Metodi Matematici e Statistici

Dr. Giovanni Nastasi
giovanni.nastasi@unict.it

Dipartimento di Matematica e Informatica
Università degli Studi di Catania, Italy

CdS in Informatica
A.A. 2023-2024

Elementi di calcolo delle probabilità

Introduzione

La statistica matematica, interpretando ed elaborando i dati raccolti su un "campione", ha fra i suoi scopi quello di fare "previsioni" sull'intera "popolazione" a cui si riferisce il fenomeno che si vuole studiare.

In quest'ambito, il ruolo del calcolo delle probabilità consiste nel misurare, sulla base dei risultati relativi al campione, l'attendibilità di circostanze o eventi non osservati.

Si fa cioè una "inferenza" al di là dei dati osservati tra questi e le possibili conclusioni mediante una relazione di probabilità.

Ad esempio, supponiamo che si voglia valutare la percentuale di pezzi difettosi in una catena di produzione. Per dire con esattezza qual è la percentuale di pezzi difettosi si dovrebbe analizzare l'intera produzione, cosa spesso impossibile, specialmente se la produzione è continua nel tempo e quindi non si conosce nulla su quanto sarà prodotto nel futuro.

Definizione classica di probabilità

Un evento E è individuato da una proposizione ben definita ed esso risulta VERO oppure FALSO.

Ad esempio, nel lancio di una moneta, la proposizione

$$E = \text{“esce testa”}$$

rappresenta un evento.

È data una famiglia di n eventi elementari (o casi possibili) a due a due incompatibili (cioè tali da non potersi verificare insieme).

Questi eventi si giudicano equiprobabili e si considera l'evento E consistente nel verificarsi di uno qualunque fra h eventi fissati scelti tra gli n possibili (cioè h è il numero di casi favorevoli ad E).

Il numero

$$P(E) = \frac{h}{n}$$

si dice **probabilità** di E che viene quindi valutata come il rapporto tra il numero di casi favorevoli ad E e il numero di casi possibili.

Esempio

Nel lancio di un dado equilibrato a 6 facce gli eventi

$$E_k = \text{“viene estratto il numero k”}, \qquad 1 \le k \le 6$$

costituiscono gli eventi elementari ed equiprobabili con $P(E_k) = \frac{1}{6}$.

Invece, l'evento

$$E = \text{“viene estratto un numero pari”}$$

non è un evento elementare. La sua probabilità si può ottenere calcolando il rapporto tra il numero di casi favorevoli e il numero di quelli possibili. Cioè esso ha probabilità $P(E) = \frac{3}{6} = \frac{1}{2}$ in quanto gli eventi elementari favorevoli sono E_2, E_4 e E_6.

Definizione frequentista di probabilità

Si considera un fenomeno ripetibile, espresso mediante una successione di eventi analoghi

$$E_1, E_2, \dots, E_i, \dots$$

Se ne osservano un numero abbastanza grande m e si conta il numero k dei successi (cioè degli eventi che si verificano).

Il numero $P(E) = \frac{k}{m}$ si assume come probabilità dell'evento $E = E_{m+1}$. Cioè si valuta la probabilità di un evento futuro mediante la frequenza relativa dei successi sugli eventi passati.

Operazioni con gli eventi

- Due eventi A e B sono **uguali**, e si scrive $A = B$, quando dall'essere vero A segue che anche B è vero e viceversa.
- Si dice che A **implica** B, e si scrive $A \subseteq B$, quando dall'essere vero A segue che è vero anche B. Pertanto $A = B$ se e solo se $A \subseteq B$ e $B \subseteq A$.
- Dato un evento E, si definisce il suo **contrario** E^c come l'evento che è vero quando E è falso e viceversa.
- Dati due eventi A e B la loro **somma logica** $A \vee B$ (o **evento unione** $A \cup B$) è quell'evento che è vero se lo è almeno uno dei due.
- Dati due eventi A e B il loro **prodotto logico** $A \wedge B$ (o **evento intersezione** $A \cap B$) è quell'evento che è vero se lo sono entrambi.

Esempio. Si effettuano due lanci di una moneta e si considerano gli eventi

$$A = \text{“testa al primo lancio”}, \qquad B = \text{“testa al secondo lancio”}.$$

L'evento $A \cup B$ è vero in corrispondenza dei risultati: TC, CT, TT.
L'evento $A \cap B$ è vero soltanto in corrispondenza del risultato TT.

Proprietà degli eventi

Si osservi che si possono considerare anche somma logica (o unione) e prodotto lofico (o intersezione) di un numero qualunque (finito o no) di eventi, e valgono le proprietà associativa, commutativa e distributive (della somma rispetto al prodotto e del prodotto rispetto alla somma).

Un evento si dice **certo**, e si indica con Ω, se la proposizione che lo descrive può essere solo VERA.
Es: Nel lancio di una moneta l'evento $A =$ "esce testa oppure croce" è certo.

Un evento si dice **impossibile**, e si indica con $\emptyset$, se la proposizione che lo descrive può essere solo FALSA.
Es: Nel lancio di un dado a sei facce l'evento $A =$ "viene estratto il numero 9" è impossibile.

Due eventi si dicono **incompatibili** se

$$A \cap B = \emptyset,$$

cioè quando il loro prodotto logico è l'evento impossibile.

Es: Nel lancio di un dado a sei facce, gli eventi

$A =$ "viene estratto il numero 2" e $B =$ "viene estratto il numero 5"

sono incompatibili. Invece, l'evento A e l'evento $C =$ "viene estratto un numero pari" sono compatibili.

Definizione soggettiva di probabilità (B. de Finetti)

Nelle situazioni reali si incontrano eventi di tipo meno schematico di quelli presi in esame fino a questo punto. Si ha quindi la necessità di valutare la probabilità come grado di fiducia nel verificarsi di un evento.

Si consideri la frase: "La probabilità che il Catania vinca la prossima partita è del 70%".

Chiaramente non si pensa ad un'urna contenente 100 esiti di partite del Catania fra le quali 70 vinte e 30 perse o pareggiate dalle quali viene estratto l'esito della prossima partita.

Non si pensa neanche alla possibilità che la partita sia ripetuta 100 volte osservando il Catania vincere 70 volte e perdere o pareggiare 30 volte.

L'atteggiamento di chi pronuncia una frase del genere può essere invece interpretato nel modo seguente. Egli, se dovesse effettuare una scommessa, sarebbe disposto ad accettare di pagare 70 euro per riceverne 100 nel caso in cui l'evento si verifichi (con la prospettiva di guadagnarne 30 a fronte di quella di perderne 70).

Definizione soggettiva di probabilità (B. de Finetti)

Questa osservazione è utile per comprendere la definizione soggettiva di probabilità.

