

Teoria dei Segnali

□ Ripasso di teoria della probabilità e variabili casuali



**Politecnico
di Torino**

Department
of Electronics and
Telecommunications

Commento importante: le slide che seguono sono da intendersi come un “elenco” degli argomenti relative alla teoria della probabilità che è necessario conoscere per questo Corso.

Tutti gli argomenti presentati in questo capitolo dovrebbero già essere noti dal corso «Metodi matematici per l'ingegneria», e vengono qui riportati (quasi) sempre senza dimostrazioni.

Il «ripasso» di teoria della probabilità presentato in queste slide è propedeutico all'argomento successivo, relativo alla descrizione probabilistica di segnali aleatori, cioè con caratteristiche casuali

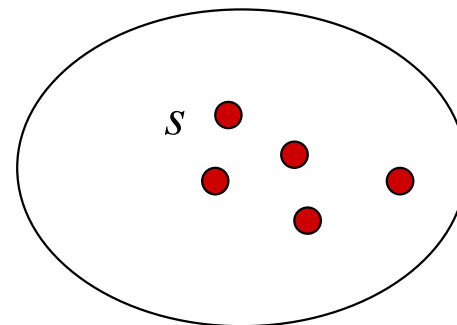
Spazio campione

Teoria assiomatica della probabilità

- Uno spazio campione S è un insieme di possibili risultati s di un esperimento casuale
- Ogni risultato ha associata una probabilità P con le seguenti proprietà:

$$P(s) \in [0,1]$$

$$\sum_{s \in S} P(s) = 1$$



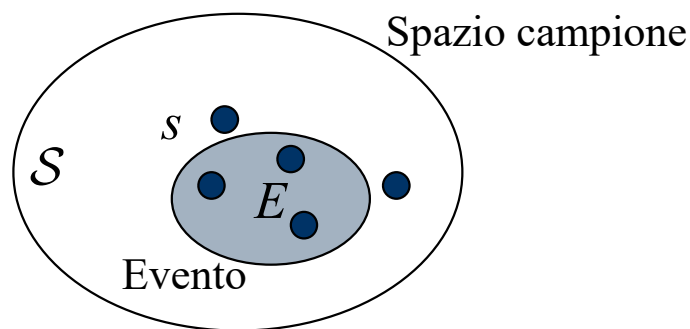
S = Spazio campione

S, P Spazio di probabilità

- I risultati dell'esperimento casuale sono sempre mutuamente esclusivi

Spazio campione

- Un evento è un insieme di possibili risultati ed è quindi un sottoinsieme dello spazio campione



$$E = \{s_i\} \subseteq \mathcal{S}$$

$$P(E) = \sum_i P(s_i)$$

- La probabilità di una unione di eventi è sempre minore o uguale della somma delle probabilità dei singoli eventi

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

*Probabilità dell'unione di due eventi
(in pratica: "OR dei due eventi")*

*Probabilità dell'intersezione di due eventi, detta anche
probabilità congiunta
(in pratica: "AND" dei due eventi")
si vedano slide successive)*

$$P\left(\bigcup_i E_i\right) \leq \sum_i P(E_i) \quad \text{limite dell'unione}$$

Probabilità congiunta

- È la probabilità dell'intersezione di due eventi A e B

$$P(A, B) \hat{=} P(A \cap B)$$

notazione

- Si tratta della probabilità che A e B avvengano "congiuntamente"
- Esempio pratico:
 - A = giornata soleggiata
 - B = giornata calda
 - $P(A, B)$ = probabilità che una giornata sia contemporaneamente soleggiata e calda

Probabilità congiunta ed eventi statisticamente indipendenti

- Due eventi A e B sono detti statisticamente indipendenti quando:

$$P(A, B) \hat{=} P(A) \cdot P(B)$$

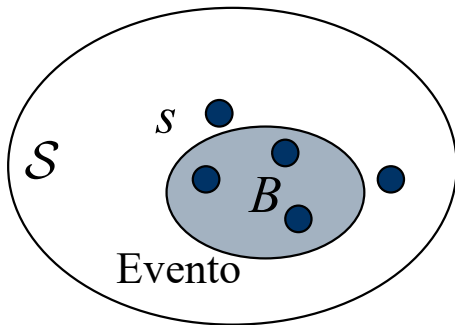
Il concetto di indipendenza statistica è molto importante: in molti problemi pratici, poter assumere che due eventi sono statisticamente indipendenti semplifica spesso in maniera sostanziale i calcoli

- In pratica: i due eventi NON si influenzano a vicenda

- L'esempio pratico della slide precedente (giornata soleggiata e calda) riguarda due eventi chiaramente NON statisticamente indipendenti
- Un semplice esempio di due eventi indipendenti
 - A=giornata soleggiata
 - B=giornata festiva

Probabilità condizionata

- Dati due generici eventi B e s , si definisce la probabilità di s condizionata all'evento B come:



$$P(s | B) \triangleq \begin{cases} \frac{1}{P(B)} P(s) & \text{se } s \in B \\ 0 & \text{se } s \notin B \end{cases}$$

da cui se $s \in B \Rightarrow P(s) = P(s | B) P(B)$

In pratica, con questa notazione si indica la probabilità dell'evento s sapendo che è "vero" l'evento B

*La formula si legge:
«probabilità di s condizionata a B » o
«probabilità di s dato B »*

- In sostanza: B diventa quindi il nuovo spazio campione, e le probabilità sono rinormalizzate a questo spazio campione
- Cioè se si sommano le probabilità di tutti i risultati s in B si ottiene:

$$\sum_{s \in B} P(s | B) = 1$$

Esempi pratici

- La probabilità di ottenere "1" al lancio di un dado è $1/6$

$$P(\text{lancio dado} = 1) = 1/6$$

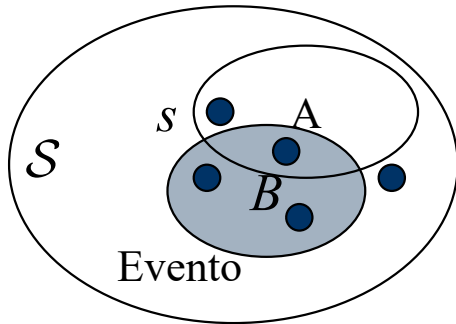
- La probabilità di ottenere "1" al lancio di un dado condizionata al fatto di sapere che il risultato sia un numero dispari è $1/3$

$$P(\text{lancio dado} = 1 \mid \text{risultato dispari}) = 1/3$$

- Si ha anche, ad esempio:

$$P(\text{lancio dado} = 1 \mid \text{risultato pari}) = 0$$

Teorema di Bayes



$$P(A \cap B) = P(A, B) = P(A | B)P(B) = P(B | A)P(A)$$

Probabilità congiunta

Questo risultato è a volte utile quando si deve calcolare la probabilità congiunta di due eventi che non sono statisticamente indipendenti. In questo caso, a volte è più facile calcolare prima le probabilità condizionate, e successivamente con la formula di Bayes calcolare le probabilità congiunte

Dimostrazione:
$$P(A \cap B) = P(A, B) = \sum_{s \in A \cap B} P(s) = P(B) \sum_{s \in A} P(s | B) = P(B)P(A | B)$$

Somma delle probabilità di tutti gli eventi s che stanno nell'intersezione tra A e B

Qui si sfrutta la formula sottostante ottenuta nelle slide precedenti, e osservando che in questo caso certamente s appartiene anche a B

$$\text{se } s \in B \Rightarrow P(s) = P(s | B)P(B)$$

Il canale di comunicazione discreto



Un esempio di applicazione dei concetti precedenti

- Consideriamo un generico sistema di comunicazione nel quale i messaggi generati da una sorgente di informazione giungono ad un ricevitore attraverso un canale di comunicazione



- **Sorgente discreta**: Produce messaggi, allineando in successione i simboli appartenenti ad un "alfabeto" di dimensione finita.
- **Caso particolare (fondamentale) Sorgente binaria discreta**: Può emettere solo i simboli X_0 ed X_1 appartenenti all'alfabeto $X = \{X_0, X_1\}$ ("bit" 0 e 1)
 - NOTA: Attraverso un codificatore, le sorgenti discrete possono essere ricondotte ad una sorgente binaria.

