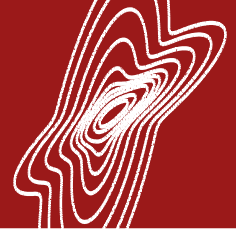


METODI STATISTICI PER LA BIOINGEGNERIA (B)

**PARTE 2: RIPASSO DI PROBABILITA' E
VARIABILI ALEATORIE**

A.A. 2025-2026

Prof. Martina Vettoretti



OUTLINE



➤ Introduzione alla probabilità

➤ Probabilità di eventi

➤ Variabili aleatorie discrete e continue

➤ Vettori di variabili aleatorie

➤ Valore atteso, varianza e deviazione standard

➤ Covarianza e correlazione

LA PROBABILITA' NEL LINGUAGGIO COMUNE



- E' probabile che fra poco piova.
- Con questo titolo di studio vi sono buone probabilità di trovare lavoro.
- E' probabile che l'incendio sia d'origine dolosa.
- Ho poche probabilità di vincere la lotteria.

Utilizziamo frequentemente il termine **probabilità** quando ci riferiamo a situazioni con **esito incerto**, a fenomeni che possono o no verificarsi.

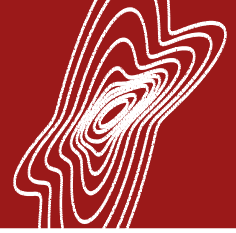
PERCHE' PARLIAMO DI PROBABILITA' ?



- Incontriamo spesso in natura fenomeni che sembrano avere caratteristiche di **casualità**.
- Spesso questa casualità è solo **apparente** e potrebbe essere dovuta a una serie di fattori che, pur essendo deterministici, possono non essere completamente noti o avere una **spiegazione troppo complessa**.
- Esempio: L'esito del lancio di una moneta è prevedibile o casuale?

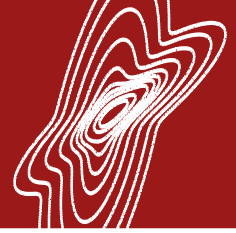


- **Interpretazione frequentista:** la probabilità di un esito è considerata una proprietà dell'esito stesso che può essere determinata operativamente ripetendo più volte l'esperimento e calcolando il **rapporto tra i casi in cui si è verificato l'esito e i casi totali**.
- **Esempio.** Un bioingegnere afferma che vi è il 90% di probabilità che l'impianto di un organo artificiale avvenga con successo senza complicazioni.
 - Interpretazione frequentista: il bioingegnere crede che osservando l'esito di molti impianti dell'organo artificiale, in situazioni simili, circa nel 90% dei casi l'impianto avverrà con successo senza complicazioni.



- **Interpretazione soggettivistica:** la probabilità di un esito è vista come il **livello di fiducia** che lo studioso ripone nel verificarsi dell'esito.

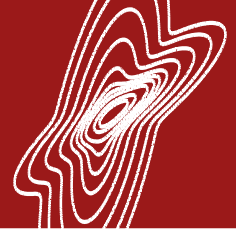
- **Esempio.** Un bioingegnere afferma che vi è il 90% di probabilità che l'impianto di un organo artificiale avvenga con successo senza complicazioni.
 - Interpretazione soggettivistica: il bioingegnere crede che sia più verosimile che l'impianto avvenga con successo senza complicazioni e 0.9 rappresenta, su una scala tra 0 e 1, il suo livello di fiducia nell'ipotesi che si verifichi questo esito.



OUTLINE



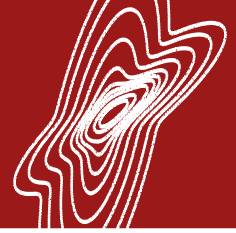
- Introduzione alla probabilità
- Probabilità di eventi
- Variabili aleatorie discrete e continue
- Vettori di variabili aleatorie
- Valore atteso, varianza e deviazione standard
- Covarianza e correlazione



SAMPLE SPACE ED EVENTI



- Consideriamo un esperimento il cui esito non sia prevedibile con certezza.
- **Sample space**, o spazio campionario, Ω : l'insieme di tutti i possibili esiti.
- **Evento E**: un qualsiasi sottoinsieme dello spazio degli esiti.
- Esempio. Consideriamo una gara tra 5 cavalli denotati dai numeri 1,2,3,4,5.
 - $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} = \{\text{tutti gli ordinamenti di } (1,2,3,4,5)\}$
 - $E = \{\text{tutti gli esiti in } \Omega \text{ che cominciano per } 3\}$

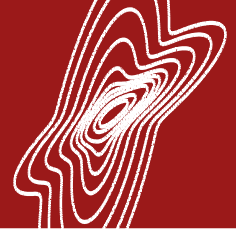


OPERAZIONI TRA EVENTI



- $E \cup F$: **unione** degli eventi E ed F , insieme degli esiti che stanno in E o in F o in tutti e due.
- $E \cap F$: **intersezione** di E ed F , insieme degli esiti che stanno sia in E che in F .
- E^c : **complementare** di E , insieme degli esiti che stanno in Ω ma non in E .
- \emptyset : **evento vuoto** che non contiene esiti.

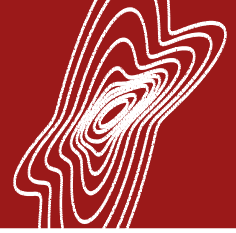
- $E \cap F = \emptyset \rightarrow E$ ed F sono **disgiunti** o mutuamente esclusivi
- $E \subset F \rightarrow$ tutti gli esiti di E sono anche in F .



PROBABILITA' DI UN EVENTO



- Se ripetiamo più volte un esperimento nelle stesse condizioni, si verifica empiricamente che la frazione di casi sul totale in cui si realizza un evento E tende, al crescere dei tentativi, ad un valore costante che dipende solo da E .
 - Questo valore è la probabilità di E .
- Ad ogni evento E sullo spazio Ω associamo un numero $P(E)$ che rappresenta la probabilità di E .



ASSIOMI DI KOLMOGOROV



➤ **Assioma 1:**

$$0 \leq P(E) \leq 1, \quad P(\emptyset) = 0$$

➤ **Assioma 2:**

$$P(\Omega) = 1$$

➤ **Assioma 3:**

Per ogni successione di eventi mutuamente esclusivi E_1, E_2, \dots, E_n :

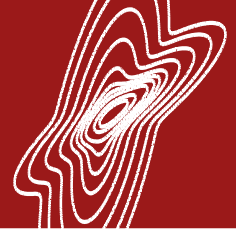
$$P(\cup_{i=1}^n E_i) = \sum_{i=1}^n P(E_i)$$

➤ **Evento certo:** evento con probabilità 1

- Esperimento: estrazione di una pallina da un'urna contenente solo palline rosse. Consideriamo come possibili esiti il colore della pallina estratta.
 - $\Omega = \{r\}$
 - $E =$ la pallina estratta è rossa. $\rightarrow P(E) = P(\{r\}) = P(\Omega) = 1$, evento certo.

➤ **Evento impossibile:** evento con probabilità 0

- Esperimento: estrazione di una pallina da un'urna contenente solo palline rosse. Consideriamo come possibili esiti il colore della pallina estratta.
 - $\Omega = \{r\}$
 - $E =$ la pallina estratta non è rossa. $\rightarrow P(E) = P(\Omega - \{r\}) = P(\emptyset) = 0$, evento impossibile.