Dato un numero reale $p \geq 0$, si dice che si effettua una scommessa di quota p su un evento E se, versando una somma pS (con $S > 0$ arbitrario), si riceve un importo S solo se E si verifica (scommessa vinta) e niente in caso contrario.

Indicando con G il guadagno si ha che

- se si verifica E allora $G_1 = S - pS = (1-p)S$;
- se si verifica E^c allora $G_2 = -pS$.

Una scommessa è coerente se i due possibili guadagni non sono dello stesso segno, cioè deve essere

$$G_1 G_2 = -p(1-p)S^2 \leq 0$$

e ciò accade se e solo se $0 \leq p \leq 1$.

Definizione. Si dice probabilità di un evento E la misura del "grado di fiducia" in E espressa da un numero reale $p = P(E)$ tale che una scommessa di quota p su E sia coerente.

Definizione assiomatica di probabilità (A.N. Kolmogorov)

Si consideri un fenomeno casuale o aleatorio e siano

- Ω l'insieme dei possibili esiti;
- $\mathcal{A}$ una famiglia di sottoinsiemi di Ω che costituiscono gli eventi;
- P una misura del grado di aspettazione o fiduciosità di accadimento degli eventi della famiglia $\mathcal{A}$.

Esempio. Nel lacio di un dado $\Omega = \{1, 2, 3, 4, 5, 6\}$. L'insieme $E = \{2, 5\}$ è un possibile evento corrispondente alla proposizione "viene estratto 2 oppure 5".

Per quanto riguarda l'insieme degli eventi $\mathcal{A}$ richiederemo per ragioni tecniche che costituisca una σ-algebra (sigma algebra) secondo la definizione seguente.

Definizione. $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ dicesi σ-algebra su Ω se

1. $\emptyset$ e $\Omega \in \mathcal{A}$;
2. se $A \in \mathcal{A}$ allora $A^C \in \mathcal{A}$ (essendo A^C il complementare di A rispetto a Ω);
3. se $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{A}$ allora anche $\cup_n A_n \in \mathcal{A}$ e $\cap_n A_n \in \mathcal{A}$ potendo l'unione e l'intersezione essere finita o numerabile.

Definizione assiomatica di probabilità (A.N. Kolmogorov)

Definizione. Sia Ω un insieme non vuoto e $\mathcal{A}$ una σ-algebra su Ω. Una **probabilità** è una applicazione $P : \mathcal{A} \to [0,1]$ tale che

1. $P(\Omega) = 1$;
2. se $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{A}$ con $A_i \cap A_j = \emptyset$ se $i \neq j$ allora $P\left(\cup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} P(A_n)$ (proprietà di σ-additività).

Osservazione. Dalla σ-additività segue la finita additività, cioè se $A_1, A_2, \ldots, A_n \in \mathcal{A}$ allora

$$P\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Si dimostra applicando la proprietà di σ-additività alla successione

$$(A_n)_{n\in\mathbb{N}} = (A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \ldots).$$

Definizione. La terna $(\Omega, \mathcal{A}, P)$ si chiama **spazio di probabilità**.

Proprietà degli spazi di probabilità

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità. Allora

1. $P(A^C) = 1 - P(A)$ per ogni $A \in \mathcal{A}$;
2. $P(A) \leq P(B)$ per ogni $A, B \in \mathcal{A}$ con $A \subseteq B$;
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ per ogni $A, B \in \mathcal{A}$.

Dimostrazione.

1. Per ogni $A \in \mathcal{A}$ si ha che $\Omega = A \cup A^C$ e quindi $1 = P(A) + P(A^C)$ per la proprietà di finita additività.
2. Per ogni $A, B \in \mathcal{A}$ si ha

$$\begin{aligned} B = B \cap \Omega &= B \cap (A \cup A^C) \\ &= (B \cap A) \cup (B \cap A^C) && \text{(per la prop. distributiva dell'inter. rispetto all'uni.)} \\ &= A \cup (B \cap A^C) && \text{(perché } A \subseteq B \text{ e quindi } B \cap A = A\text{)} \end{aligned}$$

Per la finita additività si ha $P(B) = P(A) + P(B \cap A^C) \geq P(A)$.

Proprietà degli spazi di probabilità

3. Per ogni $A, B \in \mathcal{A}$ si ha

$$\begin{aligned} A \cup B &= A \cup (B \cap \Omega) \\ &= A \cup [B \cap (A \cup A^C)] \\ &= A \cup [(B \cap A) \cup (B \cap A^C)] && \text{(per la prop. distrib. dell'inter. rispetto all'uni.)} \\ &= [A \cup (B \cap A)] \cup (B \cap A^C) && \text{(per la prop. associativa dell'unione)} \\ &= A \cup (B \cap A^C) && \text{(perché } B \cap A \subseteq A\text{)} \end{aligned}$$

Essendo A e $B \cap A^C$ insiemi disgiunti, per la finita additività si ha
$P(A \cup B) = P(A) + P(B \cap A^C)$.
Inoltre, $B = B \cap (A \cup A^C) = (B \cap A) \cup (B \cap A^C)$ e quindi
$P(B \cap A^C) = P(B) - P(A \cap B)$.
Sostituendo quest'ultima nell'espressione precedente si ha la tesi.

Proprietà degli spazi di probabilità

In generale procedendo per induzione si dimostra la seguente relazione nota come **principio di inclusione-esclusione**.

Siano $A_1, A_2, \dots, A_n \in \mathcal{A}$. Allora

$$P\left(\cup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{(i,j)\in I_2} P(A_i \cap A_j) + \sum_{(i,j,k)\in I_3} P(A_i \cap A_j \cap A_k)$$
$$+ \dots + (-1)^{n-1} P(A_1 \cap A_2 \cap \dots A_n)$$

ove I_r è l'insieme dei sottoinsiemi delle r-uple non ordinate di elementi distinti di $\{1, 2, \dots, n\}$.

Osservazione. Nel caso particolare $n = 3$ si ha

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C).$$

Eventi condizionati

Dato un evento E, spesso è di interesse considerarlo in subordine ad una data circostanza espressa mediante un altro evento H.

Definizione. Un evento condizionato $E|H$ (da leggersi "E dato H" oppure "E noto H"), con $H \neq \emptyset$, è un ente a tre valori:

- VERO quando, essendo vero H, anche E è vero;
- FALSO quando, essendo vero H, risulta falso E;
- INDETERMINATO quando H risulta falso.

Osservazione. $E|\Omega = E$.

