

# RICHIAMI DI STATISTICA E DI CALCOLO DELLE PROBABILITÀ

01/05/2020

Zafonte Mario

Appendice Tesi – 1992/1993

## Sommario

1.	Analisi e rappresentazione dei dati.....	1
1.1.	Modalità, Frequenza.....	3
1.1.1.	Poligono delle frequenze.....	8
1.1.2.	Distribuzione cumulata di frequenza.....	8
1.2.	Parametri caratteristici di una distribuzione di frequenza.....	10
1.2.1.	Indici di posizione.....	10
1.2.2.	Indici di dispersione.....	11
1.3.	Covarianza e coefficiente di correlazione lineare.....	12
2.	Eventi, insiemi e probabilità.....	13
2.1.	Esperimento.....	13
2.2.	Spazio campione.....	13
2.3.	Evento.....	14
2.4.	Evento composto.....	14
2.5.	Eventi mutuamente esclusivi o incompatibili.....	15
2.6.	Probabilità di un evento.....	15
2.7.	Probabilità condizionata ed eventi indipendenti.....	19
3.	Variabili aleatorie.....	21

3.1.	Funzione densità di probabilità.....	22
3.2.	Funzione di ripartizione.....	23
3.3.	Valor medio di una variabile aleatoria.....	25
3.4.	Varianza.....	26
3.5.	Deviazione standard.....	27
3.6.	Coefficiente di variazione.....	27
3.7.	Variabile casuale standardizzata.....	27
3.8.	Funzione di variabile aleatoria.....	28
3.9.	Momenti di una variabile aleatoria.....	30
3.10.	Funzione caratteristica.....	31
3.11.	Log-funzione caratteristica.....	33
3.12.	Relazione tra momenti e cumulanti.....	34
3.13.	Variabile aleatoria gaussiana.....	37
3.14.	Variabile aleatoria gaussiana standardizzata.....	39
3.15.	Funzione degli errori.....	41
4.	Variabili aleatorie multidimensionali.....	42
4.1.	Funzione di ripartizione congiunta.....	43
4.2.	Funzione densità di probabilità congiunta.....	43
4.3.	Densità marginale.....	45
4.4.	Densità di probabilità condizionata.....	46
4.5.	Variabili aleatorie indipendenti.....	47
4.6.	Valori caratteristici di una distribuzione bidimensionale.....	48
4.6.1.	Valor medio.....	48
4.6.2.	Varianza.....	48
4.6.3.	Covarianza.....	48
4.6.4.	Coefficiente di correlazione lineare.....	48

4.6.5.	Valor medio condizionato .....	49
4.7.	Valor medio di una funzione di variabili aleatorie .....	49
4.8.	Matrice di covarianza e di correlazione .....	51
4.9.	Momenti di una variabile aleatoria multidimensionale .....	52
4.9.1.	Prodotto di due matrici .....	53
4.9.2.	Forma vettorializzata di una matrice .....	55
4.9.3.	Potenza secondo Kronecker .....	55
4.10.	Funzione caratteristica di v.a. multidimensionale .....	56
4.11.	Log-funzione caratteristica di v.a. multidimensionale .....	57
5.	Bibliografia .....	58

## 1. Analisi e rappresentazione dei dati

Assegnato un insieme ideale di elementi nominalmente omogenei, aventi una o più caratteristiche in comune, chiamato "popolazione", si può pervenire a formulare alcune predicazioni valide sullo stesso insieme, mediante l'analisi delle informazioni raccolte su un "campione", ossia su di una parte della popolazione stessa. Il procedimento mediante il quale si costruisce il campione, detto campionamento, dev'essere tale da assicurare a ciascun membro della popolazione la stessa probabilità di essere scelto.

Affinché le informazioni raccolte sul campione possano fornire a colpo d'occhio gli aspetti salienti del fenomeno in esame riguardante la popolazione è necessario però che esse siano raggruppate in modo opportuno. Ad es. se si vogliono raccogliere informazioni sulla resistenza a compressione di un cubetto in conglomerato, appartenente alla popolazione di tutte le prove che

possono essere effettuate nel mondo, è pacifico che un campione di  
ampiezza  $N=100$ , non fornisce alcuna informazione particolarmente  
rappresentativa. Da questa banale constatazione si può intravedere  
l'importanza della buona scelta del campione quando si vuole dare una  
rappresentazione statistica ad una certa trasformazione fisica.

## 1.1. Modalità, Frequenza

### i. Caso discreto

Si supponga di lanciare un dado  $N=40$  volte e di registrare ad ogni lancio i risultati. Indicando con  $X$  il risultato del generico lancio, esso può assumere, nella singola prova, uno solo dei valori  $1, 2, 3, 4, 5, 6$ .

In generale il valore che  $X$  può assumere nella generica prova si definisce "modalità" od evento, inoltre si definisce "frequenza assoluta" il numero totale di volte che la modalità si è presentata nelle  $N$  prove e "frequenza relativa" il rapporto tra la frequenza assoluta ed il numero  $N$  di prove effettuate.

Supponiamo che i risultati ottenuti nelle  $N$  prove siano quelli riportati nella Tabella A.1; in Figura A.1 è riportata la distribuzione della frequenza assoluta delle modalità di  $X$ , che variando la scala delle ordinate può rappresentare anche la distribuzione delle frequenze relative.

TABELLA A.1

Modalità	Frequenza (f)	Frequenza relativa (f/N)
1	5	$5/40=0,125$
2	10	$10/40=0,250$
3	7	$7/40=0,175$
4	7	$7/40=0,175$
5	6	$6/40=0,150$
6	5	$5/40=0,125$
	$\Sigma f=40$	$\Sigma (f/N)=1,000$

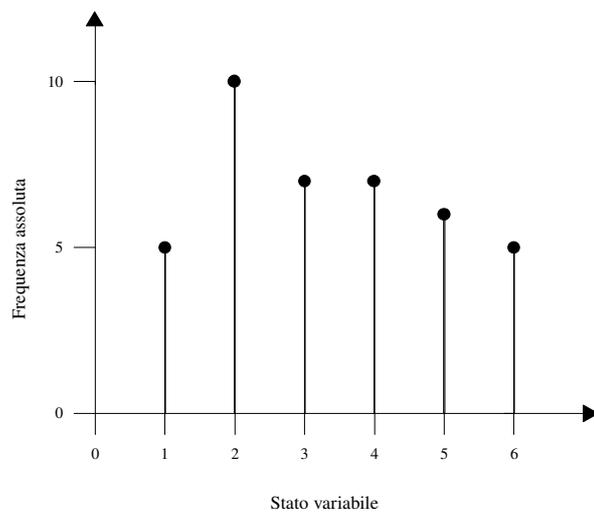


Figura A.1 - Distribuzione della frequenza assoluta delle modalità di X.

## **ii. Caso continuo**

Nel caso del dado, come abbiamo visto, i possibili risultati sono in numero finito, per tale motivo il fenomeno è detto "discreto".

Se invece consideriamo le prove di rottura a compressione dei cubetti in conglomerato, allora, indicata con  $X$  la resistenza a compressione nel generico evento, osserviamo che questa può assumere un qualsiasi valore in un certo intervallo, misurato sull'asse reale.

In generale, un fenomeno è detto "continuo" se la caratteristica misurabile dell'evento, detta "variabile" può assumere qualsiasi valore in un intervallo continuo di numeri posto sull'asse reale.

Se nella generica prova la resistenza a compressione assume ad es. il valore 35,027 MPa, non ha senso parlare di frequenza di tale preciso valore, poiché è praticamente impossibile determinare su un'altro cubetto lo stesso valore della resistenza  $X=35,027$  MPa. Per tale motivo, conviene raggruppare i dati entro una o più classi di frequenza, ossia entro intervalli opportunamente scelti, e si

definisce in questo caso una "frequenza assoluta della classe" come il numero di volte che, tra le N prove effettuate, la variabile ha assunto un valore che ricade nella classe. Dividendo la frequenza di classe per il numero totale di prove effettuate si ottiene la frequenza relativa.

Ad es. si supponga di aver effettuato N=150 prove di schiacciamento su cubetti in conglomerato, e di aver ottenuto i risultati riportati in tabella A.2. Esaminando tale tabella si possono subito trarre alcune informazioni circa i valori min e max registrati, circa la classe che si è presentata il maggior numero di volte, ecc. Per avere un quadro immediato del risultato campionario conviene riportare i dati ottenuti nelle prove sotto forma di "istogramma", ossia mediante una rappresentazione grafica costituita da una serie di rettangoli aventi come basi l'intervallo delle classi prescelte e come altezza la frequenza corrispondente. In particolare nella figura

A.2. è riportato l'istogramma di frequenza assoluta relativamente ai dati della tabella A.2.

