ROMA
TRE
Università degli studi di Roma Tre
Facoltà di Ingegneria

Appunti dalle lezioni del corso di

Teoria dei Segnali

VARIABILI ALEATORIE

Prof. Alessandro NERI

Roma, novembre 2001

INDICE

Variabile Aleatoria

L'introduzione dell'associazione tra i risultato del fenomeno aleatorio ed i punti dello spazio Ω consente di definire la probabilità come funzione d'insieme che, in linea di principio, ci consente di calcolare, tramite i teoremi che sono stati introdotti, la probabilità di ogni evento ammissibile. Nella realtà tale tipo di formulazione trova dei limiti dovuti alla scarsa operatività di tale formulazione.

L'introduzione del concetto di "variabile aleatoria" ci consente invece di esprimere le probabilità degli eventi tramite le funzioni di punto che sono le uniche rappresentabili agevolmente.

A tal scopo definiamo come variabile aleatoria X una trasformazione dallo spazio Ω in $\Re^n$ spazio euclideo n – dimensionale

$$X:\ \Omega \rightarrow \Re^n \tag{63}$$

che soddisfi le particolari condizioni che ora specificheremo.

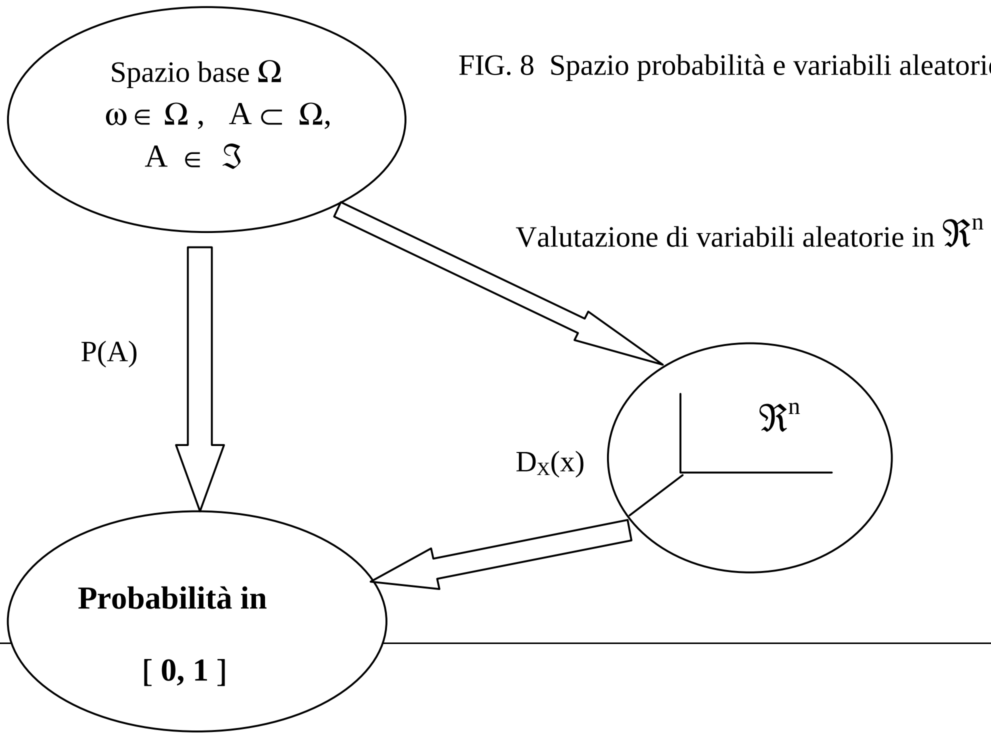


FIG. 8 Spazio probabilità e variabili aleatorie

Attraverso tale trasformazione è possibile individuare una funzione di punto $D_X(x)$

$$D_X(x): \Re^n \rightarrow [0 \ , \quad 1] \tag{64}$$

che consente di calcolare la probabilità di ogni evento ammissibile (vedi fig. 8).
Per meglio comprendere tale concetto si supponga inizialmente che lo spazio base Ω coincida con lo spazio euclideo unidimensionale $\Re$, e che la σ-algebra su cui viene definita la misura di probabilità sia il campo di Borel $\Im_B$ generato a partire dagli intervalli illimitati inferiormente del tipo

$$-\infty < X \leq \xi \tag{65}$$

già introdotti precedentemente.
Per ogni elemento E di $\Im_B$, rappresentante un evento ammissibile θ , sappiamo che la sua misura di probabilità $P_X(E)$ è ottenibile a partire da quella degli intervalli $-\infty < X \leq \xi$, poiché E può essere espresso come complementazione, unione e intersezione eventualmente eventualmente infinito numerabili, di intervalli di tale tipo. Per poter, quindi, calcolare $P_X(E)$ occorre solo conoscere la misura di probabilità di tali intervalli:

$$P_X(-\infty < X \leq \xi) \tag{66}$$

Ora per come sono definiti tali intervalli, tale probabilità risulta essere una funzione del solo estremo superiore (infatti l'altro estremo risulta fisso e pari a $-\infty$) che potremo indicare con $D_X(\xi)$:

$$D_X(\xi) \overset{\Delta}{=} P_X(-\infty < X \leq \xi) \tag{67}$$

ovvero una funzione di punto.
Noti risultati della teoria della misura consentono inoltre di affermare che, comunque si prenda un insieme $E \in \Im_B$, la sua misura di probabilità $P_X(E)$ può essere calcolata a partire dalla funzione $D_X(\bullet)$ come segue:

$$P_X(E) = \int_E d\,D_X(x) \tag{68}$$

dove l'integrale è un integrale di Lebesgue – stieltjes.