Esempio. Si consideri il lancio di un dado non truccato e gli eventi:

$$A = \text{"esce un numero non superiore a 3"}, \qquad H = \text{"esce un numero dispari"}.$$

L'evento condizionato $E|H$ può essere descritto dalla proposizione:

$$E|H = \text{"esce un numero non superiore a 3 sapendo che uscirà un numero dispari"}.$$

Probabilità condizionale

Definizione. Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e sia $H \in \mathcal{A}$ con $P(H) \neq 0$.
Dicesi **probabilità condizionale** di A noto H

$$P(A|H) = \frac{P(A \cap H)}{P(H)}, \qquad \text{per ogni } A \in \mathcal{A}.$$

Esempio. Nell'esempio precedente $A = \{1, 2, 3\}$ e $H = \{1, 3, 5\}$.
Dalla definizione si ha che

$$P(A|H) = \frac{2/6}{3/6} = \frac{2}{3}.$$

Invece, senza conoscere il verificarsi di H, la probabilità di A sarebbe $P(A) = 3/6 = 1/2$.

Definizione. Siano A e B due eventi. Essi si dicono **indipendenti** se $P(A \cap B) = P(A)P(B)$.

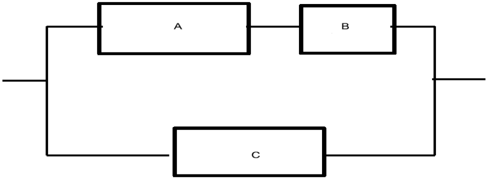
Proprietà. Se A e B sono due eventi indipendenti allora

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

cioè l'informazione aggiuntiva su A non modifica le previsioni su B.

Esercizio

Un congegno idraulico è costituito da tre valvole collegate come da figura. Le probabilità che le valvole A, B, C si guastino è rispettivamente pari a 1/3, 1/4, 1/2. Il congegno è funzionante se funziona almeno la coppia di valvole A, B o la valvola C. Si calcoli la probabilità che il congegno si guasti.



Consideriamo gli eventi:
G = "si guasta il congegno"
G_A = "si guasta la valvola A"
G_B = "si guasta la valvola B"
G_C = "si guasta la valvola C"

Possiamo inoltre ritenere che i guasti delle delle singole valvole siano tra di loro indipendenti.

Allora

$$P(G) = P((G_A \cup G_B) \cap G_C) = P(G_A \cup G_B)P(G_C) = [P(G_A) + P(G_B) - P(G_A \cap G_B)]P(G_C)$$

$$= [P(G_A) + P(G_B) - P(G_A)P(G_B)]P(G_C) = \left[\frac{1}{3} + \frac{1}{4} - \frac{1}{3} \cdot \frac{1}{4}\right]\frac{1}{2} = \frac{4+3-1}{12} \cdot \frac{1}{2} = \frac{1}{4}$$

Teorema delle probabilità totali

Siano dati n eventi $A_1, A_2, \ldots, A_n \in \mathcal{A}$ e supponiamo che essi formino una partizione di Ω, cioè $\cup_{i=1}^n A_i = \Omega$ e $A_i \cap A_j = \emptyset$ se $i \neq j$ con $P(A_i) > 0$ per ogni $i = 1, 2, \ldots, n$. Allora

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i), \qquad \text{per ogni } B \in \mathcal{A}.$$

Dimostrazione. Si ha

$$\begin{aligned} P(B) &= P(B \cap \Omega) = P(B \cap \cup_{i=1}^n A_i) \\ &= P(\cup_{i=1}^n (B \cap A_i)) && \text{(per la prop distrib. dell'uni. risp. all'inters.)} \\ &= \sum_{i=1}^{n} P(B \cap A_i) && \text{(per la finita additività)} \\ &= \sum_{i=1}^{n} P(B|A_i)P(A_i) && \text{(per def. di probabilità condizionale)} \end{aligned}$$

Esercizio

Una stazione radio riceve in maniera random segnali da due sorgenti, A e B. La probabilità di ricevere un segnale distorto da A è pari a 0.1 mentre la probabilità di ricevere un segnale distorto da B è pari a 0.2. Qual è la probabilità di ricevere un segnale distorto?

Consideriamo gli eventi:
$D =$ "si riceve un segnale distorto"
$S_A =$ "segnale da A"
$S_B =$ "segnale da B"

Utilizzando il teorema delle probabilità totali si ottiene

$$\begin{aligned} P(D) = P(D \cap \Omega) = P(D \cap (S_A \cup S_B)) = P(D|S_A)P(S_A) + P(D|S_B)P(S_B) \\ = 0.1 \cdot \frac{1}{2} + 0.2 \cdot \frac{1}{2} = 0.15 \end{aligned}$$

Teorema di Bayes

Siano $A, B \in \mathcal{A}$ con $P(A), P(B) > 0$. Allora

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}.$$

Dimostrazione. Dalla definizione di probabilità condizionale, essendo $P(A) > 0$, si ha

$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

D'altra parte, sempre dalla stessa definizione, poiché $P(B) > 0$, si ha anche

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad \text{da cui} \quad P(A \cap B) = P(A|B)P(B).$$

Sostituendo quest'ultima nell'espressione precedente si ha la tesi.

Corollario

Siano dati n eventi $A_1, A_2, \ldots, A_n \in \mathcal{A}$ e supponiamo che essi formino una partizione di Ω, cioè $\cup_{i=1}^{n} A_i = \Omega$ e $A_i \cap A_j = \emptyset$ se $i \neq j$ con $P(A_i) > 0$ per ogni $i = 1, 2, \ldots, n$ e sia inoltre $B \in \mathcal{A}$ con $P(B) > 0$. Allora, per il Teorema di Bayes,

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)},$$

ove l'ultimo passaggio si è ottenuto applicando il Teorema delle probabilità totali.

Definizione. La $P(A_i)$ viene detta **probabilità a priori** di A_i, la $P(A_i|B)$ viene detta **probabilità a posteriori** di A_i, le $P(B|A_i)$ per $i = 1, 2, \ldots, n$ costituiscono le **verosimiglianze**.

Esercizio

Una industria ha due catene di produzione, A e B, che forniscono rispettivamente il 40% e il 60% dei prodotti. Si scopre che il 25% dei pezzi prodotti dalla catena A sono difettosi contro il 7% dei pezzi prodotti nel centro B. Qual è la probabilità che un pezzo scelto a caso dall'intera produzione sia difettoso? Qual è la probabilità che un pezzo difettoso appartenga alla produzione A?