TABELLA A.2

Classe	Limiti della classe (MPa)	Frequenza (f)	Frequenza relativa (f/N)
1	30-31	3	0.020
2	31-32	8	0.053
3	32-33	15	0.100
4	33-34	25	0.167
5	34-35	28	0.187
6	35-36	22	0.147
7	36-37	29	0.193
8	37-38	9	0.060
9	38-39	5	0.033
10	39-40	6	0.040
		$\Sigma f_i = 150$	$\Sigma (f_i/N) = 1,000$

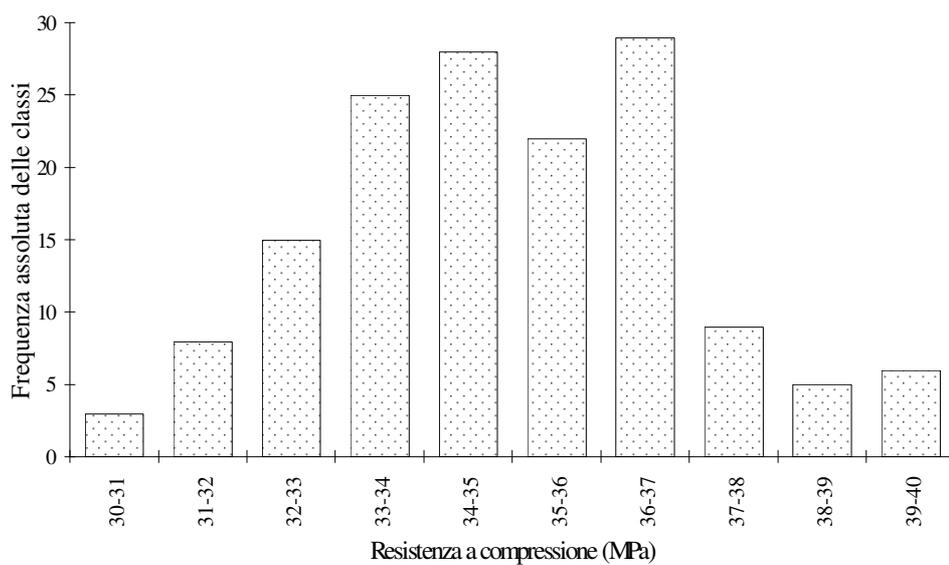


Figura A.2 - Istogramma delle frequenze assolute.

### **1.1.1. Poligono delle frequenze**

Se si congiungono i punti medi dei lati superiori dei rettangoli dell'istogramma si ottiene una diversa rappresentazione della distribuzione di frequenza, che è detta "poligono delle frequenze".

### **1.1.2. Distribuzione cumulata di frequenza**

Ad ogni distribuzione di frequenza si può associare una distribuzione cumulata di frequenza, riportando in corrispondenza delle varie modalità (caso discreto), oppure in corrispondenza dei vari intervalli di classe (caso continuo), la somma delle frequenze relative agli intervalli che precedono quello in esame (compreso l'ultimo).

Ad es. nelle figure A.3 ed A.4 sono riportati sia l'istogramma delle frequenze cumulate che il poligono delle frequenze cumulate relativamente ai dati della tabella A.2.

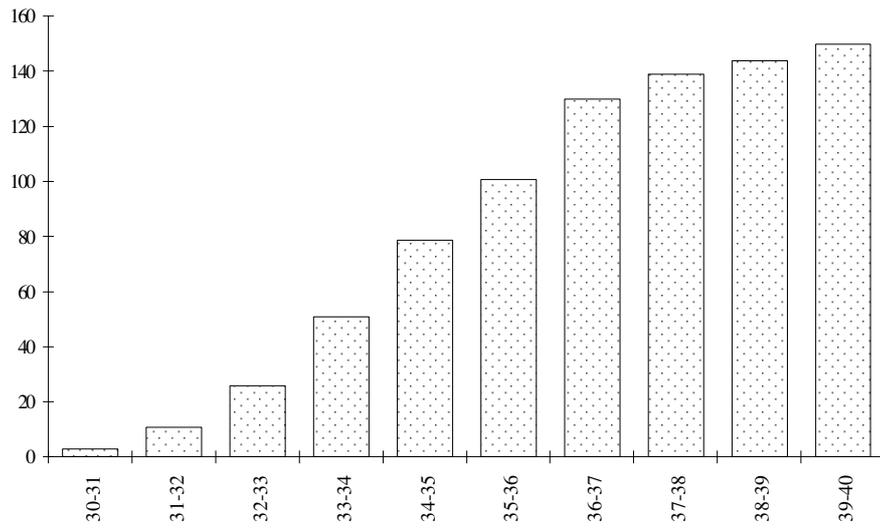


Figura A.3 - Istogramma delle frequenze cumulate

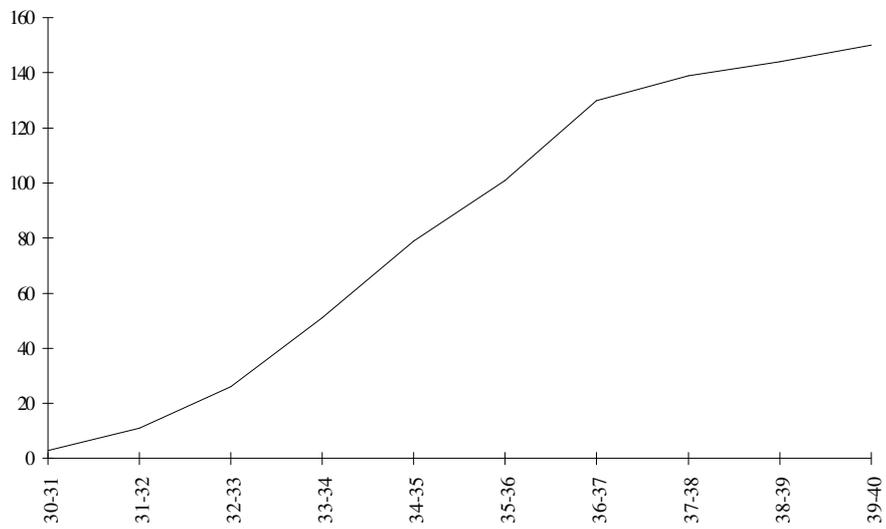


Figura A.4 - Poligono delle frequenze cumulate

## 1.2. Parametri caratteristici di una distribuzione di frequenza

Una distribuzione di frequenza riferita ad un certo campione di ampiezza numerabile può essere caratterizzata da indici che tendono a localizzarne il "centro", detti "indici di posizione", e da indici che indicano la variabilità dei dati, detti "indici di dispersione".

### 1.2.1. Indici di posizione

**Moda.** E' il valore della variabile che corrisponde alla massima frequenza. Nel caso di variabile continua, è il valor medio della classe che corrisponde alla massima frequenza.

**Mediana.** E' il valore di mezzo di una serie di dati se N è dispari, ovvero la semisomma dei due valori di mezzo per N pari.

**Media.** E' la media aritmetica delle N osservazioni sperimentali della caratteristica (o della variabile) X

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_N}{N} = \frac{\sum_{i=1}^N X_i}{N} \quad (\text{A.1})$$

## 1.2.2. Indici di dispersione

**Campo di variazione.** E' la distanza tra il valore più grande e quello più piccolo. Nell'es. riportato in tabella A.2 il campo di variazione è di 10 MPa.

**Varianza.** E' la media aritmetica del quadrato degli scarti dei valori osservati dalla media

$$s^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N} \quad (\text{A.2})$$

**Deviazione standard.** E' la radice quadrata della varianza

$$s = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N}} \quad (\text{A.3})$$

### 1.3. Covarianza e coefficiente di correlazione lineare

Se nello studio di un campione vengono rilevate due variabili X e Y, le stesse si dicono "correlate" se tra loro esiste una qualche relazione matematica. In tal caso, tracciando un grafico in cui in ascissa sono riportati i valori rilevati per X ed in ordinata i valori rilevati per Y, i punti ottenuti si adagiano su una linea retta, (correlazione lineare), od anche su una curva, (correlazione non lineare).

Una misura del grado di correlazione tra le due variabili è fornita dalla "covarianza" e cioè

$$S_{XY} = \sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y}) \quad (\text{A.4})$$

Si definiscono inoltre "scarto quadratico medio" di X e di Y le quantità

$$S_X = \sqrt{\frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^2} \quad S_Y = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (\text{A.5})$$

Si definisce infine "coefficiente di correlazione lineare" il

rapporto 
$$r = r_{XY} = \frac{S_{XY}}{S_X S_Y} \quad (\text{A.6})$$

## 2. Eventi, insiemi e probabilità

### 2.1. Esperimento

La singola osservazione su un certo fenomeno fisico viene definita "esperimento". L'esperimento si dice "casuale" se ripetuto nelle medesime condizioni, può dar luogo a risultati diversi.

### 2.2. Spazio campione

Se il fenomeno fisico è considerato, com'è d'uopo, aleatorio, l'esito dell'esperimento non sarà noto a priori; tuttavia, una volta fissato un fenomeno fisico di natura aleatoria, è sempre possibile definire un certo insieme, i cui elementi costituiscono tutti i possibili esiti dell'esperimento connesso al fenomeno. Tale insieme è chiamato "spazio campione" ed è indicato generalmente con  $\Omega$ . Ciascun elemento di tale insieme è definito "evento elementare". Ad es. nel lancio di una moneta, lo spazio campione è rappresentato dall'insieme [testa-croce], mentre l'evento elementare è rappresentato dall'elemento "testa" o dall'elemento "croce".

### 2.3. Evento

Un evento  $E$  è un sottoinsieme dello spazio campione  $\Omega$  :  $E \subset \Omega$ . La collezione  $\mathcal{F}$  di tutti i sottoinsiemi dello spazio campione prende il nome di "Spazio degli eventi". Il singolo elemento  $\omega \in \Omega$ , essendo sottoinsieme di  $\Omega$ , è un evento denominato "evento elementare". Anche l'insieme vuoto  $\Phi$  e l'insieme  $\Omega$ , essendo sottoinsiemi di  $\Omega$ , sono eventi:  $\Phi$  prende il nome di "evento impossibile",  $\Omega$  quello di "evento certo".