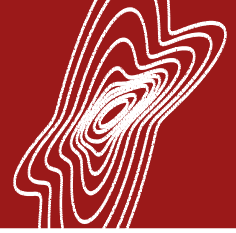


PROPOSIZIONI



➤ $P(E^c) = 1 - P(E)$

➤ $P(E \cup F) = P(E) + P(F) - P(E \cap F)$



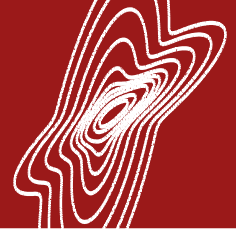
ODDS



- Gli **odds** di un evento E sono per definizione:

$$\frac{P(E)}{P(E^c)} = \frac{P(E)}{1 - P(E)}$$

- Gli odds ci dicono quanto è più facile che E si realizzi piuttosto che non lo faccia.
- Esempio: $P(E) = 3/4$
 - Quanto valgono gli odds di E ?
 - Come si interpreta questo numero?



SPAZI DI ESITI EQUIPROBABILI



- Spesso possiamo assumere che tutti i possibili esiti abbiano la stessa probabilità di realizzarsi (es. lancio di un dado non truccato).
- Esempio: Lancio di un dado non truccato. Quanto vale la probabilità che esca un numero pari?
- Supponiamo ci siano N possibili esiti equiprobabili, con probabilità p :

$$\Omega = \{1, 2, \dots, N\}$$

$$P(\{1\}) = P(\{2\}) = \dots = P(\{N\}) = p$$

→ Per gli assiomi 2 e 3, $p = 1/N$.

- Pertanto la probabilità di un evento E composto da un numero di esiti pari a $\#E$ si può calcolare come:

$$P(E) = \#E/N$$

ESEMPIO



Lancio di due dadi non truccati. Evento $E = \{\text{tutti gli esiti per cui la somma dei risultati dei due dadi è pari a } 8\}$. Quanto vale la probabilità di E ?

➤ Caso 1: non abbiamo altre informazioni.

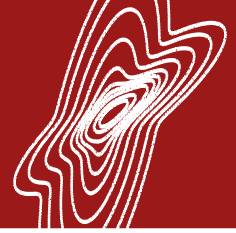
- Esiti che portano a somma pari a 8: $(2,6), (3,5), (4,4), (5,3), (6,2)$.
- Tutti gli esiti possibili sono $6 \times 6 = 36$.

➔ La probabilità di E è: $5/36$ ➔ **$P(E)$**

➤ Caso 2: sappiamo che il lancio del primo dado ha dato come esito 3 (evento F), come si modifica la probabilità di E ?

- L'esito del primo lancio è noto \rightarrow i possibili esiti diventano 6.
- E si può verificare solo quando il secondo lancio dà come risultato 5.

➔ La probabilità di E sapendo che si è verificato F è $1/6$
↳ **Probabilità condizionata di E dato F : $P(E | F)$**



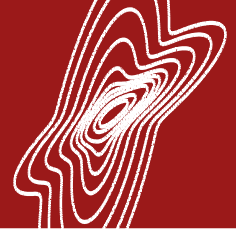
PROBABILITA' CONDIZIONATA



- La **probabilità condizionata di E dato F** rappresenta la probabilità dell'evento E, sapendo che si è verificato l'evento F, ed è definita come:

$$P(E|F) := P(E \cap F) / P(F)$$

$$\rightarrow P(E \cap F) = P(E|F) \cdot P(F)$$



FORMULA DI FATTORIZZAZIONE DEGLI EVENTI



Consideriamo N eventi, F_1, F_2, \dots, F_N , **mutuamente esclusivi** tali che:

$$\Omega = \bigcup_{i=1}^N F_i$$

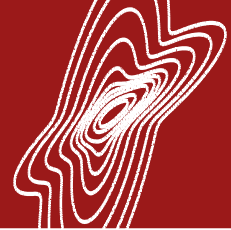
Un qualsiasi evento E si può scrivere come:

$$E = \bigcup_{i=1}^N E \cap F_i$$

Usando l'assioma 3 e la definizione di probabilità condizionata possiamo ricavare la seguente **formula di fattorizzazione degli eventi**:

$$P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + \dots + P(E|F_N)P(F_N)$$

→ A volte risulta più semplice calcolare $P(E)$ condizionando E rispetto a degli eventi mutuamente esclusivi e ricorrendo a questa formula.



FORMULA DI BAYES

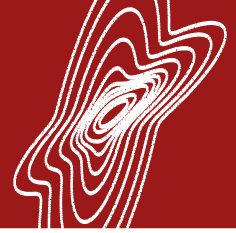
Ipotizziamo ora di conoscere che E si sia verificato. Quanto valgono le probabilità condizionate di F_j dato E ?

$$P(F_j|E) = \frac{P(F_j \cap E)}{P(E)}$$

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^N (P(E|F_i) \cdot P(F_i))}$$



**Formula
di Bayes**



EVENTI INDIPENDENTI



Due eventi E ed F si dicono **indipendenti** se vale:

$$P(E \cap F) = P(E) \cdot P(F)$$

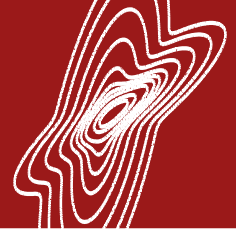
ovvero:

$$P(E|F) = P(E) \text{ e } P(F|E) = P(F)$$

altrimenti si dicono dipendenti.

Più in generale, E_1, E_2, \dots, E_n si dicono indipendenti se:

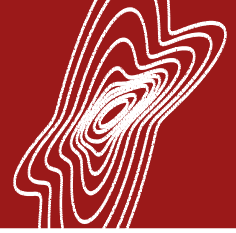
$$P\left(\bigcap_{i=1}^n E_i\right) = \prod_{i=1}^n P(E_i)$$



OUTLINE



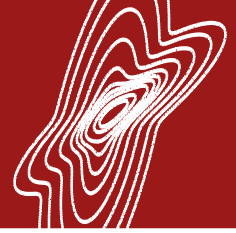
- Introduzione alla probabilità
- Probabilità di eventi
- Variabili aleatorie discrete e continue
- Vettori di variabili aleatorie
- Valore atteso, varianza e deviazione standard
- Covarianza e correlazione



VARIABILI ALEATORIE



- Quando si realizza un esperimento non sempre si è interessati a tutte le informazioni ricavabili dal suo esito. Spesso siamo interessati ad una singola quantità ricavabile dall'esito stesso.
 - Esempio. Nel lancio di due dadi potremmo essere interessati solo alla somma dei due punteggi dei dadi e non a ciascuno dei punteggi.
- **Variabile aleatoria:** una quantità di interesse determinata dal risultato di un esperimento casuale.
- Possiamo dunque attribuire una probabilità a ciascun possibile valore della variabile aleatoria.



DEFINIZIONE FORMALE DI VARIABILE ALEATORIA



- Una **variabile aleatoria**, o **variabile casuale**, X , è una funzione matematica avente come dominio uno spazio campionario Ω e come codominio \mathbb{R} .

$$X : \Omega \rightarrow \mathbb{R}$$

- In altre parole, la variabile aleatoria associa ad ogni esito in Ω uno ed un solo numero reale.
- Il valore di una variabile aleatoria è determinato dall'esito di un esperimento casuale \rightarrow possiamo attribuire delle probabilità ai possibili valori della variabile aleatoria.

ESEMPIO



Un tizio acquista due componenti elettronici, ciascuno dei quali può essere accettabile (a) o difettoso (d).

Possibili esiti: $\Omega = \{(d,d),(d,a),(a,d),(a,a)\}$.

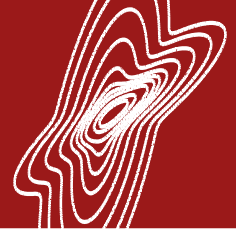
Immaginiamo che i 4 esiti abbiano queste probabilità:

- $P(\{(d,d)\}) = 0.09$
- $P(\{(d,a)\}) = 0.21$
- $P(\{(a,d)\}) = 0.21$
- $P(\{(a,a)\}) = 0.49$

Consideriamo la variabile aleatoria X pari al numero di componenti accettabili.

