METODI STATISTICI PER LA BIOINGEGNERIA (B)

PARTE 3: RIPASSO DI PROBABILITA' E VARIABILI ALEATORIE

A.A. 2024-2025

Prof. Martina Vettoretti

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 viene generalmente considerato incerto. Teoricamente si potrebbe trasformare il lancio della moneta in un fenomeno dal risultato certo se conoscessimo e tenessimo in considerazione fattori come le velocità iniziali di traslazione e di rotazione, la situazione termica ambientale, il piano d'arrivo ecc. Risulta però più comodo adottare un altro punto di vista che rinuncia alla descrizione analitica del fenomeno e attribuisce i diversi esiti dell'esperienza ad una variabilità accidentale, intesa come sintesi delle diverse condizioni che si rinunciano a specificare.

INTERPRETAZIONE FREQUENTISTA DI PROBABILITA'

- **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 DI PROBABILITA'

- **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.

SAMPLE SPACE ED EVENTI

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

- Due eventi E ed F sono disgiunti o mutuamente esclusivi se $E \cap F = \emptyset$.
- E è contenuto in F, $E \subset F$, se 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 si associa un numero P(E) che rappresenta la probabilità di E.

- P(E) deve rispettare alcune condizioni definite dagli **Assiomi di Kolmogorov**:

$$0 \leq P(E) \leq 1, \quad P(\emptyset) = 0 \qquad \text{(Assioma 1)}$$

$$P(\Omega) = 1 \qquad \text{(Assioma 2)}$$

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

$$P(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i) \qquad \text{(Assioma 3)}$$

EVENTO CERTO ED EVENTO IMPOSSIBILE

- **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. → P(E) = P({r}) = P(Ω) = 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. → 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 → odds: P(E)/(1-P(E))=3

  → è 3 volte più probabile che E si verifichi piuttosto che non si verifichi

SPAZI DI ESITI EQUIPROBABILI

- Spesso possiamo assumere che tutti i possibili esiti abbiano la stessa probabilità di realizzarsi.
  - Esempio: nel lancio di un dado non truccato, possiamo considerare equiprobabili tutti i possibili esiti, da 1 a 6.
- Supponiamo ci siano N possibili esiti equiprobabili, con probabilità p:

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

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

→ Per gli assiomi 2 e 3, ciascun esito dovrà avere probabilità $1/N$.

- Ne consegue che la probabilità di un evento E composto da un numero di esiti pari a #E si può calcolare come:

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

  - Nell'esempio del lancio del dado la probabilità che esca un numero pari sarà 3/6 perché il numero di esiti per cui E si verifica è 3 (2,4,6) e il numero di esiti possibili è 6.

PROBABILITA' CONDIZIONATA

- Spesso vogliamo calcolare la probabilità di un evento quando siamo già in possesso di alcune informazioni parziali sull'esito dell'esperimento. In questo caso le probabilità che andiamo a calcolare si dicono condizionate.
- 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:

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

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

ESEMPIO

Consideriamo il lancio di due dadi non truccati e l'evento E = {tutti gli esiti per cui la somma dei risultati dei due dadi è pari a 8}. Quanto vale la probabilità di E?

- Se non abbiamo altre informazioni, contiamo tutti i possibili esiti che portano ad una somma pari a 8. Questi sono 5: (2,6), (3,5), (4,4), (5,3), (6,2). Tutti gli esiti possibili sono 6x6=36. Quindi P(E) = 5/36.
- Se 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, quindi i possibili esiti diventano 6. L'unico esito per cui E si può verificare è quando il secondo lancio dà come risultato 5. → P(E|F) =1/6.
- Potevamo calcolare P(E|F) anche usando la definizione. P(F)=1/6, P(E∩F)=1/36 → P(E|F)=(1/36) / (1/6)=1/6.

FORMULA DI FATTORIZZAZIONE DEGLI EVENTI

Consideriamo N eventi, **F_1, F_2, …, 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:**

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

→ 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

Se pensiamo agli eventi F_j come possibili ipotesi alternative che abbiano un'influenza su un qualche esperimento, la formula di Bayes ci dice come è necessario modificare le nostre opinioni su tali ipotesi da prima a dopo l'esperimento, con le loro probabilità che passano da $P(F_j)$ a $P(F_j|E)$.