### 2.4. Evento composto

Si definisce "evento composto" un evento costituito dall'unione di eventi elementari; ad esempio nel lancio di un dado, quando si considera come evento elementare l'uscita di uno dei sei numeri sulla faccia superiore del dado, l'uscita di un numero pari rappresenta un evento composto, in quanto costituito dall'insieme degli eventi elementari rappresentati dall'uscita dei numeri 2, 4 e 6. Ad un evento composto corrisponde dunque un proprio sottoinsieme dello spazio campione.

## 2.5. Eventi mutuamente esclusivi o incompatibili

Due eventi A e B si dicono mutuamente esclusivi o incompatibili quando l'esito favorevole di uno di essi esclude l'esito favorevole dell'altro; in questo caso i due insiemi sono disgiunti, ossia  $A \cap B = \Phi$ . Ad esempio, nel lancio del dado, l'uscita di un numero pari esclude quella di un numero dispari, per cui i due eventi composti corrispondenti sono mutuamente esclusivi.

## 2.6. Probabilità di un evento

La definizione "frequentista" della probabilità di un evento è basata sul rapporto tra il numero  $n_A$  di esiti favorevoli all'evento A e il numero complessivo di esperimenti n, per n molto grande. Secondo la definizione data da R. von Mises, [43], essa è fornita dal limite della frequenza relativa per  $n \rightarrow \infty$  e cioè

$$P(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n} \quad (\text{A.7})$$

Questa definizione di probabilità detta anche empirica od a posteriori, è quella maggiormente usata esplicitamente o implicitamente dall'ingegnere.

Una definizione rigorosamente matematica della probabilità è quella fornita da A. N. Kolmogorov, (1931). La probabilità è una funzione  $P$  definita sulla famiglia  $\mathbf{F}$  di sottoinsiemi di  $\Omega$ , a valori nell'intervallo  $[0,1] \subset \mathbb{R}$

$$P : \mathcal{F} \rightarrow [0,1]$$

tale da soddisfare i seguenti assiomi

i) Condizione di non-negatività

La probabilità  $P(A)$  di un qualunque evento  $A$  è un numero reale maggiore o uguale a zero

$$P(A) \geq 0$$

ii) Condizione di normalizzazione

La probabilità dell'evento certo è pari a uno e la probabilità dell'evento impossibile è pari a zero

$$P(\Omega) = 1 \qquad P(\Phi) = 0$$

iii) Proprietà additiva

Dati due eventi  $A_1$  ed  $A_2$  disgiunti ( $A_1 \cap A_2 = \Phi$ ) la probabilità dell'evento unione  $A = A_1 \cup A_2$  è pari alla somma delle probabilità di ciascun evento

$$P(A) = P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

Sulla base dei suddetti assiomi, sfruttando le proprietà degli insiemi, si possono dimostrare i seguenti teoremi:

a) La probabilità dell'evento  $\bar{A}$ , complementare di  $A$ , risulta essere pari a:

$$P(\bar{A}) = 1 - P(A) \qquad (A.8)$$

Essendo  $\bar{A} \cap A = \Phi$  ed  $\bar{A} \cup A = \Omega$  si ha infatti

$$P(\Omega) = P(\bar{A} \cup A) = P(\bar{A}) + P(A) = 1$$

da cui segue la (A.8).

b) Dati due eventi qualsiasi A e B, risulta in generale

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (\text{A.9})$$

infatti, essendo

$$A \cup B = A \cup (\bar{A} \cap B) \quad (\text{A.10})$$

ed essendo inoltre A ed  $\bar{A} \cap B$  eventi disgiunti, (ovvero  $A \cap (\bar{A} \cap B) = \Phi$ )

per l'assioma iii) si ha

$$P(A \cup B) = P(A) + P(\bar{A} \cap B) \quad (\text{A.11})$$

Inoltre, essendo

$$B = (A \cap B) \cup (\bar{A} \cap B) \quad (\text{A.12})$$

ed essendo pure  $(A \cap B)$  ed  $(\bar{A} \cap B)$  eventi disgiunti, sempre per l'assioma

iii) si ha

$$P(B) = P[(A \cap B) \cup (\bar{A} \cap B)] = P(A \cap B) + P(\bar{A} \cap B) \quad (\text{A.13})$$

da cui si ottiene la relazione

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) \quad (\text{A.14})$$

che, sostituita nella (A.11), fornisce la (A.9).

## 2.7. Probabilità condizionata ed eventi indipendenti

Si definisce "probabilità condizionata" di due eventi A e B e si indica con  $P(A|B)$ , la probabilità che si verifichi l'evento A posto che B si sia già verificato. Pertanto, tale probabilità è definita dalla relazione

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0 \quad (\text{A.15})$$

Nell'impostazione assiomatica la (A.15) costituisce il quarto assioma della teoria della probabilità.

Due eventi A e B si dicono "indipendenti" se si verifica la condizione

$$P(A \cap B) = P(A) \cdot P(B) \quad (\text{A.16})$$

che implica necessariamente

$$P(A|B) = P(A) \quad \text{e} \quad P(B|A) = P(B) \quad (\text{A.17})$$

Si osservi che si parla di indipendenza tra due eventi A e B quando la loro intersezione non è vuota; nel caso in cui invece

$A \cap B = \Phi$  come abbiamo detto i due eventi si dicono incompatibili o disgiunti.

### 3. Variabili aleatorie

Si definisce "variabile aleatoria" (V.A.)  $X(\omega)$  o più semplicemente  $X$  una funzione che faccia corrispondere ad ogni elemento  $\omega \in \Omega$ , (dove  $\Omega$  è lo spazio campione di un qualsivoglia fenomeno fisico aleatorio) un numero reale  $X(\omega)$  e cioè

$$\begin{array}{lcl} X : & \Omega & \rightarrow \quad R \\ & \omega & \rightarrow \quad X(\omega) = x \end{array}$$

La V.A. gode della proprietà che per ogni numero reale  $r$  esiste una probabilità  $P(\omega)$  tale che  $X(\omega) \leq r$ , e cioè

$$P(\omega : X(\omega) \leq r)$$

ossia, l'insieme degli elementi  $\omega$  tali che  $X(\omega) \leq r$  è un evento.

Esistono due tipi di variabili aleatorie: le variabili aleatorie discrete e quelle continue. Le prime possono assumere soltanto un numero finito di valori, (si pensi, ad esempio, alla variabile aleatoria legata al lancio del dado); le seconde possono, invece, assumere un qualunque valore all'interno di un dato intervallo, (si

pensi, ad esempio, ai risultati di una prova di resistenza a trazione per un tondino d'acciaio).

### 3.1. Funzione densità di probabilità

Una variabile aleatoria  $X(\omega)$  risulta completamente definita, da un punto di vista della teoria matematica delle probabilità, attraverso la conoscenza della cosiddetta "funzione densità di probabilità",  $f(x)$ . Tale funzione è caratterizzata dal fatto che l'area da essa racchiusa con l'asse delle ascisse in un intervallo  $[x_1, x_2]$  fornisce la probabilità che la variabile aleatoria  $X(\omega)$  ricada in tale intervallo e cioè:

$$\int_{x_1}^{x_2} f(x)dx = P(x_1 \leq X(\omega) \leq x_2) \quad (\text{A.18})$$

Da tale definizione della funzione densità di probabilità, sortiscono le seguenti proprietà od effetti degli assiomi in precedenza assunti

1)  $f(x)dx = P(x \leq X(\omega) \leq x + dx)$

2)  $P(X(\omega) = x_1) = \int_{x_1}^{x_1} f(x)dx = 0 \Rightarrow P(X = -\infty) = 0, P(X = +\infty) = 0$

$$3) \quad P(-\infty \leq X(\omega) \leq +\infty) = \int_{-\infty}^{+\infty} f(x)dx = 1$$

Nella Figura A.5 è riportato ad es. il diagramma della funzione densità di probabilità di una certa variabile aleatoria continua,  $X(\omega)$

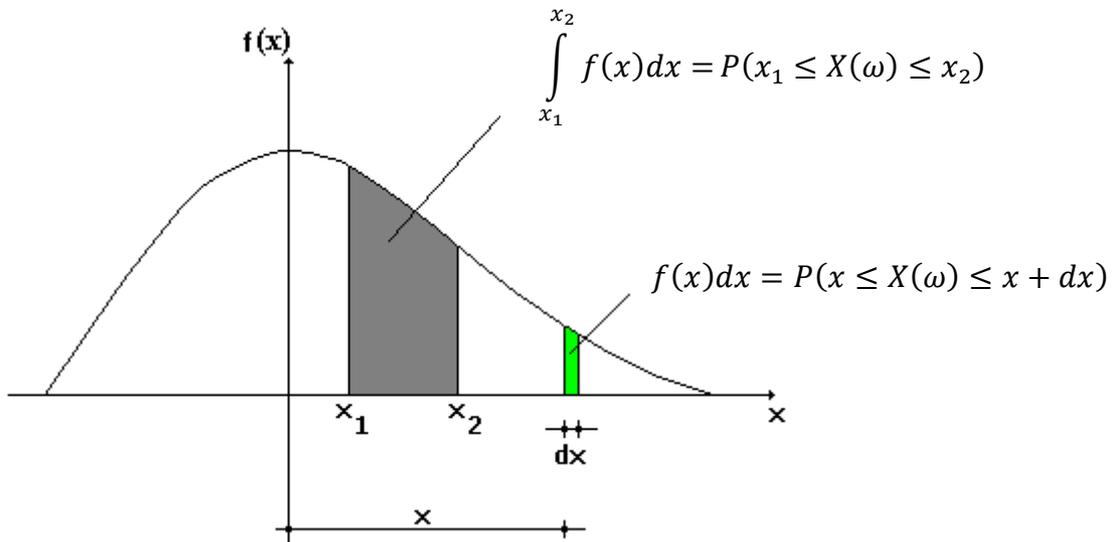


Figura A.5 - Funzione densità di probabilità di una V.A. continua

### 3.2. Funzione di ripartizione

Si definisce "funzione di ripartizione di probabilità"  $F(x)$  della v.a. continua  $X$ , la funzione definita in  $\mathbb{R}$  tale che

$$F(x) = P(X \leq x) = \int_{-\infty}^{+\infty} f(t)dt \quad 0 \leq F(x) \leq 1 \quad (\text{A.19})$$

