili aleatorie indipendenti.

Consideriamo ancora n variabili aleatorie statisticamente indipendenti X_1, X_2, \dots, X_n ; siano esse, anche se non strettamente richiesto dalla presente dimostrazione, centrate ed abbiano stessa varianza σ^2 .

Consideriamo la variabile aleatoria:

$$Y = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i \quad (2-57)$$

per la quale si ha:

$$E [Y] = 0; \quad E [Y^2] = \sigma^2 \quad (2-58)$$

Se ψ_i è la funzione caratteristica della generica variabile X_i , si ha:

$$\psi_Y(u) = \sum_{i=1}^n \psi_i \left(\frac{u}{\sqrt{n}} \right) \quad (2-59)$$

Se si sostituisce alla ψ_i lo sviluppo (2-45), ove si indichi con Δ_i i termini oltre il secondo nello sviluppo in serie, si ottiene:

$$\psi_y(u) = -\frac{\sigma^2 u^2}{2} + \sum_{i=1}^n \Delta_i \quad (2-60)$$

Ora, per $n \rightarrow \infty$, e per $u < U_0$, il termine $\sum \Delta_i$ diviene infinitesimo rispetto ad u^2 , e $\psi_y(u)$ tende perciò uniformemente al valore $-\frac{\sigma^2 u^2}{2}$, che è quello caratteristico di una distribuzione gaussiana di deviazione standard σ .

Si noti che il Teorema del limite centrale può anche esprimersi in termini di convoluzione e di densità di probabilità, ovvero, una convoluzione multipla di un gran numero di densità di probabilità ha come risultante una densità rappresentabile con la (2-51).

Si noti inoltre che la convoluzione di due densità normali dà ancora una densità normale.

In elettronica il rumore caotico ha quasi sempre un comportamento di tipo gaussiano, nel senso che, se si forma una variabile aleatoria operando dei campioni di una tensione di rumore a carattere caotico, questa è descrivibile con una densità di probabilità gaussiana di varianza σ^2 , pari alla potenza media del rumore. La ragione per la quale il rumore, soprattutto quello di origine termica, ha carattere gaussiano, deriva dal fatto che è descrivibile come la somma di un numero grande di eventi caotici. Anche per quando il rumore non abbia carattere gaussiano, lo si può ricondurre a questo tipo mediante filtraggio passa basso. E' questo il caso del rumore impulsivo, come verrà analizzato più avanti.

2-9 Legge di Gauss a due Dimensioni
Distribuzione di Rayleigh

Si consideri ora l'insieme di due variabili aleatorie continue X ed Y, di densità p(x) e q(y) gaussiane. Sia inoltre per maggior semplicità:

$$E[X] = E[Y] = 0; \quad E[X^2] = E[Y^2] = \sigma^2;$$

La densità di probabilità bidimensionale è in questo caso esprimibile nel seguente modo:

$$f(x,y) = \frac{1}{2\pi\sigma^2 \cdot \sqrt{1-c^2}} \cdot e^{-\frac{x^2 + y^2 - 2cxy}{2\sigma^2(1-c^2)}} \quad (2-61)$$

ove $c = \frac{E[X \cdot Y]}{\sigma^2} =$ coefficiente di correlazione tra X ed Y.

Nel caso in cui $c = 0$ si ha dalla (2-61):

$$f(x,y) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{x^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{y^2}{2\sigma^2}} = p(x) \cdot q(y); \quad (2-62)$$

Ovvero, nel caso di variabili aleatorie gaussiane, la condizione $c = 0$ significa anche indipendenza statistica tra le due variabili, oltrechè indipendenza lineare, come visto al paragrafo 2-4. A tale proposito si vuole qui riportare per il caso generale la condizione di indipendenza statistica tra due generiche variabili aleatorie X ed Y; questa è esprimibile con:

$$E[X^n \cdot Y^m] = E[X^n] \cdot E[Y^m] \quad (2-63)$$

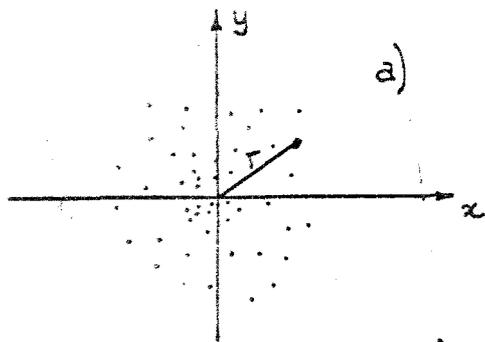
ove m ed n assumono tutti i valori interi positivi. Se si considera ora la funzione caratteristica bidimensionale $\varphi(u,v)$, trasformata della densità $f(x,y)$ nel dominio delle variabili reali u,v, la (2-63) può scriversi:

$$\varphi_{x,y}(u,v) = \varphi_x(u) \cdot \varphi_y(v) \quad (2-64)$$

Tornando ora alla legge di Gauss bidimensionale, la (2-61) è esprimibile mediante la sua trasformata nel seguente modo:

$$\varphi(u,v) = e^{-\frac{\sigma^2}{2} \cdot (u^2 + v^2 - 2cuv)} \quad (2-65)$$

Nel caso di indipendenza statistica tra X ed Y, $c = 0$, è spesso sufficiente conoscere la distribuzione della variabile $R = \sqrt{X^2 + Y^2}$. Ciò equi-



vale ad affermare che, se si considera nel piano x,y il punto $M(X,Y)$, spesso non interessa tanto la sua posizione nel piano, quanto la sua distanza dalla origine delle coordinate (Fig.14,b). E' questo ad esempio il caso della posizione dei colpi sul bersaglio del tiro a segno (Fig.14,a).

Più precisamente interesserà la probabilità $p(r).dr$ di tutti i punti compresi tra due circonferenze di raggio r ed $r + dr$.

Per ricavare la densità $p(r)$ si opera nel modo seguente. Si considera la probabilità $\int (x,y).dx.dy$, data dalla (2-62) ed in questa espressione si operano le sostituzioni $r^2 = x^2 + y^2$ e $dx.dy = r.dr.d\theta$; si ottiene infatti:

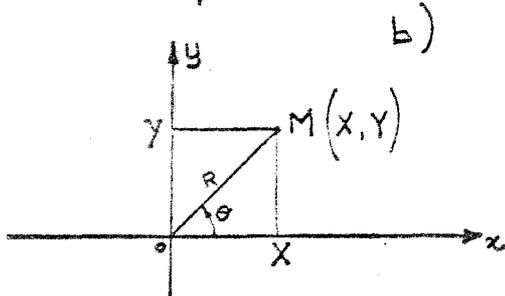


Fig. 14

$$\int (x,y).dx.dy = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}} \cdot dx.dy = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{r^2}{2\sigma^2}} \cdot r.dr.d\theta = \int (r,\theta).dr.d\theta \quad (2-66)$$

Ciò dalla $\int (x,y)$ si è passati alla $\int (r,\theta)$, ove:

$$\int (r,\theta) = \frac{1}{2\pi\sigma^2} \cdot r \cdot e^{-\frac{r^2}{2\sigma^2}} \quad (2-67)$$

Essendo ora X ed Y statisticamente indipendenti, tali saranno pure le variabili R e θ , per cui, facendo uso delle (2-20) e (2-21), si ricava:

$$p_\theta(\theta) = \int_0^\infty \frac{1}{2\pi\sigma^2} \cdot r \cdot e^{-\frac{r^2}{2\sigma^2}} dr = \frac{1}{2\pi} \quad (2-68)$$

$$p_r(r) = \int_0^{2\pi} \frac{1}{2\pi\sigma^2} \cdot r \cdot e^{-\frac{r^2}{2\sigma^2}} d\theta = \frac{r}{\sigma^2} \cdot e^{-\frac{r^2}{2\sigma^2}} \quad (2-69)$$

La (2-68) mostra che θ è uniformemente distribuita tra 0 e 2π . La (2-69) rappresenta invece la densità di probabilità corrispondente ad una distribuzione di Rayleigh.