Simbolo	Codifica in bit
A	00
B	01
C	10
D	11

I simboli possono essere (alcuni esempi):

- Flussi di bit
- M valori discreti (Esempio: caratteri ASCII)
- le uscite di un quantizzatore in un convertitore analogico-digitale (ADC)

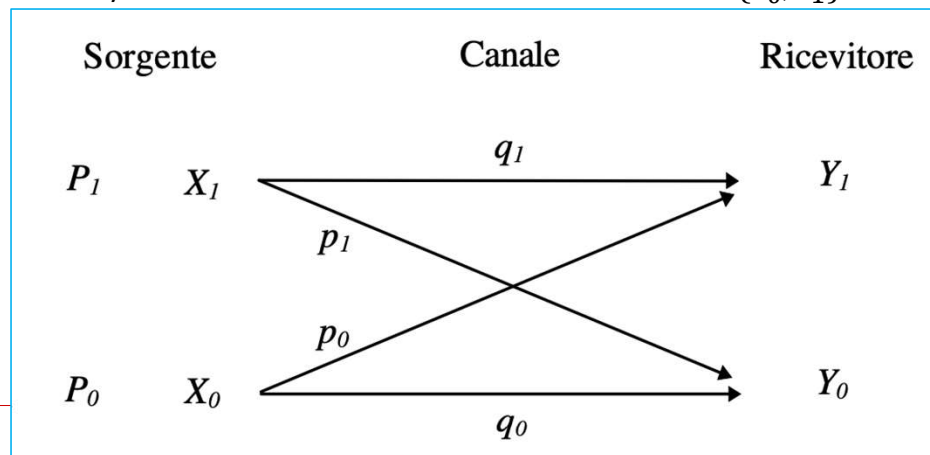
Il canale di comunicazione discreto

- L'emissione di un simbolo può essere trattata come un esperimento casuale il cui spazio campione è costituito da 2 eventi. Le probabilità associate sono note ([caratterizzazione probabilistica della sorgente](#)):

$$P_0 = P(X_0), \quad P_1 = P(X_1) \quad \text{con } P_0 + P_1 = 1$$

- **Canale di comunicazione discreto:** Trasferisce i simboli generati dalla sorgente al destinatario. I disturbi presenti su un canale reale possono generare degli errori in ricezione.
- **Canale binario discreto:** un semplice modello che descrive come simboli dell'alfabeto di ingresso $X = \{X_0, X_1\}$ in ingresso divengono, in uscita, i simboli dell'alfabeto binario $Y = \{Y_0, Y_1\}$:

Schema di un generico canale binario discreto, con le probabilità su ciascuna delle transizioni nel grafo

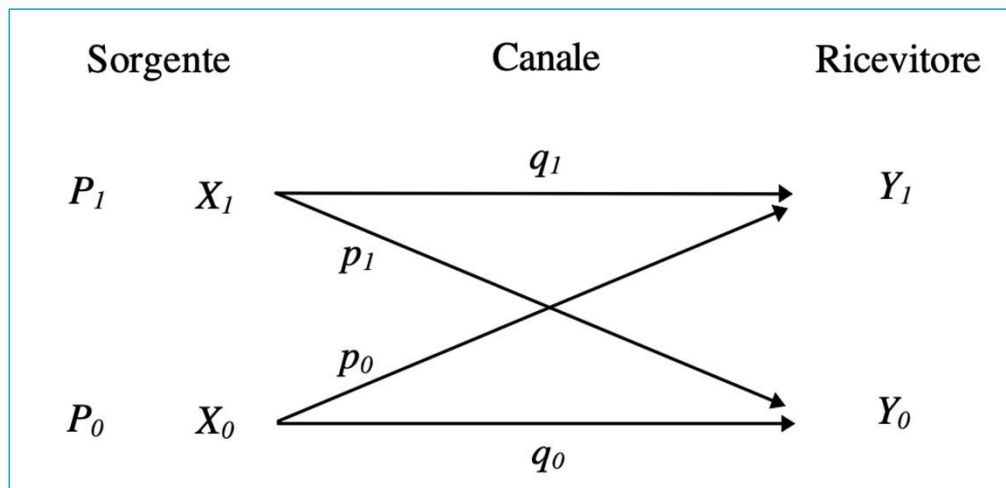


Caratterizzazione probabilistica del canale di comunicazione discreto

Gli errori provocati dal canale di trasmissione possono essere rappresentati con un modello probabilistico, definito dalle seguenti probabilità (riportate sul grafo):

- $q_1 = P\{Y_1 | X_1\} = P\{\text{"viene ricevuto } Y_1 \text{ essendo stato trasmesso } X_1\}$
- $p_1 = P\{Y_0 | X_1\} = P\{\text{"viene ricevuto } Y_0 \text{ essendo stato trasmesso } X_1\}$
- $q_0 = P\{Y_0 | X_0\} = P\{\text{"viene ricevuto } Y_0 \text{ essendo stato trasmesso } X_0\}$
- $p_0 = P\{Y_1 | X_0\} = P\{\text{"viene ricevuto } Y_1 \text{ essendo stato trasmesso } X_0\}$

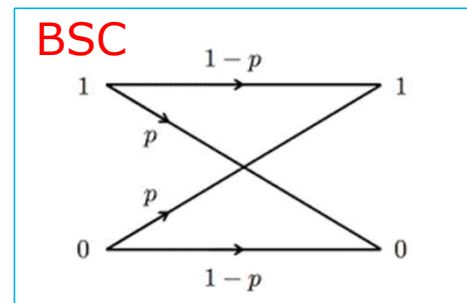
dove: $q_0 + p_0 = q_1 + p_1 = 1$



Il canale si dice binario simmetrico (binary symmetric channel BSC) SE:

$$p_0 = p_1 = p \text{ e } q_0 = q_1 = (1 - p),$$

cioé se tratta entrambi i bit "allo stesso modo"



Esempi di calcolo di probabilità sul canale di comunicazione discreto

- Calcoliamo la probabilità di ricevere Y_1 . Tale evento si verifica se uno dei due eventi si avvera:
 1. La sorgente ha trasmesso X_1 e non si è verificato un errore.
 2. La sorgente ha trasmesso X_0 e si è verificato un errore.
- Poiché i due eventi sono mutuamente esclusivi (o viene trasmesso X_0 o viene trasmesso X_1), si può calcolare la probabilità di ricevere Y_1 mediante il **teorema della probabilità totale**:

$$P\{Y_1\} = P\{(Y_1 \cap X_0) \cup (Y_1 \cap X_1)\} = P\{Y_1 | X_0\}P\{X_0\} + P\{Y_1 | X_1\}P\{X_1\} = p_0P_0 + q_1P_1$$

Analoghi risultati si ottengono per $P(Y_0)$

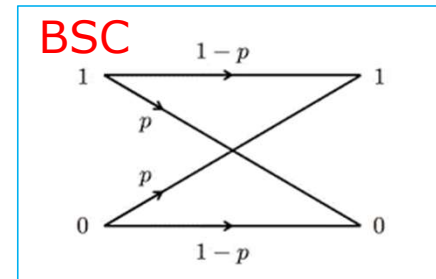
- Nel caso di BSC: $P\{Y_1\} = pP_0 + (1 - p)P_1$
 - Se anche la sorgente binaria è simmetrica, $P\{Y_1\} = 0.5p + 0.5(1 - p) = 0.5$

Il canale di comunicazione discreto: Probabilità di errore

- La principale grandezza che caratterizza la bontà del canale è la probabilità di errore, cioè la probabilità dell'evento errore E definito come:

$$E = \{X_0 \text{ trasmesso e } Y_1 \text{ ricevuto oppure } X_1 \text{ trasmesso e } Y_0 \text{ ricevuto}\}$$

$$= (X_0 \cap Y_1) \cup (X_1 \cap Y_0)$$



- Usando il teorema della probabilità totale, ed osservando che gli eventi $(X_0 \cap Y_1)$ e $(X_1 \cap Y_0)$ sono mutuamente esclusivi, si può esprimere $P(E)$ come

$$P\{E\} = P\{(X_0 \cap Y_1) \cup (X_1 \cap Y_0)\} = P\{Y_1 | X_0\}P\{X_0\} + P\{Y_0 | X_1\}P\{X_1\} = p_0P_0 + p_1P_1$$

- Nel caso di BSC: $P\{E\}_{BSC} = pP_0 + pP_1 = p(P_0 + P_1) = p$

Attendibilità del simbolo recuperato sul canale di comunicazione discreto

- Dal punto di vista del ricevitore però è importante sapere qual è **l'attendibilità del simbolo recuperato** dal meccanismo di decisione.

In altre parole, ci si pone la seguente domanda:

- Quale è la probabilità che, avendo ricevuto Y_0 , sia stato effettivamente trasmesso X_0 ?
- La risposta è data dalla probabilità "a posteriori" $P\{X_0|Y_0\}$ che si ottiene facendo uso del **teorema di Bayes**

$$P\{X_0|Y_0\} = \frac{P\{X_0, Y_0\}}{P\{Y_0\}} = \frac{P\{Y_0|X_0\}P\{X_0\}}{P\{Y_0|X_0\}P\{X_0\} + P\{Y_0|X_1\}P\{X_1\}} = \frac{q_0P_0}{q_0P_0 + p_1P_1}$$

Verifica: se il canale non fa errori, cioè se $p_i=0$ e $q_i=1$, il risultato è $P\{X_0|Y_0\} = 1$

Attendibilità del simbolo recuperato per canale BSC e sorgente simmetrica

- Consideriamo il caso particolare di canale BSC e sorgente simmetrica:

$$P\{X_0|Y_0\} = \frac{q_0 P_0}{q_0 P_0 + p_1 P_1} = q = 1 - p$$

- I sistemi di trasmissione digitale sono progettati in modo da avere p molto basso (ma ahimè mai nullo!)
 - La maggior parte degli standard trasmissivi richiedono di assicurare "all'utente finale" $p \leq 10^{-12}$
 - Anche gli hard disk creano errori sui bit nel processo di lettura/scrittura, che per la maggior parte vengono poi corretti da complessi algoritmi di "Forward error correction" (FEC)

