Capitolo I

RICHIAMI DI TEORIA DELLA PROBABILITÀ

I.1 - Lo spazio dei risultati. Gli eventi.

Spesso nella realtà ci s'imbatte in fenomeni i cui esiti non possono essere esattamente previsti, basti pensare all'estrazione del biglietto vincente di una lotteria, o al numero di chiamate in arrivo in una centrale telefonica nel corso di un'ora della giornata. Tuttavia se, osservando ripetutamente il fenomeno, si prende nota dei risultati, ci si accorge che, nella quasi totalità dei casi, essi obbediscono ad una certa *regolarità statistica*, nel senso che il rapporto tra le volte in cui si verifica un determinato risultato e il numero totale d'osservazioni tende a stabilizzarsi attorno ad un dato valore al crescere di queste ultime. Così, ad esempio, se da un'urna contenente palline nere e bianche si è estratta nel 90% dei casi una pallina bianca, e nel restante 10%, una nera, si è indotti a ritenere che l'evento "estrazione di una pallina bianca" abbia una maggiore probabilità di verificarsi dell'evento "estrazione di una pallina nera". Si è così portati ad associare ad ogni evento casuale una certa probabilità che esso accada. Tuttavia, per definire correttamente il concetto di probabilità associato ad un evento casuale, occorre richiamare alcune nozioni fondamentali concernenti il cosiddetto **spazio di probabilità**.

Per schematizzare il comportamento di un fenomeno aleatorio è opportuno introdurre il concetto di esperimento casuale che consiste in un procedimento di osservazione di un risultato, che può essere ripetuto, a partire dalle medesime condizioni iniziali, tutte le volte che si voglia. Ad esempio nell'esperimento casuale consistente nel lancio di una moneta, i possibili risultati osservabili sono *testa* (t) e *croce* (c).

Si consideri un esperimento casuale e sia ζ un suo possibile risultato. L'insieme Ω costituito da tutti i risultati che si possono manifestare prende il nome di **spazio dei risultati**.

Nell'esempio precedente del lancio di una moneta si ha:

$$\Omega = \{t, c\}$$

Nel lancio di un dado, lo spazio dei risultati è costituito dall'insieme delle sei facce e si ha:

$$\Omega = \{ f_1, f_2, f_3, f_4, f_5, f_6 \}$$

Per studiare un esperimento casuale è opportuno identificare una classe di sottoinsiemi dello spazio dei risultati, il generico elemento di tale classe è chiamato **evento**. Un evento $\boldsymbol{E}$ si dice verificato tutte le volte che l'esperimento casuale da luogo ad un risultato ω che appartiene ad $\boldsymbol{E}$. La scelta della classe, seppure nel rispetto di alcune proprietà che verranno elencate più avanti, non è obbligata e dipende dalla particolare proprietà che s'intende prendere in considerazione. Ad esempio, nel caso del lancio del dado, se si ha interesse al punteggio della faccia superiore, si devono necessariamente assumere come eventi gli insiemi contenenti i singoli risultati; mentre se si è interessati solo al fatto che il risultato sia pari o dispari ci si può limitare a prendere in considerazione gli eventi:

$$\boldsymbol{E}_p = \{ f_2, f_4, f_6 \} \quad ; \quad \boldsymbol{E}_d = \{ f_1, f_3, f_5 \}$$

L'intero spazio dei risultati $\boldsymbol{\Omega}$ è un evento, come pure lo è l'insieme vuoto $\varnothing$. Nel primo caso si parla di **evento certo**, poiché l'evento $\boldsymbol{E} = \boldsymbol{\Omega}$ si manifesta ogni volta che si compie l'esperimento; nel secondo si parla di **evento impossibile,** dato che l'evento $\boldsymbol{E} = \varnothing$ non si può manifestare.

Ovviamente agli eventi si può applicare l'algebra degli insiemi ed in particolare le note operazioni di unione, intersezione, e complementazione (vedi Fig. I.1).

- L'evento unione $\boldsymbol{E}_1 \cup \boldsymbol{E}_2$ si verifica quando almeno uno dei due eventi è verificato.
- L'evento intersezione $\boldsymbol{E}_1 \cap \boldsymbol{E}_2$ si verifica quando entrambi gli eventi sono verificati.
- L'evento $\boldsymbol{E}^C$ è verificato tutte le volte che $\boldsymbol{E}$ non è verificato.

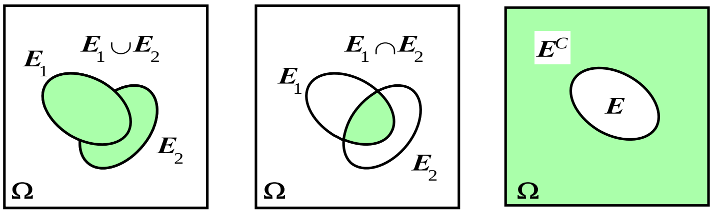

Fig.I.1 – Unione, intersezione e complementazione di eventi.

Ad esempio con riferimento al lancio di un dado, detti $\boldsymbol{E}_1$ e $\boldsymbol{E}_2$ gli eventi

$$\boldsymbol{E}_1 = \{2, 4, 6\} \quad \boldsymbol{E}_2 = \{1, 2, 3\}$$

risulta

$$\boldsymbol{E}_1 \cup \boldsymbol{E}_2 = \{1, 2, 3, 4, 6\} \quad \boldsymbol{E}_1 \cap \boldsymbol{E}_2 = \{2\}$$

Due eventi si dicono **disgiunti** quando il manifestarsi di uno implica il non manifestarsi dell'altro, in altre parole se la loro intersezione è l'evento impossibile:

(I.1.1) $$\boldsymbol{E}_1 \cap \boldsymbol{E}_2 = \varnothing$$

Si supponga adesso di ripetere N volte un certo esperimento casuale e sia ΔN il numero di volte in cui un dato evento $\boldsymbol{E}$ si è verificato. La quantità $\lim_{N\to\infty} \frac{\Delta N}{N}$ è detta probabilità associata all'evento $\boldsymbol{E}$:

(I.1.2) $$\Pr\{\boldsymbol{E}\} = \lim_{N\to\infty} \frac{\Delta N}{N}$$

In altre parole per probabilità di un evento s'intende il limite della "frequenza relativa" $\frac{\Delta N}{N}$ al tendere all'infinito del numero di ripetizioni dell'esperimento. Essendo necessariamente $0 \le \Delta N \le N$ risulta:

(I.1.3) $$0 \le \Pr\{\boldsymbol{E}\} \le 1 .$$

Nel caso dell'evento certo, essendo $\Delta N = N$, è:

(I.1.4) $$\Pr\{\boldsymbol{\Omega}\} = 1$$

Infine, se gli eventi $\boldsymbol{E}_1$ ed $\boldsymbol{E}_2$ sono disgiunti, dall'ovvia condizione $\Delta N_{\boldsymbol{E}_1 \cup \boldsymbol{E}_2} = \Delta N_{\boldsymbol{E}_1} + \Delta N_{\boldsymbol{E}_2}$ segue:

(I.1.5) $$\Pr\{\boldsymbol{E}_1 \cup \boldsymbol{E}_2\} = \Pr\{\boldsymbol{E}_1\} + \Pr\{\boldsymbol{E}_2\}$$

Si osservi che, purtroppo, poiché in ogni esperimento fisico il numero delle prove, per quanto grande, non può mai essere infinito, il limite $\Pr\{\boldsymbol{E}\}$ non può essere calcolato, né si può affermare che esso esista. Per questo motivo tale definizione, per quanto intuitiva, non può essere presa in considerazione come base per lo sviluppo di una teoria matematica della probabilità. È quindi necessario definire il concetto di probabilità per via assiomatica, prescindendo da quello di frequenza relativa. L'approccio in termini di frequenza relativa, va tuttavia tenuto in considerazione, in quanto, in molti casi, esso si rivela utile nella giustificazione intuitiva, di alcuni sviluppi teorici la cui dimostrazione sarebbe inutilmente onerosa.