Si noti che la scrittura  $(X \leq x)$ , rappresenta l'insieme dei punti  $\omega$  dello spazio campione  $\Omega$  tale che i valori  $X(\omega)$  siano sempre non superiori a  $x$ , e cioè

$$F(x) = P(\omega \in \Omega | X(\omega) \leq x) = \int_{-\infty}^{+\infty} f(t) dt \quad (A.20)$$

La F.D.P. e la F.R.P. non sono tra loro indipendenti risultando legate dalla relazione

$$f(x) = \frac{dF(x)}{dx} \quad (A.21)$$

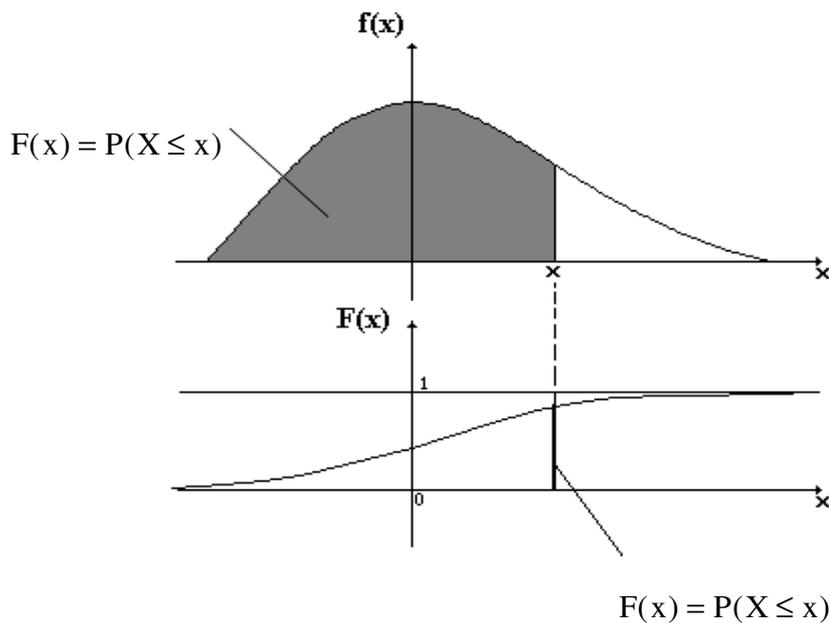


Figura A.6 - Funzione densità di probabilità e funzione ripartizione di probabilità di una V.A.

Geometricamente, il valore  $F(x)$  della funzione di ripartizione rappresenta l'area a sinistra della retta verticale passante per  $x$  compresa tra la curva di densità  $f(x)$  e l'asse  $x$  (figura A.6).

Usando la definizione (20) si possono inoltre mostrare le seguenti proprietà:

- 1)  $F(-\infty) = P(X \leq -\infty) = 0$
- 2)  $F(+\infty) = P(X \leq +\infty) = 1$
- 3)  $F(x_2) - F(x_1) = P(x_1 \leq X \leq x_2)$

### 3.3. Valor medio di una variabile aleatoria

Si definisce valor medio di una variabile aleatoria  $X$  dotata di densità  $f(x)$  e si indica con  $E[X]$  o  $\mu_x$ , l'integrale:

$$\mu(x) = E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx \quad (\text{A.22})$$

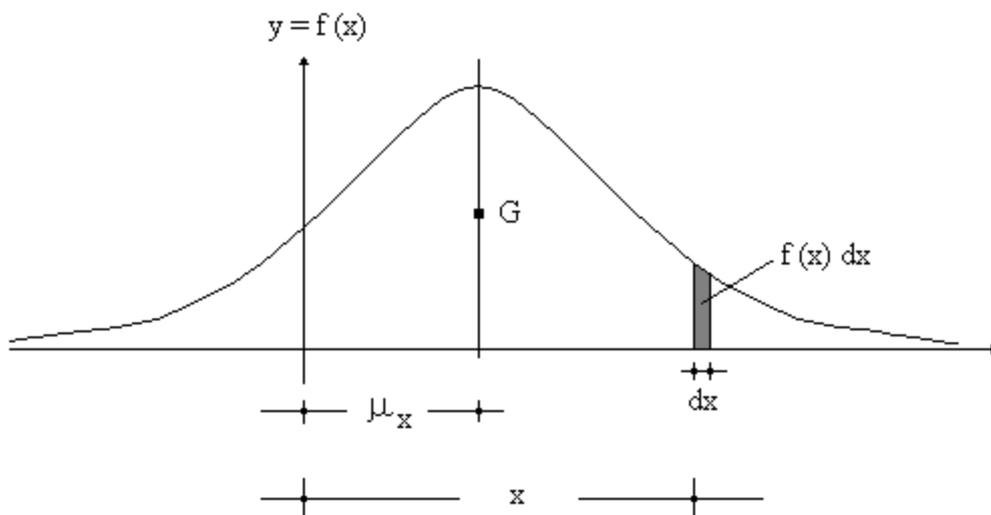


Figura A.7 - Valor medio di una V.A.

Il valor medio di una variabile aleatoria rappresenta pertanto il momento statico  $S_y$  dell'area sottesa dalla curva  $f(x)$  rispetto l'asse  $y$ . Infatti, con riferimento alla figura A.7, il momento statico dell'areola elementare  $f(x)dx$  è pari a  $dS_y = xf(x)dx$ , e quindi l'integrale della  $dS_y$  rappresenta il momento statico totale  $S_y$ . Ed essendo l'area sottesa dalla curva  $f(x)$  unitaria, si ha  $A \cdot \mu_x = \mu_x = S_y$ , ossia la (22).

### 3.4. Varianza

Si definisce varianza di una variabile aleatoria continua  $X$ , dotata di densità  $f(x)$  l'integrale:

$$\delta_x^2 = \text{var}[X] = \int_{-\infty}^{+\infty} (x - \mu_x)^2 f(x) dx \quad (\text{A.23})$$

La varianza di una variabile aleatoria rappresenta pertanto il momento d'inerzia baricentrico dell'area unitaria sottesa dalla curva  $f(x)$ .

### 3.5. Deviazione standard

La deviazione standard è la radice quadrata positiva della varianza di una variabile aleatoria e cioè

$$\delta_x = \sqrt{\text{var}[X]} \quad (\text{A.24})$$

### 3.6. Coefficiente di variazione

Il rapporto tra la deviazione standard  $\delta_x$  e il valor medio  $\mu_x$  di una variabile aleatoria si chiama coefficiente di variazione (cdv), e cioè

$$C_x = \frac{\delta_x}{\mu_x} \quad (\text{A.25})$$

### 3.7. Variabile casuale standardizzata

Se  $X$  è una variabile aleatoria con valor medio  $\mu_x$  e scarto quadratico medio  $\delta_x$ , si definisce "variabile casuale ridotta o standardizzata" associata ad  $X$ , la variabile

$$Z = \frac{X - \mu_x}{\delta_x} \quad (\text{A.26})$$

La variabile ridotta associata ad  $X$  risulta essere adimensionale con valor medio nullo e varianza unitaria e cioè  $E[Z]=0$ ,  $\text{var}[Z]=1$ .

### 3.8. Funzione di variabile aleatoria

La v.a.  $X$  è una funzione definita in  $\Omega$  a valori in  $\mathbb{R}$  la quale ad ogni elemento  $\omega \in \Omega$  fa corrispondere il numero reale  $X(\omega) = x$ :

$$\begin{array}{lcl} X : \Omega & \rightarrow & \mathbb{R} \\ \omega & \rightarrow & X(\omega) = x \end{array}$$

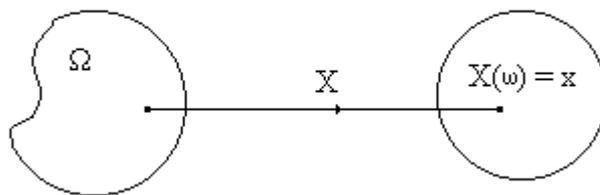


Figura A.8 - Rappresentazione grafica della V.A.

Inoltre  $X$  è v.a. solo se  $\forall r \in \mathbb{R}$  l'insieme degli elementi di  $\omega \in \Omega$  tale che  $X(\omega) \leq r$  è un evento.

Consideriamo adesso una variabile aleatoria  $Y$  legata alla v. a.