Quest'ultima è ottenibile integrando la (2-69); ovvero, considerando la probabilità che sia $R \gg r$, si ha:

$$P[R \gg r] = \frac{1}{\sigma^2} \int_r^\infty R \cdot e^{-\frac{R^2}{2\sigma^2}} dR = e^{-\frac{r^2}{2\sigma^2}} \quad (2-70)$$

Le due funzioni (2-69) e (2-70) sono rappresentate nella Fig.15. Si noti come la funzione (2-70) abbia l'andamento di una curva gaussiana. Nella curva sono riportate in forma percentuale le probabilità che sia $R >$ di $0, \sigma, 2\sigma, \text{ e } 3\sigma$ rispettivamente. E' poi interessante notare che la deviazione standard di R è uguale a $\sqrt{2 \cdot \sigma} = \sigma_R$, il valore medio è $m_1 = \sigma \cdot \sqrt{\frac{\pi}{2}}$, e che il massimo della curva $p(r)$ si ha per $r = \sigma$.

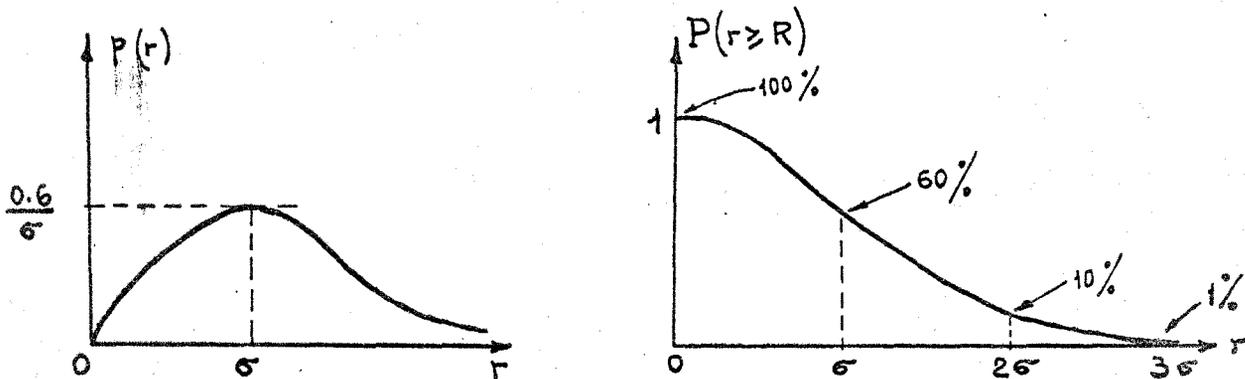


Fig.15

Si consideri ora un esempio di distribuzione di Rayleigh. Siano dati due segnali $X(t)$ ed $Y(t)$ tali che la distribuzione delle variabili aleatorie X ed Y ottenute campionando detti segnali sia gaussiana. Data poi una frequenza portante ω_0 , si crei il seguente segnale (Fig.16,a):

$$X(t) \cdot \cos \omega_0 t + Y(t) \cdot \sin \omega_0 t = \quad (2-71)$$

$$= S(t) \cos [\omega_0 t + \varphi(t)] \quad (2-72)$$

$$\text{ove } S(t) = \sqrt{X(t)^2 + Y(t)^2} \quad \text{e } \varphi(t) = \arctang \frac{-Y(t)}{X(t)} \quad (2-73)$$

La (2-72) rappresenta la forma generica di un segnale modulato di ampiezza e fase. In particolare, detta W la larghezza dello spettro di $X(t)$ ed $Y(t)$, se risulta:

$$F_0 = \frac{\omega_0}{2\pi} \gg W \quad (2-74)$$

allora il segnale (2-70) è del tipo a banda stretta (Fig.16,b). In tale caso il segnale può considerarsi involuppato dalla funzione $S(t)$, che rappresenta l'ampiezza del segnale. Se si considera ora la distribuzione statistica di detta ampiezza, contemporaneamente alla esperienza statistica fatta sui segnali $X(t)$ ed $Y(t)$, se le variabili aleatorie X ed Y sono statisticamente indipendenti,

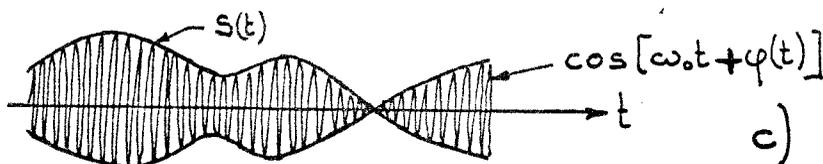
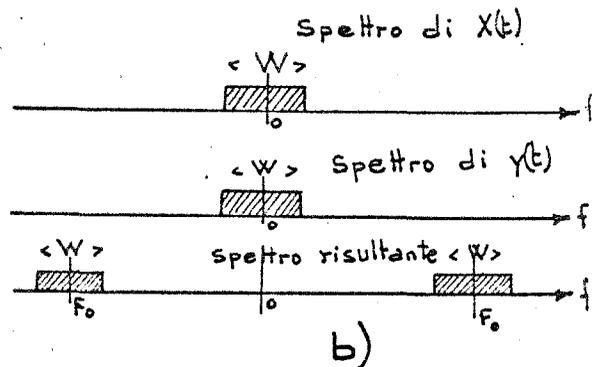
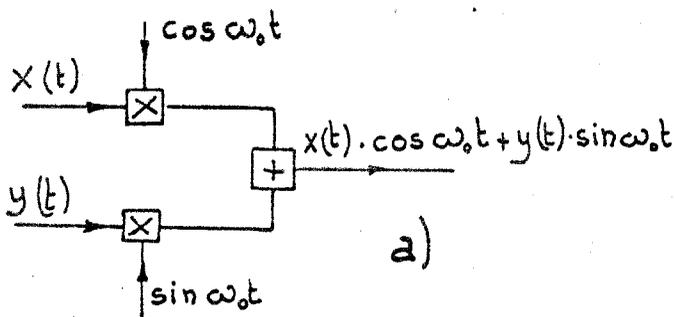


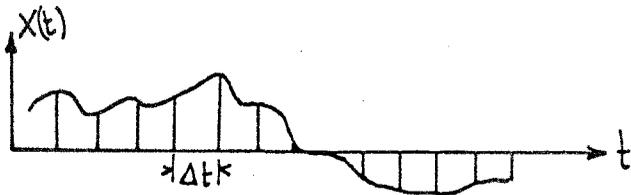
Fig. 16

dalla prima della (2-73) risulta che la distribuzione della $S(t)$ è del tipo di Rayleigh. Dalla seconda della (2-73) risulta poi che la fase $\varphi(t)$ è, come per la variabile φ , statisticamente uniformemente distribuita.