X può assumere i valori 0, 1 o 2, con le seguenti probabilità:

- $P(X = 0) = P(\{(d,d)\}) = 0.09$
- $P(X = 1) = P(\{(a,d)\} \cup \{(d,a)\}) = 0.42$
- $P(X = 2) = P(\{(a,a)\}) = 0.49$



FUNZIONE DI RIPARTIZIONE



Funzione di ripartizione o di distribuzione, F_X , di una variabile aleatoria X :

$$F_X(x) := P(X \leq x), \quad F_X: \mathbb{R} \rightarrow [0, 1]$$

La funzione di ripartizione è monotona crescente nell'intervallo $[0, 1]$.

$$\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow +\infty} F_X(x) = 1$$

VARIABILE ALEATORIA DISCRETA



➤ Una **variabile aleatoria** si dice **discreta** se può assumere solo un numero finito o infinito numerabile di valori possibili $\{x_1, x_2, \dots\}$.

➤ Se X è una variabile aleatoria discreta, si dice **funzione di massa di probabilità** o **densità discreta** la funzione:

$$p_X(a) := P(X = a), \quad p_X: \mathbb{R} \rightarrow [0, 1]$$

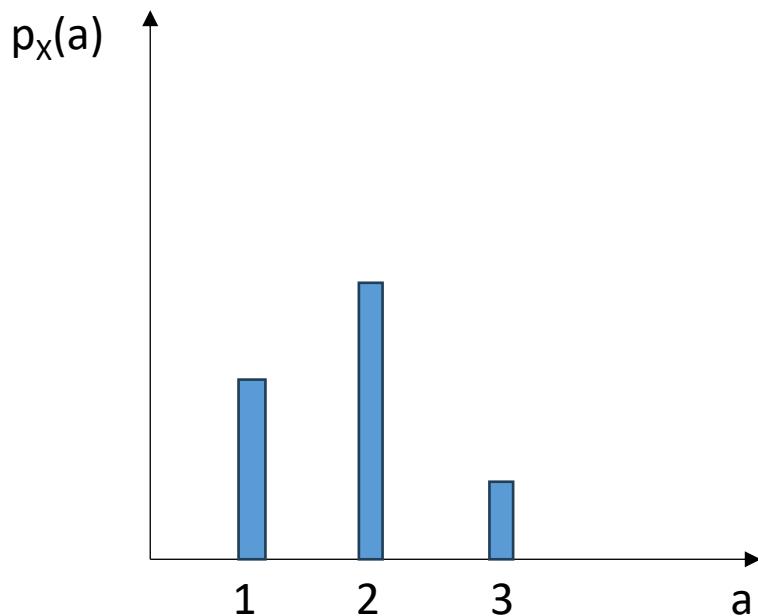
➤ Funzione di ripartizione di una variabile aleatoria discreta X :

$$F(a) = \sum_{x \leq a} p_X(x)$$

FUNZIONI DI MASSA E DI RIPARTIZIONE DI UNA VARIABILE ALEATORIA DISCRETA

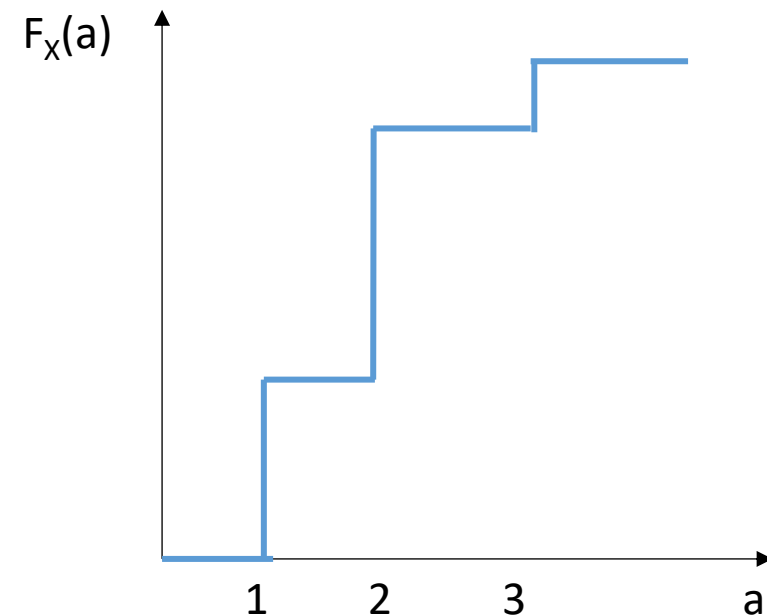


Funzione di massa

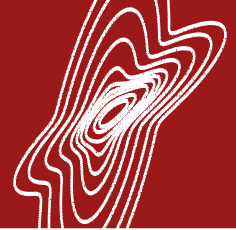


Assume valori positivi per un numero finito o infinito numerabile di valori di a . La somma di questi valori deve dare 1.

Funzione di ripartizione



Risulta una funzione a gradini.



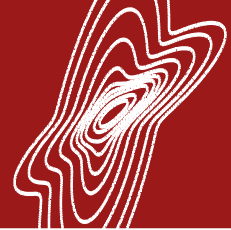
VARIABILE ALEATORIA CONTINUA



- Una **variabile aleatoria continua** può assumere un insieme continuo di valori possibili, ad es. un intervallo di numeri reali.
- Una variabile aleatoria X si dice continua se esiste una funzione f_X non negativa, definita su tutto \mathbb{R} , ed integrabile, tale che per ogni insieme B di numeri reali:

$$P(X \in B) = \int_B f_X(x) dx \quad f_X: \mathbb{R} \rightarrow [0, +\infty)$$

- f_X si chiama **funzione di densità di probabilità (ddp)**, o **densità**, di X .

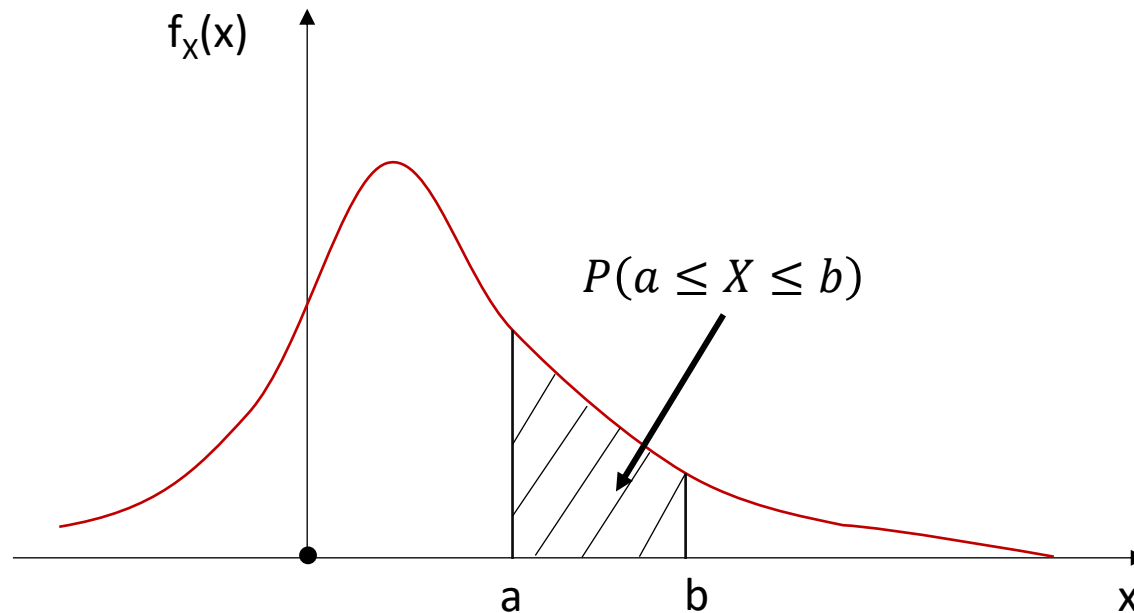


PROPRIETA' DELLA DDP

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

$$P(X = a) = \int_a^a f_X(x) dx = 0$$



Dato un valore ε sufficientemente piccolo:

$$P\left(a - \frac{\varepsilon}{2} \leq X \leq a + \frac{\varepsilon}{2}\right) = \int_{a - \frac{\varepsilon}{2}}^{a + \frac{\varepsilon}{2}} f_X(x) dx \approx \varepsilon \cdot f_X(a) \longrightarrow$$