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 n eventi E_1, E_2,…, E_n si dicono indipendenti se:

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

Molti esperimenti consistono in una successione di prove. Spesso è ragionevole assumere che gli esiti di queste prove siano indipendenti (es. lancio ripetuto di una moneta).

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.
- Una variabile aleatoria rappresenta 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

- Formalmente, 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.
- Siccome il valore di una variabile aleatoria è determinato dall'esito di un esperimento, 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).

Lo spazio campionario dei possibili esiti è: Ω = {(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)} U {(d,a)}) = 0.42
- P(X = 2) = P({(a,a)})= 0.49

FUNZIONE DI RIPARTIZIONE

La **funzione di ripartizione o di distribuzione,** F_X, di una variabile aleatoria X è definita per ogni numero reale x come:

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

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

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

VARIABILE ALEATORIA DISCRETA

- Una **variabile aleatoria** si dice **discreta** se può assumere solo un numero finito o infinito numerabile di valori possibili $\{x_1, x_2, \ldots\}$.
- Se X è una variabile aleatoria discreta, si dice **funzione di massa di probabilità** o **densità discreta** la funzione:

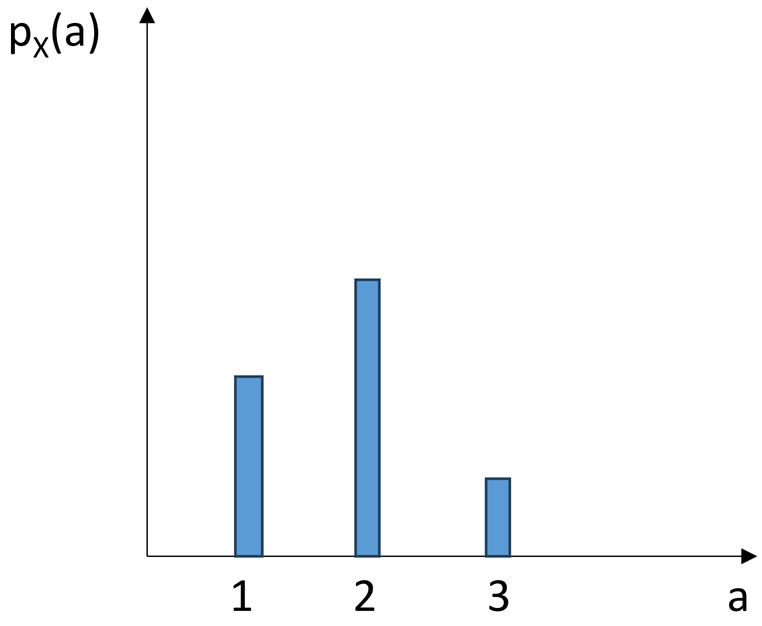
$$p_X(a) := P(X = a), \qquad p_X\colon \mathbb{R} \to [0,1]$$

- La funzione di ripartizione di una variabile aleatoria discreta X può essere calcolata a partire dalla sua funzione di massa:

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

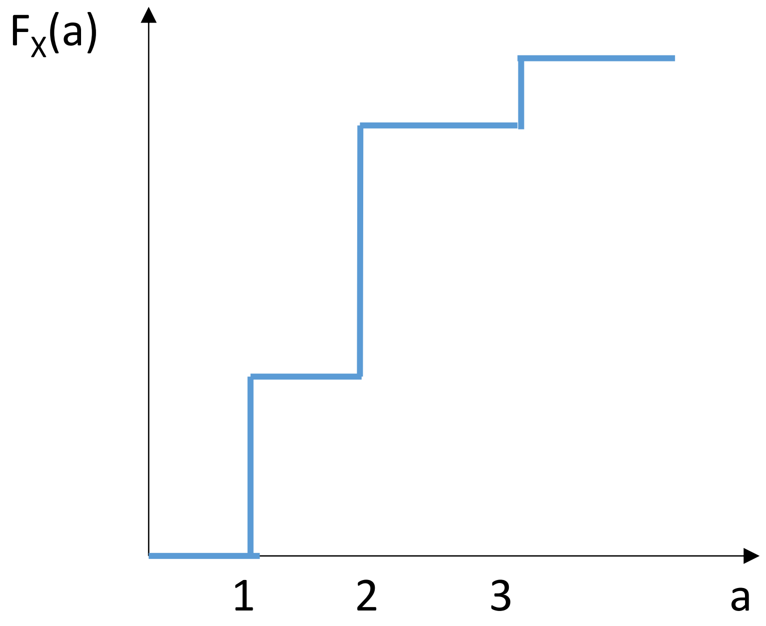
FUNZIONI DI MASSA E DI RIPARTIZIONE DI UNA VARIABILE ALEATORIA DISCRETA

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 R, ed integrabile, tale che per ogni insieme B di numeri reali:

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

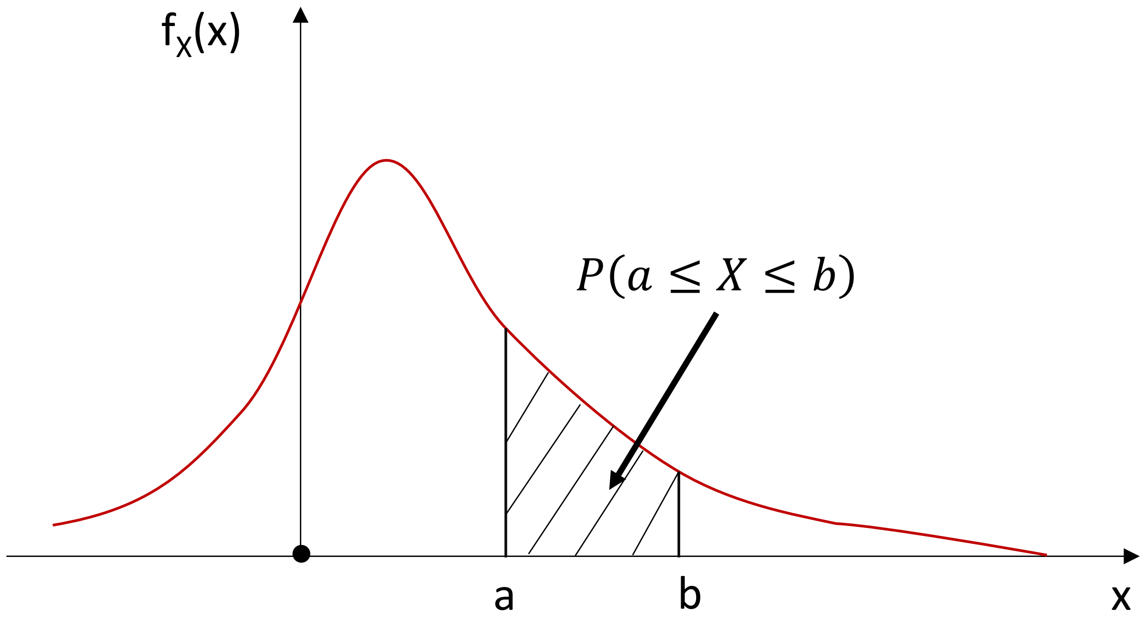
- La funzione f_X si chiama **funzione di densità di probabilità (ddp)**, o semplicemente **densità**, di X.

PROPRIETA' DELLA FUNZIONE DI DENSITA' DI PROBABILITA'

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

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

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



Dato un valore ε sufficientemente piccolo:

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

$f_X(a)$ è proporzionale alla probabilità di un intervallo sufficientemente piccolo intorno ad a, ovvero dà un'indicazione di quanto è probabile che X cada 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 densità di probabilità di una variabile aleatoria continua.
  - La densità di probabilità 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

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

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

La densità di probabilità è la derivata della funzione di ripartizione.

OSSERVAZIONE

Quando conosciamo la funzione di massa di probabilità di una variabile aleatoria discreta, o la funzione di densità di probabilità 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 $A \subset \mathbb{R}$:

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

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

FUNZIONE DI RIPARTIZIONE CONGIUNTA

Siano X e Y due variabili aleatorie che riguardano lo stesso esperimento casuale. Si dice **funzione di ripartizione congiunta** di X e Y la seguente funzione:

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

che 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.