Osservazione. Sulla determinazione sperimentale della densità di probabilità.

Si è spesso richiamato nelle pagine precedenti il passaggio da una funzione $X(t)$ ad una variabile aleatoria X , della quale si vuole poter descrivere le proprietà statistiche.

Un metodo di passaggio è quello di campionare la funzione $X(t)$ ad intervalli Δt costanti. Ogni campione sarà descrivibile con un numero e rappresenterà un valore della variabile aleatoria X . Facendo poi un numero sufficientemente grande di campioni, sarà



possibile ottenere una attendibile descrizione statistica. Va fatto però attenzione alla scelta dell'intervallo di campionatura, il cui valore minimo è legato, come vedremo in seguito, alla larghezza di banda del segnale $X(t)$. E' infatti intuibile qui come esista una distanza minima, al di sotto della quale due campioni consecutivi non sono più statisticamente indipendenti.

2-10 Legge di Gauss ad n Dimensioni

Generalizzando la legge di Gauss a due dimensioni, si può affermare che n variabili centrate X_1, X_2, \dots, X_n sono gaussiane nel loro insieme se la funzione caratteristica corrispondente è della forma:

$$\varphi(u_1, u_2, \dots, u_n) = e^{-\frac{1}{2} \sum c_{ij} \cdot u_i \cdot u_j} \quad (2-75)$$

ove $c_{ij} = E[X_i \cdot X_j]$

Perciò la legge di Gauss ad n dimensioni è interamente determinata dalla conoscenza delle $E [X_i \cdot X_j]$.

Trasformando la (2-75) si ottiene la densità multidimensionale $f(x_1, x_2, \dots, x_n)$.

La indipendenza statistica tra le variabili è assicurata dalla $C_{ij} = 0$ per $i \neq j$. In tal caso la funzione caratteristica è uguale al prodotto delle funzioni caratteristiche e similmente la densità di probabilità è uguale al prodotto delle singole densità.

Dalla $C(u_1, u_2, \dots, u_n)$ si possono ricavare i momenti che godono delle particolarità seguenti:

$$E [X_i] = 0 \quad (\text{dalla ipotesi di variabili centrate})$$

$$E [X_i \cdot X_j] = C_{ij}$$

$$E [X_i \cdot X_j \cdot X_k] = 0$$

$$E [X_i \cdot X_j \cdot X_k \cdot X_L] = E [X_i \cdot X_j] \cdot E [X_k \cdot X_L] + E [X_i \cdot X_k] \cdot E [X_j \cdot X_L] + E [X_i \cdot X_L] \cdot E [X_j \cdot X_k]$$

$$E [X_1 \cdot X_2 \dots X_{2k+1}] = 0$$

$$E [X_1 \cdot X_2 \dots X_{2k}] = \sum \prod E [X_i \cdot X_j] \quad (2-76)$$

Altre Osservazioni

- La somma di variabili aleatorie gaussiane è ancora una variabile aleatoria gaussiana.
- Il prodotto di variabili aleatorie gaussiane non ha più carattere gaussiano. In particolare nel caso di due variabili aleatorie gaussiane centrate, il loro prodotto ha la densità di probabilità:

$$p(x) = \frac{1}{\pi \sigma_1 \sigma_2} \cdot \gamma_0 \left(\frac{x}{\sigma_1 \sigma_2} \right) \quad (2-77)$$

ove $\gamma_0 \left(\frac{x}{\sigma_1 \sigma_2} \right)$ è la funzione di Bessel di ordine zero e di seconda specie.

Si esamina la prima di due tipi di distribuzioni discontinue tra loro collegate, quelle di Bernoulli e di Poisson.

Si consideri il caso in cui un evento, risultato di una esperienza aleatoria, ha due sole alternative, sì o no (+ 0 -, 1 o 0) e siano rispettivamente p ed $1-p$ probabilità dei due possibili valori.

Se si operano ora n sorteggi, ci si chiede quale è la probabilità di ottenere k volte il valore ad esempio 1, essendo 0 l'altra alternativa. Se si fa il caso in cui $n = 5$, la probabilità che si ottenga ad esempio la sequenza 01101, con $k = 3$, sarà:

$$(1-p) \cdot p \cdot p \cdot (1-p) \cdot p = p^3 \cdot (1-p)^2 \quad (2-78)$$

Ora vi siano ${}^5C_3 = \frac{5 \cdot 4 \cdot 3}{3!}$ combinazioni con lo stesso numero di 1 e 0, per cui la probabilità totale di ottenere $k = 3$ volte il valore 1 in $n = 5$ sorteggi è data da ${}^5C_3 \cdot p^3 \cdot (1-p)^2$

In generale, la probabilità di k eventi dello stesso valore in n sorteggi è data da:

$$P_n(k) = {}^nC_k \cdot p^k \cdot (1-p)^{n-k} \quad (2-79)$$

ove ${}^nC_k = \frac{n!}{k! \cdot (n-k)!}$

La (2-79) rappresenta il Teorema di Bernoulli.

Se si considera il Teorema Binomiale, che dà l'espansione del Binomio $(p+q)^n$, e si pone $1-p = q$ si ottiene:

$$(1-p)^n + \dots + {}^nC_k \cdot p^k \cdot (1-p)^{n-k} + \dots + p^n = [p+(1-p)]^n = 1 \quad (2-80)$$

La (2-80) rappresenta la probabilità, ovviamente uguale ad 1, di ottenere un qualsiasi numero di eventi k tra 0 ed n . I termini nC_k sono quindi i coefficienti della espansione binomiale.

Se si considera ora il valore medio e la varianza della distribuzione binomiale (2-79), si ha:

$$- \text{valor medio} = n \cdot p = m_{1k}$$

$$- \text{varianza} = \sigma_k^2 = n \cdot p(1-p) = m_{1k} \cdot (1-p)$$

da cui si deduce che σ_k è proporzionale alla $\sqrt{m_{1k}}$.

Ciò indica che, al crescere del numero di eventi n , a parità di p , la distribuzione diviene, relativamente al valore medio, sempre più stretta; il rapporto cioè dispersione/valore medio diventa sempre più piccolo. Conseguentemente il rapporto k/n diventa sempre più vicino al valore p , al crescere di n .

E' questo un importante risultato che rappresenta la "legge delle medie" (o dei grandi numeri).