Ciò può essere generalizzato immediatamente al caso di spazio euclideo n – dimensionale considerando che in tal caso il campo di Borel $\Im_B$ può essere generato a partire dagli intervalli del tipo

$$-\infty < X_i \leq \xi_i \quad i = 1, \ldots, n \tag{69}$$

e quindi la misura di probabilità di tali intervalli è funzione dell'estremo $\xi = (\xi_1, \xi_2, \ldots, \xi_n)$ ottenendo

$$D_X(\xi_1, \xi_2, \ldots, \xi_n) = P_X(-\infty < X_i \leq \xi_i, \; i = 1, \ldots, n) \tag{70}$$

Pertanto per ogni E $\in \Im_B$ si ha, come estensione della (68)

$$P_X(E) = \int_E d\,D_{\underline{X}}(x_1, x_2, \ldots, x_n) \tag{71}$$

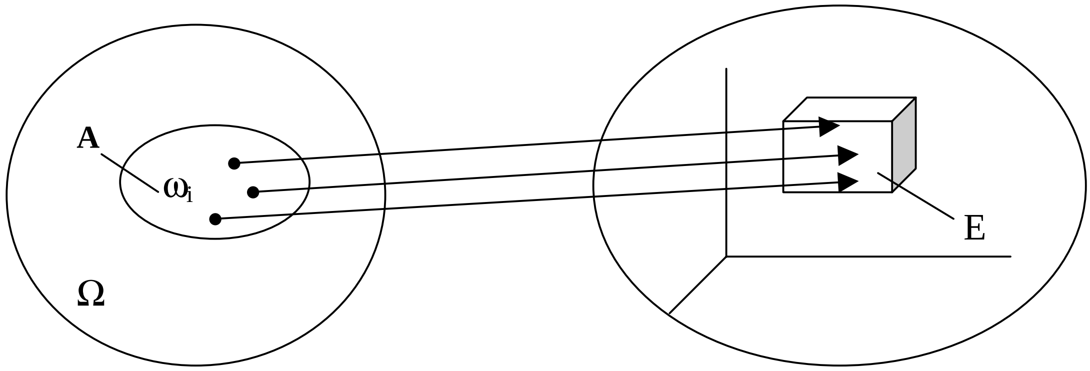
Quindi uno spazio di probabilità definito direttamente su $\Re^n$ viene ad essere del tipo

$$\left(\Re^n, \Im_B, D_{\underline{X}}(\bullet)\right) \tag{72}$$

Di conseguenza dato uno spazio di probabilità $(\Omega, \Im, P)$ sarebbe estremamente utile trovare una trasformazione X(ω) tale che:

$$(\Omega, \Im, P) \xrightarrow{X(\omega)} \left(\Re^n, \Im_B, D_{\underline{X}}(\bullet)\right) \tag{73}$$

Affinché ciò sia possibile X(ω) non può essere scelta in modo qualunque, ma occorre che, comunque si prenda un insieme E $\subseteq \Re^n$, appartenente a $\Im_B$, l'insieme A costituito dai punti $\omega \in \Omega$, che si trasformano tramite X(ω) in punti di E: $A \equiv \{\omega / \omega \in \Omega, X(\omega) \in E \in \Im_B\}$ (74), e che prende il nome di immagine inversa di E, rappresenti un evento ammissibile (ovvero A deve appartenere a $\Im$).



In realtà non è necessario verificare tale condizione per tutti gli elementi di $\Im_B$, ma solo per gli intervalli che la governano, poiché ogni altro elemento è ottenibile con operazioni di unione, intersezione e complementazione. Quindi è sufficiente verificare che gli insiemi $A \subseteq \Omega$:

$$A \equiv \{\omega / \omega \in \Omega, \quad -\infty < X_i \le \xi_i \quad i = 1, ..., n\} \tag{75}$$

appartengono a $\Im$, ovvero come si suole dire la funzione $X(\omega)$ è $\Im$ - misurabile. Ciò posto la variabile aleatoria $X(\omega)$ può essere definita formalmente come segue:

Definizione Dato uno spazio di probabilità $(\Omega, \Im, P)$ si definisce variabile aleatoria $X(\omega)$ una qualsiasi funzione X:

$$X: \Omega \rightarrow \Re^n \tag{76}$$

$\Im$ - misurabile, ovvero tale che per ogni punto $\xi \in \Re^n$ l'insieme, corrispondente in Ω, A:

$$A \equiv \{\omega / \omega \in \Omega, \quad -\infty < X_i \le \xi_i \quad i = 1, ..., n\} \tag{77}$$

appartenga a $\Im$ (cioè rappresenti un evento ammissibile).

Si noti che l'eventuale aspetto numerico dei risultati non ha nulla a che fare con il concetto di variabile aleatoria; quest'ultima infatti attribuisce valori numerici anche a risultati caratterizzati da un attributo qualitativo (p. es. il colore), ed inoltre nel caso di aspetto numerico corrispondenti possono anche essere diversi.

Per comprendere meglio il concetto di variabile aleatoria esaminiamo il caso di una funzione che non è una v. a. (variabile aleatoria).

Esempio Si consideri un generatore di tensione che può fornire ai suoi capi una tensione compresa tra 0V e 9V. In tal caso si può assumere come spazio base Ω l'intervallo (0, 9]. Si supponga inoltre di essere interessati solo al fatto che tale tensione sia inferiore a al più uguale a 6V o maggiore di tale valore. A questi due eventi corrispondono rispettivamente i due intervalli

$$I_1 \equiv (0,\ 6]\,,\quad I_1 \equiv (\ 6,\ 9] \tag{78}$$

Pertanto la σ - algebra da utilizzare per costruire lo spazio di probabilità è la σ - algebra:

$$\Im \equiv \{\emptyset,\quad \Omega = (0,9],\quad I_1 \equiv (0,\ 6]\,,\ \ I_1 \equiv (\ 6,\ 9]\} \tag{79}$$

Che rappresenta la minima σ - algebra generata da I_1 e I_2. In effetti la tensione ai capi del generatore, e quindi il valore di ω, può cadere ovunque nell'intervallo (0, 9], ma noi siamo interessati solo a conoscere se esso cade in I_1 o in I_2. Se ora si definisce la funzione $X(\omega) = \omega^2/ 10$ (80) , si può verificare immediatamente che tale funzione non è $\Im$ - misurabile. Infatti preso ad esempio ξ =2.5, l'insieme A dei punti che gli corrisponde è l'intervallo:

$$A \equiv \{\omega/\omega \in \Omega;\ \ X(\omega) \le 2.5\} = (0\ ,\ 5] \tag{75}$$

che non appartiene a $\Im$, conseguentemente tale funzione non è una v. a.

In questo caso le possibili v. a. sono solo le funzioni che assumono valori costanti sugli intervalli I_1 e I_2 (vedi fig. 11).