Consideriamo gli eventi:
$A =$ "il pezzo viene dalla catena A"
$B =$ "il pezzo viene dalla catena B"
$D =$ "il pezzo è difettoso"
Seguendo l'approccio frequentista assumiamo: $P(A) = \frac{40}{100} = \frac{2}{5}$, $P(B) = \frac{60}{100} = \frac{3}{5}$,
$P(D|A) = \frac{25}{100} = \frac{1}{4}$ e $P(D|B) = \frac{7}{100}$.
Per calcolare $P(D)$ applichiamo il Teorema delle probabilità totali

$$P(D) = P(D|A)P(A) + P(D|B)P(B) = \frac{1}{4} \cdot \frac{2}{5} + \frac{7}{100} \cdot \frac{3}{5} = 0.142$$

Per calcolare $P(A|D)$ applichiamo il corollario del Teorema di Bayes

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)} = \frac{1/4 \cdot 2/5}{0.142} \approx 0.704$$

Esempio: Il gioco di Monty-Hall

Un gioco televisivo prevede che un concorrente venga invitato a scegliere una fra tre porte dietro ad una delle quali si trova il premio. Dopo la scelta del concorrente il presentatore apre una delle due porte non scelte e dietro la quale non si trova il premio e invita poi il concorrente a decidere se mantenere la scelta iniziale o cambiare per l'altra porta ancora chiusa. Cambiare porta aumenta la probabilità di vittoria?

Introduciamo gli eventi:
$A_i =$ "premio dietro la porta i-esima", $\quad i = 1, 2, 3$
$C_i =$ "il presentatore apre la porta i-esima dopo la prima scelte del concorrente", $\quad i = 1, 2, 3$
Le probabilità a priori sono uguali, cioè $P(A_1) = P(A_2) = P(A_3) = 1/3$.

Senza ledere la generalità supponiamo che il concorrente abbia scelto la porta 3. A questo punto il conduttore aprirà la porta 1 o la porta 2. Senza ledere la generalità supponiamo che apra la porta 1.
Per il Teorema di Bayes

$$P(A_2|C_1) = \frac{P(C_1|A_2)P(A_2)}{P(C_1)} = \frac{1 \cdot 1/3}{1/2} = \frac{2}{3}$$

perché il presentatore conosce dove si trova il premio. Allora, prima che il concorrente scelga, il presentatore ha due scelte equiprobabili e pertanto $P(C_1) = 1/2$. Dopo la scelta del concorrente $P(C_1|A_2) = 1$ in quanto il conduttore ha una scelta obbligata.
Di conseguenza $P(A_3|C_1) = 1 - P(A_2|C_1) = 1 - 2/3 = 1/3$ e quindi conviene cambiare.

Richiami di calcolo combinatorio

Sia N un insieme finito di n elementi.

Definizione. Dicesi **r-disposizione** (con possibili ripetizioni) una r-upla ordinata $a_1, a_2, \ldots, a_r$ di elementi di N con gli a_i non necessariamente distinti.

Proprietà. Il numero di r-disposizioni di un insieme N di n elementi è pari a n^r.

Esempio. In un test a risposta chiusa VERO o FALSO composto da 10 domande, il numero di possibili risposte è 2^{10}.

Definizione. Una r-disposizione dicesi **semplice** se gli elementi a_i sono tutti distinti. Le r-disposizioni semplici si dicono anche disposizioni di n oggetti in classe r.

Chiaramente in questo caso $r \leq n$.

Proprietà. Il numero delle r-disposizioni semplici di n oggetti in classe r è dato da

$$D_{n,r} = \frac{n!}{(n-r)!}$$

Esempio. Il numero di parole di quattro lettere che si possono formare scegliendo quattro lettere distinte da un alfabeto di venti lettere è dato da $D_{20,4} = 20 \times 19 \times 18 \times 17 \times 16 = 1860480$.

Richiami di calcolo combinatorio

Definizione. Le disposizioni semplici di n oggetti in classe n si dicono **permutazioni**.

Proprietà. Il numero delle permutazioni di n oggetti è dato da $D_{n,n} = n!$.

Esempio. Il numero di modi in cui si possono ordinare 5 libri su uno scaffale è $D_{5,5} = 5! = 120$.

Definizione. Un insieme non ordinato di r elementi appartenenti a N, non necessariamente tutti distinti, dicesi **r-combinazione** (con possibili ripetizioni) di n elementi in classe r. Se un elemento figura m volte in una r-combinazione allora lo diremo di molteplicità m.

Definizione. Se gli elementi di una r-combinazione sono tutti distinti allora si ha una **combinazione semplice** di n elementi in classe r.

Proprietà. Il numero delle combinazioni semplici di n oggetti in classe r è dato da

$$C_{n,r} = \binom{n}{r} = \frac{n!}{r!(n-r)!}, \qquad r = 0, 1, \ldots, n$$

Esempio. Il numero delle possibili cinquine di numeri estratti in una ruota del lotto è dato da $C_{90,5} = \frac{90!}{5!85!} = 43949268$.

Richiami di calcolo combinatorio

Proprietà. Il numero delle r-combinazioni con possibili ripetizioni da un insieme di n elementi è dato da

$$\binom{n+r-1}{r}$$

Esempio. Si devono disporre r palline uguali in n scatole, ciascuna delle quali può contenere anche tutte le palline. Le possibili distribuzioni si ottengono al variare degli interi non negativi r_k tali che $r_1 + r_2 + \ldots + r_n = r$. Quindi le possibili distribuzioni sono tante quante sono le r-combinazioni con possibili ripetizioni di n oggetti in classe r.

Proprietà. Assegnati gli interi $n_1, n_2, \ldots, n_k$ con $n_1 + n_2 + \ldots + n_k = n$, il numero delle partizioni di N formate da insiemi aventi rispettivamente $n_1, n_2, \ldots, n_k$ elementi è dato da

$$C^n_{n_1,n_2,\ldots,n_k} = \frac{n!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

Esempio. Calcoliamo in quanti modi si possono distribuire 10 libri in tre scaffali in modo che il primo scaffale contenga due libri, il secondo 3 e il terzo 5. Il numero cercato è dato da $\frac{10!}{2!3!5!} = 2520$.

Esempio: Il paradosso del compleanno

Qual è la probabilità che tra $n \leq 365$ persone scelte a caso almeno due festeggino il compleanno lo stesso giorno?

Sia $S = \{1, 2, \ldots, 365\}$ l'insieme dei giorni dell'anno.

Allora prendiamo come insieme dei possibili risultati $\Omega = S^n$, cioè l'insieme delle n-uple che rappresentano i giorni di nascita delle n persone considerate.

Consideriamo l'evento

$$A = \{\omega \in \Omega \,:\, \omega \text{ ha almeno due componenti uguali}\}\,.$$

Calcoliamo

$$P(A) = 1 - P(A^C) = 1 - \frac{D_{365,n}}{365^n},$$

poiché A^C è costituito da tutte le n-uple formate da elementi distinti di S.
Si noti che è importante l'ordine con cui si succedono le componenti di ω e quindi dobbiamo considerare le disposizioni.