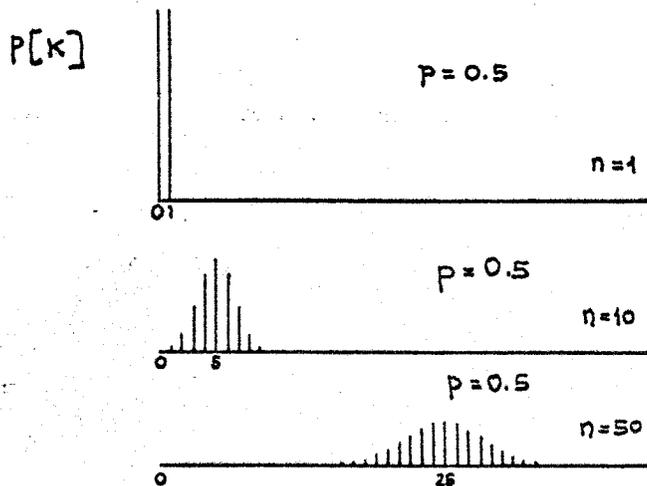


Fig. 18

Nella Fig. 18 sono rappresentati tre casi di distribuzione di Bernoulli, rispettivamente per $n = 1, 10$ e 50 sorteggi di risultato equiprobabile ($p = 0,5$). E' evidente dalla figura come il rapporto dispersione/numero di eventi diminuisca

al crescere di n . Per di più, al tendere di n ad infinito, la distribuzione di Bernoulli tende ad una distribuzione Gaussiana. Ciò può essere mostrato attraverso tediosi sviluppi, che danno come risultato (per $n \rightarrow \infty$):

$$p(k) = \frac{1}{\sqrt{2\pi} \sigma_k} \cdot e^{-\frac{(k-m_1k)^2}{2\sigma^2}} \quad (2-81)$$

ove $m_{1k} = p \cdot n$, e $\sigma_k = n \cdot p(1-p)$.

Può ancora mostrarsi come, da questa distribuzione discontinua, si possa giungere ad una densità di probabilità Gaussiana, come definito dalla (2-50).

Consideriamo ora un esempio di distribuzione di Bernoulli. Si supponga di disporre di una sorgente di informazione del tipo binario (1, 0) che fornisca ad ogni intervallo Δt un nuovo campione. Sia p la probabilità dello stato 1, e sia lo stato di ogni campione statisticamente indipendenti dagli altri. Si immagazzinino ora questi segnali sequenzialmente in una memoria digitale (shift register) della capacità di n campioni, e capace di spostare, mediante un comando ad ogni intervallo Δt , il treno di segnali.

Ad ogni comando di spostamento, un nuovo campione entra nella memoria, mentre viene eliminato quello immagazzinato $n \cdot \Delta t$ secondi prima

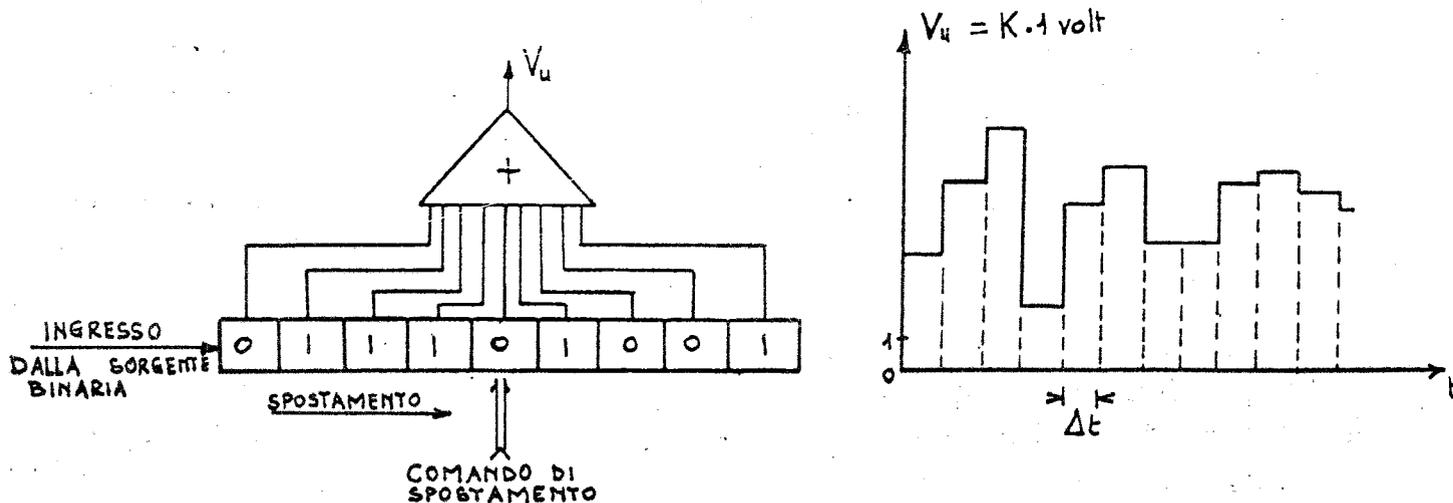


Fig. 19

Un tale sistema è schematizzato in Fig. 19, ove si intende che lo spostamento dei campioni avvenga da sinistra a destra. Si supponga per semplicità che allo stato 0 corrisponda 0 volts, ed allo stato 1, 1 volts. Tali saranno quindi le tensioni immagazzinate nella cella di memoria.

Se ora si sommano tra loro le tensioni delle n celle, la tensione V_u all'uscita sarà uguale, per ogni intervallo Δt , a $k.1$, ove k è il numero di celle nello stato 1. La distribuzione di probabilità di k è ovviamente descrivibile con la distribuzione di Bernoulli, ed in tal modo quindi è anche descrivibile la variabile aleatoria V_u .

Il risultato della V_u ad ogni istante Δt può poi considerarsi il risultato di una esperienza aleatoria. La V_u è rappresentata in Fig. 19,6.

Si può poi normalizzare la tensione V_u al numero delle celle n , ovvero si consideri la $V' = \frac{V_u}{n}$.

In tal caso il valore statistico medio della V' è dato da $\frac{n.p}{n} = p$, e la deviazione standard è uguale a $\frac{\sigma}{n} = \frac{\sqrt{n.p(1-p)}}{n} = \text{cost.} \frac{1}{\sqrt{n}}$. Quindi, al crescere di n , la $\frac{\sigma}{n}$ diminuisce come $\frac{1}{\sqrt{n}}$, e la $\frac{V_u}{n}$ tende al valore statistico medio p .

Questo sistema ora esaminato, combinazione dello "shift register" con un sistema di somma, presenta interessanti applicazioni nel campo del trattamento dei segnali, come si avrà modo di esaminare in seguito.

2-12 La Distribuzione di Poisson -

Questa distribuzione è derivata da quella di Bernoulli nel caso in cui sia n molto grande e $p \ll 1$.

Si consideri perciò il fenomeno della emissione degli elettroni dal catodo di un tubo elettronico, e sia m il numero medio di elettroni in un secondo.