I.2 - Lo spazio di probabilità.

Si consideri lo spazio dei risultati $\boldsymbol{\Omega}$ di un esperimento casuale, ad esso si associ una classe Σ di suoi sottoinsiemi i cui elementi vengono chiamati eventi. Al generico evento $\boldsymbol{E}$ si associ un numero $\Pr\{\boldsymbol{E}\}$, detto probabilità dell'evento, che soddisfi le seguenti proprietà:

$$\begin{aligned} &0 \le \Pr\{\boldsymbol{E}\} \le 1 \\ &\Pr\{\boldsymbol{\Omega}\} = 1 \qquad\qquad \text{(I.2.1)} \\ &\Pr\{\boldsymbol{E}_1 \cup \boldsymbol{E}_2\} = \Pr\{\boldsymbol{E}_1\} + \Pr\{\boldsymbol{E}_2\} \quad \text{se} \quad \boldsymbol{E}_1 \cap \boldsymbol{E}_2 = \varnothing \end{aligned}$$

In altre parole la probabilità associata all'evento $\boldsymbol{E}$ deve assumere valore non negativo e non maggiore di 1; inoltre essa deve godere della proprietà additiva rispetto all'unione di eventi disgiunti. La probabilità dell'evento certo si pone uguale ad 1.

Nel caso in cui la classe Σ contenga un numero infinito numerabile di elementi la proprietà di additività semplice (I.2.1c) deve essere riformulata. Più precisamente, data una qualsiasi famiglia al più numerabile $\{\boldsymbol{E}_i\}$ di eventi, a due a due disgiunti, vale a dire tali che $\forall i \neq j, \boldsymbol{E}_i \cap \boldsymbol{E}_j = \varnothing$, deve essere:

$$\Pr\left\{\bigcup_{i=1}^{\infty} \boldsymbol{E}_i\right\} = \sum_{i=1}^{\infty} \Pr\{\boldsymbol{E}_i\} \qquad \text{(I.2.2)}$$

Quanto sopra esposto, equivale ad affermare che la probabilità $\Pr\{\cdot\}$ è un'applicazione $\Sigma \xrightarrow{\Pr} [0,1]$.

La classe Σ degli eventi non può essere scelta in modo totalmente arbitrario; essa deve, infatti, soddisfare le seguenti proprietà:

$$\begin{aligned} &\boldsymbol{\Omega} \in \Sigma \\ &\boldsymbol{E} \in \Sigma \Rightarrow \overline{\boldsymbol{E}} \in \Sigma \qquad\qquad \text{(I.2.3)} \\ &\boldsymbol{E}_1, \boldsymbol{E}_2 \in \Sigma \Rightarrow \boldsymbol{E}_1 \cup \boldsymbol{E}_2 \in \Sigma \end{aligned}$$

dove la proprietà (I.2.3,c) è sostituita dalla

$$\boldsymbol{E}_i \in \Sigma \Rightarrow \bigcup_{i=1}^{\infty} \boldsymbol{E}_i \in \Sigma$$

se la classe contiene elementi numerabili.

La proprietà (I.2.3,a) è intuitivamente giustificata dal fatto che se ad un certo evento è associata una probabilità, resta implicitamente definita la probabilità che l'evento non si sia verificato, cioè che l'esperimento casuale dia luogo ad un qualche risultato che non appartiene all'evento considerato. Pertanto l'insieme di detti risultati deve a sua volta costituire un evento.

Dalle precedenti discendono:

a) Dalla $\varnothing = \bar{\Omega}$ è $\varnothing \in \Sigma$; $\Omega \in \Sigma$

b) Se $E_1, E_2 \in \Sigma$ è anche $\bar{E}_1, \bar{E}_2 \in \Sigma$ e per la proprietà (I.2.3,c), e poiché è $\bar{E}_1 \cup \bar{E}_2 = \overline{E_1 \cap E_2}$ ne viene $\bar{E}_1 \cup \bar{E}_2 \in \Sigma$ e $\overline{E_1 \cap E_2} \in \Sigma$ e quindi $E_1 \cap E_2 \in \Sigma$.

c) Considerazione analoghe si possono ottenere nei confronti degli eventi $\bar{E}_1 \cup E_2$, $E_1 \cap \bar{E}_2$, etc.

d) In definitiva detti E_1 e E_2 due eventi, la classe Σ contiene oltre gli eventi considerati, lo spazio dei risultati e l'evento vuoto, tutte le complementazioni, le possibili unioni o intersezioni fra gli eventi che si possono così costruire.

In conclusione si afferma che si è definito uno **spazio di probabilità** $\mathcal{S}$ se si è assegnato:

- un insieme di risultati Ω (spazio dei risultati):
- una classe Σ di sottoinsiemi di Ω (eventi) che soddisfi le (I.2.3);
- un'applicazione $\Sigma \xrightarrow{\Pr} [0,1]$ definita su Σ (probabilità) che soddisfi le (I.2.1).

Ciò viene di norma sintetizzato dalla notazione

(I.2.4) $$\mathcal{S} = (\Omega, \Sigma, \Pr)$$

Dagli assiomi (I.2.1) discendono facilmente le seguenti proprietà:

- Gli eventi $\varnothing$ e Z sono manifestamente disgiunti, pertanto in base alla (I.2.1c) si può scrivere:

(I.2.5) $$\Pr\{\varnothing \cup \Omega\} = \Pr\{\varnothing\} + \Pr\{\Omega\}$$

dalla quale, notando che è:

(I.2.6) $$\varnothing \cup \Omega = \Sigma$$

consegue:

(I.2.7) $$\Pr\{\varnothing\} = 0$$

Ciò significa che la probabilità associata all'evento impossibile è nulla.

- Più in generale essendo $\bar{E}$ ed E, disgiunti e poiché $\bar{E} \cup E = \Omega$ risulta:

(I.2.8) $$\Pr\{\bar{E} \cup E\} = \Pr\{\bar{E}\} + \Pr\{E\} = \Pr\{\Omega\} = 1$$

Discende:

(I.2.9) $$\Pr\{\bar{E}\} = 1 - \Pr\{E\}$$

Pertanto la probabilità associata al complementare di un evento E è il complemento ad 1 della probabilità associata ad E.

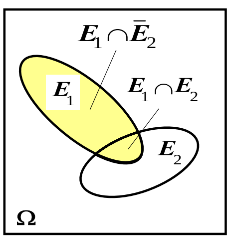

Fig.I.2 – Partizione di un evento.