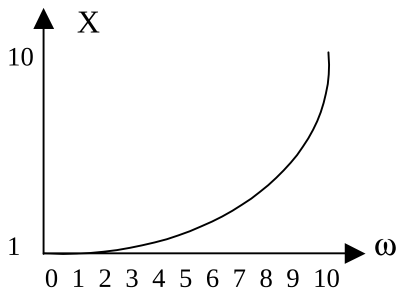


FIG. 10

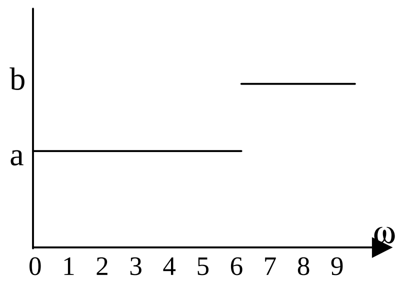


FIG. 11

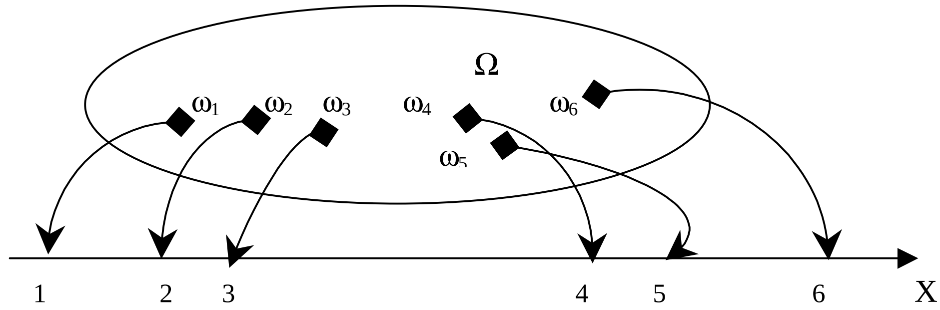
Nel seguito si indicherà con la lettera maiuscola la v. a., ovvero la funzione da Ω in $\Re^n$, (ad es. X) e con la lettera minuscola (ad es. x) la determinazione assunta da tale v. a. a seguito di una prova.

Se il numero di punti dello spazio base Ω è finito oppure infinito numerabile e la trasformazione da in $\Re^n$ è biunivoca, allora l'insieme S dei punti $\underline{x}_i \in \Re^n$ corrispondenti ai punti ω_i di Ω è ancora finito o al più infinito numerabile ed inoltre:

$$\sum_{x_i \in S} P\left(\left\{\omega : \underline{X}(\omega) = \underline{x}_i\right\}\right) = 1 \tag{82}$$

In tal caso la misura di probabilità è concentrata nei punti $\underline{x}_i$. Ad esempio nel caso di fig. 12 rappresentante il lancio di un dado.

FIG. 12



La misura di probabilità è concentrata nei punti di ascissa 1, 2, 3, 4, 5 e 6 e nel caso di dado non truccato la misura di probabilità di ciascuno di tali punti è pari ad 1/6.

Tale tipo di v. a. che prende il nome di v. a. discreta rappresenta una caso di estremo interesse in quanto risultano estremamente semplificate in tal caso le procedure per il calcolo della probabilità di eventi composti. Infatti per ogni insieme A ∈ $\mathfrak{I}_B$ si ha immediatamente che:

$$P_x(A) = \sum_{x_i \in A \cap S} P\left(\left\{\omega : \underline{X}(\omega) = x_i\right\}\right) \tag{83}$$

In realtà la condizione (83) è verificata in circostanze più ampie di quelle considerate precedentemente, pertanto viene data la seguente definizione di v. a. discreta.

Definizione Una v. a. $\underline{X}$ si dice discreta se esiste un insieme S= { x_i }finito o infinito numerabile tale che

$$\sum_{x_i \in S} P\left(\left\{\omega : \underline{X}(\omega) = x_i\right\}\right) = 1 \tag{84}$$

Commento Si noti che nella precedente definizione non si fa nessun riferimento alla numerabilità dello spazio Ω ed al tipo di trasformazione adottato si richiede soltanto che la misura di probabilità $P_x(\bullet)$ definita su $\mathfrak{R}^n$ sia concentrata in un insieme numerabile di punti.

FUNZIONE DI DISTRIBUZIONE

Nel paragrafo precedente è stato visto come il concetto di variabile aleatoria consenta di definire la probabilità di un evento di un fenomeno aleatorio per mezzo di una funzione di punto $D_{\underline{x}}(x_1, x_2, \ldots, x_n)$ che può essere interpretata come la probabilità che le componenti x_i, i=1,…, n della v. a. $\underline{X}$ siano minori o uguali di x_i , i=1,…, n:

$$D_{\underline{x}}(x_1, x_2, \ldots, x_n) \overset{\Delta}{=} \Pr\{X_i \leq x_i \, , \, i = 1, \ldots, n\} \tag{85}$$

e che prende il nome di <u>funzione di distribuzione</u> .

Si noti che il pedice di $D_{\underline{x}}(x_1, x_2, \ldots, x_n)$ serve ad indicare la v. a. a cui si riferisce la funzione di distribuzione.

La funzione di distribuzione presenta le seguenti proprietà fondamentali:

1. La funzione di distribuzione è non negativa:

$$D_{\underline{x}}(x_1, x_2, \ldots, x_n) \;\geq\; 0 \tag{86}$$

2. La funzione di distribuzione è non decrescente:

$$D_{\underline{x}}(x_1, \ldots, x_j + \Delta x_j, \ldots, x_n) \;\geq\; D_{\underline{x}}(x_1, \ldots, x_j, \ldots, x_n) \tag{87}$$

3. La funzione di distribuzione è continua da destra

$$\lim_{\substack{\Delta x_j \to 0 \\ \Delta x_j \geq 0}} D_{\underline{x}}(x_1, \ldots, x_j + \Delta x_j, \ldots, x_n) = D_{\underline{x}}(x_1, \ldots, x_j, \ldots, x_n) \tag{88}$$

4. La funzione di distribuzione tende a 0 al tendere a $-\infty$ di almeno una delle componenti x_i:

$$\lim_{x_j \to -\infty} D_{\underline{x}}(x_1, ..., x_j, ..., x_n) = 0 \tag{89}$$

5. La funzione di distribuzione tende a 1 al tendere a $+\infty$ di tutte le componenti:

$$\lim_{\substack{x_1 \to \infty \\ \cdots\cdots \\ x_j \to \infty \\ \cdots\cdots \\ x_n \to \infty}} D_{\underline{x}}(x_1, ..., x_j, ..., x_n) = 1 \tag{90}$$

La funzione di distribuzione descrive completamente le proprietà statistiche di una v. a. , infatti, come è stato già anticipato, dato un evento θ ammissibile a cui corrisponde un insieme $E \subseteq \Re^n$, la probabilità $\Pr\{\theta\}$ di tale evento può essere calcolata come:

$$\Pr\{\theta\} = \int_E d\, D_{\underline{x}}(x_1, x_2, ..., x_n) \tag{91}$$

dove l'integrale è un integrale di Lebesgue – Stieltjes.

