

Definizione 1. Sia Ω un insieme e sia $\mathcal{P}(\Omega)$ il suo insieme delle parti. Una *probabilità* su Ω è una funzione $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$ con le seguenti proprietà:

- (1) $\mathbb{P}(E) \geq 0$ per ogni evento $E \in \mathcal{P}(\Omega)$;
- (2) $\mathbb{P}(\Omega) = 1$;
- (3) per ogni famiglia $\{E_i\}_{i \in I}$ *finita o numerabile* di eventi $E_i \in \mathcal{P}(\Omega)$ tali che $E_i \cap E_j = \emptyset$ se $i \neq j$, si ha l'uguaglianza $\mathbb{P}\left(\bigcup_{i \in I} E_i\right) = \sum_{i \in I} \mathbb{P}(E_i)$.

Definizione 2. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità e siano $E, F \in \mathcal{F}$ due eventi. Supponiamo $\mathbb{P}(F) > 0$. La *probabilità di E condizionata a F* (o *probabilità di E sapendo F*) è il numero reale

$$\mathbb{P}(E \mid F) := \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Dal fatto che $E \cap F \subseteq F$ segue $\mathbb{P}(E \cap F) \leq \mathbb{P}(F)$, e dunque $\mathbb{P}(E \mid F) \in [0, 1]$. Inoltre, è facile verificare che la funzione $\mathbb{P}_F : \mathcal{F} \rightarrow \mathbb{R}$ data da

$$\mathbb{P}_F(E) = \mathbb{P}(E \mid F) \quad \text{per ogni } E \in \mathcal{F}$$

Definizione 3. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Due eventi $E, F \in \mathcal{F}$ si dicono *indipendenti* se $\mathbb{P}(E \cap F) = \mathbb{P}(E) \mathbb{P}(F)$.

Definizione 4. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità. Una collezione di n eventi $E_1, E_2, \dots, E_n \in \mathcal{F}$ si dicono *indipendenti* se

$$\mathbb{P}(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_k}) = \mathbb{P}(E_{i_1}) \mathbb{P}(E_{i_2}) \dots \mathbb{P}(E_{i_k}) \quad (1.5)$$

per ogni $k \leq n$ e per ogni sottoinsieme di indici $\{i_1, i_2, \dots, i_k\} \subseteq \{1, 2, \dots, n\}$.

Definizione 5. Supponendo per semplicità che $\mathcal{F} \equiv \mathcal{P}(\Omega)$, una qualunque funzione $X : \Omega \rightarrow \mathbb{R}$ si dice *variabile aleatoria*.

Notiamo che la controimmagine di un insieme $A \subseteq \mathbb{R}$ rispetto a una variabile aleatoria X è un *evento*:

$$X^{-1}(A) := \{\omega \in \Omega \mid X(\omega) \in A\} \in \mathcal{P}(\Omega) \quad \text{per ogni } A \subseteq \mathbb{R}.$$

Definizione 6. Sia X una variabile aleatoria. La *media* (o *speranza*) di X è il numero reale $\mathbb{E}[X]$ definito come segue:

- se X è discreta con densità $p_X : S \rightarrow [0, 1]$,

$$\mathbb{E}[X] = \sum_{x \in S} x p_X(x); \quad (1.10)$$

- se X è assolutamente continua con densità f_X ,

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x) \, dx. \quad (1.11)$$

Definizione 7. La *covarianza* di due variabili aleatorie X e Y è il numero reale $\text{Cov}(X, Y)$ definito come segue

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

Da notare che, al contrario della media, la covarianza ha come argomento *due* variabili aleatorie, e non una sola.

Dalla Proposizione 4 possiamo ricavare l'espressione esplicita di $\text{Cov}(X, Y)$ nei due casi in cui (X, Y) è un vettore aleatorio discreto

$$\text{Cov}(X, Y) = \sum_{x,y} (x - \mathbb{E}[X])(y - \mathbb{E}[Y])p_{(X,Y)}(x, y)$$

oppure assolutamente continuo

$$\text{Cov}(X, Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mathbb{E}[X])(y - \mathbb{E}[Y])f_{(X,Y)}(x, y) \, dx \, dy.$$

La proposizione seguente riassume le principali proprietà della covarianza.