$f_X(a)$ è proporzionale alla probabilità di un intervallo sufficientemente piccolo intorno ad $a \rightarrow$ indicazione di quanto è probabile che X sia vicino ad a .

PROBABILITA' VS. DENSITA' DI PROBABILITA'



- Il valore della funzione di massa di una variabile aleatoria discreta X per un certo valore a rappresenta la probabilità che X assuma il valore a :

$$p_X(a) = P(X = a)$$

- La precedente relazione non vale per la ddp di una variabile aleatoria continua.
 - La ddp di una variabile aleatoria continua NON è una probabilità!
 - Solo il suo integrale su un intervallo ha il significato di probabilità.

$$f_X(a) \neq P(X = a)$$

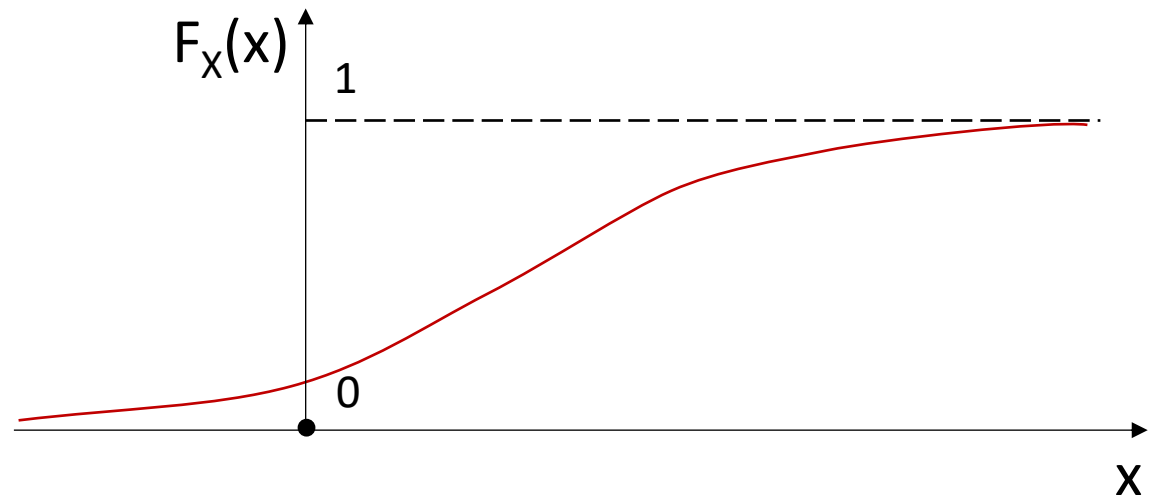
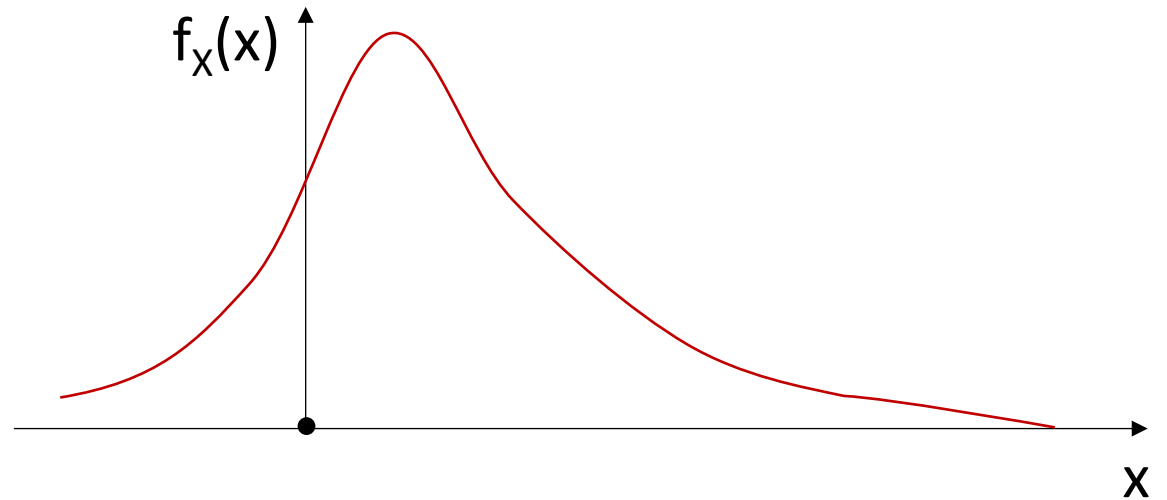
$$P(a \leq X \leq b) = \int_a^b f_X(x) dx$$

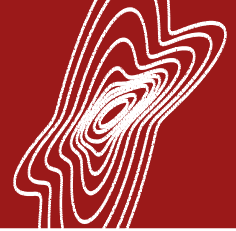
RELAZIONE TRA DENSITA' DI PROBABILITA' E FUNZIONE DI RIPARTIZIONE



$$F_X(a) = P(X \in (-\infty, a]) = \int_{-\infty}^a f_X(x) dx$$

$$\rightarrow \frac{d}{da} F_X(a) = f_X(a)$$





OSSERVAZIONE

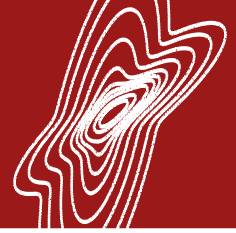


Quando conosciamo la funzione di massa di probabilità di una variabile aleatoria discreta, o la ddp di una variabile aleatoria continua, oppure quando conosciamo la funzione di ripartizione di una variabile aleatoria qualsiasi, abbiamo sufficienti informazioni per poter calcolare la probabilità di un qualsiasi evento descritto da quella variabile aleatoria.

→ Si dice che conosciamo la **distribuzione** o **legge** della variabile aleatoria.

Se due variabili aleatorie X e Y hanno la stessa funzione di ripartizione, significa che per qualsiasi insieme di numeri reali $A \subset \mathbb{R}$:

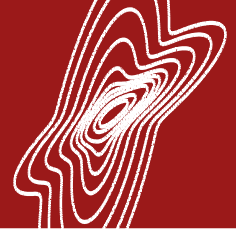
$$P(X \in A) = P(Y \in A)$$



OUTLINE



- Introduzione alla probabilità
- Probabilità di eventi
- Variabili aleatorie discrete e continue
- **Vettori di variabili aleatorie**
- Valore atteso, varianza e deviazione standard
- Covarianza e correlazione



COPPIE E VETTORI DI VARIABILI ALEATORIE



In alcuni esperimenti l'oggetto di interesse sono le relazioni presenti tra due o più quantità numeriche che possiamo rappresentare con variabili aleatorie.