- Dati E_1 e E_2, l'evento E_1 può essere scomposto nei due eventi disgiunti (v. Fig. I.2) $E_1 \cap E_2$ e $E_1 \cap \bar{E}_2$. Risulta quindi:

(I.2.10) $$\Pr\{E_1\} = \Pr\{E_1 \cap E_2\} + \Pr\{E_1 \cap \bar{E}_2\}$$

In modo analogo si ha:

(I.2.11) $$\Pr\{E_2\} = \Pr\{E_1 \cap E_2\} + \Pr\{\bar{E}_1 \cap E_2\}$$

che sommate termine a termine forniscono:

(I.2.12) $$\Pr\{E_1\}+\Pr\{E_2\}=2\Pr\{E_1\cap E_2\}+\Pr\{E_1\cap\bar{E}_2\}+\Pr\{\bar{E}_1\cap E_2\}$$

Poiché gli eventi $E_1\cap E_2$, $E_1\cap\bar{E}_2$ e $\bar{E}_1\cap E_2$ sono disgiunti e la loro unione vale $E_1\cup E_2$, si ha:

(I.2.13) $$\Pr\{E_1\cup E_2\}=\Pr\{E_1\cap E_2\}+\Pr\{E_1\cap\bar{E}_2\}+\Pr\{\bar{E}_1\cap E_2\}$$

e quindi:

(I.2.14) $$\Pr\{E_1\cup E_2\}=\Pr\{E_1\}+\Pr\{E_2\}-\Pr\{E_1\cap E_2\}$$

che costituisce una generalizzazione della proprietà (I.2.1) al caso di eventi non disgiunti.

I.3 - Probabilità condizionate. Formula di Bayes. Teorema delle probabilità composte.

Siano ΔN_{E_1} e $\Delta N_{E_1\cap E_2}$ il numero di volte che, su un totale di N ripetizioni dell'esperimento casuale, si manifestano gli eventi E_1 e $E_1\cap E_2$. Dalla relazione:

(I.3.1) $$\frac{\Delta N_{E_1\cap E_2}}{N}=\frac{\Delta N_{E_1\cap E_2}}{\Delta N_{E_1}}\cdot\frac{\Delta N_{E_1}}{N}$$

si ottiene, passando al limite per $N\to\infty$

(I.3.2) $$\Pr\{E_1\cap E_2\}=\Pr\{E_2\mid E_1\}\Pr\{E_1\}$$

dove è:

(I.3.3) $$\Pr\{E_2\mid E_1\}=\lim_{N\to\infty}\frac{\Delta N_{E_1\cap E_2}}{\Delta N_{E_1}}$$

Osservando che $\frac{\Delta N_{E_1\cap E_2}}{\Delta N_{E_1}}$ rappresenta il rapporto fra il numero di volte in cui, in un totale di N ripetizioni dell'esperimento casuale, si verifica l'evento $E_1\cap E_2$ e il numero di volte con cui si verifica l'evento E_1, la $\Pr\{E_2\mid E_1\}$ può essere interpretata come la probabilità che si verifichi l'evento E_2 sotto l'ipotesi che E_1 sia soddisfatto. Per questo motivo $\Pr\{E_2\mid E_1\}$ è detta **probabilità dell'evento E_2 condizionata da E_1**.

In modo analogo può scriversi:

(I.3.4) $$\Pr\{E_1\cap E_2\}=\Pr\{E_1\mid E_2\}\Pr\{E_2\}$$

dove $\Pr\{E_1\mid E_2\}$ denota la probabilità che si manifesti E_1 atteso che E_2 sia verificato.

Dal confronto fra le (I.3.2) e (I.3.3) si ottiene:

(I.3.5) $$\Pr\{E_1\mid E_2\}=\Pr\{E_2\mid E_1\}\frac{\Pr\{E_1\}}{\Pr\{E_2\}}$$

nota come **formula di Bayes**. Essa stabilisce la relazione tra le probabilità condizionate e quelle non condizionate degli eventi E_1 e E_2.

Se risulta:

(I.3.6) $$\Pr\{E_1\mid E_2\}=\Pr\{E_1\}\ \text{ o }\ \Pr\{E_2\mid E_1\}=\Pr\{E_2\}$$

cioè se la probabilità con cui si manifesta E_1 (E_2) è indipendente dalla circostanza che E_2 (E_1) sia verificato, i due eventi si dicono **statisticamente indipendenti**. Pertanto, in base alle (I.3.2) o (I.3.4) si ha:

(I.3.7) $$\Pr\{E_1\cap E_2\}=\Pr\{E_1\}\cdot\Pr\{E_2\}$$

La probabilità dell'intersezione di due eventi indipendenti si riduce cioè semplicemente al prodotto delle probabilità associate ai singoli eventi.

Si consideri una famiglia al più numerabile $\{E_i\}$ di sottoinsiemi di Ω a due a due disgiunti, tale che $\bigcup_{i=1}^{\infty} E_i = \Omega$. Un qualunque evento E può essere scomposto come segue: (v. Fig. I.3):

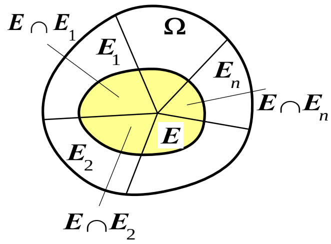

Fig.I.3 - Teorema delle probabilità composte.

(I.3.7) $$E = \bigcup_{i=1}^{\infty} E \cap E_i$$

Poiché è:

(I.3.8) $$(E \cap E_i) \cap (E \cap E_j) = \varnothing \quad \forall i \neq j$$

in base all'assioma (I.2.2) si ha:

(I.3.10) $$\Pr\{E\} = \sum_{i=1}^{\infty} \Pr\{E \cap E_i\}$$

che, utilizzando la (1.3.2), fornisce:

(I.3.11) $$\Pr\{E\} = \sum_{i=1}^{\infty} \Pr\{E \mid E_i\} \Pr\{E_i\}$$

nota come **Teorema delle probabilità composte**.

I.4 - Variabili aleatorie monodimensionali.

Si consideri un esperimento casuale caratterizzato da uno spazio di probabilità:

(I.4.1) $$\mathcal{S} = (\Omega, \Sigma, \Pr)$$

e sia $X(\cdot)$ un'applicazione che fa corrispondere ad ogni risultato $\zeta \in \Omega$ un numero reale. Il dominio di tale applicazione è quindi l'intero spazio dei risultati, il suo codominio è l'insieme $\mathbb{R}$. Sia E_x il sottoinsieme composto da tutti quei risultati a cui, tramite l'applicazione $X(\cdot)$, corrisponde un valore non superiore a x;e cioè:

(I.4.2) $$E_x \equiv \{\zeta : X(\zeta) \le x\}$$

Se qualunque sia il valore di x il sottoinsieme E_x è un evento e cioè se $E_x \subset \mathbf{F}$, si dice che X è una **variabile aleatoria** associata all'esperimento casuale.