Per meglio intendere il significato di tale integrale iniziamo con il considerare il caso unidimensionale. In tal caso, dato un intervallo aperto inferiormente del tipo:

$$T_1 = (x_1,\ x_1 + \Delta x_1] \tag{92}$$

poiché $(-\infty, x + \Delta x_1] = (-\infty, x_1] \cup (x_1, x_1 + \Delta x_1]$ in virtù del II° assioma si ha:

$$P_x(T_1) = \int_{x_1}^{x_1 + \Delta x_1} d\, D_x(x_1) = D_x(x_1 + \Delta x_1) - D_x(x_1) = \Delta_{x_1} D_x(x_1) \tag{93}$$

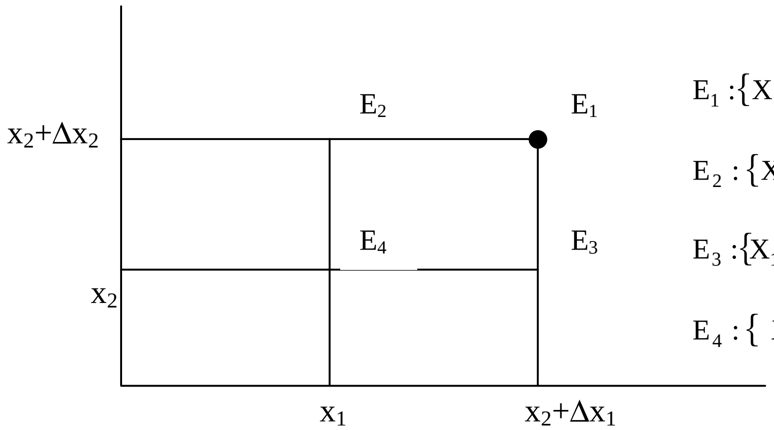
essendo $\Delta_{x_1} D_x(x_1)$ la differenza finita di ordine 1 di $D_x(x_1)$.

Nel caso bidimensionale preso un intervallo T_2 aperto inferiormente del tipo

$$T_2 \equiv \{x_1 < X_1 \le x_1 + \Delta x_1,\ x_2 < X_2 \le x_2 + \Delta x_2\} \qquad (94)$$

con riferimento alla fig. 13 si ha immediatamente che:

$$\begin{aligned} P_x(T_2) &= P_x(E_1) - P_x(E_2) - P_x(E_3) + P_x(E_4) = \\ &D_{\underline{x}}(x_1+\Delta x_1, x_2+\Delta x_2) - D_{\underline{x}}(x_1+\Delta x_1, x_2) + \\ &- D_{\underline{x}}(x_1,\ x_2+\Delta x_2) + D_{\underline{x}}(x_1,\ x_2) \end{aligned} \qquad (95)$$



$E_1 : \{X_1 \le x_1 + \Delta x_1,\ X_2 \le x_2 + \Delta x_2\}$

$E_2 : \{X_1 \le x_1,\ \ X_2 \le x_2 + \Delta x_2\}$

$E_3 : \{X_1 \le x_1 + \Delta x_1,\ \ X_2 \le x_2\}$

$E_4 : \{X_1 \le x_1,\ \ X_2 \le x_2\}$

FIG. 13

che può essere riscritta più sinteticamente come

$$P_x(T_2) = \Delta^2 D_x(x_1,\ x_2) \qquad (96)$$

facendo ricorso alla differenza finita di ordine 2 di $D_{\underline{x}}(x_1,\ x_2)$ definita come:

$$\begin{aligned} \Delta^2 D_x(x_1,\ x_2) &= \Delta_{x_1 x_2} D_x(x_1,\ x_2) \overset{\Delta}{=} \Delta_{x_2}(\Delta_{x_1} D_x(x_1,\ x_2)) = \\ &= \Delta_{x_1}(\Delta_{x_2} D_x(x_1,\ x_2)) \end{aligned} \qquad (97)$$

Nel caso generale n – dimensionale definita (sempre tramite la differenza finita di ordine 1) come differenza finita di ordine n la quantità

$$\begin{aligned} \Delta^n D_{\underline{x}}(x_1, x_2, \ldots, x_n) &= \Delta_{x_1 x_2 \ldots x_n} D_{\underline{x}}(x_1, x_2, \ldots, x_n) \overset{\Delta}{=} \\ &\overset{\Delta}{=} \Delta_{x_n} \Delta_{x_{n-1}} \ldots \Delta_{x_2} \Delta_{x_1} D_x(x_1,\ x_2, \ldots, x_n) \end{aligned} \qquad (98)$$

si può mostrare facilmente che la misura di probabilità $P_x(T_n)$ dell'intervallo aperto inferiormente:

$$T_n \equiv \{x \mid x_i < X_i \le x_i + \Delta x_i, \quad i = 1,\dots,n\} \tag{99a}$$

risulta:

$$P_x(T_n) = \int_{x_1}^{x_1+\Delta x_1} \dots\dots \int_{x_n}^{x_n+\Delta x_n} d\,D_{\underline{x}}(x_1, x_2, \dots, x_n) = \Delta^n\, D_{\underline{x}}(x_1, x_2, \dots, x_n) \tag{99b}$$

Pertanto, in sintesi si può affermare che la differenza sostanziale tra un integrale del tipo

$$\int_E d\,D_{\underline{x}}\,(x_1, x_2, \dots, x_n) \tag{100}$$

ed un integrale di Lebesgue del tipo

$$\int_E d\,x_1\, d\,x_2 \dots dx_n \tag{101}$$

risiede nell'aver sostituito alla misura di Lebesgue di un intervallo T_n:

$$\Delta x_1\, \Delta x_2 \dots\dots \Delta x_n \tag{102}$$

che ne rappresenta il volume, la quantità

$$\Delta_{x_1} \Delta_{x_2} \dots \Delta_{x_n}\; D_{\underline{x}}\,(x_1,\; x_2, \dots, x_n) \tag{103}$$

In realtà la relazione

$$\Pr\{\theta\} = \int_E d\,D_{\underline{x}}\,(x_1, x_2, \dots, x_n) \tag{104}$$

risulta troppo formale per poter essere impiegata direttamente. Sarebbe in effetti molto più comodo poter esprimere tale probabilità tramite un integrale di Lebesgue. Allo scopo di analizzare in quali condizioni ciò possa accadere è utile introdurre i due concetti seguenti.