Si divida ora l'intervallo di tempo di un secondo in n intervalli molto piccoli, tali che sia trascurabile la probabilità di avere più di un elettrone in ogni intervallo. La probabilità di avere affatto un elettrone in un intervallo è ora data da $p = \frac{m}{n}$. Quindi la probabilità di avere k elettroni in n intervalli segue la distribuzione di Bernoulli. Se ora si è fatto n molto grande sarà anche $p \ll 1$. In tal caso la (2-79) può scriversi nel seguente modo:

$$p_n(k) = \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k} = \frac{1}{k!} \left[n(n-1)(n-2) \dots (n-k+1) \right] \cdot \left(\frac{m}{n}\right)^k \cdot \left(1 - \frac{m}{n}\right)^n \cdot \left(1 - \frac{m}{n}\right)^{-k} =$$

$$\approx \frac{1}{k!} \cdot n^k \cdot \left(\frac{m}{n}\right)^k \cdot e^{-m} = e^{-m} \cdot \frac{m^k}{k!}$$

E' questa la distribuzione di Poisson o funzione di frequenza che gode delle seguenti proprietà:

$$E[k] = m; \quad E[k]^2 = m^2 + m; \quad \sigma^2 = m = E[k];$$

$$\varphi(u) = \exp[m(e^{ju} - 1)]; \quad \psi(u) = m(e^{ju} - 1) \quad (2-83)$$

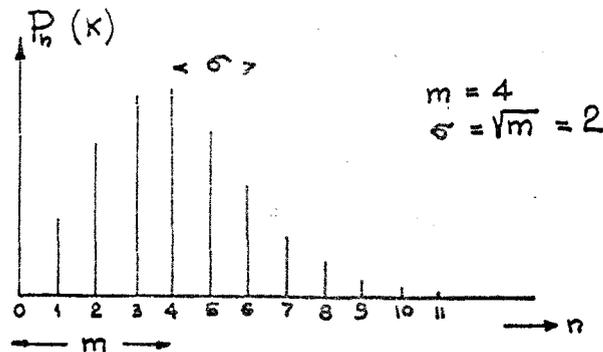


Fig. 20

Nella Fig. 20 è illustrato un esempio di distribuzione di Poisson, con $m = 4$. Quando $m \gg 1$, allora la distribuzione di Poisson diviene molto stretta. Ad esempio, tornando all'esempio della emissione di elettroni dal catodo, una corrente di 100 mA corrisponde ad $m = 6,3 \times 10^8$ elettroni per 10^{-9} sec. Ora la possibilità di avere nel tempo 10^{-9} sec il numero m esatto di elettroni è solo di $1,59 \times 10^{-5}$. Tuttavia, la probabilità di avere un numero che differisca da m del 1% è di 10^{-13700} .

2-13 Distribuzione Esponenziale

Un'altra distribuzione, derivabile da quella di Poisson, è la distribuzione esponenziale. Consideriamo ancora l'esempio visto prima della emissione elettronica, con m elettroni al secondo, e si sia operata la suddivisione in intervalli $\Delta t = \frac{1 \text{sec}}{n}$ uguali, e molto piccoli.

Si rappresenti con 1 o 0 la presenza o meno di un elettrone libero in tale intervallo. Il fenomeno della emissione sarà allora descrivibile con una successione di zeri, intercalati con alcuni 1:

0010000.....001000.....

0.01234..... $\frac{1}{s}$

Sia ora $s \cdot \Delta t$ l'intervallo tra due 1. Questo conterrà $s-1$ zeri. La probabilità di tale intervallo sarà quindi data dal prodotto della probabilità di avere $s-1$ zeri, con la probabilità di avere un 1; ovvero:

$$P(s) = (1-p)^{s-1} \cdot p \quad (2-84)$$

Se ora si sostituisce nella (2-84) $p = m \cdot \Delta t$, e $t = s \cdot \Delta t$, si ottiene:

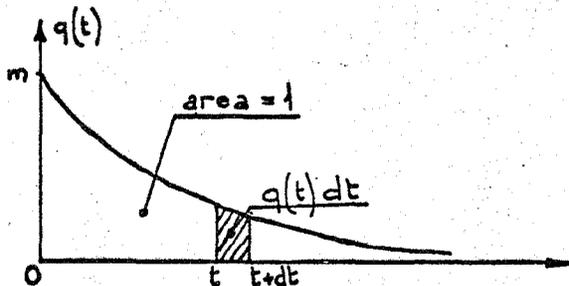
$$P(s) = q(t) \cdot \Delta t = \left(1 - \frac{mt}{s}\right)^{s-1} \cdot m \cdot \Delta t \quad (2-85)$$

ove $q(t)$ è la densità di probabilità definita da:

$$q(t) = \frac{P(s)}{\Delta t}$$

Se ora si fa $s \rightarrow \infty$, ovvero $n \rightarrow \infty$, nella (2-85), si ottiene:

$$q(t) \cdot dt = m \cdot e^{-mt} \cdot dt \quad (t > 0) \quad (2-86)$$



La (2-86) rappresenta la distribuzione di probabilità esponenziale. Infatti tale è il tipo della funzione $q(t)$, Fig.21. La probabilità $q(t) \cdot dt$ rappresenta quindi la probabilità di avere due eventi successivi distanti il tempo t , con la tolleranza dt .

Fig. 21

Si noti come, a parità di intervallo dt , tale probabilità diminuisca al crescere di t . Ciò può essere spiegato nel seguente modo. La probabilità $q(t).dt = P(s)$ rappresenta la probabilità che l'intervallo tra due 1 abbia una certa dimensione t . E' ora evidente che, più grande è t , minore è la probabilità che in tale intervallo non si verifichi nessun altro evento di tipo 1 . Inoltre al tendere di t a zero, riducendo t ad un intervallo infinitesimale dt successivo al primo evento 1 di partenza per $t = 0$, la probabilità di un evento 1 in tale intervallo dt è ancora data da $p = m.dt$, essendo questa la probabilità di un evento in ogni intervallo infinitesimale. Ciò giustifica il fatto che $q(0) = m$.

2-14 Le Funzioni Aleatorie - Processi Aleatorii o Stocastici

Finora si è trattato delle variabili aleatorie, definite su di un vasto numero di esperienze di un dato fenomeno, il risultato di ognuna delle quali è una grandezza numerica. Se ora il risultato di ogni esperienza è invece una intera funzione, determinata od aleatoria, si tratta di un processo aleatorio o stocastico, e l'insieme dei membri del processo stocastico, risultato ognuno di una data esperienza, costituisce la funzione aleatoria. Generalmente le funzioni, risultato di ogni esperienza, sono funzioni del tempo t , per cui si può indicare la funzione aleatoria come $X(t, \mathcal{E})$, funzione del tempo t e della esperienza \mathcal{E} .

Con tale simbolismo $X(t_0, \mathcal{E})$ è la variabile aleatoria definita sull'insieme dei valori dei vari membri del processo al tempo $t = t_0$.

Analogamente $X(t, \mathcal{E}_0)$ rappresenta uno dei membri del processo stocastico, che ha perduto il carattere di aleatorietà relativo alla esperienza \mathcal{E} generica.