Per $n = 30$ si trova $P(A) \approx 0.507$, per $n = 50$ si ha $P(A) \approx 0.974$.

Esercizi proposti

Esercizio 1. In una rete idrica vi sono quattro valvole. Ciascuna valvola ha probabilità $1/4$ di guastarsi in un anno. Quando una valvola si guasta impedisce il flusso di acqua. Le valvole funzionano indipendentemente l'una dall'altra. Due valvole sono collegate in serie e queste due sono collegate a loro volta in parallelo con le altre due. Qual è la probabilità che la rete smetta di fornire acqua?

Esercizio 2. Per determinare una password di 3 caratteri si hanno a disposizione le lettere A, B, C, D. Provando a caso con tali lettere, qual è la probabilità di indovinare la password in un singolo tentativo? Qual è tale probabilità se le lettere devono essere distinte? Qual è tale probabilità se le lettere devono essere distinte e non conta l'ordine?

Esercizio 3. Per determinare una password di 3 caratteri si hanno a disposizione le lettere A, B, C, D, E. Provando a caso con tali lettere, qual è la probabilità di indovinare la password in un singolo tentativo? Qual è tale probabilità se si sa che una delle lettere che figurano nella password è la A? Qual è tale probabilità se si sa che nella password figurano sia la A che la B?

Variabili aleatorie

Variabili aleatorie

Definzione. Sia E un evento. Si chiama indicatore di E

$$|E| = \begin{cases} 1, & \text{se } E \text{ è VERO} \\ 0, & \text{se } E \text{ è FALSO} \end{cases}$$

Il concetto di variabile aleatoria generalizza quello di evento.
Intuitivamente è un numero univocamente definito ma non (o non ancora) conosciuto.
Supponiamo per il momento che Ω sia un insieme finito.

Esempio. Sia X il primo numero estratto al lotto sulla ruota di Roma di martedì 20 febbraio 2024. Consideriamo l'insieme

$$\Omega = \{1, 2, \ldots, 90\}$$

e gli eventi

$$E_i = \text{“il primo numero estratto è } i\text{”}, \qquad i = 1, 2, \ldots, 90.$$

Allora si può scrivere

$$X = 1|E_1| + 2|E_2| + \ldots + 90|E_{90}|.$$

Quindi $X = k$ se e solo se E_k è VERO.

Variabili aleatorie

In generale, dati n numeri reali $x_1, x_2, \ldots, x_n$ e n eventi $E_1, E_2, \ldots, E_n$ e posto

$$X = x_1|E_1| + x_2|E_2| + \ldots + x_n|E_n|,$$

X definisce una funzione reale che si chiama variabile aleatoria.

Esempio. Supponiamo di effettuare n lanci con una moneta e poniamo

$$E_i = \text{“TESTA all'}i\text{-esimo lancio”}$$

e $x_1 = x_2 = \ldots = x_n = 1$.
Allora la variabile aleatoria

$$X = x_1|E_1| + x_2|E_2| + \ldots + x_n|E_n| = |E_1| + |E_2| + \ldots + |E_n|$$

rappresenta il numero di volte che esce TESTA su n lanci.

Operazioni con le variabili aleatorie

Data la variabile aleatoria

$$X = x_1|E_1| + x_2|E_2| + \ldots + x_n|E_n|$$

e dato $\alpha \in \mathbb{R}$ si definisce il loro **prodotto** ponendo

$$\alpha X = \alpha x_1|E_1| + \alpha x_2|E_2| + \ldots + \alpha x_n|E_n|.$$

Data inoltre la variabile aleatoria

$$Y = y_1|A_1| + y_2|A_2| + \ldots + y_k|A_k|$$

si definisce la **somma** tra X e Y ponendo

$$X + Y = x_1|E_1| + x_2|E_2| + \ldots + x_n|E_n| + y_1|A_1| + y_2|A_2| + \ldots + y_k|A_k|.$$

Date due variabili aleatorie X e Y come sopra si definisce il loro **prodotto** ponendo

$$XY = \sum_{i=1}^{n} \sum_{j=1}^{k} x_i|E_i|y_j|A_j|.$$

Variabili aleatorie e loro distribuzioni

Più in generale si ha la seguente definizione.

Definizione. Dato uno spazio di probabilità $(\Omega, \mathcal{A}, P)$ si dice **variabile aleatoria** (abbr. v.a.) un'applicazione $X : \Omega \to \mathbb{R}$ tale che, per ogni $t \in \mathbb{R}$, l'insieme $\{\omega : X(\omega) \leq t\}$ appartenga ad $\mathcal{A}$.

Osservazione. Più in generale, sarà importante considerare insiemi del tipo $\{\omega : X(\omega) \in A\}$ dove A è un sottoinsieme di $\mathbb{R}$.

Osservazione. Per snellire la notazione, per ogni $a, b \in \mathbb{R}$ con $a < b$, porremo

$$\{X \leq a\} = \{\omega : X(\omega) \leq a\}$$
$$\{X > a\} = \{\omega : X(\omega) > a\}$$
$$\{a < X \leq b\} = \{\omega : a < X(\omega) \leq b\}$$

e scriveremo anche, ad esempio, $P(\{a < X \leq b\}) = P(a < X \leq b)$.

Variabili aleatorie e loro distribuzioni

Proprietà. Sia X una variabile aleatoria e $a, b \in \mathbb{R}$ con $a < b$. Allora si ha

1. $\{X > a\} \in \mathcal{A}$;
2. $\{a < X \leq b\} \in \mathcal{A}$;
3. $\{X = a\} \in \mathcal{A}$.

Definizione. Data una variabile aleatoria X, l'applicazione

$$A \to P\Big(\{\omega \, : \, X(\omega) \in A\}\Big)$$

per ogni $A \subseteq \mathbb{R}$ si dice **legge** o **distribuzione** di X.

Si osservi che tutti gli insieme del tipo $(a < X < b)$ per ogni $a, b \in \mathbb{R}$ devono essere eventi. Dato che tali insiemi costituiscono una base per gli aperti, è naturale richiedere che $\mathcal{A}$ contenga la famiglia degli aperti. La più piccola σ-algebra contenente la famiglia degli insiemi aperti di $\mathbb{R}$ è la corrispondente σ-algebra di Borel e si indica con $\mathcal{B}$.

Variabili aleatorie e loro distribuzioni

Definizione. Al variare di $A \in \mathcal{B}$ l'applicazione $P(X \in A)$ definisce la legge di X.

Nello studio delle leggi delle variabili aleatorie distingueremo due casi: nel primo X può assumere un numero finito o una infinità numerabile di valori e parleremo di **variabili aleatorie discrete**; nel secondo caso X potrà assumere valori su tutto $\mathbb{R}$ o su in suo intervallo e parleremo di **variabili aleatorie continue**.