Attendibilità del simbolo recuperato sul canale di comunicazione discreto

- Consideriamo un canale BSC e sorgente simmetrica in due casi estremi

$$p = 1$$

probabilità di ricevere il simbolo Y_0 $P\{Y_0\} = p_1P_1 + q_0P_0 = \frac{1}{2}$

probabilità di errore $P\{E\}_{BSC} = p = 1$

probabilità a posteriori $P\{X_0|Y_0\} = \frac{(1-p)P_0}{(1-p)P_0 + pP_1} = 0$

la probabilità a posteriori vale 0 quindi si è certi che, avendo ricevuto Y_0 , sia stato trasmesso X_1 (errore sistematico)

$$p = \frac{1}{2}$$

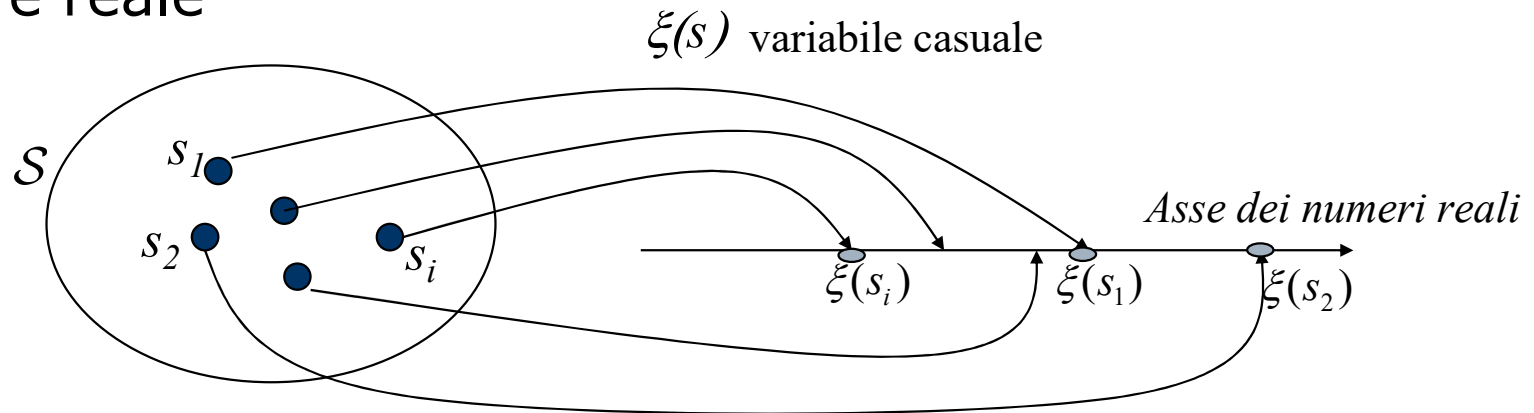
$$P\{Y_0\} = p_1P_1 + q_0P_0 = \frac{1}{2}$$
$$P\{E\}_{BSC} = p = \frac{1}{2}$$
$$P\{X_0|Y_0\} = \frac{(1-p)P_0}{(1-p)P_0 + pP_1} = \frac{1}{2}$$

quando l'incertezza sul simbolo trasmesso vale $\frac{1}{2}$ siamo nel caso peggiore perché significa che l'uscita risulta indipendente dall'ingresso

In entrambi i casi si riceve il simbolo Y_0 con probabilità $\frac{1}{2}$ ma mentre nel primo caso c'è sempre errore (e quindi si può recuperare perché basta invertire la decisione), nell'altro caso il simbolo ricevuto è totalmente casuale come nel caso del lancio di una moneta

Variabile casuale

- Una variabile casuale $\xi(s)$ è una trasformazione (mapping) che associa ad ogni elemento di uno spazio campione un valore reale

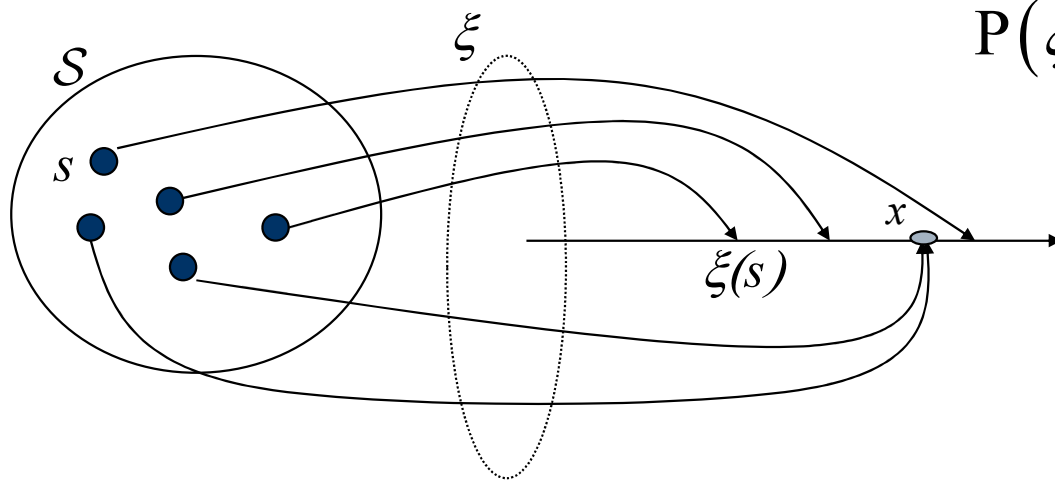


Nella maggior parte dei casi, l'associazione tra spazio campione e numeri reali è «ovvia». Si pensi ai seguenti esempi

- Temperatura in una stanza: se considerata come evento casuale, è ovvio che la relativa variabile casuale è semplicemente il valore reale assunto dalla temperatura
- Tensione casuale ai capi di un componente elettronico
- I sei numeri che può assumere il lancio di un dado

Variabile casuale

- Una variabile casuale è quindi uno strumento che ci consente di rappresentare sull'asse reale uno spazio campione.
- La probabilità indotta dallo spazio campione sull'asse reale può essere così espressa:



$$P(\xi=x) = \sum_{s \in \mathcal{S}: \xi(s)=x} P(s)$$

Somma delle probabilità di tutti i risultati s che sono mappati in x

Questa formula dà il caso generale. In pratica, in realtà, nella maggior parte dei casi pratici, ogni evento dello spazio campione è mappato ad un singolo valore reale, e dunque la sommatoria di fatto sparisce.

Quando la variabile casuale assume valori reali, è spesso molto più rilevante valutare la probabilità di un intervallo di valori reali

$$P(\xi \in [a, b])$$

Funzione di distribuzione cumulativa

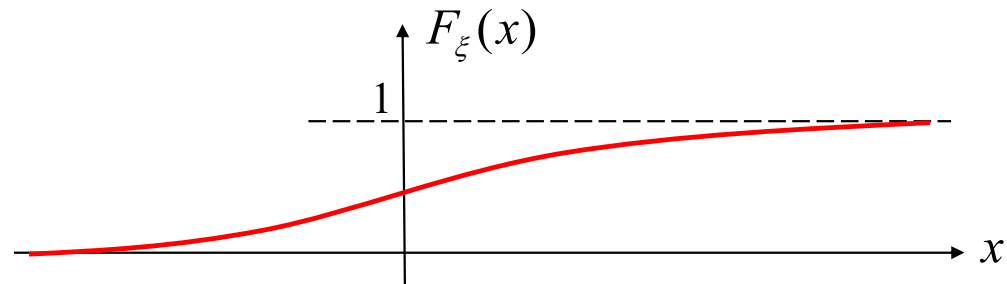
□ Definizione:

$$F_{\xi}(x) \triangleq P(\xi \leq x) = \sum_{s: \xi(s) \leq x} P(s)$$

□ Proprietà

- È una funzione monotona non decrescente
- Essendo coincidente con una probabilità, è compresa tra zero e 1

$$\begin{aligned} F_{\xi}(\infty) &= 1 \\ F_{\xi}(-\infty) &= 0 \end{aligned}$$



Esempio qualitativo di andamento di una funzione di distribuzione cumulativa per una variabile casuale che assume valori da $-\infty$ a $+\infty$

Proprietà fondamentale

- La conoscenza della funzione di distribuzione cumulativa permette di calcolare la probabilità di un qualunque intervallo di valori $[a, b]$ tramite:

$$P(\xi \in [a, b]) = F_{\xi}(b) - F_{\xi}(a)$$

- Ne deriva che la funzione di distribuzione cumulativa caratterizza completamente una variabile casuale

- Dimostrazione:

$$P(\xi \in [a, b]) = P(\xi \in [-\infty, b]) - P(\xi \in [-\infty, a]) = F_{\xi}(b) - F_{\xi}(a)$$

Densità di probabilità

□ Definizione:

$$f_{\xi}(x) \triangleq \frac{\partial}{\partial \mathbf{x}} F_{\xi}(x)$$

■ Derivata della funzione di distribuzione cumulativa

□ Proprietà

■ È sempre maggiore o uguale a zero

■ Ha area unitaria

$$\int_{-\infty}^{\infty} f_{\xi}(x) dx = F_{\xi}(+\infty) - F_{\xi}(-\infty) = 1$$

■ Proprietà fondamentale:

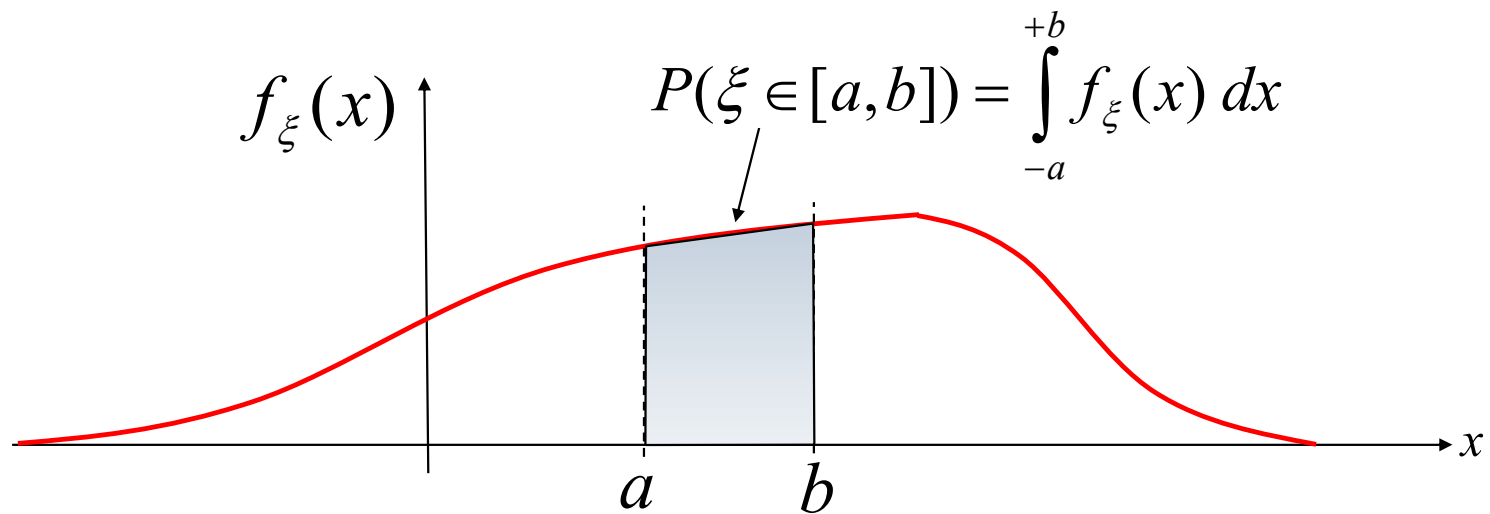
■ Dimostrazione:

$$P(\xi \in [a, b]) = \int_a^b f_{\xi}(x) dx$$

$$P(\xi \in [a, b]) = P(\xi \in [-\infty, b]) - P(\xi \in [-\infty, a]) = F_{\xi}(b) - F_{\xi}(a) = \int_a^b f_{\xi}(x) dx$$

Densità di probabilità

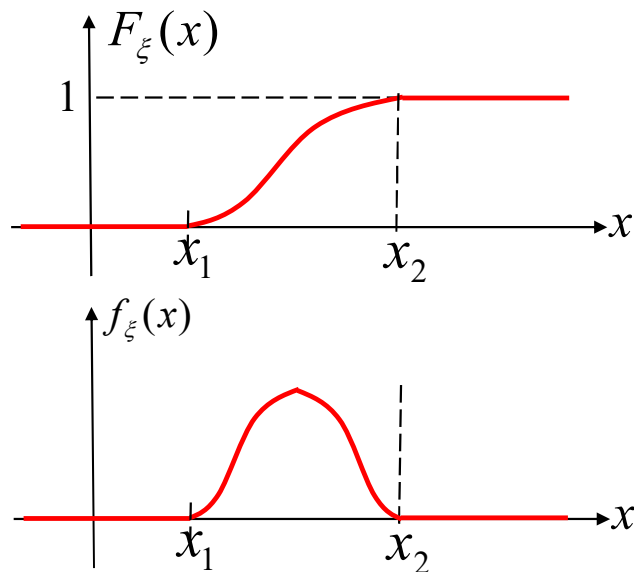
□ Graficamente:



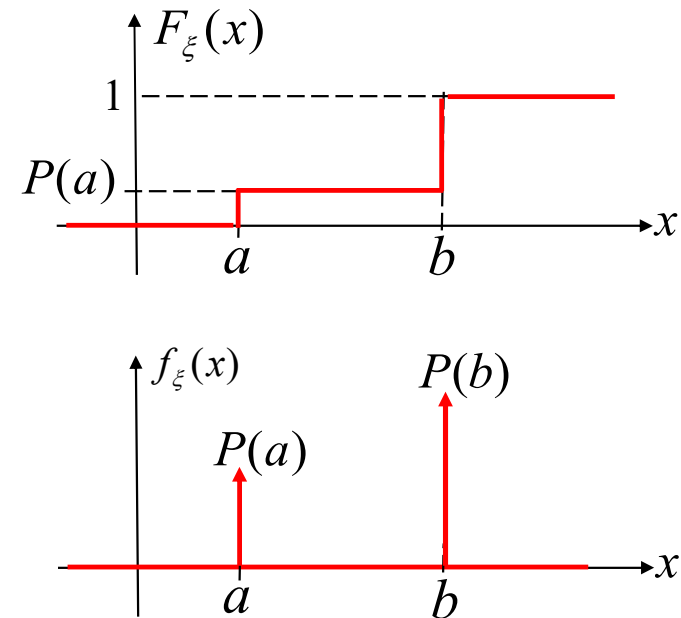
Esempio qualitativo di andamento di una densità di probabilità per una variabile casuale che assume valori da $-\infty$ a $+\infty$

Esempi grafici

- Variabile casuale che assume valori continui solo all'interno di un certo intervallo $[x_1, x_2]$



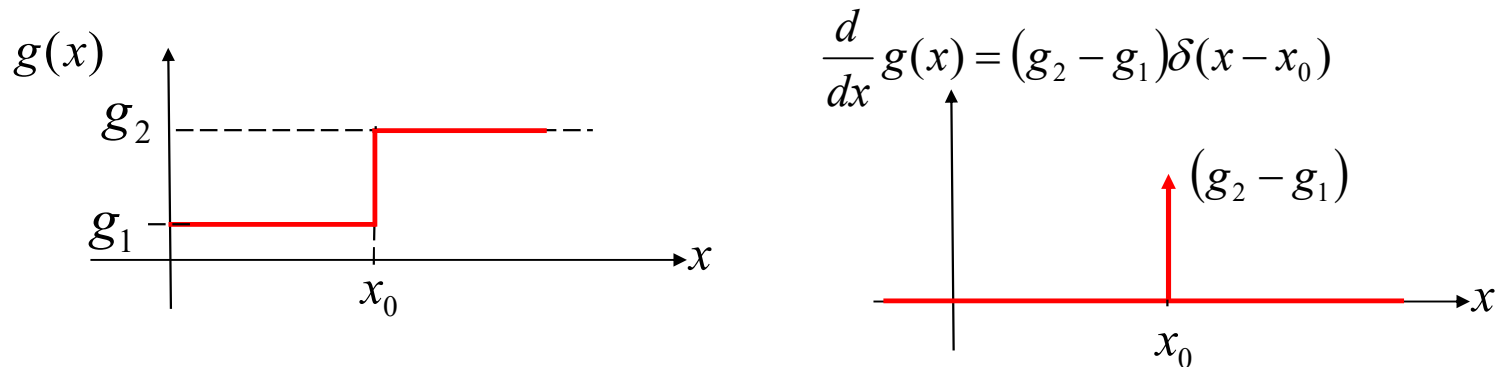
- Variabile casuale che assume solo due valori discreti a e b



(si veda slide successiva per la derivazione di questo risultato)