Ad esempio se nell'esperimento casuale "lancio di una moneta", assumendo $\mathcal{S} = \{\varnothing, t, c, \Omega\}$, si definisce l'applicazione $X(\cdot)$ che associa al risultato "testa" il valore 0 e al risultato "croce" il valore 1, si è definita una variabile aleatoria. Infatti:

- ogni semiretta $(-\infty, x]$ con $x < 0$ ha come controimmagine l'insieme vuoto che è un evento;
- ogni semiretta $(-\infty, x]$ con $0 \le x < 1$ ha come controimmagine l'insieme $E_x = \{t\} \in \Sigma$;
- ogni semiretta $(-\infty, x]$ con $x \ge 1$ ha come controimmagine l'insieme Ω che è anch'esso un evento.

In sostanza definire una variabile aleatoria equivale a costruire un nuovo esperimento casuale che ha come spazio dei risultati l'insieme $\mathbb{R}$, e come classe di eventi la classe $\mathbf{B}$, ottenuta effettuando, quante volte si voglia, operazioni di unione, intersezione, complementazione degli eventi del tipo $(-\infty, x]$. Tale classe, per quel che interessa l'analisi dei segnali aleatori, contiene tutti gli intervalli (chiusi, aperti e semiaperti), tutte le semirette (aperte e chiuse all'origine) e tutti i punti di $\mathbb{R}$.

I.5 - Funzione di distribuzione di probabilità.

Si consideri un esperimento casuale caratterizzato da uno spazio di probabilità $\mathbf{S}=(\Omega,\mathbf{F},\Pr_E)$. Ci si rende facilmente conto che è possibile calcolare la probabilità $\Pr_B$ da attribuire al generico evento $B \in \mathbf{B}$, se si definisce la seguente funzione:

(I.5.1) $$P_X(x) = \Pr\{E_x\} = \Pr\{\zeta : X(\zeta) \le x\}$$

detta **funzione di distribuzione di probabilità**, associata alla variabile aleatoria X.

Risulta infatti:

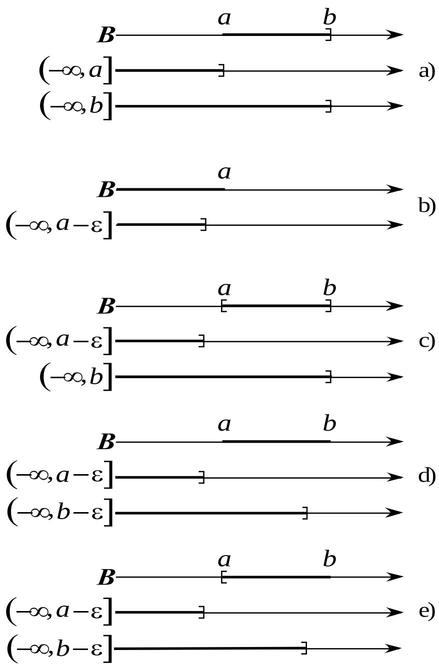

Fig.I.4 – Insiemi in $\mathbb{R}$.

a) intervallo semiaperto a sinistra (Fig.I.4,a).

Sia:

(I.5.1) $$B = (a,b] = \{a < X \le b\}$$

Si ha:

(I.5.2) $$\{X \le b\} = \{X \le a\} \cup \{a < X \le b\}$$

essendo gli eventi $\{X \le a\}$ e $\{a < X \le b\}$ disgiunti, risulta:

(I.5.3) $$\Pr\{a < X \le b\} = P_X(b) - P_X(a)$$

b) semiretta d'origine destra aperta (Fig.I.4,b).

Sia

(I.5.4) $$B = (-\infty, a) = \{X < a\}$$

Poiché si può scrivere:

(I.5.5) $$B = \lim_{\varepsilon\to 0}\{X \le a-\varepsilon\}$$

risulta:

(I.5.6) $$\Pr\{X < a\} = \lim_{\varepsilon\to 0} P_X(a-\varepsilon) = P_X(a^-)$$

c) intervallo chiuso (Fig.I.4,c).

Sia

(I.5.7) $$B = [a,b] = \{a \le X \le b\}$$

Dalla condizione:

(I.5.8) $$\{X \le b\} = \lim_{\varepsilon\to 0}\{X \le a-\varepsilon\} \cup \{a \le X \le b\}$$

essendo $\{X \le a-\varepsilon\} \cap \{a \le X \le b\} = \varnothing$ si ottiene:

(I.5.9) $$\Pr\{a \le X \le b\} = P_X(b) - P_X(a^-)$$

d) intervallo aperto (Fig.I.4,d).

Sia

(I.5.10) $$B = (a,b) = \{a < X < b\}$$

Dalla condizione

(I.5.11) $$\lim_{\varepsilon\to 0}\{X \le b-\varepsilon\} = \{X \le a\} \cup \{a < X < b\}$$

essendo $\{X \le a\} \cap \{a < X < b\} = \varnothing$, si ottiene:

(I.5.12) $$\Pr\{a < X < b\} = P_X(b^-) - P_X(a)$$

e) intervallo semiaperto a destra (Fig.I.4,e).

Sia

(I.5.13) $$B = [a,b) = \{a \le X < b\}$$

Dalla condizione

(I.5.14) $$\lim_{\varepsilon\to 0}\{X \le b-\varepsilon\} = \lim_{\varepsilon\to 0}\{X \le a-\varepsilon\} \cup \{a \le X < b\}$$

Eessendo $\{X \le a-\varepsilon\} \cap \{a \le X < b\} = \varnothing$, è:

(I.5.15) $$\Pr\{a \le X < b\} = P_X(b^-) - P_X(a^-)$$

Quanto sopra esposto, evidenzia chiaramente che la distribuzione di probabilità fornisce una descrizione statistica completa della variabile aleatoria X; cosicché, normalmente, si fa riferimento allo spazio di probabilità indotto in $\mathbb{R}$ dalla variabile aleatoria, piuttosto che allo spazio di probabilità originario **S**.

Esempio E.I.5

Nel caso della variabile aleatoria X, che nel lancio di una moneta associa 0 al risultato testa e 1 al risultato croce, si ottiene, assumendo gli eventi $\{\text{testa}\}$ e $\{\text{croce}\}$ equiprobabili:

$$P_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

il cui andamento è riportato in Fig. E.I.1.

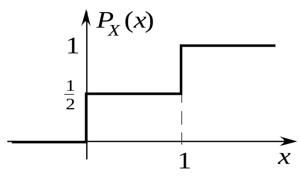

Fig.E.I.1

I.6 - Proprietà della distribuzione di probabilità.