$X$  attraverso il seguente legame funzionale

$$Y = \varphi(X) \tag{A.27}$$

il quale, ad ogni elemento  $X(\omega) = x$ , fa corrispondere un ben definito numero reale

$$Y(\omega) = y = \varphi[X(\omega)] = \varphi(x) \tag{A.28}$$

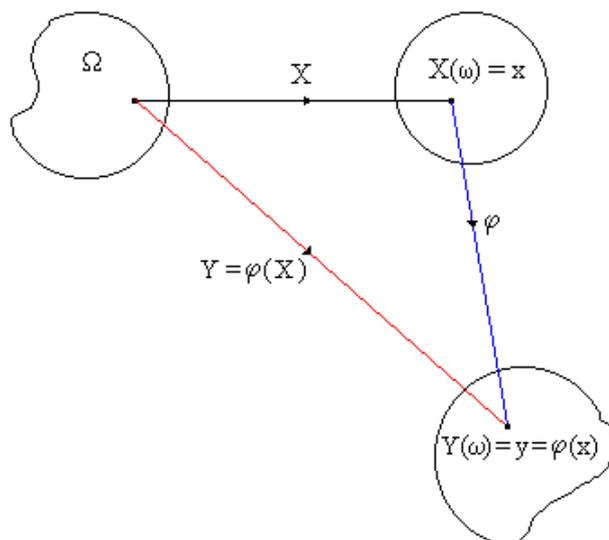


Figura A.9 - Rappresentazione grafica di una funzione di V.A.

Se la funzione reale  $\varphi(\cdot)$  è monotona, appare evidente che la probabilità che la variabile aleatoria  $X$  cada nell'intervallo  $[x, x+dx]$  coincide con la probabilità che la variabile  $Y$  cada nell'intervallo  $[y, y+dy]$ , avendo posto

$$y = \varphi(x) \quad (\text{A.29})$$

Allora è possibile scrivere la seguente relazione

$$f(x) dx = h(y) dy \quad (\text{A.30})$$

dalla quale si ottiene

$$h(y) = f(x) dx/dy \quad (\text{A.31})$$

Il valor medio di  $Y$  è fornito dalla seguente espressione

$$E[Y] = \int_{-\infty}^{+\infty} yh(y)dy \quad (\text{A.32})$$

che, ricordando le (29) e (30), può essere riscritta come segue

$$E[Y] = E[\varphi(x)] = \int_{-\infty}^{+\infty} \varphi(x)f(x)dx \quad (\text{A.33})$$

Quindi, per valutare il valor medio di una funzione di una variabile aleatoria  $X$  è sufficiente eseguire l'integrale del prodotto di tale funzione per la funzione densità di probabilità di  $X$ .

### 3.9. Momenti di una variabile aleatoria

Si definisce momento di ordine  $r$  di una variabile aleatoria  $X$  e si indica con  $m_r[X]$  il valor medio della variabile aleatoria  $X^r$ . Data  $Y=X^r$ , con  $r$  intero, con riferimento alla (33) si ha dunque:

$$m_r(X) = E[X^r] = \int_{-\infty}^{+\infty} x^r f(x)dx \quad (\text{A.34})$$

in particolare risulta:

$$m_0(X) = E[X^0] = \int_{-\infty}^{+\infty} f(x)dx \quad (\text{A.35})$$

$$m_1(X) = E[X] = \int_{-\infty}^{+\infty} xf(x)dx = \mu_x \quad (\text{A.36})$$

$$m_2(X) = E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x)dx \quad (\text{A.37})$$

### 3.10. Funzione caratteristica

Oltre che dalla funzione densità di probabilità, una variabile aleatoria può essere rappresentata attraverso la conoscenza di un'altra importante funzione, detta "funzione caratteristica", che è pari alla media della variabile aleatoria  $Y = \exp(-i\theta X)$ , e cioè

$$M_x(\theta) = \int_{-\infty}^{+\infty} \exp(-i\theta x) f(x) dx = E[\exp(-i\theta x)] \quad (\text{A.38})$$

essendo  $\theta$  un parametro reale ed avendo indicato con  $i$  l'unità immaginaria. La funzione caratteristica rappresenta, quindi, la trasformata di Fourier della funzione densità di probabilità. Nota la funzione caratteristica, mediante la trasformata inversa di Fourier si può quindi ricavare la F.D.P. della v.a.  $X$  e cioè

$$f(x) = \int_{-\infty}^{+\infty} \exp(i\theta x) M_x(\theta) d\theta \quad (\text{A.39})$$

Espandendo in serie di McLaurin la funzione caratteristica si ottiene

$$M_x(\theta) = M_x(\theta)|_{\theta=0} + \frac{dM_x(\theta)}{d\theta}|_{\theta=0}\theta + \frac{1}{2!} \frac{d^2 M_x(\theta)}{d\theta^2}|_{\theta=0}\theta^2 + \dots \quad (\text{A.40})$$

da cui, essendo

$$\frac{d^k M_x(\theta)}{d\theta^k} = (-i)^k = \int_{-\infty}^{+\infty} x^k \exp(-i\theta x) f(x) dx \quad (\text{A.41})$$

ed in conseguenza

$$\left. \frac{d^k M_x(\theta)}{d\theta^k} \right|_{\theta=0} = (-i)^k \int_{-\infty}^{+\infty} x^k f(x) dx = (-i)^k m_k[X] \quad (\text{A.42})$$

si ottiene l'importante risultato

$$M_x(\theta) = 1 + (-i)^1 m_1[X]\theta + \frac{(-i)^2}{2!} m_2[X]\theta^2 + \frac{(-i)^3}{3!} m_3[X]\theta^3 + \dots \quad (\text{A.43})$$

ossia

$$M_x(\theta) = \sum_{j=0}^{\infty} \frac{(-i)^j}{j!} m_j[X]\theta^j. \quad (\text{A.44})$$

Da tale espressione si evince che i momenti di una variabile aleatoria  $X$  rappresentano, a meno dei termini  $(-i)^j$ , i coefficienti dello sviluppo in serie di McLaurin della funzione caratteristica. Per cui, se sono noti i momenti (fino all'ordine infinito) di una variabile aleatoria, è possibile risalire alla funzione caratteristica e da questa, attraverso la trasformata inversa di Fourier, alla funzione densità di probabilità. Ciò implica che una

variabile aleatoria può essere caratterizzata pienamente, da un punto di vista probabilistico, dalla conoscenza di tutti i suoi momenti.

### 3.11. Log-funzione caratteristica

Un'altra funzione che permette di caratterizzare una variabile aleatoria è la cosiddetta log-funzione caratteristica data dal logaritmo naturale della funzione caratteristica e cioè  $\ln M_x(\theta)$ . Tale funzione, sviluppata in serie di McLaurin, fornisce

$$\ln M_x(\theta) = \ln M_x(\theta)|_{\theta=0} + \frac{d \ln M_x(\theta)}{d\theta} \Big|_{\theta=0} \theta + \frac{1}{2!} \frac{d^2 \ln M_x(\theta)}{d\theta^2} \Big|_{\theta=0} \theta^2 + \dots \quad (\text{A.45})$$

da cui, posto

$$k_j[X] = \frac{1}{(-i)^j} \frac{d^j \ln M_x(\theta)}{d\theta^j} \Big|_{\theta=0} \quad (\text{A.46})$$

detto "cumulante di ordine j" della variabile aleatoria X, si trae

la relazione

$$\ln M_x(\theta) = \sum_{j=1}^{\infty} \frac{(-i)^j}{j!} k_j[X] \theta^j \quad (\text{A.47})$$

Applicando ora la funzione esponenziale ad ambo i membri, si ottiene la relazione

$$M_x(\theta) = \exp\left(\sum_{j=1}^{\infty} \frac{(-i)^j}{j!} k_j[X] \theta^j\right) \quad (\text{A.48})$$

la quale permette di esprimere la funzione caratteristica attraverso i cumulanti fino all'ordine infinito della variabile aleatoria. Tale espansione in serie della funzione caratteristica, rispetto alla (40), ha il vantaggio di essere convergente.

### 3.12. Relazione tra momenti e cumulanti

I cumulanti e i momenti di una variabile aleatoria sono tra loro legati da semplici relazioni che consentono di ottenere il cumulante di ordine  $j$ , noti tutti i momenti fino all'ordine  $j$ , e viceversa (Di Paola et al., 1991, [31]). Tali relazioni hanno la seguente forma ricorsiva

$$\left\{ \begin{array}{l} m_1[X] = k_1[X] \\ m_2[X] = k_2[X] + k_1[X] \cdot m_1[X] \\ m_3[X] = k_3[X] + 2k_2[X] \cdot m_1[X] + k_1[X] \cdot m_2[X] \\ m_4[X] = k_4[X] + 3k_3[X] \cdot m_1[X] + 3k_2[X] \cdot m_2[X] + k_1[X] \cdot m_3[X] \\ \dots\dots\dots \end{array} \right. \quad (\text{A.49})$$

e sono facili da ricordare, se si osserva che i coefficienti a secondo membro sono quelli del triangolo di Tartaglia.

I cumulanti rivestono una notevole importanza nell'ambito della descrizione di una variabile aleatoria. Infatti i cumulanti fino al quarto ordine hanno un significato fisico molto preciso che permette di caratterizzare, almeno qualitativamente, la funzione densità di probabilità della variabile considerata. Infatti:

- il cumulante del primo ordine coincide con il valor medio della variabile aleatoria

$$k_1[X] = m_1[X] = E[X] = \mu_x \quad (\text{A.50})$$

- il cumulante del secondo ordine coincide con la varianza della variabile aleatoria

$$k_2[X] = m_2[X] - m_1^2[X] = \text{var}[X] = \sigma_x^2 \quad (\text{A.51})$$

dove  $\sigma_x$  è la deviazione standard.