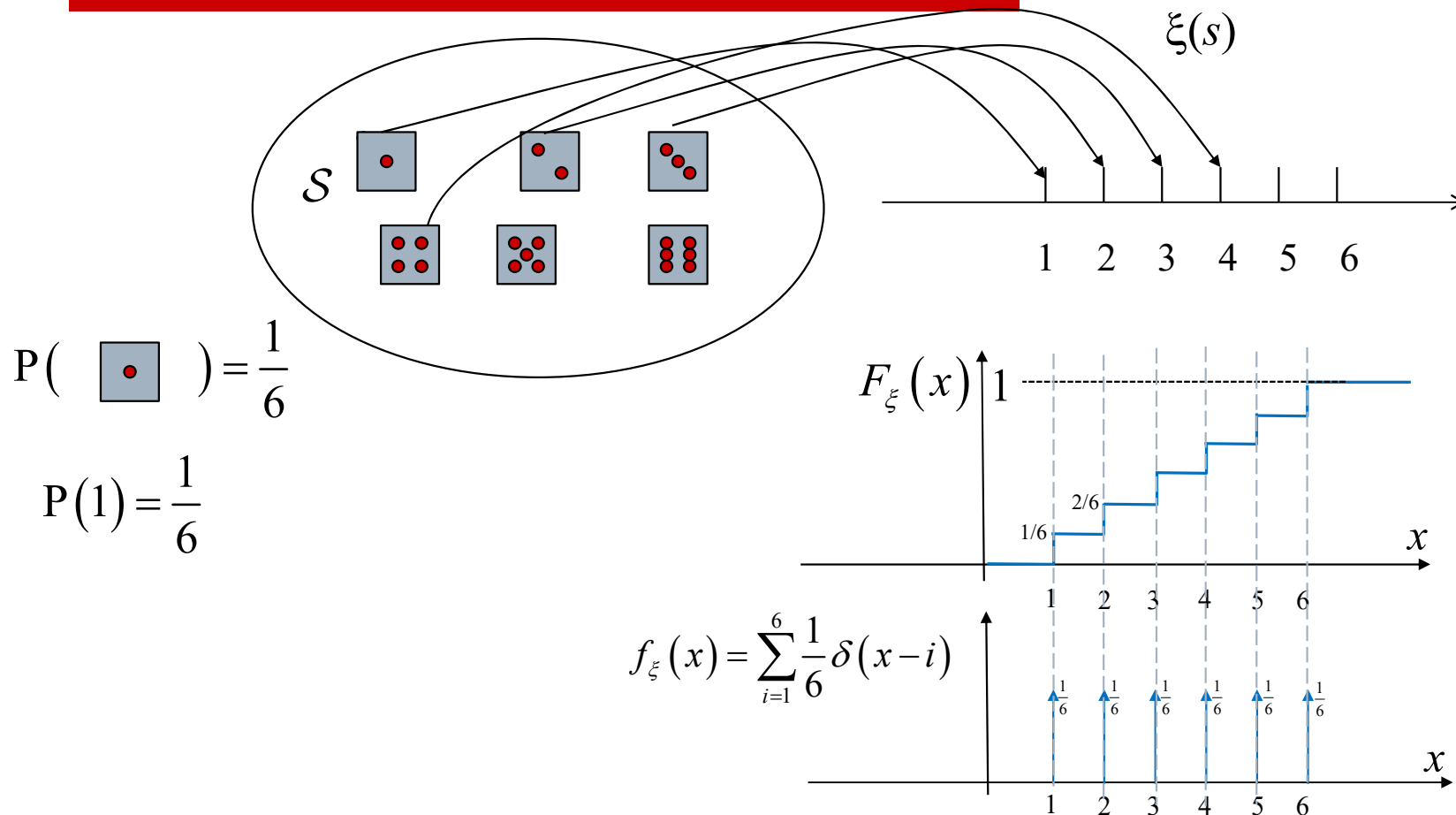
Derivata di una funzione a gradini

- Come visto in precedenza, la funzione delta è utile anche per esprimere la derivata di una funzione con delle discontinuità "a gradini"
- Si ricorda che tramite funzioni "nel senso tradizionale" la derivate NON esiste sulla discontinuità



Dimostrazione intuitiva: si provi a fare il passaggio inverso, integrando la delta da $-\infty$ a x

Esempio variabile casuale discreta: lancio di un dado



Insiemi di variabili casuali

- ❑ In moltissimi problemi ingegneristici e fisici, è necessario trattare più di una variabile casuale contemporaneamente
- ❑ Questo risulta solitamente abbastanza complesso dal punto di vista matematico, soprattutto se le varie variabili casuali in gioco NON sono indipendenti
- ❑ Nelle prossime slide, si riportano alcuni cenni sulla caratterizzazione probabilistica di insiemi di variabili casuali

Caratterizzazione di insiemi di variabili casuali

□ Distribuzione cumulativa congiunta di variabili casuali

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) \triangleq P(\xi_1 \leq x_1 \cap \dots \cap \xi_n \leq x_n) \quad \text{probabilità congiunta}$$

□ Densità di probabilità congiunta

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) \triangleq \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$$

□ Indipendenza statistica

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = f_{\xi_1}(x_1) \cdots f_{\xi_n}(x_n)$$

Prodotto delle densità di
probabilità

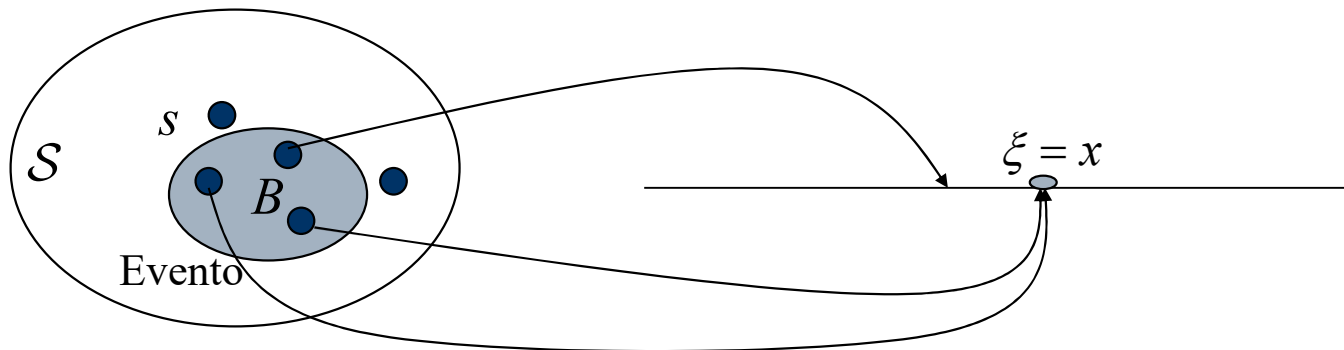
Il concetto di indipendenza statistica, che avevamo introdotto nelle slides precedenti a livello di probabilità, e la relativa formula sulle densità di probabilità sono molto importanti.

Nell'ambito dei processi casuali (prossimo argomento del corso) faremo spesso l'ipotesi di indipendenza statistica

Distribuzione cumulativa condizionate

- Dato il verificarsi di un evento B nello spazio campionario, è possibile definire una distribuzione condizionata.

$$F_{\xi}(x | B) \triangleq P(\xi \leq x | s \in B) = \frac{1}{P(B)} \sum_{s \in B: \xi(s) \leq x} P(s)$$



- Conseguentemente, si può definire la densità di probabilità condizionata come:
- $$f_{\xi}(x | B) \triangleq \frac{d}{dx} (F_{\xi}(x | B))$$

Distribuzioni condizionate

- Dal teorema di Bayes discendono le seguenti proprietà:

$$f_{\xi,\eta}(x,y) = f_{\xi|\eta}(x|y)f_{\eta}(y) = f_{\eta|\xi}(y|x)f_{\xi}(x)$$

- Da cui, siccome

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f_{\xi,\eta}(x,y)dy$$

- Allora

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f_{\xi|\eta}(x|y)f_{\eta}(y)dy$$

In alcune situazioni relative ai processi casuali, questa formula risulterà utile

Valore atteso e momenti

- Data una funzione $g(\cdot)$ che operi su variabile casuale ξ , si definisce il suo valore atteso (*Expected Value*) come:

$$E\{g(\xi)\} \triangleq \int_{-\infty}^{\infty} g(x) f_{\xi}(x) dx$$

Nota: Il risultato è un numero reale che dipende sia dalla densità di probabilità che dalla funzione $g(\cdot)$

- A partire da questa definizione generale, si definiscono poi i cosiddetti «Momenti di ordine k »:

$$\mu_k \triangleq E\{\xi^k\}$$

- In particolare:

$$E\{\xi\} = \mu = \mu_1 = \int_{-\infty}^{+\infty} x \cdot f_{\xi}(x) dx$$

Il momento di ordine 1 è denominato «media»

$$E\{\xi^2\} = \mu_2 = \int_{-\infty}^{+\infty} x^2 \cdot f_{\xi}(x) dx$$

Il momento di ordine 2 è denominato «Valore quadratico medio»

Momenti centrali

□ Momenti centrali $m_k \triangleq E\left\{(\xi - \mu)^k\right\}$

□ Varianza di una variabile casuale

$$m_2 = E\left\{(\xi - \mu)^2\right\} \triangleq \sigma_\xi^2$$

□ Commenti "qualitativi" sui momenti:

- La media di una variabile casuale indica qualitativamente il "centro" della densità di probabilità
- La varianza dà una indicazione della larghezza della densità di probabilità attorno alla media
- Si definisce deviazione standard la radice quadrata della varianza

$$\sigma_\xi^2 = \int_{-\infty}^{+\infty} (x - \mu_1)^2 \cdot f_\xi(x) dx \quad \Rightarrow \quad \sigma_\xi = \sqrt{\int_{-\infty}^{+\infty} (x - \mu_1)^2 \cdot f_\xi(x) dx}$$

Definizione di varianza *Definizione di deviazione standard*

Momenti definiti su due variabili casuali

Coefficiente di correlazione

- Date due variabili casuali, si possono definire momenti e momenti centrali congiunti

$$\mu_{kn} = E\{\xi^k \eta^n\}$$
$$m_{kn} = E\left\{\left(\xi - \mu_\xi\right)^k \left(\eta - \mu_\eta\right)^n\right\}$$

- La covarianza è definita come: $\sigma_{\xi\eta} = m_{1,1}$

- Il coefficiente di correlazione è: $\rho_{\xi\eta} = \frac{\sigma_{\xi\eta}}{\sigma_\xi \sigma_\eta}$

Variabili linearmente indipendenti e statisticamente indipendenti

□ Indipendenza statistica

$$f_{\xi_1, \xi_2}(x_1, x_2) = f_{\xi_1}(x_1) f_{\xi_2}(x_2)$$

Terminologia: quando è verificata questa condizione si usa dire che le due variabili casuali sono statisticamente indipendenti

□ Indipendenza lineare ("scorrelazione")

$$E\{\xi_1 \xi_2\} = E\{\xi_1\} E\{\xi_2\}$$

Terminologia: quando è verificata questa condizione si usa dire che le due variabili casuali sono scorrelate

Dimostrazione

$$\begin{aligned} E\{\xi_1 \xi_2\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2 \\ &\quad \downarrow \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{\xi_1}(x_1) f_{\xi_2}(x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} x_1 f_{\xi_1}(x_1) dx_1 \int_{-\infty}^{\infty} x_2 f_{\xi_2}(x_2) dx_2 \\ &= E\{\xi_1\} E\{\xi_2\} \end{aligned}$$

Combinazione lineare di variabili casuali

□ Se $Z = \sum_{i=1}^N \alpha_i X_i$ vale sempre \longrightarrow $\mu_Z = \sum_{i=1}^N \alpha_i \mu_i$ *Importante: l'operatore di media è dunque un operatore lineare*

□ se ξ_i sono scorrelate, allora vale anche $\sigma_Z^2 = \sum_{i=1}^N \alpha_i^2 \sigma_i^2$