La distribuzione di probabilità caratterizza completamente una variabile aleatoria. È quindi importante studiarne le proprietà. In quel che segue se ne elencano alcune.

a) **Valori limite**

Sia $P_X(x)$ la distribuzione di probabilità di una variabile aleatoria X. Se si fa tendere il suo argomento a $-\infty$ l'insieme $(-\infty, x]$ si identifica con l'insieme vuoto. È allora:

(I.6.1) $$\lim_{x\to -\infty} P_X(x) = 0$$

Se viceversa si fa tendere x a $+\infty$ l'insieme $(-\infty, x]$ si identifica con $\mathbb{R}$ la cui controimmagine, secondo X, è $\boldsymbol{\Omega}$; cosicché risulta:

(I.6.2) $$\lim_{x\to\infty} P_X(x) = 1$$

b) **Monotonia e limitatezza**

Se $x_1 \le x_2$ risulta $\{X \le x_1\} \subseteq \{X \le x_2\}$ e questo comporta:

(I.6.3) $$P_X(x_1) \le P_X(x_2)$$

la funzione $P_X(x)$ è una funzione non decrescente di x; in aggiunta, tenendo presente la (I.6.2), è

(I.6.4) $$P_X(x) \le P_X(+\infty) = 1$$

c) **Continuità**

Tenendo presente la (I.5.3) si può scrivere:

(I.6.5) $$\Pr\{x_0 < X \le x_0 + \varepsilon\} = P_X(x_0+\varepsilon) - P_X(x_0)$$

e quindi al limite, per $\varepsilon \to 0$:

(I.6.6) $$\lim_{\varepsilon\to 0} \Pr\{x_0 < X \le x_0 + \varepsilon\} = \lim_{\varepsilon\to 0}\left[P_X(x_0+\varepsilon) - P_X(x_0)\right] = P_X(x_0^+) - P_X(x_0)$$

D'altra parte è evidente che $\lim_{\varepsilon\to 0}\{x_0 < X \le x_0 + \varepsilon\} = \varnothing$, per cui dalla precedente si ottiene:

(I.6.7) $$P_X(x_0^+) = P_X(x_0)$$

La funzione $P_X(x)$ è continua a destra.

Ponendo nella (I.5.9) $b = a = x_0$ si ha:

(I.6.8) $$\Pr\{X = x_0\} = P_X(x_0) - P_X(x_0^-)$$

La funzione $P_X(x)$ non è continua a sinistra a meno che non sia $\Pr\{X = x_0\} = 0$.

L'andamento della funzione di distribuzione di probabilità associata ad una data variabile aleatoria, suggerisce una possibile classificazione delle variabili aleatorie. Precisamente, se la $P_X(x)$ è continua in $\mathbb{R}$, la variabile aleatoria cui essa è associata si dice **di tipo continuo**, se la $P_X(x)$ è una funzione costante a tratti la variabile si dice **di tipo discreto**, nei restanti casi si parla di variabile aleatoria **di tipo misto**.

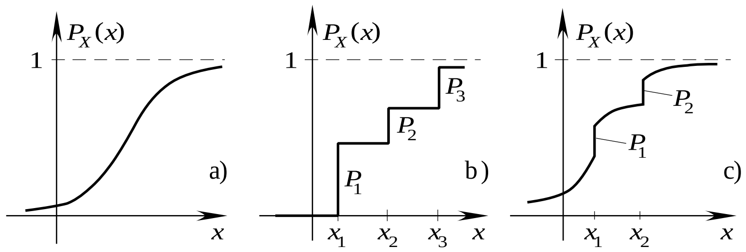

Fig.I.5 – Funzioni di distribuzione per variabili aleatorie di tipo continuo a), discreto b) e misto c).

Per maggior chiarezza, gli andamenti tipici della funzione di distribuzione di probabilità per i diversi tipi di variabili aleatorie sono mostrati in Fig. I.5.

I.7 - Densità di probabilità di una variabile aleatoria continua.

Una variabile aleatoria continua che ammetta una distribuzione di probabilità può essere caratterizzata anche mediante la cosiddetta **densità di probabilità** $p_X(\cdot)$:

(I.7.1) $$p_X(x) = \frac{dP_X(x)}{dx}$$

dove la derivata può contenere distribuzioni delte di Dirac.

Dalle proprietà viste nel paragrafo precedente, relative alla distribuzione di probabilità, si deducono facilmente le corrispondenti proprietà che caratterizzano la funzione densità di probabilità di una variabile aleatoria.

Poiché la distribuzione di probabilità è una primitiva della rispettiva densità deve risultare:

(I.7.2) $$\int_a^b p_X(x)dx = P_X(b) - P_X(a) \equiv \Pr\{a < X \le b\}$$

Se l'integrale si estende all'intero asse reale si ottiene:

(I.7.3) $$\int_{-\infty}^{\infty} p_X(x)dx = P_X(\infty) - P_X(-\infty) = 1$$

che costituisce la **condizione di normalizzazione**. Essa significa che l'area sottesa dalla densità di probabilità di una variabile aleatoria è sempre unitaria.

Inoltre poiché la $P_X(x)$ è una funzione non decrescente del suo argomento, la densità di probabilità di una variabile aleatoria non può assumere valori negativi.

(I.7.4) $$p_X(x) \ge 0$$

Il concetto di densità di probabilità può essere esteso facilmente anche al caso di variabili aleatorie di tipo discreto. Una variabile di tipo discreto ha una densità di probabilità costituita da un insieme di delta di Dirac localizzate nei punti di discontinuità e di ampiezza pari ai relativi salti della corrispondente distribuzione di probabilità.

È facile rendersi conto che la condizione di normalizzazione vale anche per le variabili aleatorie di tipo discreto. Se P_i denota la probabilità associata all'evento $\{X = x_i\}$ si ha:

(I.7.5)
$$\sum_{i=1}^{n} \Pr\{X = x_i\} = \sum_{i=1}^{n} P_i = 1$$

dove la sommatoria può estendersi fino a ∞.

I.8 – Esempi di funzioni di probabilità.

In quel che segue sono riportate le più comuni funzioni di probabilità (densità e distribuzione) sia per variabili aleatorie continue che discrete.

I.8.1 - Distribuzione uniforme.

Una variabile aleatoria X si dice uniformemente distribuita nell'intervallo (a,b) se la sua densità di probabilità si mantiene costante in detto intervallo ed è nulla in tutti i punti esterni ad esso. Essa si denota con $X \sim \mathbf{U}(a,b)$. Naturalmente, dovendo essere soddisfatta la condizione di normalizzazione si ha (v. Fig. I.6, a):

(I.8.1)
$$p_X(x) = \frac{1}{b-a}\,\mathrm{rect}\left(\frac{x - \frac{a+b}{2}}{b-a}\right)$$

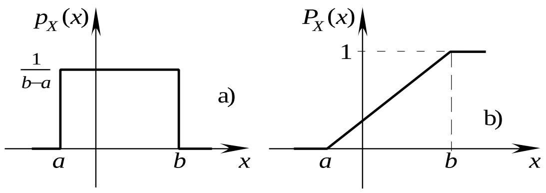

Fig.I.6 - Densità e distribuzione uniforme

Essa descrive il comportamento di una quantità aleatoria che assume con eguale probabilità un qualsiasi valore appartenente ad (a,b).

La funzione di distribuzione associata ad una variabile aleatoria uniformemente distribuita in (a,b) vale (v. Fig. I.4, b):

(I.8.2)
$$P_X(x) = \begin{cases} 0 & x \le a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases}$$

I.8.2 - Distribuzione normale o gaussiana.

Una variabile aleatoria si dice gaussiana o normale di parametri m e σ^2 se la sua densità di probabilità è del tipo:

(I.8.3)
$$p_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

qualunque sia $m \in \mathbb{R}$ e $\sigma \in \mathbb{R}$. Essa si denota con $X \sim \mathbf{N}(m,\sigma^2)$. L'andamento della densità di probabilità di una variabile gaussiana con $m = 0$ e per alcuni valori del parametro σ è mostrato in Fig. I.7.