- il cumulante del terzo ordine fornisce un'indicazione sull'asimmetria della F.D.P. della variabile aleatoria. Infatti

esso è legato al "coefficiente di asimmetria"  $\gamma_a$ , il quale è

definito dal rapporto

$$\gamma_a = \frac{k_3[X]}{\sqrt{k_2^3[X]}} \quad (\text{A.52})$$

E' immediato notare che tale coefficiente è nullo solo quando la funzione densità di probabilità è simmetrica rispetto al valor medio della variabile; esso, quindi, misura il livello di asimmetria della funzione densità di probabilità;

- il cumulante del quarto ordine ha il significato connesso con il cosiddetto "coefficiente d'eccesso"  $\gamma_e$ , definito dal rapporto:

$$\gamma_e = \frac{k_4[X]}{k_2^2[X]} \quad (\text{A.53})$$

tale coefficiente fornisce infatti una stima sulla acutezza del picco della funzione densità di probabilità in prossimità del suo valore massimo.

I cumulanti di ordine maggiore non hanno significati altrettanto evidenti in confronto ai primi quattro.

Dalle espressioni di  $M_x(\theta)$  in funzione di  $f(x)$  o di  $m_j[X]$  o di  $k_j[X]$  risulta evidente come la completa descrizione della variabile aleatoria  $X$  possa essere ottenuta indifferentemente dalla F.D.P. o dalla sua funzione caratteristica od anche dai suoi momenti ovvero dai suoi cumulanti di ogni ordine.

In particolare si osservi ad es. che se

$$\begin{cases} k_1[X] \neq 0 \\ k_s[X] = 0 \forall s \geq 2 \end{cases} \Rightarrow X \text{ è deterministica;}$$

$$\begin{cases} k_1[X] \neq 0; k_2[X] \neq 0 \\ k_s[X] = 0 \forall s \geq 3 \end{cases} \Rightarrow X \text{ è simmetrica e non ha eccesso per cui,}$$

si dice che è normalmente distribuita.

### 3.13. Variabile aleatoria gaussiana

Una variabile aleatoria si dice "gaussiana", o, "normale" se tutti i suoi cumulanti di ordine maggiore del secondo sono nulli. La funzione caratteristica di una variabile aleatoria gaussiana assume, quindi, la seguente espressione ricavabile dalla (44) ponendovi  $k_j[X]=0 \forall j \geq 3$ :

$$M_x[\theta] = \exp\left(-ik_1[X]\theta - \frac{1}{2}k_2[X]\theta^2\right) \quad (\text{A.54})$$

ossia

$$M_x[\theta] = \exp\left(-i\mu_x\theta - \frac{1}{2}\sigma^2\theta^2\right) \quad (\text{A.54a})$$

Dalla precedente, effettuando l'antitrasformata di Fourier, si ottiene la funzione densità di probabilità della variabile aleatoria gaussiana nella forma ben nota

$$f(x) = \frac{1}{\sqrt{2\pi k_2[X]}} \exp\left[-\frac{1}{2} \frac{(x-k_1[X])^2}{k_2[X]}\right] \quad (\text{A.55})$$

od anche

$$f(x) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x-\mu_x}{\sigma_x}\right)^2\right] \quad (\text{A.55a})$$

Le (54) e (55) mostrano che la completa descrizione di una variabile aleatoria gaussiana è ottenibile attraverso la sola conoscenza dei cumulanti del primo e del secondo ordine, o, viste le relazioni intercorrenti fra i cumulanti e i momenti, anche attraverso la conoscenza del valor medio e della varianza della variabile stessa. E' importante sottolineare che, mentre i cumulanti di ordine maggiore del secondo di una variabile aleatoria gaussiana sono nulli, la stessa cosa non può dirsi per i momenti di ordine maggiore del secondo.

La relativa funzione di ripartizione, applicando la (19), è data

da

$$F(x) = P(X \leq x) = \frac{1}{\sigma_x \sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2} \left(\frac{t-\mu_x}{\sigma_x}\right)^2\right] dt \quad (\text{A.56})$$

### 3.14. Variabile aleatoria gaussiana standardizzata

Se si introduce la v.a. ridotta (normalizzata o standardizzata)

corrispondente a  $X$  e cioè

$$Z = \frac{X - \mu_x}{\sigma_x} \quad (\text{A.57})$$

poiché risulta

$$E[Z] = \frac{E[X] - \mu_x}{\sigma_x} = \frac{\mu_x - \mu_x}{\sigma_x} = 0 \quad \text{valor medio nullo}$$

$$\text{var}[Z] = \text{var}\left(\frac{X - \mu_x}{\sigma_x}\right) = \frac{\sigma_x}{\sigma_x} = 1 \quad \text{deviazione standard unitaria}$$

si può esprimere la funzione densità di probabilità della nuova

variabile aleatoria nella forma seguente

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad (\text{A.58})$$

mentre per la funzione di ripartizione si ha la forma

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{t^2}{2}\right) dt \quad (\text{A.59})$$

La rappresentazione grafica della (58) prende il nome di "curva a campana", ed in opportune tabelle, [30], sono riportati i valori dell'area al di sotto di tale curva limitata fra le ordinate  $z=0$  e qualunque valore positivo  $z$ .

In particolare si hanno ad es. i seguenti risultati

- $P(-1 \leq Z \leq 1) = P\left(-1 \leq \frac{X-\mu_x}{\sigma_x} \leq 1\right) = P(\mu_x - \sigma_x \leq X \leq \mu_x + \sigma_x) = 0.6827$
- $P(-2 \leq Z \leq 2) = P\left(-2 \leq \frac{X-\mu_x}{\sigma_x} \leq 2\right) = P(\mu_x - 2\sigma_x \leq X \leq \mu_x + 2\sigma_x) = 0.9545$
- $P(-3 \leq Z \leq 3) = P\left(-3 \leq \frac{X-\mu_x}{\sigma_x} \leq 3\right) = P(\mu_x - 3\sigma_x \leq X \leq \mu_x + 3\sigma_x) = 0.9973$

Utilizzando inoltre la (A.57) e cioè

$$X = \mu_x + \sigma_x Z$$

si ottengono le seguenti relazioni

$$P(X \leq x_0) = P(X - \mu_x \leq x_0 - \mu_x) = P\left(\frac{X-\mu_x}{\sigma_x} \leq \frac{x_0-\mu_x}{\sigma_x}\right) = P\left(Z \leq \frac{x_0-\mu_x}{\sigma_x}\right) \Rightarrow$$

- $$\Phi\left(\frac{x_0-\mu_x}{\sigma_x}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x_0-\mu_x}{\sigma_x}} \exp\left(-\frac{t^2}{2}\right) dt$$

- $$P(a \leq X \leq b) = \Phi\left(\frac{b-\mu_x}{\sigma_x}\right) - \Phi\left(\frac{a-\mu_x}{\sigma_x}\right)$$

### 3.15. Funzione degli errori

Si definisce "funzione degli errori" e si indica con  $\text{erf}(z)$  la funzione:

$$\text{erf}(z) = \frac{1}{\sqrt{2\pi}} \int_0^z \exp\left(-\frac{t^2}{2}\right) dt \quad (\text{A.60})$$

Tale funzione è legata alla funzione F.R.P.S.  $\Phi(z)$ ; per  $z \geq 0$  si ha:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 \exp\left(-\frac{t^2}{2}\right) dt + \frac{1}{\sqrt{2\pi}} \int_0^z \exp\left(-\frac{t^2}{2}\right) dt = 0.5 + \text{erf}(z) \quad (\text{A.61})$$

## 4. Variabili aleatorie multidimensionali

Per definire e per descrivere matematicamente gli eventi che dipendono da  $n$  parametri casuali, occorre introdurre la nozione di variabile aleatoria n-dimensionale, denominata anche vettore aleatorio ad  $n$  componenti. Ad es. nel caso in cui le variabili aleatorie sono due si parla di variabile aleatoria doppia o bidimensionale, e tutti i concetti espressi nei paragrafi precedenti con riferimento ad una singola variabile aleatoria possono essere facilmente estesi al caso multidimensionale.

Si considerino  $n$  variabili aleatorie  $X_1, X_2, X_3, \dots, X_n$ ; indichiamo con  $\bar{X}$  il vettore di tali v.a. e con  $\bar{x}$  il vettore le cui componenti  $x_1, x_2, x_3, \dots, x_n$  rappresentano i domini delle v.a.  $X_1, X_2, X_3, \dots, X_n$ :

$$\bar{X}^T = [X_1 \ X_2 \ X_3 \ \dots \ X_n] \quad (\text{A.62})$$

$$\bar{x}^T = [x_1 \ x_2 \ x_3 \ \dots \ x_n] \quad (\text{A.63})$$

## 4.1. Funzione di ripartizione congiunta

Si definisce Funzione di ripartizione congiunta  $F(x_1, x_2, x_3, \dots, x_n)$  delle  $n$  variabili aleatorie  $X_1, X_2, X_3, \dots, X_n$  la funzione definita in  $\mathbb{R}$  con valori dati dalla probabilità degli eventi  $(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3, \dots, X_n \leq x_n)$  e cioè

$$F(x_1, x_2, x_3, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3, \dots, X_n \leq x_n) \quad (\text{A.64})$$

Tale funzione rappresenta quindi la probabilità del verificarsi congiunto degli  $n$  eventi:  $(X_1 \leq x_1), (X_2 \leq x_2), (X_3 \leq x_3) \dots (X_n \leq x_n)$ .