Definizione 8. La *varianza* di una variabile aleatoria X è il numero reale $\text{Var}(X)$ dato da

$$\text{Var}(X) = \text{Cov}(X, X) = \mathbb{E}[(X - \mathbb{E}[X])^2] .$$

L'espressione esplicita di $\text{Var}(X)$ nei due casi in cui X è discreta o assolutamente continua può ancora essere ricavata tramite la Proposizione 4, ottenendo

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \begin{cases} \sum_x (x - \mathbb{E}[X])^2 p_X(x) & \text{se } X \text{ è discreta} \\ \int_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f_X(x) dx & \text{se } X \text{ è assolutamente continua} \end{cases}$$

Notiamo in particolare che in entrambi i casi $\text{Var}(X)$ è *sempre un numero positivo*, in quanto somma o integrale di quantità positive. Possiamo prenderne pertanto la radice quadrata: la quantità $\sqrt{\text{Var}(X)}$ si chiama *deviazione standard* di X .

La varianza gode delle seguenti proprietà fondamentali, che vanno confrontate con le analoghe proprietà della media.

Definizione 9. Un *campione aleatorio* di numerosità n è una successione di variabili aleatorie X_1, X_2, \dots, X_n tali che

- (a) le variabili aleatorie X_1, X_2, \dots, X_n sono indipendenti;
- (b) tutte le variabili aleatorie X_1, X_2, \dots, X_n hanno la stessa densità.

Un campione aleatorio è pertanto una successione di variabili aleatorie *indipendenti e identicamente distribuite* (i.i.d.). In particolare, per l'uguaglianza delle loro densità, tutte le X_i hanno la stessa media e la stessa varianza: $\mathbb{E}[X_i] = \mathbb{E}[X_j]$ e $\text{Var}(X_i) = \text{Var}(X_j)$ per ogni $i \neq j$, in quanto sia la media che la varianza dipendono solo dalla densità.

Per chiarire ancora meglio la definizione, supponiamo per fissare le idee che le X_i siano tutte assolutamente continue e ciascuna abbia densità f_{X_i} . Allora il punto (b) richiede che $f_{X_i} = f_{X_j} =: f$ per ogni $i \neq j$, dove f è la densità comune. Il punto (a) significa invece che la densità congiunta è

$$f_{(X_1, X_2, \dots, X_n)}(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \dots f_{X_n}(x_n) = f(x_1)f(x_2) \dots f(x_n).$$

Definizione 10. Sia X_1, \dots, X_n un campione aleatorio. Una *statistica* $T_n = t_n(X_1, \dots, X_n)$ è una funzione del vettore aleatorio $\vec{X} = (X_1, \dots, X_n)$.

Sono esempi di statistiche:

(i) la *media campionaria*

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i$$

in cui la funzione t_n è

$$t_n(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^n x_i;$$

(ii) la *varianza campionaria*

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

in cui la funzione t_n è

$$t_n(x_1, \dots, x_n) = \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{i=1}^n x_i \right)^2;$$

(iii) la *deviazione standard campionaria* $S_n := \sqrt{S_n^2}$.

Definizione 11. Sia X_1, \dots, X_n un campione aleatorio. Supponiamo di aver fatto una misura di X_1, \dots, X_n e aver trovato i valori x_1, \dots, x_n ($x_i \in \mathbb{R}$). Allora diremo che x_1, \dots, x_n è la *realizzazione* del campione ottenuta nella nostra misura.