Lo studio della funzione aleatoria $X(t, \mathcal{E})$ è lo studio dell'insieme delle variabile aleatorie $X(t_0, \mathcal{E})$, ottenibili per tutti i possibili valori di t_0 .

Si ometterà nel seguito, riferendoci alla funzione aleatoria, il simbolo \mathcal{E} e si indicherà questa con $X(t)$.

Un esempio di processo stocastico è quello costituito dai segnali d'uscita $X_1(t), X_2(t), \dots, X_n(t)$, di n generatori di rumore, Fig.22.

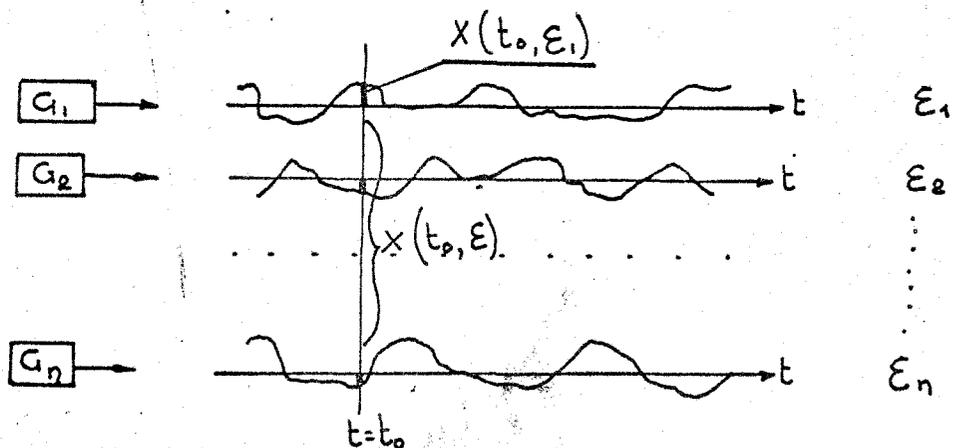


Fig.22

L'esperienza \mathcal{E} consiste qui nella scelta di un particolare generatore.

Le proprietà statistiche della funzione aleatoria $X(t)$ sono note quando si conoscono tutte le possibili distribuzioni di probabilità a k dimensioni, ottenute campionando la funzione aleatoria in un numero finito k di punti, t_1, t_2, \dots, t_k , con k qualunque. Ad ogni scelto istante t_i , tutti i membri del processo stocastico vengono campionati a formare la variabile aleatoria $X(t_i)$.

Ad ogni insieme t_1, t_2, \dots, t_k corrisponde un insieme di k variabili aleatorie $X(t_1), X(t_2), \dots, X(t_k)$.

Se si ammette che ogni variabile sia continua e che esista una densità di probabilità a k dimensioni, le proprietà stocastiche di $X(t)$ saranno note quando si conosce, per tutti gli insiemi t_1, t_2, \dots, t_k la funzione:

$$p(X_1, X_2, \dots, X_k; t_1, t_2, \dots, t_k) \quad (2-87)$$

ovvero la funzione:

$$\varphi(u_1, u_2, \dots, u_k; t_1, t_2, \dots, t_k) = E \left\{ e^{j(u_1 X_1 + \dots + u_k X_k)} \right\} \quad (2-88)$$

Similmente al caso delle variabili aleatorie, è spesso possibile descrivere le proprietà statistiche della funzione aleatoria $X(t)$ mediante i momenti.

Se si considera la variabile aleatoria $X(t_i)$, ottenuta campionando i membri del processo al tempo $t = t_i$, si potrà per questa considerare i vari momenti, generalmente funzione del tempo di campionatura t_i , questi si indicano con $E[X(t)] = \overline{X(t)}$, $E[X^2(t)] = \overline{X^2(t)}$, etc. ove al simbolo t_i si sostituisce genericamente t . Si intende infatti che la operazione di media statistica $E[\]$ viene operata sui vari membri del processo, allo stesso tempo t .

Se si considera ora l'insieme di due variabili aleatorie $X(t_1)$ e $X(t_2)$, si potrà considerare il momento del secondo ordine:

$$E [X(t_1) \cdot X(t_2)] = \overline{X(t_1, t_2)} \quad (2-89)$$

E' questa generalmente funzione dei tempi t_1 e t_2 .

La (2-89) prende generalmente il nome di funzione di autocorrelazione statistica, e coincide con la covarianza nel caso di variabili centrate. Tali le si considereranno in seguito, salvo avviso.

Si consideri poi l'insieme t_1, \dots, t_k ; è possibile similmente introdurre il momento d'ordine k :

$$E [X(t_1) \cdot X(t_2) \dots \dots X(t_k)]$$

La condizione di indipendenza tra due funzioni aleatorie $X(t)$ ed $Y(t)$ è inoltre assicurata dalla indipendenza dei due insiemi $X(t_1), \dots, X(t_k)$ ed $Y(t_1'), \dots, Y(t_{k'})$, ottenuti campionando le due funzioni aleatorie agli istanti t_1, \dots, t_k e t_1', \dots, t_{k}' rispettivamente, con k e k' qualunque.

2-15 Stazionarietà ed Ergodicità

Si consideri ora il caso in cui tutte le proprietà statistiche della funzione aleatoria $X(t)$ sono invarianti per uno spostamento dell'origine dei tempi. In tal caso, il processo si dice stazionario in senso stretto, e si ha che:

$$p(X_1 \dots X_k; t_1, \dots, t_k) = p(X_1, \dots, X_k; t_1 + \tau, \dots, t_k + \tau) \quad (2-90)$$

ovvero le proprietà statistiche di $X(t)$ sono invarianti per una traslazione τ .

In particolare:

$E[X^n(t)] = \overline{X^n(t)}$, con n intero qualunque, è indipendente da t . Inoltre:

$$\Gamma(t_1, t_2) = \Gamma(t, t-\tau) = E[X(t) \cdot X(t-\tau)] \quad (2-91)$$

Ovvero la covarianza è solo funzione di $\tau = t_2 - t_1$, e non dei valori di t_1 e t_2 .

Il caso della stazionarietà del processo aleatorio riveste grande importanza in quanto permette notevoli semplificazioni.

Di tale tipo possono considerarsi molti fenomeni fisici, specialmente quando il tempo di osservazione non è troppo grande. Un esempio di non stazionarietà può considerarsi quello degli n generatori di rumore della figura 22 nel periodo successivo alla accensione, durante il raggiungimento della temperatura di regime.

Tra i processi stazionari riveste notevole importanza il caso di quei processi nei quali è possibile dedurre le proprietà statistiche del processo da quelle di un suo qualunque membro. E' questo il caso dei processi Ergodici. Si può sostituire alla osservazione simultanea di un gran numero di fenomeni simili quella di un solo di essi, per un tempo molto lungo. Se si indica con f una generica funzione certa e si crea la funzione di funzione aleatoria $f[X(t)]$, si ha che per i processi ergodici la media statistica e quella temporale si equivalgono, ovvero:

$$\overline{f[X(t)]} = E\{f[X(t)]\} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f[X(t)] \cdot dt = \overline{f[X(t)]} \quad (2-92)$$

Ove i simboli $\overline{\quad}$ ed $\overline{\quad}$ rappresentano media statica e temporale rispettivamente.