Dimostrazione

$$\begin{aligned} E[|Z - \mu_Z|^2] &= E\left[\sum_{j=1}^N \alpha_j (X_j - \mu_j) \sum_{i=1}^N \alpha_i (X_i - \mu_i)\right] \\ &= E\left[\sum_{j=1}^N \sum_{i=1}^N \alpha_i \alpha_j (X_j - \mu_j) (X_i - \mu_i)\right] \\ &= \sum_{j=1}^N \sum_{i=1}^N \alpha_i \alpha_j E[(X_j - \mu_j) (X_i - \mu_i)] \\ &= \sum_{i=1}^N \alpha_i^2 E[(X_i - \mu_i)^2] + \sum_{i=1}^N \sum_{j \neq i}^N \alpha_i \alpha_j E[(X_j - \mu_j) (X_i - \mu_i)] \\ &= \sum_{i=1}^N \alpha_i^2 \sigma_i^2 + \sum_{j=1}^N \sum_{i \neq j}^N \alpha_i \alpha_j \rho_{ij} \sigma_j \sigma_i \end{aligned}$$

Funzione caratteristica

- La funzione caratteristica è definita come:

$$C_{\xi}(p) \triangleq E\{e^{jp\xi}\} = \int_{-\infty}^{\infty} e^{jpx} f_{\xi}(x) dx$$

In questa espressione, p è un numero reale

È legata alla Trasformata di Fourier della densità di probabilità

- Dalla funzione caratteristica è possibile calcolare i momenti come:
$$\mu_k = j^{-k} C_{\xi}^{(k)}(0)$$

Derivata k -esima calcolata in zero
(si veda la pagina successiva per la dimostrazione)

Dimostrazione dell'ultima formula

$$\begin{aligned}C_{\xi}^{(k)}(p) &= \frac{d^k}{dp^k} \int_{-\infty}^{\infty} e^{jpx} f_{\xi}(x) dx \\&= \int_{-\infty}^{\infty} \frac{d^k}{dp^k} e^{jpx} f_{\xi}(x) dx \\&= \int_{-\infty}^{\infty} (jx)^k e^{jpx} f_{\xi}(x) dx \\&= j^k \int_{-\infty}^{\infty} x^k e^{jpx} f_{\xi}(x) dx\end{aligned}$$



$$C_{\xi}^{(k)}(0) = j^k \mu_k$$

Somma di variabili casuali

- Se $Z = X + Y$ e X e Y sono statisticamente indipendenti allora:

$$f_Z(z) = f_X(z) * f_Y(z)$$

Prodotto di convoluzione tra le
densità di probabilità

funzioni caratteristiche:

$$C_Z(p) = C_X(p) \cdot C_Y(p)$$

Dimostrazione

$$F_Z(z) \triangleq P(X + Y \leq z)$$

$$= \int P(X \leq z - y | Y = y) f_Y(y) dy$$

$$= \int P(X \leq z - y) f_Y(y) dy$$

$$= \int F_X(z - y) f_Y(y) dy$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) = \int \frac{d}{dz} F_X(z - y) f_Y(y) dy$$

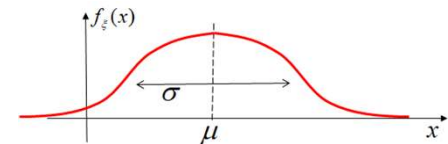
$$f_Z(z) = \int f_X(z - y) f_Y(y) dy$$

Variabile Gaussiana o “Normale”

- La distribuzione di probabilità Gaussiana riveste un enorme importanza in moltissime applicazioni, ed è definita come:

Perché la distribuzione Gaussiana è così rilevante? Teorema limite centrale, si veda qualche slide più avanti...

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$



Nota: una gaussiana ha due parametri liberi μ e σ

- La sua cumulativa è: $F_{\xi}(x) = 1 - \frac{1}{2} \operatorname{erfc}\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)$

Equivalentemente: $F_{\xi}(x) = 1 - Q\left(\frac{x-\mu}{\sigma}\right)$

- Dove si sono definite le due funzioni speciali:

Funzione di errore complementare $\operatorname{erfc}(x) \triangleq \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$

Funzione “Q” $Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$

Formule di conversione

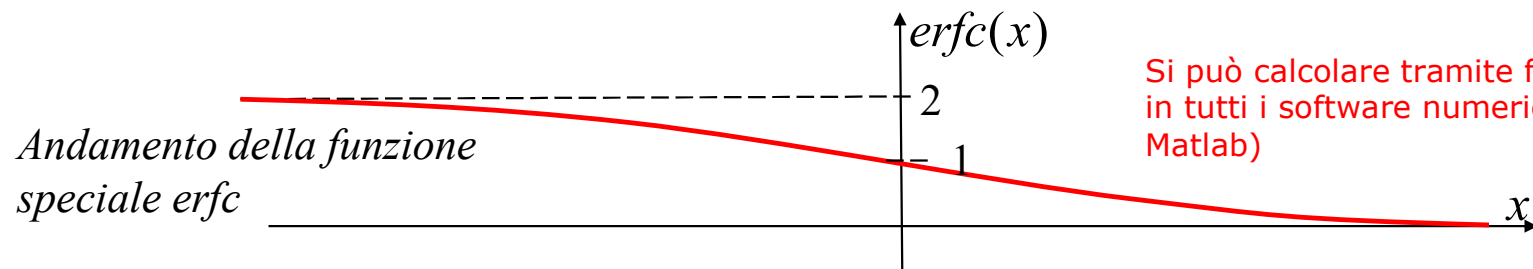
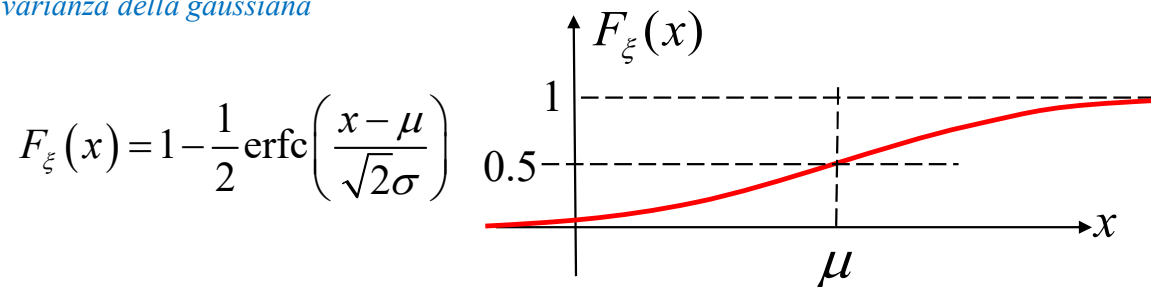
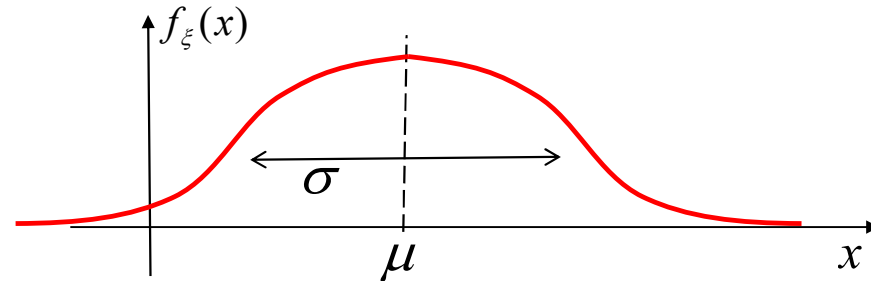
$$Q(x) \triangleq \frac{1}{2} \operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right)$$

$$\operatorname{erfc}(x) \triangleq 2Q(\sqrt{2}x)$$

Variabile Gaussiana o “Normale”

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

IMPORTANTE: in questa espressione della densità di probabilità, i due parametri liberi μ e σ^2 corrispondono direttamente alla media e alla varianza della gaussiana



Andamento della funzione speciale erfc

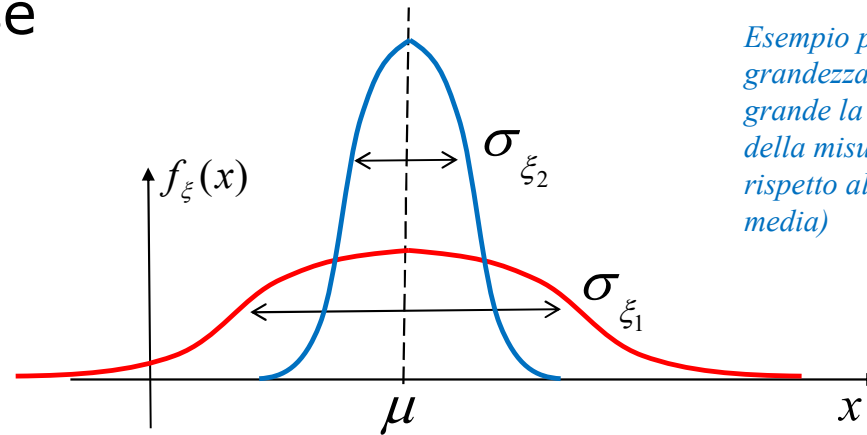
Si può calcolare tramite funzioni disponibili in tutti i software numerici (ad esempio Matlab)

Variabile Gaussiana o “Normale”

- Confronto tra due Gaussiane con la stessa media ma con varianze diverse

$$f_{\xi}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

IMPORTANTE: in questa espressione della densità di probabilità, i due parametri liberi μ e σ^2 corrispondono direttamente alla media e alla varianza della gaussiana



Esempio pratico misura sperimentale di una grandezza fisica tramite misure ripetute: più è grande la varianza, maggiore è l'incertezza della misura (da intendersi come "errore" rispetto al valore nominale, che coincide con la media)

$$media \triangleq \int_{-\infty}^{\infty} x \cdot f_{\xi}(x) dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \mu$$

$$varianza \triangleq \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f_{\xi}(x) dx = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx = \sigma^2$$

Questi due integrali NON sono per nulla ovvi...