Definizione 12. Sia X una variabile aleatoria continua con densità f_X . Dato $\alpha \in (0, 1)$, il *quantile* (destro) di ordine α della densità di X è il numero $x_\alpha \in \mathbb{R}$ t.c.

$$\mathbb{P}(X \geq x_\alpha) = \alpha.$$

Equivalentemente, il quantile x_α è il numero reale definito da

$$F_X(x_\alpha) = 1 - \alpha, \quad \int_{x_\alpha}^{+\infty} f_X(x) \, dx = \alpha,$$

dove F_X è la funzione di ripartizione di X . Si osservi che la funzione $\alpha \mapsto x_\alpha$ è *decescente*, in quanto al crescere di α deve aumentare l'area sottesa dalla densità alla destra di x_α , e quindi x_α deve spostarsi a sinistra.

Definizione 13. Sia X_1, \dots, X_n un campione aleatorio, dove le X_i sono variabili aleatorie assolutamente continue con densità

$$f_{X_i}(x) \equiv f(x \mid \theta)$$

dipendente da un parametro incognito θ . Siano $A_n = a_n(X_1, \dots, X_n)$ e $B_n = b_n(X_1, \dots, X_n)$ due statistiche per le quali valga sempre $A_n < B_n$, e tali che la probabilità dell'evento $\{A_n \leq \theta \leq B_n\}$ sia

$$\mathbb{P}(A_n \leq \theta \leq B_n) = 1 - \alpha,$$

dove $\alpha \in (0, 1)$ è un numero reale *che non dipende dal parametro θ* . Supponiamo di aver fatto una misura di X_1, \dots, X_n e aver trovato la realizzazione x_1, \dots, x_n . Allora diremo che l'intervallo

$$(a_n(x_1, \dots, x_n), b_n(x_1, \dots, x_n))$$

è un *intervallo di confidenza di livello $1 - \alpha$* per il parametro incognito θ (scriveremo per brevità $IC_\theta(1 - \alpha)$). Equivalentemente, si dice che

$$\theta \in (a_n(x_1, \dots, x_n), b_n(x_1, \dots, x_n))$$

con confidenza $1 - \alpha$.

Definizione 14. Sia X_1, \dots, X_n un campione. Un'*ipotesi statistica* è un'affermazione su uno o più parametri incogniti della densità delle X_i .

Definizione 15. Un *test d'ipotesi* è una regola di decisione tra H_0 e H_1 , e consiste nella seguente procedura:

1. fisso una *statistica test* $T = t(X_1, \dots, X_n)$;
2. stabilisco una *regione di rifiuto* (o *regione critica*) $C \subset \mathbb{R}$;
3. *rifiuto* H_0 se con le mie misure x_1, \dots, x_n trovo $t(x_1, \dots, x_n) \in C$; in caso contrario, la accetto.

Definizione 16. In un test d'ipotesi, commetto un errore:

- (a) *di I specie*, se rifiuto H_0 quando H_0 in realtà è vera;
- (b) *di II specie*, se accetto H_0 quando H_0 in realtà è falsa.

La definizione dei due tipi di errore è chiarita dalla seguente tabella.

	Accetto H_0	Rifiuto H_0
H_0 è vera	OK	Errore I specie
H_0 è falsa	Errore II specie	OK

Definizione 17. Il *livello di significatività* α di un test è la probabilità di commettere l'errore di I specie, cioè

$$\alpha := \mathbb{P}_{H_0} (\text{“rifiuto } H_0\text{”}) = \mathbb{P}_{H_0} (T \in C) .$$

Definizione 18. In un test d'ipotesi, il *p-value* è il più piccolo livello di significatività α che impone di rifiutare H_0 in base ai dati x_1, \dots, x_n che abbiamo misurato:

$$\begin{cases} \alpha > p\text{-value} & \Rightarrow & \text{rifiuto } H_0 \\ \alpha < p\text{-value} & \Rightarrow & \text{accetto } H_0 \end{cases} .$$

Quanto più il *p-value* è piccolo, tanto più l'ipotesi nulla H_0 è poco verosimile. Infatti, H_0 può essere accettata solo per valori di significatività ancora più piccoli del *p-value*.