- **Esempio.** In uno studio sulle cause di tumore vogliamo indagare la relazione tra il numero di sigarette fumate al giorno e l'età in cui viene diagnosticato il tumore.

Estendiamo i concetti visti finora alle coppie e poi, più in generale, ai vettori di variabili aleatorie.

FUNZIONE DI RIPARTIZIONE CONGIUNTA



X e Y : due variabili aleatorie che riguardano lo stesso esperimento casuale.

Funzione di ripartizione congiunta di X e Y :

$$F_{X,Y}(x, y) := P(X \leq x, Y \leq y)$$

Essa rappresenta la probabilità che $X \leq x$ e contemporaneamente $Y \leq y$.

$F_{X,Y}(x, y)$ consente di calcolare la probabilità di tutti gli eventi descritti singolarmente o congiuntamente da X e da Y .

➤ $F_X(x) = F_{X,Y}(x, \infty)$

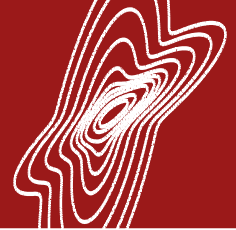
➤ $F_Y(y) = F_{X,Y}(\infty, y)$



X e Y : due variabili aleatorie discrete che assumono i valori x_1, x_2, \dots , e y_1, y_2, \dots , rispettivamente.

Funzione di massa di probabilità congiunta o densità discreta congiunta:

$$p_{X,Y}(x_i, y_j) := P(X = x_i, Y = y_j), \quad i = 1, 2, \dots, \quad j = 1, 2, \dots$$



FUNZIONI DI MASSA MARGINALI



Da $p_{X,Y}(x_i, y_j)$ è possibile ricavare le funzioni di massa di probabilità di X e di Y \rightarrow **funzioni di massa marginali**

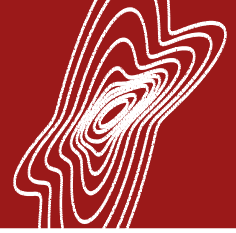
Per l'assioma 3 \rightarrow

$$p_X(x_i) = P(X = x_i) = P\left(\bigcup_j \{X = x_i, Y = y_j\}\right)$$
$$= \sum_j P(X = x_i, Y = y_j) = \sum_j p_{X,Y}(x_i, y_j)$$

Analogamente:

$$p_Y(y_j) = \sum_i p_{X,Y}(x_i, y_j)$$

Attenzione! Non è vero il viceversa: conoscere $p_X(x_i)$ e $p_Y(y_j)$ non consente di conoscere $p_{X,Y}(x_i, y_j)$.



ESERCIZIO (1 / 2)

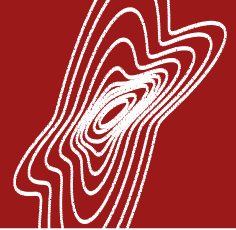


In una certa popolazione:

- il 15% delle coppie non ha figli
- il 20% delle coppie ha 1 figlio
- il 35% ne ha 2
- il 30% ne ha 3.

Ogni bambino ha pari probabilità di essere maschio o femmina, indipendentemente dagli altri.

Se selezioniamo una famiglia a caso e indichiamo con X il numero di figlie femmine e Y il numero di figli maschi, **qual è la funzione di massa congiunta delle variabili aleatorie X e Y ?**



ESERCIZIO (2/2)



La funzione di massa congiunta si può rappresentare con questa tabella:

$p_{X,Y}(X=i,Y=j)$		j				Totali per righe $p_X(X=i)$
		0	1	2	3	
i	0	0,15	0,1	0,0875	0,0375	0,375
	1	0,1	0,175	0,1125	0	0,3875
	2	0,0875	0,1125	0	0	0,2
	3	0,0375	0	0	0	0,0375
Totali per colonne $p_Y(Y=j)$		0,375	0,3875	0,2	0,0375	

p_Y e p_X si chiamano **marginali** proprio perché possono essere calcolate a margine della tabella della funzione di massa congiunta.

DENSITA' DI PROBABILITA' CONGIUNTA



Due variabili aleatorie X e Y si dicono **congiuntamente continue** se esiste una funzione non negativa $f_{X,Y}(x,y)$ tale che per ogni sottoinsieme C del piano cartesiano si ha che:

$$P((X, Y) \in C) = \iint_{(x,y) \in C} f_{X,Y}(x, y) dx dy$$

$f_{X,Y}(x,y)$: **densità di probabilità congiunta** di X e Y .

Se C è il prodotto cartesiano di A e B :

$$P(X \in A, Y \in B) = \int_B \int_A f_{X,Y}(x, y) dx dy$$

Il valore $f_{X,Y}(a,b)$ indica quanto è probabile che (X,Y) cada vicino ad (a,b) .

LEGAME CON LA FUNZIONE DI RIPARTIZIONE, DENSITA' MARGINALI



$$F_{X,Y}(a, b) = P(X \leq a, Y \leq b) = \int_{-\infty}^b \int_{-\infty}^a f_{X,Y}(x, y) dx dy$$

$$f_{X,Y}(a, b) = \frac{\partial^2 F_{X,Y}(a, b)}{\partial a \partial b}$$

Densità di probabilità marginali:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dy$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx$$

VARIABILI ALEATORIE INDIPENDENTI



Due variabili aleatorie X e Y che riguardano lo stesso esperimento si dicono **indipendenti** se per ogni coppia di insiemi di numeri reali A e B si ha che:

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$$

viceversa si dicono dipendenti.

Se X e Y sono indipendenti allora per ogni x e y si ha che:

- $F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$
- $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$, se X e Y sono variabili aleatorie discrete
- $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$, se X e Y sono variabili aleatorie continue

Nella sostanza, due variabili aleatorie sono indipendenti se conoscere il valore di una variabile non influenza il valore dell'altra.

GENERALIZZAZIONE AL CASO DEI VETTORI ALEATORI (1/2)



Vettore aleatorio: n variabili aleatorie riguardanti uno stesso esperimento.

Le definizioni viste per le coppie di variabili aleatorie si estendono facilmente ai vettori aleatori.

➤ Funzione di ripartizione congiunta di n variabili aleatorie X_1, X_2, \dots, X_n :

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) := P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

➤ Funzione di massa congiunta di n variabili aleatorie **discrete** X_1, X_2, \dots, X_n :

$$p_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) := P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$$

GENERALIZZAZIONE AL CASO DEI VETTORI ALEATORI (2/2)



- Funzione di densità di probabilità congiunta di n variabili aleatorie continue X_1, X_2, \dots, X_n :

$$\begin{aligned} &P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) \\ &= \int_{A_n} \dots \int_{A_2} \int_{A_1} f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned}$$

- X_1, X_2, \dots, X_n sono indipendenti se:

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n) = \prod_{i=1}^n P(X_i \in A_i)$$

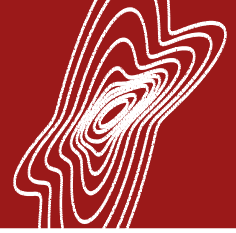


Le relazioni tra due variabili aleatorie possono essere studiate mediante la **distribuzione condizionale** di una delle due dato il valore dell'altra.

X e Y : due variabili aleatorie discrete con funzione di massa congiunta $p_{X,Y}(x,y)$. **Funzione di massa di probabilità condizionata:**

$$p_{X|Y}(x|y) := P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p_{X,Y}(x, y)}{p_Y(y)}$$

$$\forall x, \forall y, \text{ con } p_Y(y) > 0$$



ESERCIZIO (1 / 2)



In una certa popolazione:

- il 15% delle coppie non ha figli
- il 20% delle coppie ha 1 figlio
- il 35% ne ha 2
- il 30% ne ha 3.