Variabili aleatorie discrete

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità.

Definizione. Una v.a. $X : \Omega \to \mathbb{R}$ si dice **discreta** se l'insieme $X(\Omega)$ è finito o al più numerabile ed è privo di punti di accumulazione.

Osservazione. Dato $A \in \mathcal{B}$ si ha

$$P(X \in A) = P\left(X \in \cup_{a \in A}\{a\}\right) = \sum_{a \in A} P(X = a)$$

ove la somma è finita oppure è una serie.

Definizione. L'applicazione $p(x) = P(X = x)$ al variare di $x \in \mathbb{R}$ si dice **densità discreta** di X.

Proprietà. La densità discreta $p(x)$ gode delle seguenti proprietà

1. $p(x) \geq 0$ per ogni $x \in \mathbb{R}$ e $p(x) = 0$ tranne al più una infinità numerabile di elementi di $\mathbb{R}$;
2. $\sum_{x \in \mathbb{R}} p(x) = P(X \in \mathbb{R}) = 1$, ove la somma è intesa sui valori per cui $p(x) \neq 0$.

Variabili aleatorie discrete

Definizione. Dicesi **funzione di ripartizione** (f.r.) della v.a. X l'applicazione $F_X : \mathbb{R} \to [0,1]$ definita da

$$F_X(t) = P(X \leq t).$$

Osservazione. Nota la densità $p(x)$ si ha $F_X(t) = \sum_{x \leq t} p(x)$.

Definizione. Sia X una v.a. discreta e $x_1, x_2, \ldots, x_i, \ldots$ i valori che essa assume. Si dice che X ammette speranza matematica finita se

$$\sum_{i=1}^{\infty} |x_i| p(x_i) < +\infty$$

in tal caso si definisce **speranza matematica** (o **media** o **valore atteso**) di X la quantità

$$E[X] = \sum_{i=1}^{\infty} x_i p(x_i).$$

Variabili aleatorie discrete

Proprietà. Sia X una v.a. e $F : \mathbb{R} \to \mathbb{R}$ tale che $F(X)$ sia pure una v.a.. $Z = F(X)$ ammette speranza matematica finita se e solo se

$$\sum_{i=1}^{\infty} |F(x_i)|p(x_i) < +\infty$$

e in tal caso

$$E[F(X)] = \sum_{i=1}^{\infty} F(x_i)p(x_i).$$

Esempio. A partire dalla densità di una v.a. X, grazie alla precedente proprietà, è possibile calcolare direttamente $E[X^2]$.

Variabili aleatorie discrete

Definizione. Data una v.a. discreta X e k un intero positivo, si dice che X ammette momento finito di ordine k se esiste finita la speranza matematica di X^k e in tal caso $E[X^k]$ si dice **momento** di ordine k di X. Analogamente X ammette momento finito centrato di ordine k se esiste finita la speranza matematica di $(X - E[X])^k$ e in tal caso $E[(X - E[X])^k]$ si dice **momento centrato** di ordine k di X.

Dalla proprietà precedente si ha che i momenti, se esistono, si calcolano tramite le formule

$$E[X^k] = \sum_i x_i^k p(x_i)$$

$$E[(X - E[X])^k] = \sum_i (x_i - E[X])^k p(x_i)$$

Definizione. Sia X una v.a. avente momento centrato del secondo ordine finito. In tal caso tale momento dicesi **varianza** di X e si scrive

$$\sigma^2 = \mathrm{VAR}(X) = E[(X - E[X])^2]$$

e la radice aritmetica della varianza, $\sigma = \sqrt{\mathrm{VAR}(X)}$, viene detta **deviazione standard**.

Proprietà delle v.a. discrete

Proprietà. Siano X e Y due v.a. che ammettono speranza matematica finita. Allora

1. $E[cX] = cE[X]$ per ogni $c \in \mathbb{R}$;
2. $E[X + Y] = E[X] + E[Y]$;
3. Se $P(X \leq Y) = 1$ allora $E[X] \leq E[Y]$;
4. $|E[X]| \leq E[|X|]$.

Proprietà. Se la v.a. X ammette momento finito di ordine k allora ammette anche momento fiito di ordine r per ogni $0 < r \leq k$.

Proprietà. Sia X una v.a. che ammette varianza finita. Allora

1. $\mathrm{VAR}(X) = E[X^2] - (E[X])^2$;
2. $\mathrm{VAR}(aX) = a^2\mathrm{VAR}(X)$ per ogni $a \in \mathbb{R}$;
3. $\mathrm{VAR}(a + X) = \mathrm{VAR}(X)$ per ogni $a \in \mathbb{R}$.

Esempio: Distribuzione discreta uniforme

Una variabile aleatoria X segue una distribuzione discreta uniforme su un insieme finito Ω di numerosità n se ha come densità discreta $p(x)$ quella che attribuisce a tutti gli elementi x di Ω la stessa probabilità p di verificarsi, cioè

$$p(x) = \frac{1}{n}, \qquad \forall x \in \Omega.$$

Esempio. Nel lancio di un dado, la probabilità di uscita di ciascuna faccia è 1/6.

Senza perdita di generalità, si può identificare Ω con l'insieme dei numeri da 1 a n,

$$\Omega = \{1, 2, \ldots, n\}.$$

In questo caso, determiniamo la funzione di ripartizione $F_X(t)$. Per ogni $t \in \mathbb{R}$ si ha

$$F_X(t) = \begin{cases} 0 & \text{se } t < 1 \\ 1/n & \text{se } 1 \le t < 2 \\ 2/n & \text{se } 2 \le t < 3 \\ \vdots & \\ 1 & \text{se } t \ge n \end{cases}$$

Esempio: Distribuzione discreta uniforme

Media.

$$E[X] = \sum_{i=1}^{n} i p(i) = \sum_{i=1}^{n} i \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Varianza.

$$E[X^2] = \sum_{i=1}^{n} i^2 p(i) = \sum_{i=1}^{n} i^2 \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}.$$

$$\mathrm{VAR}(X) = E[X^2] - (E[X])^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

Alcune importanti distribuzioni discrete:

Vedi blocco appunti 2

Variabili aleatorie continue

Passiamo adesso a considerare le v.a. continue, rimuovendo l'ipotesi che $X(\Omega)$ sia al più numerabile. Introdotta la funzione di ripartizione come nel caso discreto $F(t) = P(X \leq t)$, si dimostrano le seguenti proprietà.