Definizione Data una misura di probabilità $P_x(\bullet)$ definita sul campo di Borel $\Im_B$, essa è detta singolare (rispetto alla misura di Lebesgue) se esiste un insieme $S \subseteq \Re^n$ tale che la sua misura di Lebesgue sia nulla e risulti

$$P_x(S)=1 \tag{105}$$

Definizione Data una misura di probabilità $P_x(\bullet)$ definita sul campo di Borel $\Im_B$, essa è detta assolutamente continua (rispetto alla misura di Lebesgue) se per ogni insieme $E \in \Im_B$ con misura di Lebesgue nulla si ha:

$$P_x(E)=0 \tag{106}$$

Commento ovviamente se la v. a. X è discreta allora la misura di probabilità risulta singolare. In tal caso, infatti, la v. a. trasforma i punti dello spazio base Ω in un numero di punti al più infinito numerabili di $\Re^n$. Di conseguenza l'insieme di tali punti, a cui corrisponde una misura di Lebesgue nulla. In realtà tale condizione è solo sufficiente per la singolarità, per convincersene basta osservare che ogni sottospazio di $\Re^n$ di dimensione minore di n presenta un'infinità non numerabile di elementi, ma misura di Lebesgue nulla.

Ciò posto è possibile dimostrare che se la misura di probabilità $P_x(\bullet)$ è assolutamente continua, allora esiste una funzione non negativa $p_{\underline{x}}(x_1, x_2, \dots, x_n)$ tale che:

$$\Pr\{\theta\}=\int_E d\,D_{\underline{x}}\,(x_1, x_2, \dots, x_n)=\int_E p_{\underline{x}}\,(x_1, x_2, \dots, x_n)\,dx_1\,dx_2 \dots dx_n \quad \forall E \in \Im_B \tag{107}$$

dove l'integrale è un integrale di lebesgue.

La funzione $p_{\underline{x}}(x_1, x_2, \dots, x_n)$ prende il nome di funzione di densità di probabilità .

Si noti che in base alla (107) e alla definizione di funzione di distribuzione si ha immediatamente che

$$D_{\underline{x}}(x_1, x_2, ..., x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} .... \int_{-\infty}^{x_n} p_{\underline{x}}(x_1, x_2, ..., x_n)\, dx_1\, dx_2 ... dx_n \quad (108)$$

ovvero che la f. d. p. $p_{\underline{x}}(x_1, x_2, \ldots, x_n)$ risulta essere:

$$p_{\underline{x}}(x_1, x_2, ..., x_n) = \frac{\partial^n D_{\underline{x}}(x_1, x_2, ..., x_n)}{\partial x_1\, \partial x_2 ... \partial x_n} \quad (109)$$

laddove tale derivata esista.

Poiché la funzione di distribuzione è non decrescente, dalla precedente relazione segue immediatamente che relazione segue immediatamente che la f. d. p. è non negativa:

$$p_{\underline{x}}(x_1, x_2, ..., x_n) \geq 0 \quad (110)$$

Inoltre dal IV° assioma segue immediatamente la seguente condizione di normalizzazione

$$\boxed{\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} .... \int_{-\infty}^{x_n} p_{\underline{x}}(x_1, x_2, ..., x_n)\, dx_1\, dx_2 ... dx_n = 1} \quad (111)$$

Per comprendere meglio il significato di $p_{\underline{x}}(x_1, x_2, \ldots, x_n)$ si consideri che dato un intervallo T_n del tipo

$$T_n \equiv \{x \mid x_i < X_i \leq x_i + \Delta x_i, \quad i = 1, ..., n\} \quad (112)$$

in virtù del teorema della media la probabilità che $\underline{x}$ appartenga a T_n può essere espresso come segue:

$$\Pr\{x_i < X_i \leq x_i + \Delta x_i, i = 1, ..., n\} = p_{\underline{x}}(x'_1, x'_2, ..., x'_n)\, \Delta x_1\, \Delta x_2 ... \Delta x_n \quad (113)$$

essendo $\underline{x}' = (x'_1, x'_2, \dots, x'_n)$ un punto interno a tale intervallo. Pertanto se ora si fa tendere l'intervallo ad un intervallo infinitesimo di volume $dx_1\, dx_2 \dots dx_n$ si ha:

$$\Pr\{x_i < X_i \le x_i + \Delta x_i, i = 1, \dots, n\} = p_{\underline{x}}(x'_1, x'_2, \dots, x'_n)\, dx_1\, dx_2 \dots dx_n \tag{114}$$

Tale risultato mette in rilievo come la funzione $p_{\underline{x}}(x_1, x_2, \dots, x_n)$ non sia una probabilità, ma solo una densità e che per ottenere una probabilità occorre moltiplicarla per il volume dell'ipercubo [$dx_1\, dx_2 \dots dx_n$]. Da ciò segue inoltre che se $P_x(\bullet)$ è assolutamente continua, la probabilità associata al singolo punto è nulla (in realtà tale proprietà è proprio ciò che determina l'assoluta continuità e non una conseguenza della (114)).

Ovviamente non si riesce ad esprimere la probabilità $\Pr\{\theta\}$ di un evento in termini di integrale di $p_{\underline{x}}(x_1, x_2, \dots, x_n)$, almeno in un senso ordinario, se esiste un insieme di misura (di Lebesgue), nulla a cui corrisponda una probabilità non nulla.

Ciò accade ad esempio quando la v. a. è discreta e quindi la misura di probabilità è singolare.