- $F_X(x) = F_{X,Y}(x, \infty)$
- $F_Y(y) = F_{X,Y}(\infty, y)$

FUNZIONE DI MASSA DI PROBABILITA' CONGIUNTA

Se X e Y sono due variabili aleatorie discrete che assumono i valori $x_1, x_2, \ldots,$ e $y_1, y_2, \ldots,$ rispettivamente, la funzione:

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

È la loro **funzione di massa di probabilità congiunta** o **densità discreta congiunta.**

FUNZIONI DI MASSA MARGINALI

E' possibile ricavare da $p_{X,Y}(x_i, y_i)$ le funzioni di massa di probabilità di X e di Y. Esse vengono dette funzioni di massa marginali.

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

Per l'assioma 3

$$= \sum_j P(X = x_i, Y = y_j) = \sum_j p_{X,Y}(x_i, y_j)$$

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 $p_{X,Y}(x_i, y_j)$.

ESEMPIO (1/2)

All'interno di una certa popolazione, il 15% delle coppie non ha figli, il 20% delle coppie ha 1 figlio, il 35% ne ha 2 e 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?

- P(X=0,Y=0)=0.15
- P(X=1,Y=0)=P(1 figlio,1 femmina)=P(femmina|1 figlio)P(1 figlio)=0.5 x 0.20=0.1
- P(X=2,Y=0)=P(2 figli,2 femmine)=P(2 femmine|2 figli)P(2 figli)=0.5^2 x 0.35=0.0875
- P(X=3, Y=0)=P(3 figli,3 femmine)=P(3 femmine|3 figli)P(3 figli)=0.5^3 x 0.30=0.0375

Analogamente si possono calcolare le probabilità congiunte quando Y è 1, 2 e 3.

ESEMPIO (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	0.1500	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	

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:

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

La funzione $f_{X,Y}(x,y)$ si dice **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 $f_{X,Y}(a,b)$ fornisce un'indicazione di quanto è probabile che (X,Y) cada vicino ad (a,b).

LEGAME CON LA FUNZIONE DI RIPARTIZIONE, DENSITA' MARGINALI

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

$$f_{X,Y}(a,b) = \frac{\partial^2 F(a,b)}{\partial a\,\partial b}$$ $\longrightarrow$ La densità di probabilità congiunta è la derivata seconda della funzione di ripartizione.

A partire dalla densità di probabilità congiunta $\mathrm{f}_{X,Y}(x,y)$ possiamo ricavare la **densità di probabilità marginali** $\mathrm{f}_X(x)$ e $\mathrm{f}_Y(y)$.

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

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

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.

Ne deriva che 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)

In generale potremmo avere N variabili aleatorie riguardanti uno stesso esperimento, che formano un cosiddetto **vettore aleatorio**.

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,…, X_n:

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

- Funzione di massa congiunta di n variabili aleatorie X_1, X_2,…, X_n:

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

GENERALIZZAZIONE AL CASO DEI VETTORI ALEATORI (2/2)

- Funzione di densità di probabilità congiunta di n variabili aleatorie X_1, X_2,…, X_n:

$$P(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n)$$
$$= \iint\limits_{(x_1,x_2 \dots x_n \in C)} \cdots \int f_{X_1,X_2,\dots,X_n}(x_1, x_2, \dots, x_n)\, dx_1 dx_2 \cdots dx_n$$

- Le n variabili aleatorie X_1, X_2,…, X_n sono indipendenti se:

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

FUNZIONE DI MASSA DI PROBABILITA' CONDIZIONATA

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

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

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

$$\forall\, x, \forall y, con\; p_Y(y) > 0$$

ESEMPIO

All'interno di una certa popolazione, il 15% delle coppie non ha figli, il 20% delle coppie ha 1 figlio, il 35% ne ha 2 e 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 $p_{Y|X}(y|1)$?