In particolare si ha, per processi ergodici:

$$\overline{X(t)} = \overline{X(t)} \quad (2-93)$$

$$\Gamma(t, t-\tau) = \overline{X(t) \cdot X(t-\tau)} = \overline{X(t) \cdot X(t-\tau)} = c(\tau), \quad (2-94)$$

ove

$$\overline{X(t) \cdot X(t-\tau)} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T X(t) \cdot X(t-\tau) dt = C(\tau) \quad (2-95)$$

è la funzione di autocorrelazione di $X(t)$.

Si noti che un processo ergodico è sempre stazionario, ma non viceversa.

Come esempio di funzione aleatoria stazionaria ma non ergodica si può considerare la funzione $X(t) = A \cos \varphi(t)$, ove $\varphi(t)$ varia in modo da produrre un segnale a modulazione di fase aleatoria, mentre A è una variabile aleatoria che assume i valori 1 e 2.

Si noti che, sia il processo stazionario o non, è in generale:

$$\overline{X(t)} = \overline{X(t)} \quad (2-96)$$

E' spesso sufficiente che il processo sia stazionario del secondo ordine per considerarlo stazionario, vale a dire che siano invarianti i momenti del primo e del secondo ordine per uno spostamento dell'asse dei tempi. E' cioè sufficiente che siano:

$$\overline{X(t)} \text{ indipendente da } t, \text{ e } \overline{X(t_1) X(t_2)} = C(\tau) \text{ funzione solo di } \tau.$$

Come esempio di funzione aleatoria stazionaria al second'ordine si può considerare la funzione:

$$X(t) = A \cos \omega t + B \sin \omega t \quad (2-97)$$

ove A e B sono variabili aleatorie centrate ed indipendenti ed ω è una costante. Si può dimostrare in tal caso che è:

$$E[X(t)] = 0; \quad E[X(t) \cdot X(t-\tau)] = E[A^2] \cos \omega \tau.$$

vale a dire la funzione è stazionaria del second'ordine, mentre risulta $E[X^4(t)]$ funzione di t .

2-16 Le Funzioni Aleatorie Gaussiane

La funzione aleatoria $X(t)$ è gaussiana quando, dato l'insieme t_1, t_2, \dots, t_k , con k qualunque, l'insieme di k variabili aleatorie $X(t_1), X(t_2), \dots, X(t_k)$, obbedisce ad una legge di Gauss a k dimensioni. Nel caso in cui

la funzione $X(t)$ è centrata, la funzione caratteristica è della forma:

$$\varphi(u_1, u_2, \dots, u_k) = e^{\frac{1}{2} \sum_i \sum_j \Gamma(t_i, t_j) \cdot u_i \cdot u_j} \quad (2-98)$$

Cioè le proprietà statistiche di una funzione aleatoria gaussiana centrata $X(t)$ sono completamente determinate dall'insieme delle covarianze $\Gamma(t_i, t_j)$ con i ed j qualunque.

Nel caso di ergodicità la covarianza è uguale alla funzione di autocorrelazione $C(\tau)$, che è perciò sufficiente a stabilire le proprietà statistiche della funzione aleatoria, ovvero a determinare la densità di probabilità di qualsiasi ordine.

Si è visto, per le variabili aleatorie unidimensionali, che la somma di un gran numero di variabili aleatorie indipendenti dà luogo con buona approssimazione ad una variabile gaussiana.

Analogamente si può dimostrare che se si considera un gran numero di variabili multidimensionali X' (X_1' , X_2' , ..., X_k'), X'' (X_1'' , X_2'' , ..., X_k'')..., indipendenti e le si sommano assieme, la variabile risultante $X(\sum_i X_1^i, \sum_i X_2^i, \dots, \sum_i X_k^i)$ seguirà una legge di Gauss a k dimensioni.

Tutte le volte quindi che $X(t)$ può considerarsi come la somma di un gran numero di contributi elementari indipendenti e dello stesso peso sull'insieme, allora $X(t)$ sarà una funzione aleatoria Gaussiana.

Si consideri ad esempio il rumore di fondo di un amplificatore dovuto alla corrente d'ingresso. Questa può considerarsi costituita da cariche elementari distribuite secondo Poisson, ad intervallo medio $\Delta t_i = 1/\rho$. Se q è la carica elementare ed I la corrente d'ingresso, sarà $\rho = I/q$.

Sia θ la costante di tempo dell'amplificatore.

La tensione d'uscita dell'amplificatore $X(t)$ può considerarsi come un membro di una funzione aleatoria, le cui caratteristiche differiranno perciò a seconda del valore di $\rho \cdot \theta$. Infatti, se $\rho \cdot \theta \gg 1$ allora la tensione d'uscita $X(t)$ sarà la risultante di un gran numero di impulsi elementari prodotti dalle singole cariche, di modo che presenterà carattere gaussiano. Se invece non è $\rho \cdot \theta \gg 1$, allora sarà possibile per l'amplificatore risolvere parzialmente o non i vari impulsi elementari e la tensione d'uscita perderà il carattere gaussiano. Si noti che in tal caso una ulteriore operazione di filtraggio

passa-basso, oppure di integrazione, che riporti il sistema alla condizione $\beta \theta \gg 1$, trasformerà il segnale nuovamente in un segnale gaussiano. Generalmente, la misura di fenomeni microscopici con apparati di misura, la cui frequenza di risposta è notevolmente inferiore a quella media di occorrenza dei fenomeni, dà luogo alla misura di funzioni aleatorie gaussiane.

E' interessante nel caso di processo ergodico centrato, mettere in evidenza la relazione esistente tra la funzione di autocorrelazione $C(\tau)$ e la densità di probabilità del secondo ordine $p(X_1, X_2)$.

Si consideri infatti le due variabili aleatorie $X_1 = X(t_1)$ ed $X_2 = X(t_2)$. Per essa è valida la relazione (2-61), ovvero:

$$p(X_1, X_2) = \frac{1}{2\pi\sigma^2\sqrt{1-c^2}} \cdot e^{-\frac{X_1^2 + X_2^2 - 2cX_1X_2}{2\sigma^2(1-c^2)}} = p(X_1, X_2; \tau) \quad (2-99)$$

ove ora, grazie alla ipotesi di ergodicità:

$$c = \frac{E[X_1 \cdot X_2]}{\sigma^2} = \frac{C(\tau)}{C(0)} \quad (2-100)$$

Si noti che $C(0) = \overline{X^2(t)} = \overline{X^2(t)} = \sigma^2$;

$$E' \text{ poi: } p(X_1) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-X_1^2/2\sigma^2} \quad (2-101)$$

da cui si ha per la probabilità condizionale:

$$p(X_2/X_1) = \frac{p(X_1, X_2; \tau)}{p(X_1)} = \frac{1}{\sqrt{2\pi} \cdot \sigma_z} \cdot e^{-\frac{(X_2 - cX_1)^2}{2\sigma_z^2}} \quad (2-102)$$

$$\text{ove } \sigma_z = \sigma \sqrt{1-c^2}.$$

Quindi la densità di probabilità della X_2 , noto il valore della X_1 , è gaussiana, con valore medio cX_1 e varianza $\sigma_z^2 = \sigma^2 \sqrt{1-c^2}$.