In tal caso ci si può ricondurre formalmente ad una espressione del tipo (107) facendo ricorso all'impulso matematico, il quale consente di ottenere, formalmente, un contributo finito all'integrale da parte di un punto, e quindi di concentrare la misura di probabilità in tale punto. E' sufficiente osservare che in generale una misura di probabilità $P_x(\bullet)$ non è necessariamente o singolare o assolutamente continua, ma può comunque essere decomposta nella forma (nota come decomposizione di Lebesgue):

$$P_x(E) = \alpha P_x^{(1)}(E) + (1-\alpha)\, P_x^{(2)}(E) \qquad \forall E \in \Im_B \tag{119}$$

con $0 \le \alpha \le 1$ e $P_x^{(1)}(\bullet)$ e $P_x^{(1)}(\bullet)$ rispettivamente assolutamente continua e singolare.

VARIABILI ALEATORIE UNIDIMENSIONALI

Nel caso in cui la v. a. sia unidimensionale si può dimostrare che le proprietà (86) e (90) della funzione di distribuzione implicano che la funzione di distibuzione è continua quasi ovunque cioè a meno di un insieme di misura (di Lebesque) nulla in cui presenta discontinuità di 1ª specie. Tale proprietà implica che nel caso unidimensionale possono verificarsi solamente tre casi:

- La funzione di distribuzione è continua ovunque (fig. 14a)
- La funzione di distribuzione è costante a tratti (fig. 14b)
- La funzione di distribuzione presenta un numero finito (non nullo) o un'infinità numerabile di discontinuità di 1ª specie e risulta in generale crescente tra due punti di discontinuità consecutivi (fig. 14c)

La v. a. corrispondente prende il nome di

A. V. a. continua
B. V. a. discreta
C. V. a. mista

FIG. 14a – v. a. continua

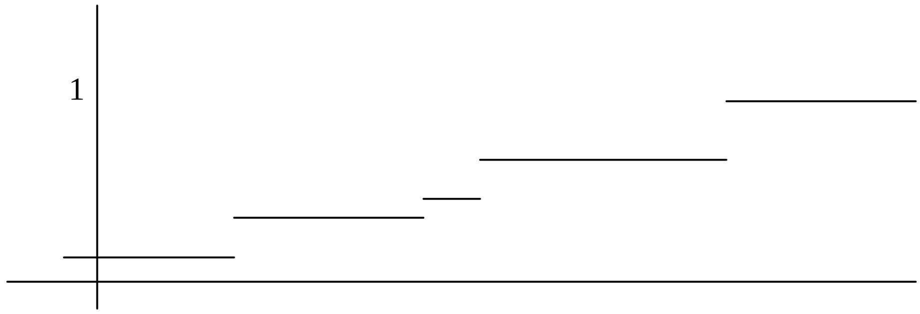


FIG. 14b – v. a. discreta

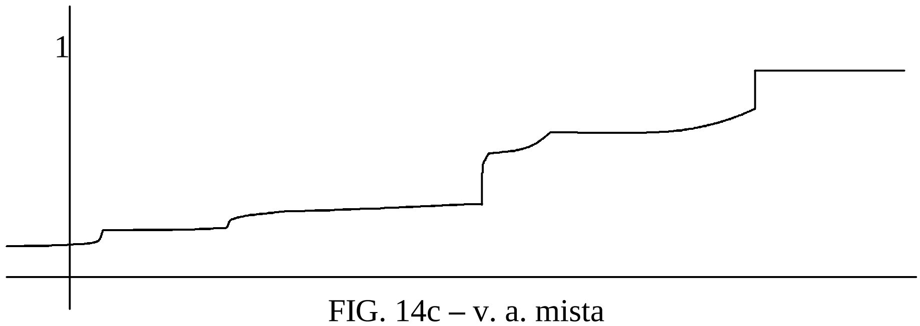


FIG. 14c – v. a. mista

1. VARIABILE ALEATORIA CONTINUA

Poiché nel caso di v. a. continua la funzione di distribuzione è continua ovunque essa è derivabile quasi ovunque, pertanto risulta definibile la f. d. p. $p_X(x)$ per la quale si ha:

$$p_x(x) = \frac{d D_X(x)}{dx} \tag{116}$$

$$D_X(x) = \int_{-\infty}^{x} p_X(\xi) d\xi \tag{117}$$

e che gode delle proprietà (110) e (111) che nel caso unidimensionale si riscrivono.

$$p_X(x) \geq 0 \tag{118}$$

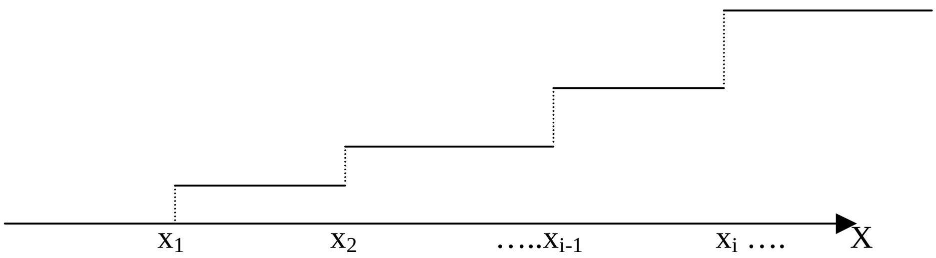
$$\int_{-\infty}^{+\infty} p_X(x) dx = 1 \tag{119}$$

Inoltre in base alla (114) si ha che la quantità $p_X(x)$ rappresenta la probabilità che la v. a. X assuma una determinazione appartenente all'intervallo (x, x+dx].

2. VARIABILE ALEATORIA DISCRETA

E' stato affermato, senza dimostrarlo, che nel caso in cui la funzione di distribuzione sia costante a tratti la v. a. unidimensionale corrispondente è discreta. Nel presente paragrafo verrà data, per prima cosa, la prova di tale proprietà dopodiché verranno sviluppati gli strumenti analitici per lo studio di tali classi di v. a.

Allo scopo di verificare, quindi che la v. a. in esame è del tipo discreto secondo la definizione data, consideriamo una funzione di distribuzione $D_x(x)$ costante a tratti come quella di fig. 15 e indichiamo con x_i , i=1, 2, …, h ascisse dei punti di discontinuità di 1ª specie.