Ogni bambino ha pari probabilità di essere maschio o femmina,
indipendentemente dagli altri.

Selezioniamo una famiglia a caso. Ipotizziamo di sapere che la famiglia selezionata ha una e una sola figlia femmina ($X=1$). Quanto vale la funzione di massa condizionata del numero di maschi nella famiglia selezionata, ovvero $p_{Y|X}(y | 1)$?

ESERCIZIO 2/2



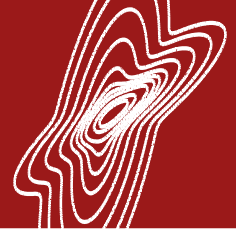
➤ $p_{Y|X}(0 | 1) = \dots$

➤ $p_{Y|X}(1 | 1) = \dots$

➤ $p_{Y|X}(2 | 1) = \dots$

➤ $p_{Y|X}(3 | 1) = \dots$

$p_{X,Y}(X=i,Y=j)$		j				Totali per righe $p_X(X=i)$
		0	1	2	3	
i	0	0.1500	0.1000	0.0875	0.0375	0.3750
	1	0.1000	0.1750	0.1125	0	0.3875
	2	0.0875	0.1125	0	0	0.2000
	3	0.0375	0	0	0	0.0375
Totali per colonne $p_Y(Y=j)$		0.3750	0.3875	0.2000	0.0375	



DENSITA' CONDIZIONALE



X e Y: due variabili aleatorie continue con densità di probabilità congiunta $f_{X,Y}(x,y)$.

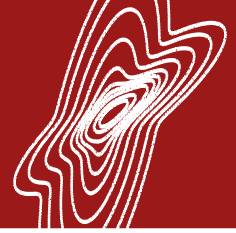
Densità condizionale di X rispetto a Y:

$$f_{X|Y}(x|y) := \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad \forall x, \forall y: f_Y(y) > 0$$

$f_{X|Y}(x|y)$ indica quanto è probabile che X sia vicino a x, noto che Y è vicino a y.

L'integrale della densità condizionale rappresenta la probabilità che una variabile aleatoria assuma valori in un certo insieme, noto il valore di un'altra variabile aleatoria:

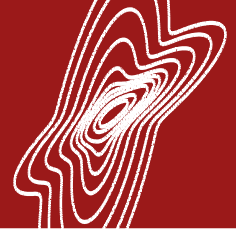
$$P(X \in A | Y = y) = \int_A f_{X|Y}(x|y) dx$$



OUTLINE



- Introduzione alla probabilità
- Probabilità di eventi
- Variabili aleatorie discrete e continue
- Vettori di variabili aleatorie
- **Valore atteso, varianza e deviazione standard**
- Covarianza e correlazione



VALORE ATTESO DI UNA VARIABILE ALEATORIA DISCRETA



X : variabile aleatoria discreta che può assumere i valori x_1, x_2, \dots .

Il **valore atteso** di X è, se esiste, la quantità:

$$E[X] := \sum_i x_i \cdot p_X(x_i)$$

- $E[X]$ è una media pesata dei valori di X , usando come pesi le loro probabilità.
- Per questo esso viene anche detto **media** di X o **aspettazione** di X .
- Domanda: che unità di misura ha $E[X]$?

INTERPRETAZIONE DEL VALORE ATTESO



- Il valore atteso NON È il valore che ci aspettiamo di osservare per una variabile aleatoria.
- Osserviamo k valori distinti di X , x_1, x_2, \dots, x_k , con frequenze assolute N_1, N_2, \dots, N_k ($N = N_1 + N_2 + \dots + N_k$) e calcoliamo la media di questi valori:

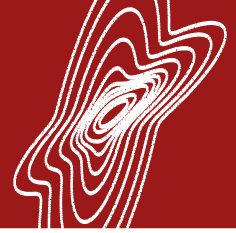
$$\frac{x_1 N_1 + x_2 N_2 + \dots + x_k N_k}{N} = x_1 \frac{N_1}{N} + x_2 \frac{N_2}{N} + \dots + x_k \frac{N_k}{N} \approx$$

Se N è molto grande

↓

$$\approx x_1 p_X(x_1) + x_2 p_X(x_2) + \dots + x_k p_X(x_k) = \sum_{i=1}^k [x_i \cdot p_X(x_i)] = E[X]$$

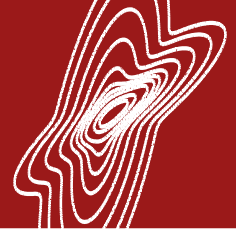
Il valore atteso rappresenta quel valore a cui tende la media di N valori osservati di X per N molto grande.



ESERCIZIO



Sia X il punteggio che si ottiene lanciando un dado non truccato. Quanto vale il valore atteso di X ?



VALORE ATTESO DI UNA VARIABILE ALEATORIA CONTINUA



- X : variabile aleatoria continua con ddp $f_X(x)$
- Valore atteso, o media o aspettazione di X :

$$E[X] := \int_{-\infty}^{+\infty} x \cdot f_X(x) dx$$

VALORE ATTESO DI UNA FUNZIONE DI UNA VARIABILE ALEATORIA

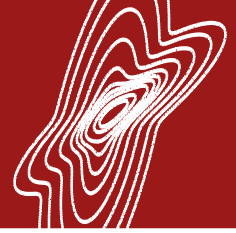


- Sia $g(X)$ una funzione di una variabile aleatoria X .
- Se X è discreta con funzione di massa $p_X(x)$:

$$E[g(X)] = \sum_i g(x_i) \cdot p_X(x_i)$$

- Se X è continua con ddp $f_X(x)$:

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x) dx$$



CASI PARTICOLARI



➤ $g(X) = a \cdot X + b \rightarrow E[g(X)] = a \cdot E[X] + b$

➤ $g(X) = a \cdot X \rightarrow E[g(X)] = a \cdot E[X]$

➤ $g(X) = b \rightarrow E[g(X)] = b$

VALORE ATTESO DI UNA FUNZIONE DI DUE VARIABILI ALEATORIE



Siano X e Y sono due variabili aleatorie e $g(X, Y)$ una qualsiasi loro funzione.

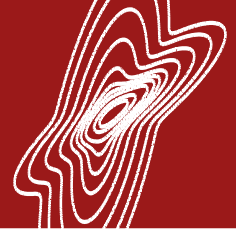
➤ Se X e Y sono discrete:

$$E[g(X, Y)] = \sum_j \sum_i g(x_i, y_j) \cdot p_{X, Y}(x_i, y_j)$$

➤ Se X e Y sono continue:

$$E[g(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) \cdot f_{X, Y}(x, y) dx dy$$

➤ Caso particolare: $E[X + Y] = E[X] + E[Y]$



MOMENTI DI ORDINE N



- Il valore atteso di una variabile aleatoria X viene anche detto **momento di primo ordine** di X .
- In generale, il **momento di ordine n** di una variabile aleatoria X è definito come $E[X^n]$.

- Se X è discreta:

$$E[X^n] = \sum_i x_i^n \cdot p_X(x_i)$$

- Se X è continua:

$$E[X^n] = \int_{-\infty}^{+\infty} x^n \cdot f_X(x) dx$$

MOMENTI CENTRALI DI ORDINE N



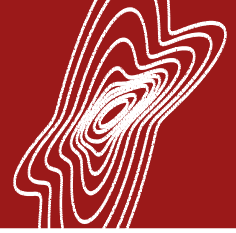
➤ Il **momento centrale di ordine n** di una variabile aleatoria X è definito come $E[(X - E[X])^n]$.