- $p_{Y|X}(0|1) = P(Y=0|X=1) = P(Y=0,X=1) / P(X=1) = 0.1/0.3875$
- $p_{Y|X}(1|1) = P(Y=1|X=1) = P(Y=1,X=1) / P(X=1) = 0.175/0.3875$
- $p_{Y|X}(1|1) = P(Y=2|X=1) = P(Y=2,X=1) / P(X=1) = 0.1125/0.3875$
- $p_{Y|X}(3|1) = P(Y=3|X=1) = P(Y=3,X=1) / P(X=1) = 0$

DENSITA' CONDIZIONALE

Siano X e Y due variabili aleatorie continue con densità di probabilità congiunta $f_{X,Y}(x,y)$. Si dice **densità condizionale** di X rispetto a Y la funzione:

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

- Interpretazione analoga alle altre densità di probabilità: $f_{X|Y}(x|y)$ rappresenta un'indicazione di quanto è probabile che X sia vicino a x, sapendo che Y è vicino a y.
- L'integrale della densità condizionale rappresenta la probabilità che una variabile aleatoria assuma valori in un certo insieme, essendo noto il valore di un'altra variabile aleatoria:

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

VALORE ATTESO DI UNA VARIABILE ALEATORIA DISCRETA

Sia X una variabile aleatoria discreta che può assumere i valori x_1, x_2, … .

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

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

- Il valore atteso non è altro che una media pesata dei valori X, usando come pesi le loro probabilità.
- Per questo esso viene anche detto **media** di X o **aspettazione** di X.
- E[X] ha la stessa unità di misura di X.

INTERPRETAZIONE DEL VALORE ATTESO

- Il valore atteso non è il valore che ci aspettiamo di osservare per una variabile aleatoria.
- Piuttosto, possiamo interpretare il valore atteso come quel valore a cui tende la media di N valori osservati di X per N molto grande. Infatti se osserviamo n valori distinti di X, $x_1, x_2, \ldots, x_n$, con frequenze assolute $N_1, N_2, \ldots, N_n$:

$$\frac{x_1 N_1 + x_2 N_2 + \cdots + x_n N_n}{N} = x_1 \frac{N_1}{N} + x_2 \frac{N_2}{N} + \cdots + x_n \frac{N_n}{N} \approx$$

$$\approx x_1 p_X(x_1) + x_2 p_X(x_2) + \cdots x_n p_X(x_n) = \sum_{i=1}^{n} [x_i \cdot p_X(x_i)] = \mathrm{E}[X]$$

ESEMPIO

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

$$E[X]=1\cdot 1/6 + 2\cdot 1/6 + 3\cdot 1/6 + 4\cdot 1/6 + 5\cdot 1/6 + 6\cdot 1/6 = 21/6 = 7/2 = 3.5$$

In questo caso E[X] non è neanche un possibile valore di X. Infatti E[X] non è un valore che ci aspettiamo di osservare per X, ma il valore a cui tende la media dei valori di X per un numero di valori osservati molto grande.

VALORE ATTESO DI UNA VARIABILE ALEATORIA CONTINUA

Sia X una variabile aleatoria continua con densità di probabilità $f_X(x)$, si dice valore atteso, o media o aspettazione di X la quantità:

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

VALORE ATTESO DI UNA FUNZIONE DI UNA VARIABILE ALEATORIA

- Se X è una variabile aleatoria discreta, che assume valori $x_1, x_2, \ldots$, con funzione di massa $p_X(x)$, e g(X) è una sua funzione, allora il valore atteso di g(X) è:

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

- Se X è una variabile aleatoria continua con funzione di densità di probabilità $f_X(x)$, e g(X) è una sua funzione, allora il valore atteso di g(X) è:

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

CASI PARTICOLARI

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

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

- $g(X) = b \quad \rightarrow \quad E[g(X)] = b$

VALORE ATTESO DI UNA FUNZIONE DI DUE VARIABILI ALEATORIE

Se X e Y sono due variabili aleatorie e g(X,Y) una qualsiasi loro funzione allora il valore atteso di g(X,Y) è:

- 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

- Esiste anche il **momento centrale di ordine n** di una variabile aleatoria X, 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:

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

LA VARIABILITA’

Il valore atteso rappresenta il ‘baricentro’ dei valori assunti dalla variabile aleatoria, però non coglie la variabilità o la dispersione di questi valori.

Per esempio le seguenti variabili aleatorie hanno tutte valore atteso nullo, ma la variabilità dei loro valori è molto differente.

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

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 coincide con il **momento centrale di ordine 2**.
- La varianza si può calcolare anche come:

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

- La varianza rappresenta quanto i valori di X sono dispersi attorno al valore atteso.
- L'unità di misura di Var(X) è quella di X al quadrato.