Nella figura 23 è rappresentata la variazione della $p(X_2/X_1)$, per un dato valore di X_1 , in funzione di τ , e per una tipica funzione $c(\tau)$.

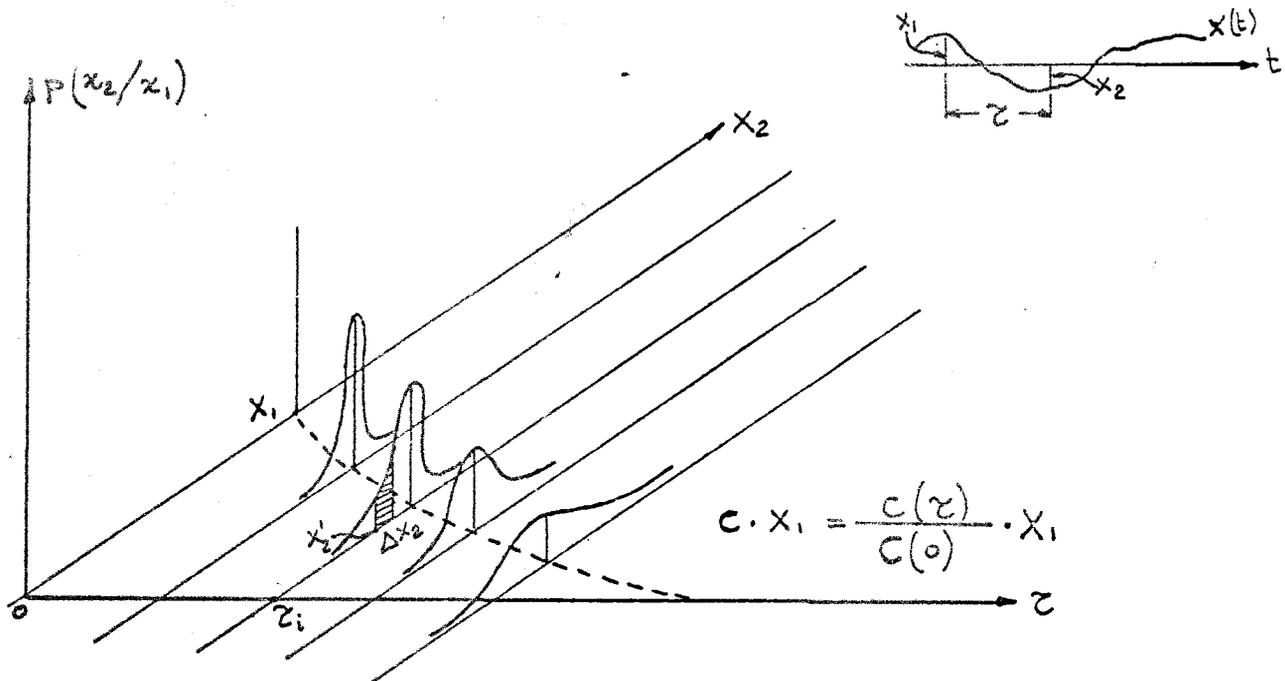


Fig. 23

Per $\tau = 0$ si conosce con probabilità 1 il valore di $X_2 = X_1$. Perciò la densità di probabilità è rappresentata da una funzione $\delta(X)$.

Per $\tau = \tau$: la probabilità che sia $X_2^1 \leq X_2 \leq X_2^1 + \Delta X_2$ è data dall'area tratteggiata di figura. A parità di ΔX_2 tale probabilità è massima intorno al valore $\frac{c(\tau)}{c(0)} \cdot X_1$.

La variazione di tale probabilità al crescente di τ dipende solo dalla $c(\tau)$, secondo la (2-102). Generalmente al crescere di τ i valori della X_2 tendono all'indipendenza rispetto ad X_1 . Più rapida è la pendenza di $c(\tau)$, più rapidamente diminuisce la dipendenza di X_2 da X_1 .

La indipendenza tra X_2 ed X_1 è assicurata quando $c(\tau) = 0$. In tal caso la (2-102) diviene uguale alla (2-101); ovvero $\sigma_2 = \sigma$, e $c \cdot X_1 = 0$.

E' quindi evidente l'intima relazione esistente tra la $p(X_2/X_1)$ e la funzione di auto correlazione $c(\tau)$. Quest'ultima rappresenta quindi una misura della rapidità con la quale i successivi valori del processo divergono indipendenti dai precedenti. E' qui intuitivo, e ci riserviamo di dimostrarlo nel prossimo capitolo, come a maggior ripidità della funzione $c(\tau)$, ovvero più rapida indipendenza, debba corrispondere la presenza di componenti a frequenza maggiore nello spettro del segnale.

Infine nella figura 24 sono riportati i contorni a densità di probabilità bidimensionale $p(X_1, X_2) =$ costante, per i valori di $C = 0$, $C = 1$ e $0 < C < 1$.

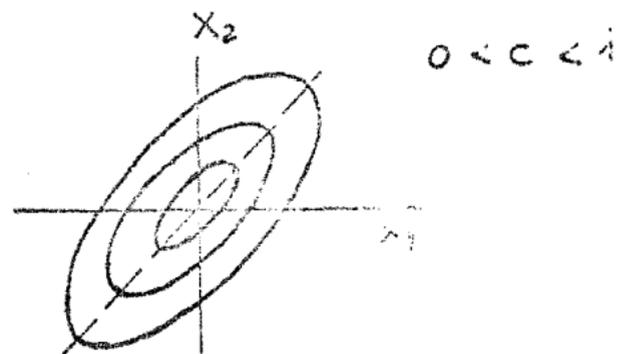
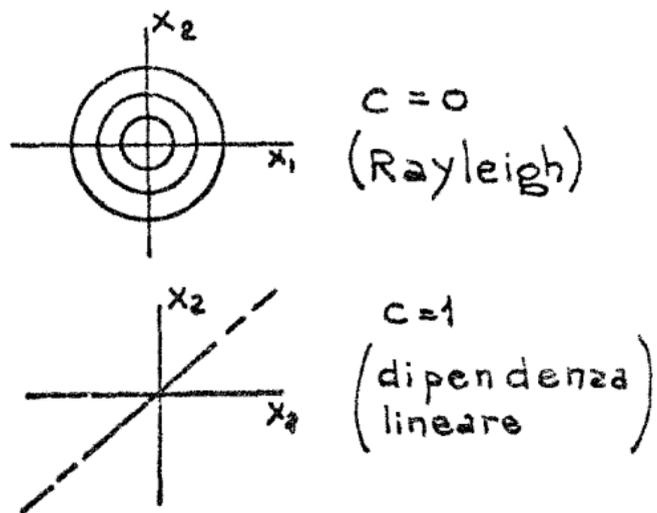


Fig. 24