## 4.2. Funzione densità di probabilità congiunta

Si definisce Funzione densità di probabilità congiunta la funzione  $f(x_1, x_2, x_3, \dots, x_n)$  non negativa e normalizzata

$$\begin{aligned} 1) & f(x_1, x_2, x_3, \dots, x_n) \geq 0 \\ 2) & \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, x_2, x_3, \dots, x_n) dx_1 dx_2 dx_3 \dots dx_n = 1 \end{aligned} \quad (\text{A.65})$$

tale che per la  $F(x_1, x_2, x_3, \dots, x_n)$  risulti

$$F(x_1, x_2, x_3, \dots, x_n) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(\zeta, \eta, \dots, \xi) d\zeta d\eta \dots d\xi \quad (\text{A.66})$$

Tale funzione rappresenta, nell'iperspazio a  $n+1$  dimensioni, una superficie che definisce la probabilità che, congiuntamente, le  $n$

variabili aleatorie assumano valori definiti all'interno di determinati intervalli dei corrispondenti domini. Ad esempio, per una variabile bidimensionale, la funzione densità di probabilità congiunta, relativa alle due variabili  $X_1$  e  $X_2$ , è una superficie bidimensionale  $f(x_1, x_2)$  definita in  $x_1, x_2$ , che è detta superficie di probabilità. Il volume racchiuso da questa superficie è uguale a 1, in accordo con l'assioma della normalizzazione. La probabilità che la v.a.  $X_1$  sia compresa tra  $a_1$  e  $b_1$  e che  $X_2$  sia compresa tra  $a_2$  e  $b_2$  è rappresentata dal volume sotteso da tale funzione in corrispondenza del dominio  $[a_1, b_1] \cap [a_2, b_2]$ , vedi figura A.10, ed è fornita da:

$$P(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f(x_1, x_2) dx_1 dx_2 \quad (\text{A.67})$$

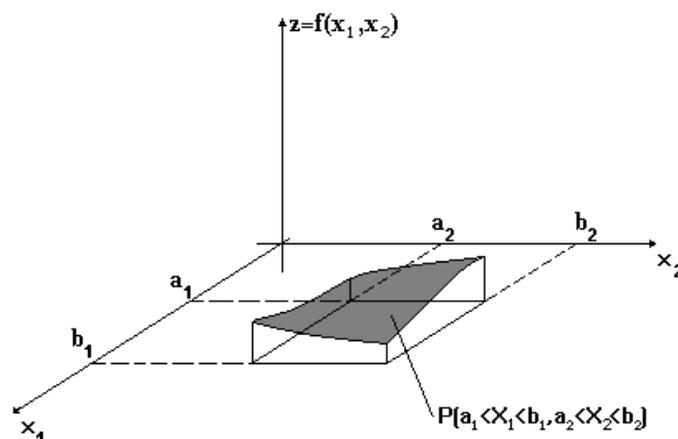


Figura A.10 - Funzione densità di probabilità congiunta di una v.a. bidimensionale.

### 4.3. Densità marginale

Data una variabile aleatoria a  $n$  dimensioni, si consideri la funzione di probabilità congiunta relativa alle prime  $m$  componenti del vettore  $\bar{X}$  (con  $m < n$ ) senza considerare le successive  $n-m$  variabili. Tale funzione viene definita "funzione densità di probabilità marginale" della variabile  $\bar{X}$ , e può essere ottenuta a partire dalla funzione densità di probabilità congiunta di  $\bar{X}$ , saturando le variabili  $x_{m+1}, x_{m+2}, \dots, x_n$  tramite integrazione, e cioè

$$f(x_1, x_2, x_3, \dots, x_m) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, x_2, x_3, \dots, x_n) dx_{m+1} dx_{m+2} \dots dx_n \quad (\text{A.68})$$

Ad esempio, nel caso di variabile bidimensionale, la funzione densità di probabilità marginale relativa ad una delle due variabili può essere ottenuta a partire da quella congiunta, attraverso la seguente integrazione

$$f(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2) dx_2 \quad (\text{A.69})$$

Dalle relazioni precedenti appare evidente che il numero di informazioni contenute nella funzione densità di probabilità congiunta è superiore al numero di quelle contenute nelle funzioni densità di probabilità marginale e che, mentre è possibile ottenere una qualunque funzione densità di probabilità marginale a partire da quella congiunta, non è possibile il viceversa.

#### 4.4. Densità di probabilità condizionata

Sia  $f(x_1, x_2, x_3, \dots, x_n)$  la funzione densità di probabilità congiunta delle v.a.  $X_1, X_2, X_3, \dots, X_n$ , si definisce "funzione densità di probabilità condizionata" delle prime  $m$  variabili (con  $m < n$ ) rispetto alle altre  $n-m$  variabili la funzione

$$f(x_1, x_2, \dots, x_m \mid x_{m+1}, x_{m+2}, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_m, \dots, x_n)}{f(x_{m+1}, x_{m+2}, \dots, x_n)} \quad (A.70)$$

Da tale relazione è possibile ottenere la funzione densità di probabilità congiunta relativa alle  $n$  variabili a partire dalla conoscenza della funzione densità di probabilità condizionata.

## 4.5. Variabili aleatorie indipendenti

Le variabili aleatorie  $X_1, X_2, X_3, \dots, X_n$  si dicono indipendenti se, e solo se, la funzione densità di probabilità congiunta di quante e quali si vogliano variabili tra le  $n$  assegnate è uguale al prodotto delle distribuzioni marginali unidimensionali delle variabili considerate, ossia ad es.

$$f(x_1, x_2, x_3, \dots, x_n) = f(x_1) f(x_2) f(x_3) \dots f(x_n) \quad (\text{A.71})$$

E' questo l'unico caso in cui è possibile costruire la funzione densità di probabilità congiunta a partire dalle funzioni densità relative alle singole variabili. E' ovvio che in questo caso la funzione densità di probabilità condizionata si può esprimere più semplicemente mediante la relazione

$$f(x_1, x_2, \dots, x_m | x_{m+1}, x_{m+2}, \dots, x_n) = f(x_1, x_2, \dots, x_m) \quad (\text{A.72})$$

## 4.6. Valori caratteristici di una distribuzione bidimensionale

### 4.6.1. Valor medio

Siano  $X_1$  e  $X_2$  due variabili casuali continue la cui funzione densità di probabilità congiunta è  $f(x_1, x_2)$ . I valori medi di  $X_1$  e  $X_2$ ,

denominati anche medie marginali, sono definiti rispettivamente da

$$E[X_1] = \mu_{x_1} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{+\infty} x_1 f(x_1) dx_1 \quad (\text{A.73})$$

$$E[X_2] = \mu_{x_2} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_2 f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{+\infty} x_2 f(x_2) dx_2$$

dove  $f(x_1)$  ed  $f(x_2)$  sono le funzioni densità di probabilità marginali. Le espressioni delle altre rilevanti misure sono quelle appresso riportate.

### 4.6.2. Varianza

$$\text{var}[X_1] = \sigma_{x_1}^2 = E[(X_1 - \mu_{x_1})^2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 - \mu_{x_1})^2 f(x_1, x_2) dx_1 dx_2 \quad (\text{A.74})$$

### 4.6.3. Covarianza

$$\text{cov}[X_1, X_2] = \sigma_{x_1, x_2} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_1 - \mu_{x_1})(x_2 - \mu_{x_2}) f(x_1, x_2) dx_1 dx_2 \quad (\text{A.75})$$

### 4.6.4. Coefficiente di correlazione lineare

$$\rho = \rho_{x_1, x_2} = \frac{\text{cov}[X_1, X_2]}{\sigma_{x_1} \sigma_{x_2}} \quad (\text{A.76})$$

#### 4.6.5. Valor medio condizionato

$$E[X_2|X_1] = \int_{-\infty}^{+\infty} x_2 f(x_2|x_1) dx_2 \quad (\text{A.77})$$

avendo indicato con

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f(x_1)} \quad (\text{A.78})$$

la funzione densità di probabilità condizionata di  $X_2$  rispetto  $X_1$ .

Si osservi che se le v.a. sono indipendenti risulta:  $E[X_2|X_1] = E[X_2]$

#### 4.7. Valor medio di una funzione di variabili aleatorie

Si consideri una funzione vettoriale  $\varphi$ , di dimensione  $l$ , che faccia corrispondere ad una variabile aleatoria  $\bar{X}$  ad  $n$  dimensioni una variabile aleatoria  $\bar{Y}$  ad  $l$  dimensioni, e cioè

$$\bar{Y} = \varphi(X_1, X_2, \dots, X_n) \quad (\text{A.79})$$

Il valor medio di tale v.a. può essere determinato generalizzando la (33) e cioè

$$E[\bar{Y}] = E[\varphi(\bar{X})] = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \varphi(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (\text{A.80})$$

E' facile mostrare che se la funzione vettoriale dipende solo dalle prime m componenti del vettore  $\bar{X}$  (con  $m < n$ ), allora all'interno dell'integrale, che bisogna calcolare per ottenere la media di tale funzione, compare soltanto la funzione densità di probabilità marginale corrispondente alle prime m componenti; infatti, tenendo presente la (68), si ha

$$\begin{aligned}
 E[\bar{Y}] &= E[\varphi(X_1, X_2 \dots X_m)] = \int_{-\infty \dots n \text{ volte}}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x_1, \dots, x_m) f(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n = \\
 &\int_{-\infty \dots m \text{ volte}}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x_1, \dots, x_m) dx_1 dx_2 \dots dx_m \int_{-\infty \dots m \text{ volte}}^{+\infty} f(x_1, \dots, x_n) dx_{m+1} dx_{m+2} \dots dx_n = (A.81) \\
 &\int_{-\infty \dots m \text{ volte}}^{+\infty} \int_{-\infty}^{+\infty} \varphi(x_1, \dots, x_m) f(x_1, \dots, x_m) dx_1 dx_2 \dots dx_m
 \end{aligned}$$

E' facile, allora verificare che la media di una funzione dipendente da una sola delle componenti del vettore  $\bar{X}$  è fornita dalla seguente relazione

$$E[X_i] = \int_{-\infty}^{+\infty} x_i f(x_i) dx_i \quad (A.82)$$

dove compare soltanto la funzione densità di probabilità della componente considerata.