Alcuni valori numerici "importanti" su variabile casuale gaussiana

□ Probabilità entro $\pm\sigma$ attorno alla media

In generale abbiamo visto che:

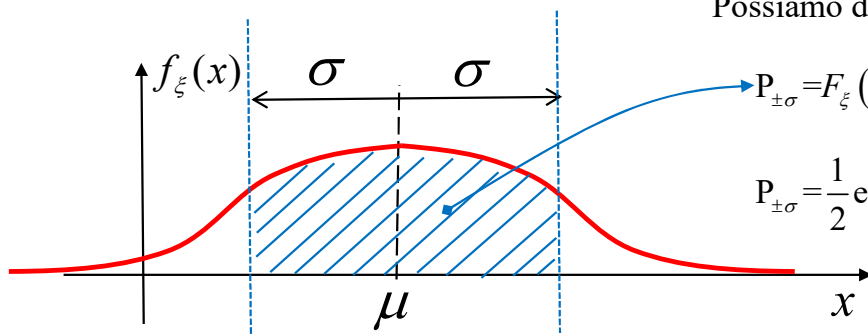
$$F_{\xi}(x) = 1 - \frac{1}{2} \operatorname{erfc}\left(\frac{x - \mu}{\sqrt{2}\sigma}\right)$$

Possiamo dunque calcolare la probabilità in questione come:

$$P_{\pm\sigma} = F_{\xi}(\mu + \sigma) - F_{\xi}(\mu - \sigma) = \left(1 - \frac{1}{2} \operatorname{erfc}\left(\frac{\sigma}{\sqrt{2}\sigma}\right)\right) - \left(1 - \frac{1}{2} \operatorname{erfc}\left(\frac{-\sigma}{\sqrt{2}\sigma}\right)\right)$$

$$P_{\pm\sigma} = \frac{1}{2} \operatorname{erfc}\left(\frac{-1}{\sqrt{2}}\right) - \frac{1}{2} \operatorname{erfc}\left(\frac{+1}{\sqrt{2}}\right) = 0.6827$$

$$\Rightarrow P_{\pm\sigma} \approx 68.2 \%$$



□ Analogamente si ottiene:

$$P_{\pm 2\sigma} \approx 95.4 \%$$

$$P_{\pm 3\sigma} \approx 99.7 \%$$

$$P_{\pm 4\sigma} \approx 99.99 \%$$

$$P_{\pm 5\sigma} \approx 99,99994\% \quad (1 - P_{\pm 5\sigma} \approx 5.7e-07)$$

*Provare a casa; la funzione
erfc() si deve calcolare in
Matlab*

https://it.wikipedia.org/wiki/Regola_68-95-99,7

Funzione caratteristica di una variabile casuale Gaussiana

□ Richiami dalle tavole delle trasformate di Fourier

$$\mathcal{F}\left\{e^{-\pi t^2}\right\} = e^{-\pi f^2}$$

$$\begin{aligned}\mathcal{F}\left\{e^{-t^2/2}\right\} &= \sqrt{2\pi}e^{-(2\pi f)^2/2} \\ \mathcal{F}\left\{e^{-\left[\frac{t}{\sigma}\right]^2/2}\right\} &= \sigma\sqrt{2\pi}e^{-(2\pi f\sigma)^2/2} \quad \sigma_f^2 = \left(\frac{1}{2\pi\sigma_t}\right)^2 \\ \mathcal{F}\left\{e^{-\left[\frac{t-\mu}{\sigma}\right]^2/2}\right\} &= \sigma\sqrt{2\pi}e^{-(2\pi f\sigma)^2/2}e^{-j2\pi f\mu}\end{aligned}$$

□ Funzione caratteristica di una Gaussiana

$$\mathcal{F}\left\{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\left[\frac{t-\mu}{\sigma}\right]^2/2}\right\} = e^{-(2\pi f\sigma)^2/2}e^{-j2\pi f\mu}$$

Proprietà variabili gaussiane

- Qualsiasi combinazione lineare di variabili casuali gaussiane è ancora una gaussiana
 - Anche in presenza di dipendenza statistica
 - In particolare:
 - La somma di due v.c. Gaussiane è una Gaussiana
 - Il prodotto di convoluzione di due distribuzioni gaussiane è una distribuzione gaussiana
- Due v.c. gaussiane scorrelate sono anche statisticamente indipendenti
- La funzione caratteristica di una Gaussiana è una Gaussiana

Variabile gaussiana

- Per una variabile gaussiana a valor medio nullo i momenti valgono:

$$m_k = \begin{cases} (k-1)!!\sigma^k & k \text{ pari} \\ 0 & k \text{ dispari} \end{cases}$$

Il “doppio fattoriale” è definito come:

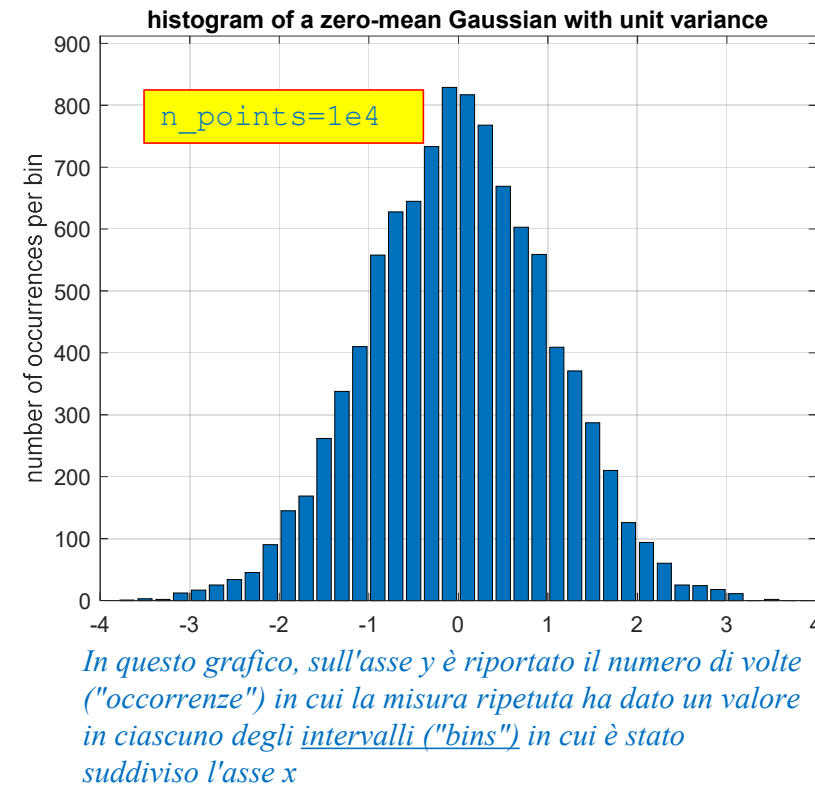
- per k dispari: $k!! = 1 \cdot 3 \cdot 5 \cdots k$
- per k pari: $k!! = 2 \cdot 4 \cdot 6 \cdots k$

Abbiamo dunque:

$$\begin{aligned} m_2 &= (2-1)!!\sigma^2 = 1!!\sigma^2 = \sigma^2 \\ m_4 &= (4-1)!!\sigma^4 = 3!!\sigma^4 = 3 \cdot 1 \cdot \sigma^4 = 3\sigma^4 \\ m_6 &= (6-1)!!\sigma^6 = 5!!\sigma^6 = 5 \cdot 3 \cdot 1 \cdot \sigma^6 = 15\sigma^6 \end{aligned}$$