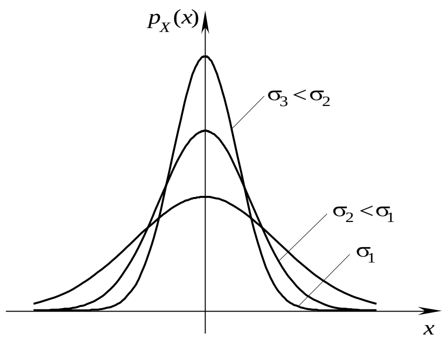

Fig. I.7- Densità di probabilità a) e funzione di distribuzione b) gaussiana.

La funzione di distribuzione di probabilità vale:

(I.8.4)
$$P_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(z-m)^2}{2\sigma^2}}\,dz$$

L'integrale che compare nella precedente non può essere calcolato in forma chiusa. Tuttavia esso si può esprimere in termini della cosiddetta **funzione Q**, definita come segue:

(I.8.5) $$Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^{\infty}e^{-u^2/2}du$$

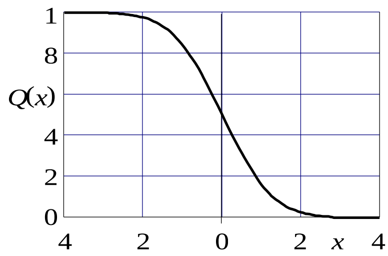

Fig.I.9 – Funzione $Q(x)$.

(vedi Fig. I.9). A tale scopo basta effettuare nell'integrale che compare nella (I.8.4) la trasformazione di variabili $u=(z-m)/\sigma$. Si ottiene:

(I.8.6) $$\begin{aligned}P_X(x)&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}e^{-\frac{(z-m)^2}{2\sigma^2}}dz-\frac{1}{\sqrt{2\pi\sigma^2}}\int_x^{\infty}e^{-\frac{(z-m)^2}{2\sigma^2}}dz=1-\frac{1}{\sqrt{2\pi\sigma^2}}\int_x^{\infty}e^{-\frac{(z-m)^2}{2\sigma^2}}dz=\\&=1-\frac{1}{\sqrt{2\pi}}\int_{(x-m)/\sigma}^{\infty}e^{-u^2/2}du=1-Q\left(\frac{x-m}{\sigma}\right)\end{aligned}$$

La funzione $P_X(x)$ è rappresentata nella Fig.I.8.

I.9 – Funzioni di probabilità condizionate.

Sia $\mathcal{S}=(\Omega,\Sigma,\Pr)$ uno spazio di probabilità e $X(\zeta)$ una variabile aleatoria definita su **S**. Se $\boldsymbol{E}$ denota un evento non vuoto per modo che si abbia $\Pr\{\boldsymbol{E}\}\neq 0$, la quantità

(I.9.1) $$P_{X|\boldsymbol{E}}\left(X\le x\,\middle|\,X\in\boldsymbol{E}\right)=\frac{\Pr\{X\le x;X\in\boldsymbol{E}\}}{\Pr\{X\in\boldsymbol{E}\}}$$

dove $\{X\le x;X\in\boldsymbol{E}\}$ denota l'intersezione fra gli insiemi $\{X\le x\}$ e $\{X\in\boldsymbol{E}\}$, rappresenta la probabilità che la variabile aleatoria $X(\zeta)$ assuma valori non superiori a x atteso che X appartiene ad $\boldsymbol{E}$. Essendo ovviamente:

(I.9.2) $$\begin{aligned}&\{X\le-\infty\}\cap\{X\in\boldsymbol{E}\}=\varnothing\cap\{X\in\boldsymbol{E}\}=\varnothing\\&\{X\le\infty\}\cap\{X\in\boldsymbol{E}\}=\boldsymbol{\Omega}\cap\{X\in\boldsymbol{E}\}=\{X\in\boldsymbol{E}\}\\&\forall x_1,x_2\text{ con }x_1\le x_2\Rightarrow\{X\le x_1\}\cap\{X\in\boldsymbol{E}\}\subseteq\{X\le x_2\}\cap\{X\in\boldsymbol{E}\}\end{aligned}$$

è:

(I.9.3) $$\begin{aligned}&P_{X|\boldsymbol{E}}\left(X\le-\infty\,\middle|\,X\in\boldsymbol{E}\right)=0\\&P_{X|\boldsymbol{E}}\left(X\le\infty\,\middle|\,X\in\boldsymbol{E}\right)=1\\&\forall x_1,x_2\text{ con }x_1\le x_2\Rightarrow P_{X|\boldsymbol{E}}\left(X\le x_1\,\middle|\,X\in\boldsymbol{E}\right)\le P_{X|\boldsymbol{E}}\left(X\le x_2\,\middle|\,X\in\boldsymbol{E}\right)\end{aligned}$$

La $P_{X|\boldsymbol{E}}\left(X\le x\,\middle|\,X\in\boldsymbol{E}\right)$ quindi presenta le stesse proprietà di una distribuzione di probabilità e per questo motivo la funzione:

(I.9.4) $$P_{X|\boldsymbol{E}}\left(x\,\middle|\,X\in\boldsymbol{E}\right)=\frac{\Pr\{X\le x;X\in\boldsymbol{E}\}}{\Pr\{X\in\boldsymbol{E}\}}$$

definisce la **funzione distribuzione di probabilità** condizionata dall'evento $(X\in\boldsymbol{E})$.

Derivando la (I.9.4) rispetto a x si ottiene:

(I.9.5) $$p_{X|\boldsymbol{E}}\left(x\,\middle|\,X\in\boldsymbol{E}\right)=\frac{d}{dx}P_{X|\boldsymbol{E}}\left(x\,\middle|\,X\in\boldsymbol{E}\right)$$

che costituisce la **densità di probabilità condizionata** dall'evento $(X\in\boldsymbol{E})$.

Si ha:

(I.9.6) $$\int_{-\infty}^{\infty}p_{X|\boldsymbol{E}}\left(x\,\middle|\,X\in\boldsymbol{E}\right)dx=1$$

che costtituisce la condizione di normalizzazione per la densità di probabilità condizionata.

I.10 - Variabili aleatorie bidimensionali. Funzioni di probabilità congiunte.

Sia $\mathcal{S}=(\Omega,\Sigma,\Pr)$ uno spazio di probabilità e siano $X(\zeta)$ ed $Y(\zeta)$ due variabili aleatorie definite sull'insieme dei risultati. Se, per ogni coppia (x,y) di numeri reali, gli insiemi $\boldsymbol{E}_x=\{\zeta : X(\zeta)\le x\}$ e $\boldsymbol{E}_y=\{\zeta : Y(\zeta)\le y\}$ sono eventi in $\boldsymbol{\Omega}$, il sottoinsieme $\boldsymbol{E}_{xy}$ definito dalla:

(I.10.1) $$\boldsymbol{E}_{xy}=\boldsymbol{E}_x\cap\boldsymbol{E}_y=\{\zeta : X(\zeta)\le x\,;Y(\zeta)\le y\}$$

costituisce un evento in $\boldsymbol{\Omega}$ in quanto ottenuto dall'intersezione di eventi in $\boldsymbol{\Omega}$. Le grandezze X, Y definiscono allora una coppia di variabili aleatorie o, alternativamente, una variabile aleatoria bidimensionale associata ad **S** che, come nel caso monodimensionale, può essere caratterizzata dalla funzione $P_{XY}(x,y)$:

(I.10.2) $$P_{XY}(x,y)=\Pr\{\zeta : X\le x;Y\le y\}$$

che prende il nome di **distribuzione di probabilità congiunta**.