Preso un intervallo del tipo $(x_i - \varepsilon, x_i]$, in base alla definizione di funzione di distribuzione si ha immediatamente che

$$\Pr\{x_i - \varepsilon < X \le x_i\} = D_X(x_i) - D_X(x_i - \varepsilon) \tag{120}$$

Da cui per $\varepsilon \to 0$ si ha:

$$\Pr\{X = x_i\} = D_X(x_i) - D_X(x_i^-) \tag{121}$$

avendo indicato con $D_X(x_i^-)$ il limite da sinistra:

$$D_X(x_i^-) = \lim_{x \to x_i^-} D_X(x_i) \qquad (122)$$

Quindi il salto della funzione di distribuzione $D_X(x_i)$ - $D_X(x_i^-)$ costituisce proprio la probabilità che la v. a. X assuma la determinazione x_i. Si osservi, inoltre, che essendo $D_X(x)$ costante a tratti si ha che $D_X(x_i^-) = D_X(x_{i-1})$ e quindi

$$\Pr\{X = x_i\} = D_X(x_i) - D_X(x_i^-) \qquad (122.a)$$

Poiché inoltre $D_X(x) \to 0$ per $x \to -\infty$ e $D_X(x) \to 1$ per $x \to +\infty$, indicato con $S=\{x_i,\ i=1,\ 2,\ \ldots\}$ l'insieme dei punti di discontinuità di 1ª specie, si ha immediatamente che (posto $x_0 = -\infty$)

$$1 = D_X(+\infty) - D_X(-\infty) = \sum_{x_i \in S} D_X(x_i) - D_X(x_{i-1}) = \sum_{x_i \in S} \Pr\{X = x_i\} \qquad (123)$$

che rappresenta proprio la condizione adottata nella definizione di v. a. discreta. Nel caso di v. a. discreta unidimensionale, la v. a. è completamente descritta dalle probabilità $\Pr\{X = x_i\}$, che verranno indicate nel seguito con P_i e con $P(x_i)$:

$$P_i = P(x_i) \stackrel{\Delta}{=} \Pr\{X = x_i\} \qquad (124)$$

in accordo con la notazione più diffusa. In base a tale notazione la condizione di normalizzazione si riscrive come

$$\sum_i P_i = 1 \qquad (125)$$

La funzione $P(x_i)$, i=1, 2, … prende il nome di <u>funzione di probabilità</u>.

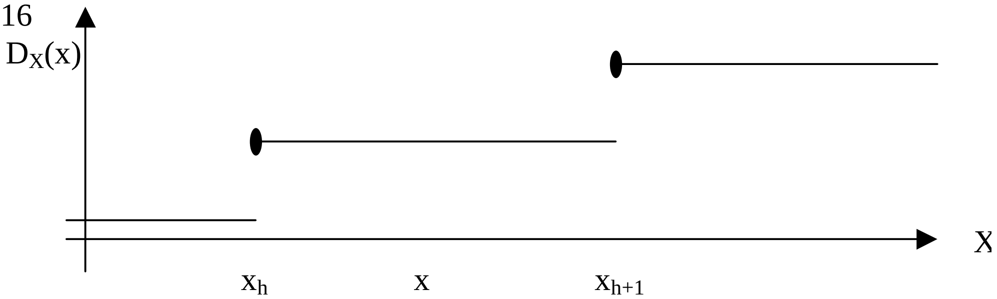
La funzione di distribuzione $D_X(x)$ è legata alla funzione di probabilità della relazione

$$D_X(x) = \sum_{i=1}^{h} P(x_i) \qquad (126)$$

in cui x_h è la determinazione per la quale $x_h \le x < x_{h+1}$.

Infatti, con riferimento alla fig. 16, la funzione di distribuzione di x, $D_X(x)$, coincide con il valore in x_h, $D_X(x_h)$ (essendo costante nell'intervallo [x_h , x_{h+1})).

FIG. 16



In realtà la funzione di probabilità è completamente sufficiente per descrivere la probabilità di un qualsiasi evento ammissibile. Infatti, per ogni insieme $A \in \Im_B$ si ha che la probabilità Pr{*A*} dell'evento corrispondente *A* è pari a:

$$\Pr\{A\} = \sum_{A \cap S} P(x_i) \tag{127}$$

Esempio Consideriamo il lancio di un dado non truccato e associamo ai possibili risultati i punti della retta di ascissa 1,2,3,4,5 e 6. Poiché i risultati sono tutti equiprobabili (dado non truccato) si ha:

$$P_i = \frac{1}{6} \quad , \quad i = 1, \dots, 6 \tag{128}$$

a cui corrisponde la funzione di distribuzione di fig. 17.

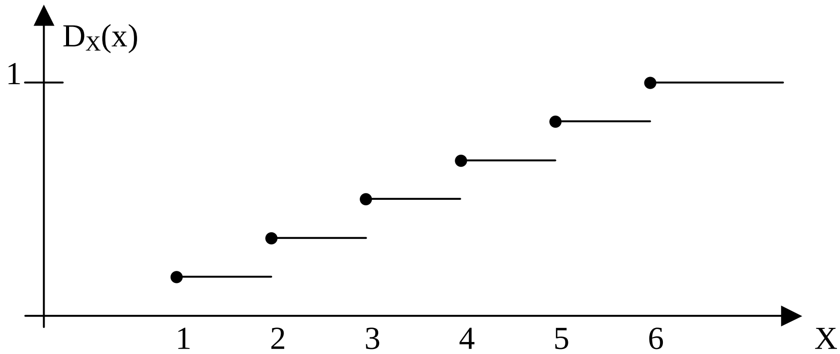


FIG. 17

La proprietà dell'evento A = "il risultato è pari" a cui corrisponde l'insieme A={2, 4, 6} è pertanto pari a :

$$\sum_{i\in A\cap S} P_i = P_2 + P_4 + P_6 = \frac{3}{6} = \frac{1}{2} \tag{129}$$

Allo scopo però di unificare la trattazione tra le v. a. continue e le v. a. discrete si può introdurre formalmente una f. d. p. per v. a. discrete facendo ricorso all'impulso matematico.