## 4.8. Matrice di covarianza e di correlazione

Per una variabile aleatoria multidimensionale le covarianze (75)

e i coefficienti di correlazione lineare (76) definiscono due insiemi

ordinati di  $n \times n$  costanti, denominati rispettivamente

matrice di covarianza, [V]

$$[V] = \begin{bmatrix} v_{11} & v_{12} & \cdot & \cdot & \cdot & v_{1n} \\ v_{12} & v_{22} & \cdot & \cdot & \cdot & v_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{1n} & v_{n2} & \cdot & \cdot & \cdot & v_{nn} \end{bmatrix} \quad (\text{A.83})$$

essendo

$$v_{ij} = \text{cov}[X_i, X_j]$$

$$v_{ii} = \text{var}[X_i]$$

matrice di correlazione, [R]

$$[R] = \begin{bmatrix} \rho_{11} & \rho_{12} & \cdot & \cdot & \cdot & \rho_{1n} \\ \rho_{12} & \rho_{22} & \cdot & \cdot & \cdot & \rho_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{1n} & \rho_{n2} & \cdot & \cdot & \cdot & \rho_{nn} \end{bmatrix} \quad (\text{A.84})$$

essendo

$$\rho_{ij} = \frac{\text{cov}[X_i, X_j]}{\sigma_{x_i} \sigma_{x_j}} = \frac{v_{ij}}{\sqrt{v_{ii} v_{jj}}}$$

La matrice di covarianza  $[V]$  è simmetrica (in quanto  $v_{ij}=v_{ji}$ ) ed è definita positiva. Se le v.a. sono indipendenti, essendo  $\text{cov}[X_i, X_j]=0$  la matrice  $[V]$  è inoltre diagonale.

Anche la matrice  $[R]$  è simmetrica e definita positiva e se le v.a. sono indipendenti essa è uguale alla matrice identità  $I_n$ .

#### 4.9. Momenti di una variabile aleatoria multidimensionale

Data una variabile aleatoria  $\bar{X}$  ad  $n$  dimensioni, si definiscono momenti di ordine  $r$  della variabile  $\bar{X}$  tutte le possibili medie del tipo

$$E[X_1^t X_2^u \dots X_n^v] \quad \text{con } t+u+\dots+v=r \quad (\text{A.85})$$

Ad esempio, nel caso di una variabile aleatoria bidimensionale  $\mathbf{x}^T=[X_1 X_2]$ , i possibili momenti del secondo ordine sono tre, e precisamente:  $E[X_1^2]$ ,  $E[X_1 X_2]$  ed  $E[X_2^2]$ .

I momenti di ordine  $r$  di una variabile aleatoria ad  $n$  dimensioni possono essere espressi sinteticamente attraverso l'introduzione

dell'algebra di Kronecker [31,39]. Infatti, è facile verificare che tutti i possibili momenti di ordine  $r$  di  $\bar{X}$  possono essere raggruppati in un vettore attraverso la relazione:

$$m_r[\bar{X}] = E[\bar{X}^{[r]}] = E[\bar{X} \otimes \dots \text{ r volte } \dots \otimes \bar{X}] \quad (\text{A.86})$$

dove il simbolo  $\otimes$  indica il prodotto di Kronecker e l'esponente all'interno delle parentesi quadre indica l'elevazione a potenza secondo Kronecker.

Si richiamano di seguito alcune delle operazioni fondamentali dell'algebra di Kronecker.

#### 4.9.1. Prodotto di due matrici

Date due matrici  $A$  e  $B$  di ordine  $(p \times q)$  e  $(s \times t)$ , si definisce "prodotto di Kronecker" e si indica con  $A \otimes B$  la matrice di ordine  $(ps \times qt)$  che si ottiene moltiplicando ciascun elemento  $a_{ij}$  di  $A$

per l'intera matrice  $B$  e cioè

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{12}B & a_{22}B & \dots & a_{2n}B \\ \dots & \dots & \dots & \dots \\ a_{1n}B & a_{n2}B & \dots & a_{nn}B \end{bmatrix} \quad (\text{A.87})$$

E' importante notare che, per potersi effettuare il prodotto di Kronecker, non è necessario che le due matrici A e B siano conformi, così come accade quando se ne vuole effettuare il prodotto classico; quindi la precedente definizione può essere considerata come una generalizzazione dell'operazione di moltiplicazione tra matrici.

Il prodotto di Kronecker è caratterizzato dalle seguenti proprietà:

$$A \otimes (B \otimes C) = (A \otimes B) \otimes C \quad (\text{A.88})$$

$$(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D \quad (\text{A.89})$$

$$(A \otimes B)(C \otimes D) = AC \otimes BD \quad (\text{A.90})$$

$$(A \otimes B)^T = A^T \otimes B^T \quad (\text{A.91})$$

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \quad (\text{A.92})$$

## 4.9.2. Forma vettorializzata di una matrice

Per forma vettorializzata di una matrice  $A$  si intende il vettore colonna ottenuto scrivendo semplicemente le colonne della matrice una dietro l'altra, esso si indica con  $\text{Vec}(A)$ .

Se  $D$  è una matrice esprimibile attraverso il prodotto:  $D=ABC$ , allora è valida la relazione:

$$\text{Vec}(D) = (C^T \otimes A)\text{Vec}(B) \quad (\text{A.93})$$

## 4.9.3. Potenza secondo Kronecker

Le potenze secondo Kronecker di una matrice  $A$  possono essere definite recursivamente come segue

$$A^{[1]} = A; \quad \dots A^{[k+1]} = A \otimes A^{[k]} \quad (\text{A.94})$$

Le potenze di Kronecker sono caratterizzate dalle seguenti proprietà

$$A^{[k+j]} = A^{[k]} \otimes A^{[j]} \quad (\text{A.95})$$

$$(AB)^{[k]} = A^{[k]}B^{[k]} \quad (\text{A.96})$$

#### 4.10. Funzione caratteristica di v.a. multidimensionale

Analogamente al caso monodimensionale, un vettore di variabili aleatorie  $\bar{X}$  può essere caratterizzato dalla funzione caratteristica multidimensionale  $M_{\bar{X}}(\bar{\theta})$  che è posta in relazione con la funzione densità di probabilità congiunta dall'operatore trasformata di Fourier mediante la posizione

$$M_{\bar{X}}(\bar{\theta}) = \int_{-\infty \dots n \text{ volte}}^{+\infty} \int_{-\infty}^{+\infty} \exp(-i\bar{\theta}\bar{x}) f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (\text{A.97})$$

in cui  $\bar{\theta}$  è un vettore di  $n$  parametri reali e cioè  $\bar{\theta}^T = [\theta_1 \ \theta_2 \ \dots \ \theta_n]$ .

Sviluppando in serie di McLaurin, ed utilizzando la notazione di Kronecker [31,39] si ha inoltre

$$M_{\bar{X}}(\bar{\theta}) = \sum_{j=0}^{\infty} \frac{(-i)^j}{j!} m_j^T[\bar{X}] \bar{\theta}^{[j]} \quad (\text{A.98})$$

#### 4.11. Log-funzione caratteristica di v.a. multidimensionale

Anche in questo caso la log-funzione caratteristica può essere definita come il logaritmo naturale della corrispondente funzione caratteristica, e cioè

$$\ln[M_{\bar{X}}(\bar{\theta})] = \ln\left[\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \exp(-i\bar{\theta}\bar{x})f(x_1, x_2, \dots, x_n)dx_1 dx_2 \dots dx_n\right] \quad (\text{A.99})$$

Dalla precedente, sviluppando in serie di McLaurin ed utilizzando la notazione di Kronecker [31,39], si ricava

$$M_{\bar{X}}(\bar{\theta}) = \exp\left\{\sum_{j=0}^{\infty} \frac{(-i)^j}{j!} k_j^T[\bar{X}]\bar{\theta}^{[j]}\right\} \quad (\text{A.100})$$

in cui compaiono i cumulanti  $k[X]$ , legati ai momenti da relazioni analoghe alle (49).

## 5. Bibliografia

1. C. Meyer, "Random vibration analysis" - CISM Course of modeling of R/C structures. Udine, 1993.
2. L. Gotusso, "Calcolo numerico," Clup, Milano, 1978.
3. E. Kreyszig, "Advanced Engineering Mathematics," Wiley, N.Y., 1972.
4. S.C. Chapra, R.P. Canale, "Metodi numerici per l'ingegneria" McGraw-Hill, Milano, 1988.
5. G. Falsone, "Calcolo differenziale stocastico per l'analisi di sistemi strutturali soggetti a forzanti non gaussiane," Palermo, 1991.
6. Isaac Elishakoff, "Probabilistic methods in the theory of structures" - John Wiley & Sons, New York, 1983.
7. Heinz Parkus, "Random processes in mechanical sciences" - Springer-Verlag, Wien-NewYork, 1969.
8. J.L.Zeman, "Approximate analysis of stochastic processes in mechanics" - Springer-Verlag, Wien-NewYork, 1971.
9. M.R. Spiegel, "Statistica" - Etas, Milano, 1973.