Dal momento che si ha $\{\zeta : X(\zeta)<\infty\}=\{\zeta : Y(\zeta)<\infty\}=\boldsymbol{\Omega}$ e $\boldsymbol{\Omega}\cap\boldsymbol{E}=\boldsymbol{E}$ valgono le:

(I.10.3) $$\begin{aligned}P_{XY}(\infty,y)&=P_Y(y)\\ P_{XY}(x,\infty)&=P_X(x)\end{aligned}$$

cioè le distribuzioni di probabilità (monodimensionali) associate alle variabili aleatorie X ed Y possono essere dedotte dalla distribuzione di probabilità congiunta associata alla variabile aleatoria bidimensionale (X,Y). In omaggio a questa circostanza, talvolta le funzioni $P_X(x)$ e $P_Y(y)$ sono denominate **distribuzioni marginali**.

La funzione distribuzione di probabilità congiunta gode delle proprietà qui sotto elencate:

a) Se si fa tendere x o y a $-\infty$, l'insieme $\boldsymbol{E}_{xy}$ tende all'insieme vuoto. Si ha pertanto:

(I.10.4) $$\begin{aligned}P_{XY}(-\infty,y)&=0\\ P_{XY}(x,-\infty)&=0\\ P_{XY}(-\infty,-\infty)&=0\end{aligned}$$

b) Se si fanno tendere entrambe le quantità x e y a $+\infty$ l'insieme $\boldsymbol{E}_{xy}$ tende a $\boldsymbol{\Omega}$, ciò comporta

(I.10.5) $$P_{XY}(\infty,\infty)=1$$

c) Poiché $x_1\le x_2$ e $y_1\le y_2$ implica $\boldsymbol{E}_{x_1y_1}\subseteq\boldsymbol{E}_{x_2y_2}$ la $P_{XY}(\cdot,\cdot)$ deve soddisfare la proprietà:

(I.10.6) $$P_{XY}(x_1,y_1)\le P_{XY}(x_2,y_2)$$

La caratterizzazione statistica di una variabile bidimensionale (X,Y) può essere ottenuta per mezzo della cosiddetta **densità di probabilità congiunta** definita dalla:

(I.10.7) $$p_{XY}(x,y)=\frac{\partial^2 P_{XY}(x,y)}{\partial x\partial y}$$

Se la derivata che compare nella (I.10.7) è intesa in senso distribuzionale, il concetto di densità di probabilità può essere esteso al caso di variabili aleatorie bidimensionali discrete o miste per le quali la funzione $P_{XY}(x,y)$ può presentare dei salti di ampiezza finita.

Le proprietà della funzione $P_{XY}(x,y)$ sopra riportate, si traducono facilmente in termini della funzione $p_{XY}(x,y)$. Si ha così:

(I.10.8) $$p_{XY}(x,y) \geq 0$$

nei punti un cui la densità di probabilità congiunta è continua. Altrimenti deve presentare distribuzioni delta con pesi positivi. Inoltre è:

(I.10.9) $$\int_{x_1}^{x_2}\int_{y_1}^{y_2} p_{XY}(\xi,\eta)d\xi d\eta = \Pr\{\zeta : x_1 < X(\zeta) \leq x_2 ; y_1 < Y(\zeta) \leq y_2\}$$

In particolare dalla precedente, tenendo conto della terza delle (I.10.4), si deduce:

(I.10.10) $$P_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} p_{XY}(\xi,\eta)d\xi d\eta$$

La condizione (I.10.5) si traduce nella seguente condizione di normalizzazione:

(I.10.11) $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_{XY}(\xi,\eta)d\xi d\eta = 1$$

Per quanto riguarda le densità di probabilità marginali dalle (I.10.4) tenendo conto della (I.10.10), si ottengono le:

(I.10.12) $$P_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} p_{XY}(\xi,\eta)d\xi d\eta$$
$$P_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} p_{XY}(\xi,\eta)d\xi d\eta$$

le quali derivate rispettivamente rispetto a x e a y forniscono:

(I.10.13) $$p_X(x) = \int_{-\infty}^{\infty} p_{XY}(x,y)dy$$
$$p_Y(y) = \int_{-\infty}^{\infty} p_{XY}(x,y)dx$$

Le variabili aleatorie X e Y si dicono **statisticamente indipendenti** se la loro distribuzione di probabilità congiunta si può esprimere come segue:

(I.10.14) $$P_{XY}(x,y) = P_X(x)\cdot P_Y(y)$$

o, in modo equivalente, la densità congiunta si può scrivere:

(I.10.15) $$p_{XY}(x,y) = p_X(x)\cdot p_Y(y)$$

cioè due variabili aleatorie associate ad uno stesso esperimento casuale si dicono statisticamente indipendenti se le funzioni di probabilità congiunte si fattorizzano in termini delle rispettive funzioni di probabilità marginali.

Tutte le considerazioni fin qui esposte nel caso di due variabili aleatorie possono essere facilmente estese al caso di n variabili aleatorie definite su uno stesso esperimento casuale o, che è lo stesso, al caso di un vettore aleatorio n-dimensionale.

Nella caratterizzazione statistica dei segnali aleatori ci si imbatte in funzioni di probabilità condizionate del tipo

$$\Pr\left\{Y \in \boldsymbol{E}_y \middle| X = x\right\} \text{ o } \Pr\left\{X \in \boldsymbol{E}_x \middle| Y = y\right\}$$

che presentano un qualche difficoltà dal momento che l'applicazione della formula di Bayes conduce ad una forma indeterminata del tipo $\frac{0}{0}$. Ciononostante si può ottenere il risultato se si considerano le probabilità condizionate ottenute dal limite seguente:

(I.10.16) $$\Pr\left\{Y \in \boldsymbol{E}_y \middle| X = x\right\} = \lim_{\Delta x \to 0} \Pr\left\{Y \in \boldsymbol{E}_y \middle| x < X \leq x + \Delta x\right\}$$

e analogamente per la $\Pr\left\{X \in \boldsymbol{E}_x \middle| Y = y\right\}$. Infatti utilizzando la formula di Bayes si può scrivere:

(I.10.17) $$\Pr\left\{Y \in \boldsymbol{E}_y \middle| x < X \le x+\Delta x\right\} = \frac{\Pr\left\{Y \in \boldsymbol{E}_y ; x < X \le x+\Delta x\right\}}{\Pr\left\{x < X \le x+\Delta x\right\}} = \frac{\int_x^{x+\Delta x}\int_{\boldsymbol{E}_y} p_{XY}(\xi,\eta)d\xi d\eta}{\int_x^{x+\Delta x} p_X(\xi)d\xi}$$

essendo $p_X(x)$ la densità di probabilità marginale della variabile X. Effettuando il limite si ha:

(I.10.18) $$\Pr\left\{Y \in \boldsymbol{E}_y \middle| X = x\right\} = \frac{\int_{\boldsymbol{E}_y} p_{XY}(x,\eta)d\eta}{p_X(x)}$$

Si noti che la precedente è definita per tutti i valori di x per cui risulta $p_X(x) \neq 0$.