A tale scopo con riferimento alla fig. 18 si osservi che

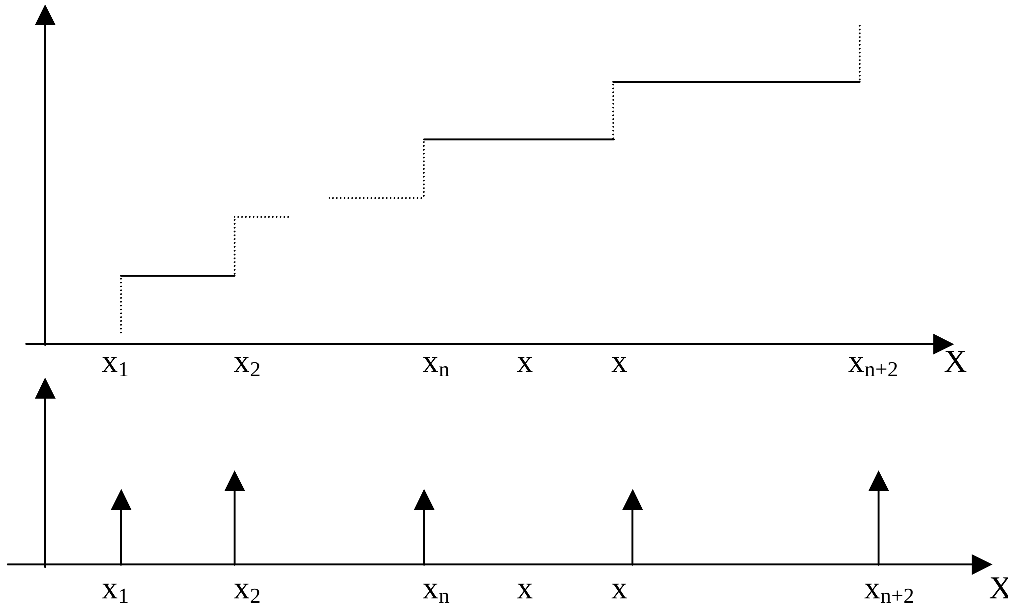


FIG. 18 a – b

La relazione (126) che fornisce la funzione di distribuzione $D_X(x)$ in funzione di $P(x_i)$ può essere riscritta come segue facendo ricorso alla proprietà dell'impulso matematico:

$$D_X(x)=\sum_{i=1}^{h}P(x_i)=\int_{-\infty}^{x}\sum_{i=1}^{h}P(x_i)u_0(x-x_i)dx=\int_{-\infty}^{x}\sum_{x_i\in S}P(x_i)u_0(x-x_i)dx \tag{130}$$

infatti solo gli impulsi matematici che cadono nell'intervallo $(-\infty\ ,\ x]$ danno contributo all'integrale.

Il confronto tra la (126) e la (130) suggerisce di definire formalmente una f. d. p. per v. a. discrete come

$$p_x(x) = \sum_{x_i \in S} P(x_i)\, u_0(x - x_i) \tag{131}$$

ovvero la f. d. p. di una v. a. discreta è costituita da una successione di impulsi matematici centrati nei punti x_i di area pari a $P(x_i)$. Ovviamente la f. d. p. data dalla (131) non rappresenta la derivata della $D_X(x)$ a meno di considerare l'impulso matematico come la derivata formale del gradino matematico. C'è infine da rilevare che l'utilizzo dell'impulso matematico nella definizione della f. d. p. è reso possibile dal fatto che nel calcolo delle probabilità di eventi tali impulsi matematici compaiono sempre sotto il segno d'integrale.

3. VARIABILE ALEATORIA MISTA

Nel caso di v. a. mista la funzione di distribuzione $D_X(x)$ può essere decomposta nella somma di una funzione continua ovunque $g_x^{(1)}(x)$ e di una funzione costante a tratti $g_x^{(2)}(x)$:

$$D_X(x) = g_x^{(1)}(x) + g_x^{(2)}(x) \tag{132}$$

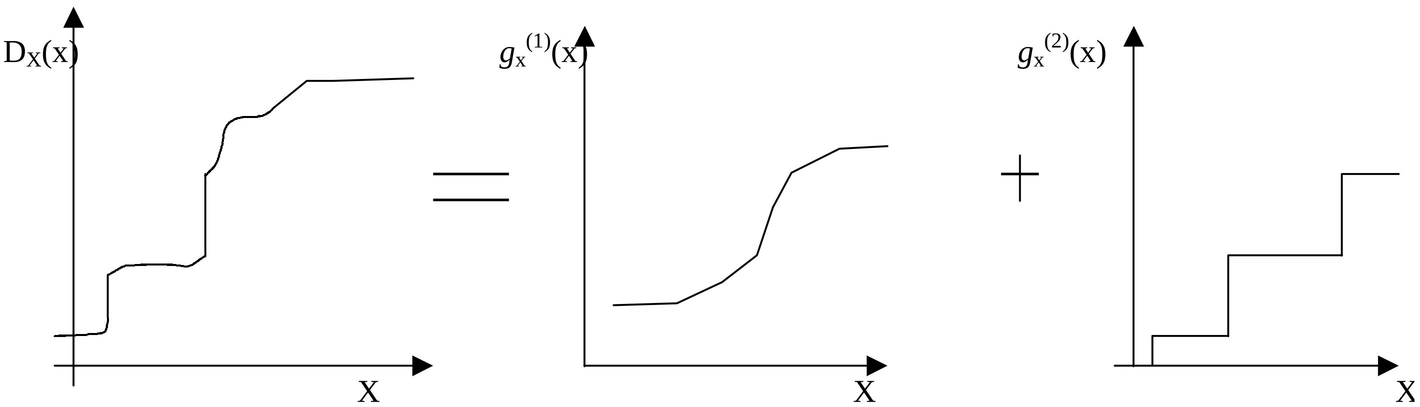


FIG. 19 Decomposizione della funzione di distribuzione di variabile aleatoria mista.

Ovvero normalizzando $g_x^{(1)}(x)$ rispetto a $g_x^{(1)}(+\infty)$ e $g_x^{(2)}(x)$ rispetto a $g_x^{(2)}(+\infty)$, posto

$$\alpha^{-1} = \frac{1}{g_x^{(1)}(+\infty)} \quad \left[\alpha = g_x^{(1)}(+\infty)\right]$$

si ha:

$$D_x(x) = \alpha D_x^{(1)}(x) + (1-\alpha) D_x^{(1)}(x) \tag{133}$$

che rappresenta appunto la decomposizione di Lebesgue introdotta nel paragrafo.

Pertanto in base alla (133) ed ai risultati dei paragrafi precedenti si ha che la f. d. p. (in senso limite) di una v. a. mista può essere posta nella forma

$$p_x(x) = \alpha\, p_x^{(1)}(x) + (1-\alpha) \sum_{x_i \in S} P(x_i)\, u_0(x - x_i)$$