▪ Se X è discreta:

$$E[(X - E[X])^n] = \sum_i (x_i - E[X])^n \cdot p_X(x_i)$$

▪ Se X è continua:

$$E[(X - E[X])^n] = \int_{-\infty}^{+\infty} (x - E[X])^n \cdot f_X(x) dx$$



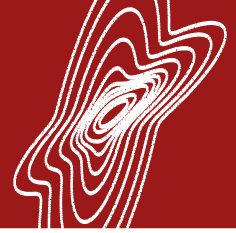
ESERCIZIO



Calcola il valore atteso per le seguenti variabili aleatorie:

- $X = 0$ con probabilità 1
- $Y = \begin{cases} -1 & \text{con probabilità } 0.5 \\ +1 & \text{con probabilità } 0.5 \end{cases}$
- $Z = \begin{cases} -100 & \text{con probabilità } 0.5 \\ +100 & \text{con probabilità } 0.5 \end{cases}$

Cosa puoi osservare?



VARIANZA



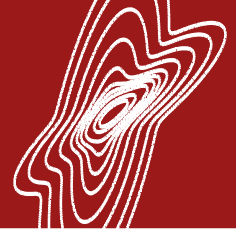
Se X è una variabile aleatoria con valore atteso μ , la sua **varianza**, indicata con $Var(X)$ o σ^2 , è definita come:

$$Var(X) := E[(X - \mu)^2]$$

- La varianza è il **momento centrale di ordine 2**.
- Rappresenta quanto i valori di X sono dispersi attorno al valore atteso.
- Si può calcolare anche come:

$$Var(X) = E[X^2] - \mu^2$$

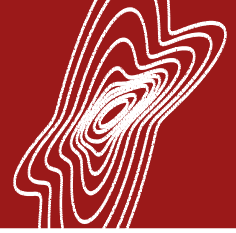
- Domanda: che unità di misura ha $Var(X)$?



ESERCIZIO



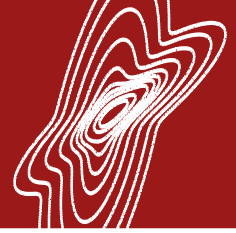
Calcola la varianza del punteggio di un dado non truccato (X).



PROPRIETA' DELLA VARIANZA



- $Var(X) \geq 0$
- $Var(a \cdot X + b) = a^2 \cdot Var(X)$
- $Var(b) = 0$
- $Var(X + b) = Var(X)$
- $Var(a \cdot X) = a^2 \cdot Var(X)$



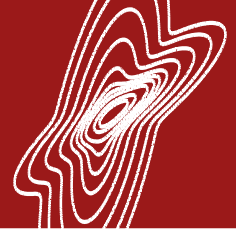
DEVIAZIONE STANDARD



- **Deviazione standard**, o *standard deviation*, della variabile aleatoria X :

$$\sigma(X) =: \sqrt{\text{Var}(X)}$$

- Domanda: che unità di misura ha $\sigma(X)$?



VARIABILE ALEATORIA SCARTO



Data una variabile aleatoria X , si dice **scarto di X** la variabile aleatoria:

$$Y := X - E[X]$$

- Domanda 1: Quanto vale il valore atteso di Y ?
- Domanda 2: Quanto vale la varianza di Y ?

Suggerimento: applicare le proprietà di valore atteso e varianza.

VARIABILE ALEATORIA STANDARDIZZATA



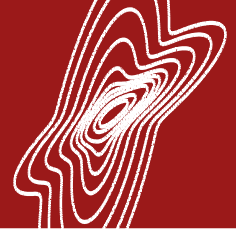
Data una variabile aleatoria X , la seguente trasformazione di X si dice **variabile aleatoria standardizzata**:

$$Z := \frac{X - E[X]}{\sigma(X)}$$

Domanda: quanto valgono **valore atteso** e **varianza** di una variabile aleatoria standardizzata?

$$E[Z] = E\left[\frac{X - E[X]}{\sigma(X)}\right] = \frac{1}{\sigma(X)} E[X] - \frac{E[X]}{\sigma(X)} = 0$$

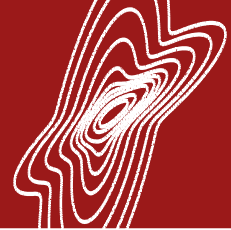
$$\text{Var}(Z) = \text{Var}\left(\frac{X - E[X]}{\sigma(X)}\right) = \frac{1}{[\sigma(X)]^2} \text{Var}(X) = 1$$



OUTLINE



- Introduzione alla probabilità
- Probabilità di eventi
- Variabili aleatorie discrete e continue
- Vettori di variabili aleatorie
- Valore atteso, varianza e deviazione standard
- **Covarianza e correlazione**



COVARIANZA



X e Y : due variabili aleatorie con valore atteso μ_X e μ_Y .

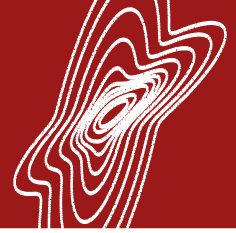
Si definisce **covarianza** di X e Y la quantità:

$$\text{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Formulazione equivalente:

$$\text{Cov}(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y]$$

- E' indicata anche con σ_{XY} .
- Intuitivamente, $\text{Cov}(X, Y)$ ci dà un'indicazione della relazione tra X e Y .



PROPRIETA' DELLA COVARIANZA



- A differenza della varianza, la covarianza può assumere valori negativi
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, X) = Var(X)$
- $Cov(a \cdot X, Y) = Cov(X, a \cdot Y) = a \cdot Cov(X, Y)$

- Siano X , Y e Z tre variabili aleatorie qualsiasi:

$$\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$$

- Siano X_1, X_2, \dots, X_n e Y variabili aleatorie qualsiasi:

$$\text{Cov}\left(\sum_{i=1}^n X_i, Y\right) = \sum_{i=1}^n \text{Cov}(X_i, Y)$$

- Siano X_1, X_2, \dots, X_n e Y_1, Y_2, \dots, Y_m variabili aleatorie qualsiasi:

$$\text{Cov}\left(\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j\right) = \sum_{j=1}^m \sum_{i=1}^n \text{Cov}(X_i, Y_j)$$



Attenzione!

$$\text{Var} \left(\sum_{i=1}^n X_i \right) \neq \sum_{i=1}^n \text{Var}(X_i)$$

Si ha invece che:

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \text{Cov}(X_i, X_j)$$

Nel caso $n=2 \rightarrow \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y)$

Siano X e Y due variabili aleatorie indipendenti.

Si può dimostrare che:

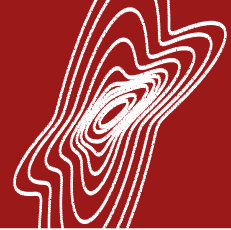
$$E[X \cdot Y] = E[X] \cdot E[Y]$$

Questo implica che:

$$\text{Cov}(X, Y) = E[X \cdot Y] - E[X] \cdot E[Y] = 0$$

Pertanto, solo nel caso di variabili aleatorie indipendenti si ha che:

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i)$$



ESEMPIO (1 / 2)

Consideriamo due variabili aleatorie X e Y indicatrici di due eventi A e B :

$$X = \begin{cases} 1 & \text{se } A \text{ si verifica} \\ 0 & \text{altrimenti} \end{cases} \quad Y = \begin{cases} 1 & \text{se } B \text{ si verifica} \\ 0 & \text{altrimenti} \end{cases}$$

Il prodotto tra X e Y sarà la funzione:

$$XY = \begin{cases} 1 & \text{se } X=1 \text{ e } Y=1 \\ 0 & \text{altrimenti} \end{cases}$$

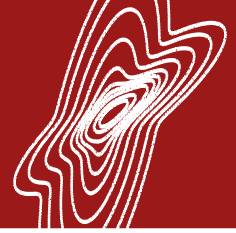
➤ $E[X]=P(X=1)$

➤ $E[Y]=P(Y=1)$

➤ $E[XY]=P(X=1, Y=1)$



$$\begin{aligned} \text{Cov}(X, Y) &= E[XY] - E[X] \cdot E[Y] = \\ &= P(X=1, Y=1) - P(X=1) \cdot P(Y=1) \end{aligned}$$



ESEMPIO (2/2)



$$\text{Cov}(X,Y) = P(X=1,Y=1) - P(X=1) \cdot P(Y=1)$$

Proviamo a dare un'interpretazione a questo valore di covarianza.