Proprietà. Valgono le seguenti relazioni:

1. $0 \leq F(t) \leq 1$ per ogni $t \in \mathbb{R}$;
2. $F(t_1) \leq F(t_2)$ per ogni $t_1, t_2 \in \mathbb{R}$ con $t_1 \leq t_2$;
3. $\lim_{t \to -\infty} F(t) = 0$, $\lim_{t \to +\infty} F(t) = 1$;
4. F è continua a destra, cioè $\lim_{\varepsilon \to 0^+} F(t + \varepsilon) = F(t)$.

Osservazione. In generale la funzione di ripartizione non è continua, ma essendo monotona esistono finiti

$$\lim_{\varepsilon \to 0^-} F(t + \varepsilon) = F(t^-), \qquad \lim_{\varepsilon \to 0^+} F(t + \varepsilon) = F(t^+).$$

Il salto di $F(t)$ ha il significato seguente

$$F(t^+) - F(t^-) = P(X = t).$$

Se t è un punto di continuità per la funzione di ripartizione allora $P(X = t) = 0$.

Variabili aleatorie continue

Definizione. Una v.a. dicesi **continua** se tale è la sua funzione di ripartizione.

Definizione. Una applicazione $f : \mathbb{R} \to \mathbb{R}$ rappresenta una **densità** se gode delle seguenti proprietà:

1. $f(x) \geq 0$ per ogni $x \in \mathbb{R}$;
2. $\int_{-\infty}^{+\infty} f(x)\, dx = 1$.

Definizione. Una v.a. X dicesi assolutamente continua se ammette densità, cioè se esiste una densità $f(x)$ tale che $F(t) = \int_{-\infty}^{t} f(x)\, dx$.

Osservazione. La densità non è unica ma definita a meno di un sottoinsieme di misura nulla.

Proprietà. Se X è una v.a. continua allora

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = F(b) - F(a).$$

Variabili aleatorie continue

Definizione. Sia X una v.a. continua di densità $f(x)$. Si dice che X ammette speranza matematica finita se

$$\int_{-\infty}^{+\infty} |x| f(x)\, dx < +\infty$$

in tal caso si definisce **speranza matematica** (o **media** o **valore atteso**) di X la quantità

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\, dx.$$

Variabili aleatorie continue

Definizione. Data una v.a. discreta X di dentistà $f(x)$ e k un intero positivo, si dice che X ammette momento finito di ordine k se esiste finita la speranza matematica di X^k e in tal caso $E[X^k]$ si dice **momento** di ordine k di X. Analogamente X ammette momento finito centrato di ordine k se esiste finita la speranza matematica di $(X - E[X])^k$ e in tal caso $E[(X - E[X])^k]$ si dice **momento centrato** di ordine k di X.

Si ha che i momenti, se esistono, si calcolano tramite le formule

$$E[X^k] = \int_{-\infty}^{+\infty} x^k f(x)\, dx$$

$$E[(X - E[X])^k] = \int_{-\infty}^{+\infty} (x - E[X])^k f(x)\, dx$$

Definizione. Sia X una v.a. avente momento centrato del secondo ordine finito. In tal caso tale momento dicesi **varianza** di X e si scrive

$$\sigma^2 = \mathrm{VAR}(X) = E[(X - E[X])^2]$$

e la radice aritmetica della varianza, $\sigma = \sqrt{\mathrm{VAR}(X)}$, viene detta **deviazione standard**.

Proprietà delle v.a. continue (analoghe al caso discreto)

Proprietà. Siano X e Y due v.a. che ammettono speranza matematica finita. Allora

1. $E[cX] = cE[X]$ per ogni $c \in \mathbb{R}$;
2. $E[X+Y] = E[X] + E[Y]$;
3. Se $P(X \leq Y) = 1$ allora $E[X] \leq E[Y]$;
4. $|E[X]| \leq E[|X|]$.

Proprietà. Se la v.a. X ammette momento finito di ordine k allora ammette anche momento fiito di ordine r per ogni $0 < r \leq k$.

Proprietà. Sia X una v.a. che ammette varianza finita. Allora

1. $\mathrm{VAR}(X) = E[X^2] - (E[X])^2$;
2. $\mathrm{VAR}(aX) = a^2\mathrm{VAR}(X)$ per ogni $a \in \mathbb{R}$;
3. $\mathrm{VAR}(a+X) = \mathrm{VAR}(X)$ per ogni $a \in \mathbb{R}$.

Quantili

Definizione. Assegnato $\alpha \in [0,1]$, si definisce **quantile** di ordine α di una v.a. X avente f.r. $F_X(t)$ il valore

$$q_\alpha = \sup\left\{r \in \mathbb{R} \,:\, F_X(r) \leq \alpha\right\}.$$

Osservazione. Nel caso in cui $F_X(t)$ sia strettamente crescente allora esiste un unico valore $r \in \mathbb{R}$ tale che $F(r) = \alpha$. Se X è assolutamente continua e la sua densità è strettamente positiva allora $F(t)$ è certamente monotona e in tal caso

$$\alpha = \int_{-\infty}^{q_\alpha} f(x)\,dx,$$

ove $f(x)$ è la densità di X.

Osservazione. Il quantile di ordine 0.5 dicesi mediana m_e e rappresenta il valore per cui $P(X \leq m_e) = P(X \geq m_e) = 1/2$. I tre quantili $q_{1/4}$, $q_{2/4}$, $q_{3/4}$ costituiscono rispettivamente il primo, il secondo e il terzo quartile. Si osservi che il secondo quartile coincide con la mediana.

Alcune importanti distribuzioni continue:

Vedi blocco appunti 3

Indipendenza e correlazione di variabili aleatorie

Variabili aleatorie discrete multidimensionali

Definizione. Una v.a. discreta **m-dimensionale** è una applicazione $X : \Omega \to \mathbb{R}^m$ che associa

$$\omega \longmapsto (X_1(\omega), X_2(\omega), \dots, X_m(\omega))$$

e tale che le sue componenti X_i sono v.a. discrete.

Definizione. Data una v.a.discreta m-dimensionale, dicesi **densità congiunta** di X la funzione $p : \mathbb{R}^m \to [0, 1]$ definita da

$$p(x) = P(X = x) = P(X_1 = x_1, X_2 = x_2, \dots, X_m = x_m)$$

per ogni $x = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$.

Definizione. Le densità $p_1(x_1), p_2(x_2), \dots, p_m(x_m)$ delle componenti $X_1, X_2, \dots, X_m$ si dicono **densità marginali**.