ESEMPIO

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

$$\mu = E[X] = \frac{7}{2}$$

$$E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{91}{6}$$

$$Var(X) = E[X^2] - \mu^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

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

- Si dice **deviazione standard**, o *standard deviation*, della variabile X la radice quadrata della sua varianza:

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

- L'unità di misura di σ(X) è la stessa di X.

VARIABILE ALEATORIA SCARTO

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

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

La variabile aleatoria scarto ha **valore atteso nullo** e **varianza pari alla varianza di X**:

$$E[Y] = E\big[X - E[X]\big] = E[X] - E[X] = 0$$

$$Var(Y) = Var(X - E[X]) = Var(X)$$

VARIABILE ALEATORIA STANDARDIZZATA

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

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

La variabile aleatoria standardizzata ha **valore atteso nullo** e **varianza unitaria:**

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

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

COVARIANZA

Siano X e Y due variabili aleatorie con valore atteso μ_X e μ_Y. Si definisce **covarianza** di X e Y la quantità:

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

Formulazione equivalente:

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

Intuitivamente la covarianza di X e Y, indicata anche come σ_{XY}, ci dà un'indicazione della relazione tra X e Y.

Proprietà:

- $Cov(X,Y) = Cov(Y,X)$
- $Cov(X,X) = Var(X)$
- $Cov(a \cdot X, Y) = Cov(X, a \cdot Y) = a \cdot Cov(X,Y)$

Nota: mentre la varianza è per forza ≥0, la covarianza può anche assumere valori negativi.

COVARIANZA DELLA SOMMA DI VARIABILI ALEATORIE

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

$$Cov(X+Y,Z) = Cov(X,Z) + Cov(Y,Z)$$

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

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

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

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

VARIANZA DELLA SOMMA DI VARIABILI ALEATORIE

In generale la varianza della somma di variabili aleatorie è diversa dalla somma delle varianze delle variabili aleatorie:

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

Si ha invece che:

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

Nel caso n=2 → $Var(X+Y) = Var(X) + Var(Y) + 2 \cdot Cov(X, Y)$

CASO CON VARIABILI ALEATORIE INDIPENDENTI

Siano X e Y due variabili aleatorie indipendenti.

Si può dimostrare che:

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

Questo implica che:

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

E pertanto, <u>solo nel caso di variabili aleatorie indipendenti</u> si ha che:

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

ESEMPIO (1/2)

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

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

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)

➡ Cov(X,Y)=E[XY]-E[X]·E[Y]=

=P(X=1,Y=1)-P(X=1)·P(Y=1)

ESEMPIO (2/2)

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

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

- $\mathrm{Cov}(X,Y)=0 \Leftrightarrow$ X e Y sono indipendenti
- $\mathrm{Cov}(X,Y)>0 \Leftrightarrow P(X=1,Y=1) > P(X=1)\cdot P(Y=1)$

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

  $P(X=1 \mid Y=1) > P(X=1)$

  ovvero se quando $Y=1$ è più probabile che anche $X=1$
- $\mathrm{Cov}(X,Y)<0 \Leftrightarrow P(X=1 \mid Y=1) < P(X=1)$

  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.

- $Cov(X,Y)=0$ se X e Y sono indipendenti.

Inoltre si può dimostrare che:

- $Cov(X,Y)>0$ se X e Y tendono ad assumere valori grandi o piccoli contemporaneamente.
- $Cov(X,Y)<0$ 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:

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

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 Corr(X,Y):

- Se Corr(X,Y) >0, al crescere (decrescere) di X cresce (decresce) anche Y.

  → X e Y si dicono **positivamente o direttamente correlate.**

- Se Corr(X,Y) <0, al crescere (decrescere) di X, Y decresce (cresce).