➤ $\text{Cov}(X,Y)=0 \iff X$ e Y sono indipendenti

➤ $\text{Cov}(X,Y)>0 \iff P(X=1,Y=1) > P(X=1) \cdot P(Y=1)$

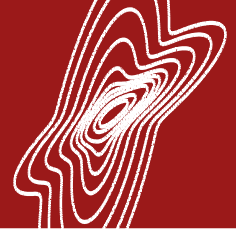
$$P(X=1,Y=1) / P(Y=1) > P(X=1)$$

$$P(X=1 | Y=1) > P(X=1)$$

ovvero se quando $Y=1$ è più probabile che anche $X=1$

➤ $\text{Cov}(X,Y)<0 \iff P(X=1 | Y=1) < P(X=1)$

ovvero se quando $Y=1$ è meno probabile che anche $X=1$



INTERPRETAZIONE DELLA COVARIANZA



La covarianza rappresenta un'importante indicatore della relazione tra due variabili aleatorie.

➤ $\text{Cov}(X,Y)=0$ se X e Y sono indipendenti.

Inoltre si può dimostrare che:

➤ $\text{Cov}(X,Y)>0$ se X e Y tendono ad assumere valori grandi o piccoli contemporaneamente.

➤ $\text{Cov}(X,Y)<0$ se quando X tende ad assumere valori grandi, Y tende ad assumere valori piccoli e viceversa.

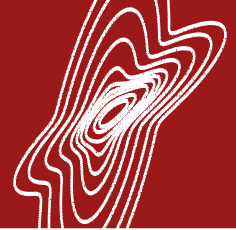
COEFFICIENTE DI CORRELAZIONE LINEARE



Più propriamente la forza della relazione tra X e Y si può misurare con il **coefficiente di correlazione lineare**, o coefficiente di correlazione di **Pearson**, un numero puro (senza unità di misura) che tiene conto anche delle deviazioni standard di X e Y :

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$$

$\text{Corr}(X, Y)$ varia tra -1 e 1 ed esprime il grado di relazione lineare tra X e Y .



INTERPRETAZIONE DEL COEFFICIENTE DI CORRELAZIONE LINEARE



Segno di $\text{Corr}(X,Y)$:

- Se $\text{Corr}(X,Y) > 0$, al crescere (decrescere) di X , cresce (decesce) anche Y .
→ X e Y si dicono **positivamente o direttamente correlate**.
- Se $\text{Corr}(X,Y) < 0$, al crescere (decrescere) di X , Y decresce (cresce).
→ X e Y si dicono **negativamente o inversamente correlate**.

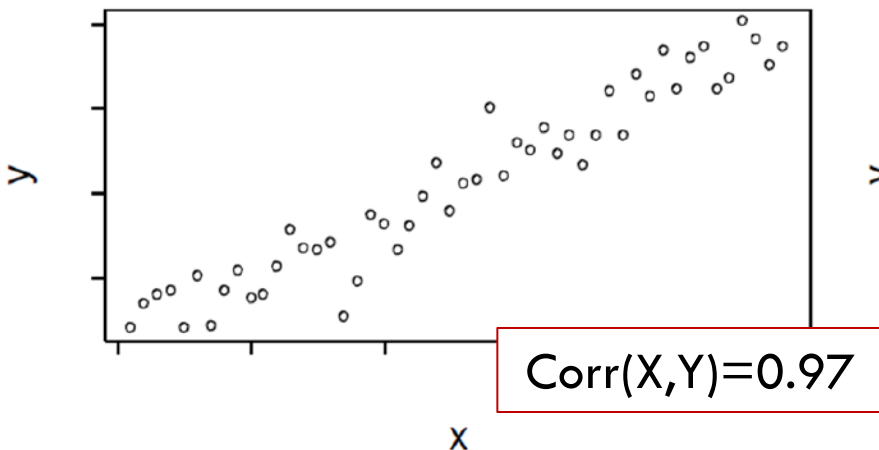
Valore assoluto di $\text{Corr}(X,Y)$:

- E' tanto maggiore quanto **più** è forte la relazione lineare **tra X e Y** .
- E' pari a 1 quando la relazione tra X e Y è perfettamente lineare: $Y=a \cdot X+b$.
- E' 0 quando non c'è relazione lineare tra X e Y → X e Y **scorrelate o incorrelate**.

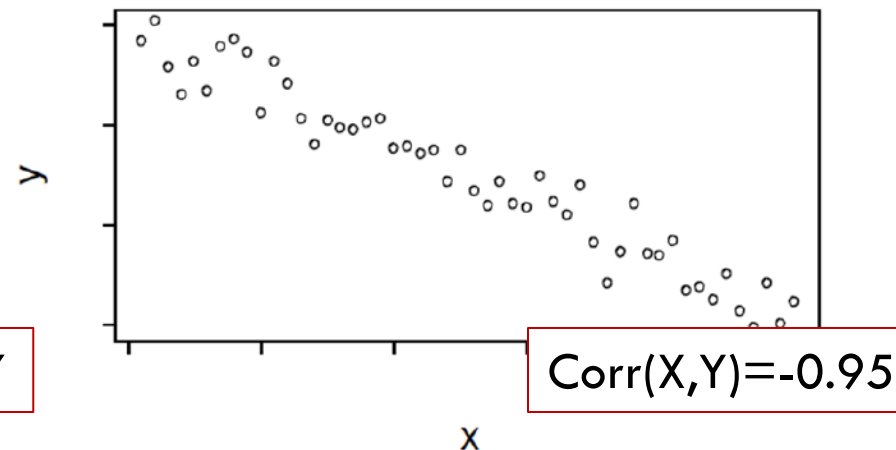
ESEMPI



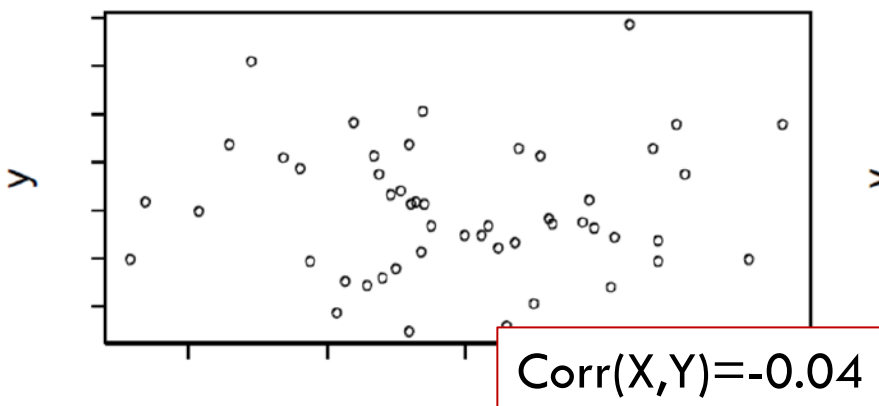
relazione lineare positiva



relazione lineare negativa



nessuna relazione



relazione non lineare