Variabili aleatorie indipendenti

Osservazione. Se è nota la densità congiunta restano determinate quelle marginali. Consideriamo, infatti, per semplicità il caso bidimensionale $X = (X_1, X_2)$. Si ha

$$\begin{aligned} p_1(x_1) &= P(X_1 = x_1) = P(X_1 = x_1, X_2 \in \mathbb{R}) = P(X_1 = x_1, \cup_{z\in\mathbb{R}} \{X_2 = z\}) \\ &= \sum_{z\in\mathbb{R}} P(X_1 = x_1, X_2 = z) = \sum_{z\in\mathbb{R}} p(x_1, z). \end{aligned}$$

Tuttavia in generale dalla conoscenza delle densità marginali non è possibile determinare univocamente quella congiunta.

Definizione. Siano $X_1, X_2, \ldots, X_n$ v.a definite su di uno stesso spazio di probabilià. Esse si dicono **indipendenti** se comunque si scelgono $A_1, A_2, \ldots, A_n \in \mathcal{B}$ si ha che

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdot P(X_2 \in A_2) \cdot \ldots \cdot P(X_n \in A_n).$$

Variabili aleatorie indipendenti

Proprietà. Siano X e Y due v.a. indipendenti aventi speranza matematica finita. Allora anche XY ha speranza matematica finita e vale $E[XY] = E[X]E[Y]$.

Proprietà. Le v.a. $X_1, X_2, \ldots, X_n$ definite su di uno stesso spazio di probabilità sono indipendenti se e solo se

$$p(x_1, x_2, \ldots, x_n) = p_1(x_1) \cdot p_2(x_2) \cdot \ldots \cdot p_n(x_n), \qquad \forall (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n,$$

ove $p(x_1, x_2, \ldots, x_n)$ rappresenta la densità congiunta e le $p_i(x_i)$ sono le densità marginali.

Definizione. Date due v.a. X e Y di densità congiunta $p(x, y)$, dicesi **densità condizionale** di X per $Y = y$, la quantità

$$p_{X|Y}(x|y) = \begin{cases} \dfrac{p(x,y)}{p_Y(y)} & \text{se } p_Y(y) > 0 \\ 0 & \text{altrimenti} \end{cases}$$

ove $p_Y(y)$ è la densità marginale di Y.

Proprietà. Se X e Y sono v.a. indipendenti allora $p_{X|Y}(x|y) = p_X(x)$.

Somma di variabili aleatorie

Proprietà. Siano X e Y due v.a. di densità congiunta $p(x, y)$. Allora la v.a. $Z = X + Y$ ha densità $p_Z(z)$ data da

$$p_Z(z) = \sum_{t \in \mathbb{R}} p(t, z - t) = \sum_{t \in \mathbb{R}} p(z - t, t), \qquad \forall t \in \mathbb{R}.$$

Corollario. Nelle stesse ipotesi del teorema precedente, se X e Y sono v.a. indipendenti allora

$$p_Z(z) = \sum_{t \in \mathbb{R}} p_X(t) p_Y(z - t) = \sum_{t \in \mathbb{R}} p_X(z - t) p_Y(t), \qquad \forall t \in \mathbb{R}$$

ove p_X e p_Y sono rispettivamente le densità marginali di X e Y.

Covarianza

Definizione. Date due v.a. X e Y, si definisce, se esiste finita, **covarianza** tra X e Y

$$\mathrm{COV}(X,Y) = E[(X - E[X])(Y - E[Y])].$$

Proprietà. Se esiste la covarianza tra le v.a. X e Y allora

$$\mathrm{COV}(X,Y) = E[XY] - E[X]E[Y].$$

In particolare se X e Y sono v.a. indipendenti allora $\mathrm{COV}(X,Y) = 0$.

Proprietà. Se X e Y ammettono varianza finita allora

$$\mathrm{VAR}(X+Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y) + 2\mathrm{COV}(X,Y).$$

In particolare se X e Y sono v.a. indipendenti allora

$$\mathrm{VAR}(X+Y) = \mathrm{VAR}(X) + \mathrm{VAR}(Y).$$

Correlazione lineare

Definizione. Siano X e Y due v.a. aventi varianza finita. Si definisce **coefficiente di correlazione lineare** tra X e Y la quantità

$$\rho(X,Y) = \frac{\mathrm{COV}(X,Y)}{\sqrt{\mathrm{VAR}(X)\mathrm{VAR}(Y)}}.$$

Proprietà. Si ha $-1 \leq \rho(X,Y) \leq 1$ e $\rho(X,Y) = \pm 1$ se e solo se X e Y sono legate da una dipendenza lineare, cioè $Y = aX + b$ con $a, b \in \mathbb{R}$.

Variabili aleatorie continue multidimensionali

Definizione. La v.s. $X = (X_1, X_2, \ldots, X_m)$ costituisce una v.a. m-dimensionale se ciascuna X_i è una v.a. monodimensionale. Le densità delle singole X_i, qualora esistano, costituiscono le **densità marginali** e si indicano con f_{X_i}.

Definizione. Diremo che X ammette densità, ovvero che X è assolutamente continua come v.a. m-dimensionale, se esiste una applicazione $f : \mathbb{R}^m \to \mathbb{R}$ non negativa tale che

$$\begin{aligned} F_X(x_1, x_2, \ldots, x_m) &= P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_m \leq x_m) \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_m} f(t_1, t_2, \ldots, t_m)\, dt_1\, dt_2 \cdots dt_m \end{aligned}$$

per ogni $(x_1, x_2, \ldots, x_m) \in \mathbb{R}^m$, ove $F_X(x_1, x_2, \ldots, x_m)$ rappresenta la f.r. di X e f dicesi densità congiunta di X.

Variabili aleatorie continue multidimensionali

Consideriamo il caso $m = 2$, cioè $X = (X_1, X_2)$.

Analogamente al caso discreto, nota la densità congiunta $f(x_1, x_2)$ le densità marginali sono date da

$$f_{X_1}(x_1) = \int_{\mathbb{R}} f(x_1, t)\, dt, \qquad f_{X_2}(x_2) = \int_{\mathbb{R}} f(t, x_2)\, dt.$$

Dalla conoscenza delle singole densità marginali non si può in generale risalire alla densità congiunta tranne che le componenti siano v.a. indipendenti.

Proprietà. Sia $X = (X_1, X_2)$ una v.a. bidimensionale di densità congiunta $f(x_1, x_2)$. Allora X_1 e X_2 sono v.a. indipendenti se e solo se $f(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2)$, essendo f_{X_1} e f_{X_2} le densità marginali relative a X_1 e X_2 rispettivamente.

Esempio: Legge normale bivariata (vedi blocco appunti 3).

Riferimenti bibliografici

1. V. Romano, Metodi matematici per i corsi di ingegneria, Città Studi, 2018.
2. P. Baldi, Calcolo delle probabilità e statistica , Mc Graw-Hill, Milano, 1992.
3. R. Scozzafava, Incertezza e probabilità, Zanichelli, 2001.