In modo analogo si ha:

(I.10.19) $$\Pr\left\{x \in \boldsymbol{E}_x \middle| Y = y\right\} = \frac{\int_{\boldsymbol{E}_x} p_{XY}(\xi,y)d\xi}{p_y(y)}$$

valida per valori di y per cui risulta $p_Y(y) \neq 0$.

Se si particolarizzano gli eventi $\boldsymbol{E}_x$ e $\boldsymbol{E}_y$ come segue:

$$\boldsymbol{E}_x = \left\{X \le x\right\} \text{ e } \boldsymbol{E}_y = \left\{Y \le y\right\}$$

le (I.10.18) e (I.10.19) definisco le distribuzioni di probabilità condizionate definite dalle:

(I.10.20) $$P_{X|Y=y}\left(x\middle|Y=y\right) = \frac{\int_{-\infty}^{x} p_{XY}(\xi,y)d\xi}{p_Y(y)}$$
$$P_{Y|X=x}\left(y\middle|X=x\right) = \frac{\int_{-\infty}^{y} p_{XY}(x,\eta)d\eta}{p_X(x)}$$

da cui derivando si deducono le seguenti densità di probabilità condizionate:

(I.10.21) $$p_{X|Y=y}\left(x\middle|Y=y\right) = \frac{p_{XY}(x,y)}{p_Y(y)} \qquad p_{Y|X=x}\left(y\middle|X=x\right) = \frac{p_{XY}(x,y)}{p_X(x)}$$

od anche:

(I.10.22) $$p_{XY}(x,y) = p_{X|Y=y}\left(x\middle|Y=y\right)\cdot p_Y(y) = p_{Y|X=x}\left(y\middle|X=x\right)\cdot p_X(x)$$

La densità di probabilità congiunta di una coppia di variabili aleatorie può essere espressa in termini di delle densità di probabilità condizionate e della corrispondente densità di probabilità marginali.

I.11 – Medie statistiche.

Ad ogni variabile aleatoria $X(\zeta)$ definite su uno spazio di probabilità $\mathcal{S}$ si può associare un importante operatore detto media statistica. Sia $y = f(x)$ una funzione continua, la quantità $Y(\zeta) = f\left[X(\zeta)\right]$, come è facile riconoscere, individua un'altra variabile aleatoria. La quantità

(I.11.1) $$E\left\{f(x)\right\} \equiv \overline{f(x)} \equiv \int_{-\infty}^{\infty} f(x)p_X(x)dx$$

costituisce la media statistica di y.

Particolare importanza rivestono i cosiddetti **momenti** della variabile aleatoria definiti ponendo $f(\cdot) = (\cdot)^n$. Essi valgono:

(I.11.2) $$m_n \equiv E\left\{x^n\right\} \equiv \overline{x^n} \equiv \int_{-\infty}^{\infty} x^n p_X(x)dx \qquad n = 0,1,2,\ldots$$

Di particolare interesse sono:

a) il momento $m_X^{(1)}$ che costituisce il *valore medio* di $X(\zeta)$:

(I.11.3) $$m_X^{(1)} \equiv E\{x\} \equiv \overline{x} \equiv \int_{-\infty}^{\infty} x p_X(x) dx$$

b) il momento $m_X^{(2)}$ che costituisce il *valore quadratico medio* di $X(\zeta)$:

(I.11.4) $$m_X^{(2)} \equiv E\left\{x^2\right\} \equiv \overline{x^2} \equiv \int_{-\infty}^{\infty} x^2 p_X(x) dx$$

Analogamente ai momenti $m_X^{(n)}$ sono definiti i **momenti centrali** definiti dalla:

(I.11.5) $$\mu_X^{(n)} \equiv E\left\{\left(x - m_X^{(1)}\right)^n\right\} \equiv \overline{\left(x - m_X^{(1)}\right)^n} \equiv \int_{-\infty}^{\infty} \left(x - m_X^{(1)}\right)^n p_X(x) dx$$

Di particolare interesse è il momento centrale σ^2 detto **varianza** della variabile aleatoria $X(\zeta)$:

(I.11.6) $$\sigma_X^2 \equiv E\left\{\left(x - m_X^{(1)}\right)^2\right\} \equiv \overline{\left(x - m_X^{(1)}\right)^2} \equiv \int_{-\infty}^{\infty} \left(x - m_X^{(1)}\right)^2 p_X(x) dx$$

Dalla (I.11.6), essendo $\left(x - m_X^{(1)}\right)^2 = x^2 - 2x m_X^{(1)} + \left[m_X^{(1)}\right]^2$, si deduce:

(I.11.7) $$\sigma_X^2 = m_X^{(2)} - \left[m_X^{(1)}\right]^2$$

Analoghi operatori di media statistica possono essere associati ad una coppia di variabili aleatorie $X(\zeta)$ e $Y(\zeta)$ definite su uno spazio di probabilità $\mathcal{S}$. Sia $z = f(x, y)$ una funzione continua, la quantità $Z(\zeta) = f\left[X(\zeta), Y(\zeta)\right]$ individua un'altra variabile aleatoria. La quantità

(I.11.8) $$E\{f(x, y)\} \equiv \overline{f(x, y)} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y) p_{XY}(x, y) dx dy$$

costituisce la media statistica di z. Anche in questo caso risultano di particolare interesse i cosiddetti **momenti incrociati** definiti dalle:

(I.11.9) $$m_{XY}^{(p,q)} \equiv E\left\{x^p y^q\right\} \equiv \overline{x^p y^q} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q p_{XY}(x, y) dx dy \qquad p, q = 0, 1, 2, \ldots$$

Di particolare interesse è il momento $m_{XY}^{(1,1)}$

(I.11.10) $$m_{XY}^{(1,1)} \equiv E\{xy\} \equiv \overline{xy} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy p_{XY}(x, y) dx dy$$

e la **varianza mutua** o **incrociata** σ_{XY} definita dalla:

(I.11.11) $$\sigma_{XY} \equiv E\left\{\left(x - m_X^{(1)}\right)\left(y - m_Y^{(1)}\right)\right\} \equiv \overline{\left(x - m_X^{(1)}\right)\left(y - m_Y^{(1)}\right)} \equiv \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left(x - m_X^{(1)}\right)\left(y - m_Y^{(1)}\right) p_{XY}(x, y) dx dy$$

Poiché è $\left(x - m_X^{(1)}\right)\left(y - m_Y^{(1)}\right) = xy - x m_Y^{(1)} - y m_X^{(1)} + m_X^{(1)} m_Y^{(1)}$, risulta:

(I.11.12) $$\sigma_{XY} = m_{XY}^{(1,1)} - m_X^{(1)} m_Y^{(1)}$$