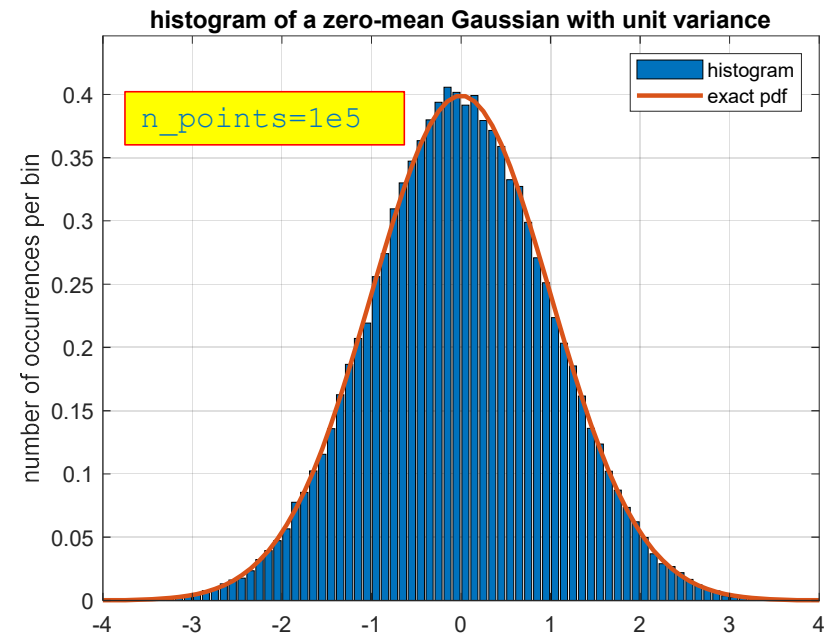
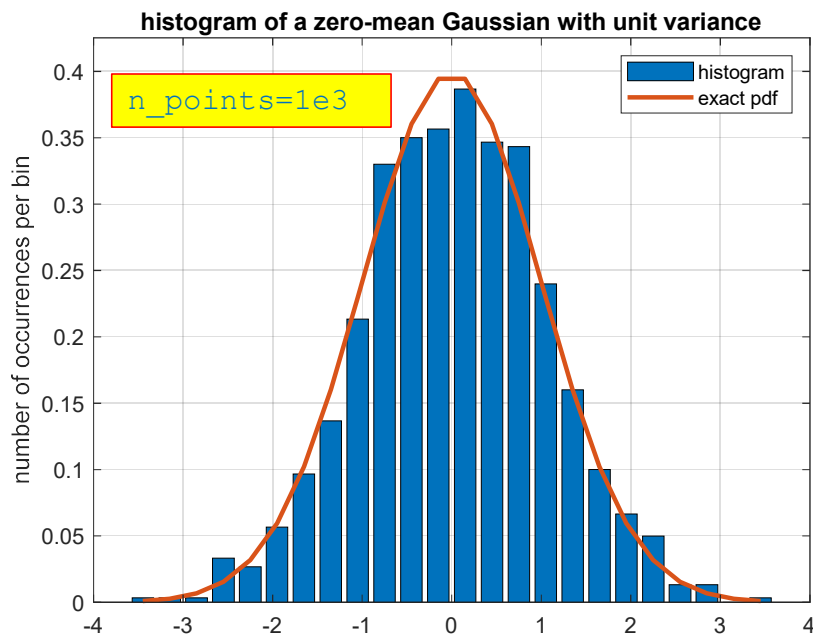
Densità di probabilità e istogrammi

- Come si possono stimare sperimentalmente le densità di probabilità?
- Se si effettua una stima per misure ripetute di una certa grandezza fisica, la densità di probabilità è stimabile con l'istogramma delle occorrenze nelle misure ripetute
 - Esempio pratico (ottenuto per simulazione in Matlab)
 - Supponiamo che una certa grandezza fisica abbia una distribuzione Gaussiana
 - Ne otteniamo i relativi istogrammi



Densità di probabilità e istogrammi

- Se si normalizza l'istogramma in modo che abbia "area unitaria" si ottiene una stima della d.d.p
 - La stima è tanto migliore quanto più è elevato
 - il numero di prove ripetute
 - Il numero di intervalli utilizzati



Densità di probabilità e istogrammi

Nella pratica, si ha solitamente il seguente problema:

- dato un istogramma sperimentale, si cerca di approssimarlo con una pdf nota che approssimi meglio possibile l'istogramma

- **Esempio: aspettativa di vita**

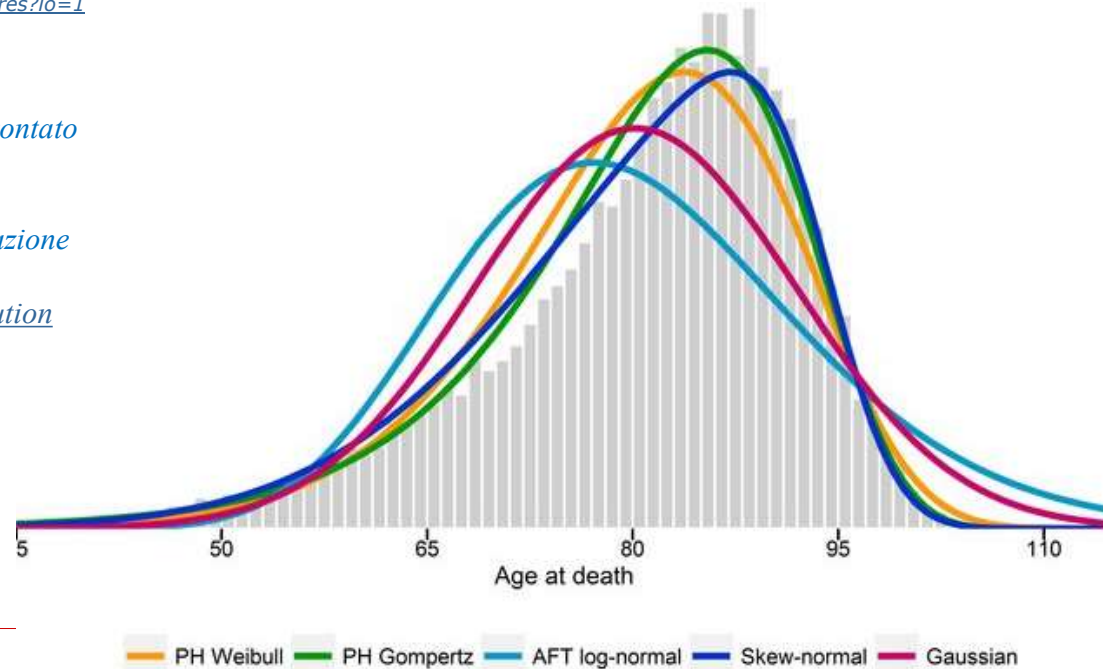
- https://www.researchgate.net/publication/281513046_Modeling_absolute_differences_in_life_expectancy_with_a_censored_skew-normal_regression_approach/figures?lo=1

In questo studio, l'istogramma sperimentale è confrontato con alcune d.d.p. presenti in letteratura.

Nel caso specifico, si ottiene una buona approssimazione con la d.d.p. denominata "skew normal"

https://en.wikipedia.org/wiki/Skew_normal_distribution che ha la seguente d.d.p.

$$\frac{2}{\omega\sqrt{2\pi}} e^{-\frac{(x-\xi)^2}{2\omega^2}} \int_{-\infty}^{\alpha\left(\frac{x-\xi}{\omega}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



Teorema limite centrale

- combinando linearmente un numero elevato di **v.c. indipendenti con qualunque distribuzione** la v.c. risultante tende ad avere una distribuzione gaussiana con i seguenti parametri:

- Media:

$$\mu_{Z_N} = \sum_{i=1}^N \alpha_i \mu_i$$

- Varianza:

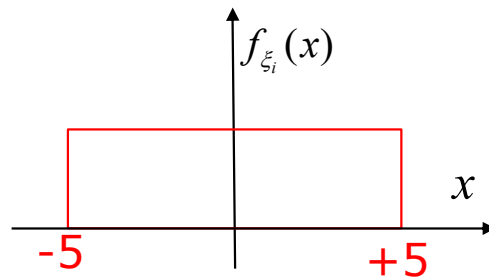
$$\sigma_{Z_N}^2 = \sum_{i=1}^N \alpha_i^2 \sigma_i^2$$

$$Z_N = \sum_{i=1}^N \alpha_i \cdot \xi_i$$

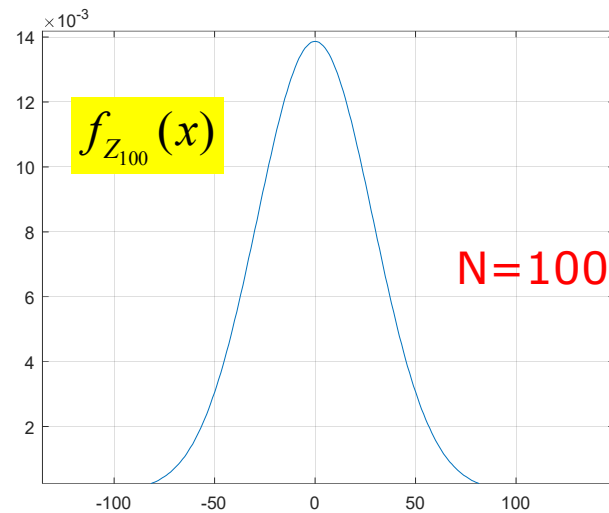
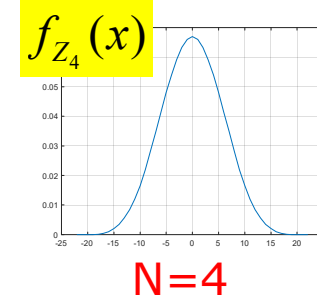
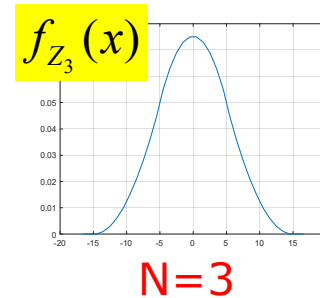
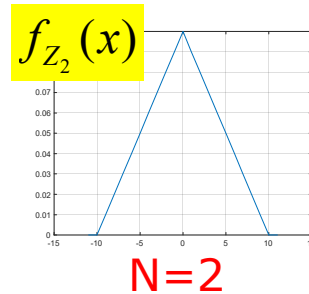
In pratica: la somma di N variabili casuali qualunque, a patto che siano indipendenti tra di loro, tende ad una Gaussiana per N elevati.

- Questo teorema giustifica il motivo per cui moltissimi fenomeni aleatori (fisici, economici, sociali) sono modellizzabili con ottima approssimazione come v.c. gaussiane

Esempio: somma di N variabili casuali uniformi e statisticamente indipendenti



$$Z_N = \frac{1}{N} \sum_{i=1}^N \xi_i$$



Esempio: somma di N variabili statisticamente indipendenti e con densità di probabilità generica