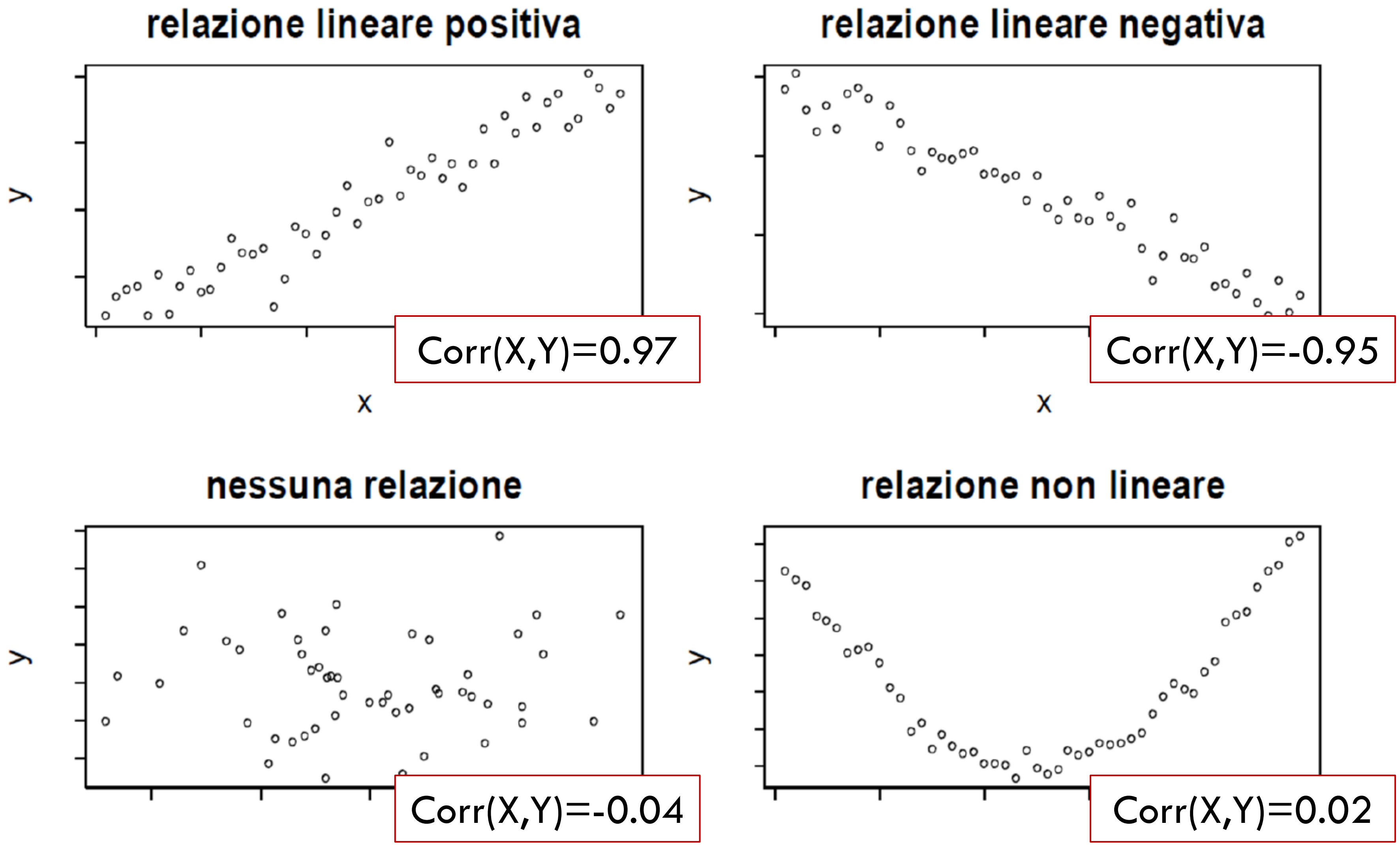
  → X e Y si dicono **negativamente o inversamente correlate.**

Valore assoluto di Corr(X,Y):

- E' tanto maggiore quanto è forte la relazione lineare.
- E' pari a 1 quando la relazione tra X e Y è perfettamente lineare ovvero: Y=a·X+b.
- E' pari a 0 quando non c'è relazione tra X e Y.

  → X e Y si dicono **scorrelate** o **incorrelate.**

ESEMPI


relazione lineare positiva
y
x
Corr(X,Y)=0.97
relazione lineare negativa
y
x
Corr(X,Y)=-0.95
nessuna relazione
y
Corr(X,Y)=-0.04
relazione non lineare
y
Corr(X,Y)=0.02