

Introduzione alla probabilità

Definizione

Un **sample space** Ω di un esperimento randomico è l'insieme di tutti i possibili esiti dell'esperimento la cui occorrenza o non occorrenza può essere stabilita in modo non ambiguo da un osservatore.

Definizione

Un **evento** è un qualunque insieme di esiti che si può realizzare; è dunque un qualunque sottoinsieme del sample space.

Definizione

Un'informazione (o **σ -algebra degli eventi**) è una particolare famiglia di eventi che gode di particolari proprietà (che sono elencate di seguito).

Proprietà di un'algebra degli eventi \mathcal{E}

- 1) $\mathcal{E} \neq \emptyset \iff \emptyset \in \mathcal{E}$
- 2) Se $E_1, E_2 \in \mathcal{E} \implies E_1 \cup E_2 \in \mathcal{E}$
- 3) Se $E \in \mathcal{E} \implies E^c \in \mathcal{E}$

Proprietà di una σ -algebra degli eventi \mathcal{E}

- 1) $\mathcal{E} \neq \emptyset \iff \emptyset \in \mathcal{E}$
- 2) $(E_n)_{n \in \mathbb{N}} : E_n \in \mathcal{E} \implies \bigcup_{n=1}^{+\infty} E_n \in \mathcal{E}$
- 3) Se $E \in \mathcal{E} \implies E^c \in \mathcal{E}$

Proposizione

Consideriamo la partizione di eventi $\mathcal{P} \equiv \{E_j\}_{j \in N}$, dove E_j appartiene a una certa σ -algebra \mathcal{E} e $E_{j_1} \cap E_{j_2} = \emptyset \quad \forall j_1 \neq j_2$. Assumiamo che la partizione \mathcal{P} sia numerabile. Allora la σ -algebra generata da \mathcal{P} (che indichiamo con $\sigma(\mathcal{P})$) è la famiglia di tutti gli eventi che possiamo scrivere come l'unione di un insieme di pezzi di \mathcal{P} . In simboli:

$$\sigma(\mathcal{P}) = \left\{ E = \bigcup_{j \in J} E_j, \quad J \subseteq N \right\}$$

Funzione di probabilità

È una funzione $P: \mathcal{E} \rightarrow \mathbb{R}^+$ che, a seconda se \mathcal{E} è un'algebra o una σ -algebra, gode di determinate proprietà.

- Se \mathcal{E} è un'algebra:
 - 1) $P(\emptyset) = 0$
 - 2) $P(\Omega) = 1$
 - 3) $P(A \cup B) = P(A) + P(B) \quad \forall A, B \in \mathcal{E} : A \cap B = \emptyset$
- Se \mathcal{E} è una σ -algebra:
 - 1) $P(\emptyset) = 0$
 - 2) $P(\Omega) = 1$
 - 3) $P(\bigcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n) \quad \forall (A_n)_{n \in \mathbb{N}} : A_{n_1} \cap A_{n_2} = \emptyset \quad \forall n_1 \neq n_2$

Indipendenza

Due eventi E, F si dicono indipendenti se:

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

Inoltre, due famiglie \mathcal{E}, \mathcal{F} di eventi si dicono indipendenti se:

$$P(E \cap F) = P(E) \cdot P(F) \quad \forall E \in \mathcal{E} \quad \forall F \in \mathcal{F}$$

Insieme di eventi indipendenti

Supponiamo di avere tanti eventi $(E_j)_{j \in J}$. Questi sono:

- Indipendenti pairwise se $P(E_{j_1} \cap E_{j_2}) = P(E_{j_1}) \cdot P(E_{j_2})$
 $\forall j_1, j_2 \in J : j_1 \neq j_2$
- Indipendenti totalmente se $P(\bigcap_{k=1}^n E_{j_k}) = \prod_{k=1}^n P(E_{j_k})$
 $\forall \{j_1, \dots, j_n\} \subseteq J$

Se gli eventi sono indipendenti totalmente, allora sono anche indipendenti pairwise. Tuttavia, non vale il viceversa.

Probabilità condizionata

Definizione

$$P(F) > 0 \implies P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Formula di simmetria

$$P(E), P(F) > 0 \implies P(E|F) = \frac{P(F|E) \cdot P(E)}{P(F)}$$

Formula della probabilità totale

Sia $N \subseteq \mathbb{N}$ e sia $(F_n)_{n \in N}$ una partizione di Ω . Abbiamo:

$$P(E) = \sum_{n \in N} P(E|F_n) \cdot P(F_n) \quad \forall E \in \mathcal{E}$$

Teorema di Bayes

$$P(A_i|E) = \frac{P(E|A_i)P(A_i)}{\sum_{j=1}^n P(E|A_j)P(A_j)}$$

Variabile aleatoria reale

È una funzione $X : \Omega \rightarrow \mathbb{R}$ definita su uno spazio di probabilità (Ω, \mathcal{E}, P) .
Dato lo spazio $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, per una variabile aleatoria deve valere:

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad \{X \in B\} = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{E}$$

Distribuzione

È una funzione di probabilità $P_X : \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}^+$ definita sull'asse reale:

$$P_X(B) = P(X \in B)$$

Funzione di distribuzione

È una funzione $F_X : \mathbb{R} \rightarrow \mathbb{R}$ definita come:

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x)$$

Proprietà della funzione di distribuzione

- 1) $F_X(x) \geq 0 \quad \forall x \in \mathbb{R}$
- 2) $\lim_{x \rightarrow +\infty} F_X(x) = 1$
- 3) $\lim_{x \rightarrow -\infty} F_X(x) = 0$
- 4) $F_X(x) \leq F_X(y) \quad \forall x \leq y$
- 5) $\lim_{x \rightarrow x_0^+} F_X(x) = F_X(x_0)$
- 6) $\lim_{x \rightarrow x_0^-} F_X(x) \in \mathbb{R}$
- 7) Il numero di punti di discontinuità di F_X è al più numerabile.

Funzione di densità

La funzione di distribuzione F_X di una variabile aleatoria X è **assolutamente continua** se:

$$F_X(x) = \int_{(-\infty, x]} F'_X(u) d\mu_L(u)$$

Quando ciò è verificato, possiamo definire la funzione di densità $f_X : \mathbb{R} \rightarrow \mathbb{R}$ come:

$$\begin{cases} F'_X(x) & \text{dove } F'_X(x) \text{ è differenziabile} \\ \text{Valore arbitrario altrove} \end{cases}$$

Proprietà della funzione di densità

- 1) $f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$
- 2) $\int_{\mathbb{R}} f_X(x) d\mu_L(x) = 1$

Condizione sufficiente per l'assoluta continuità di F_X

- $F'_X(x)$ esiste ovunque
- $|F'_X(x)| < L \quad \forall x \in \mathbb{R}$

Mediana di una variabile aleatoria

È un valore $x \in \mathbb{R}$ tale che $P(X \leq x) \geq \frac{1}{2}$ e $P(X \geq x) \geq \frac{1}{2}$
Se $F_X(x)$ è continua \implies la mediana x è unica e $P(X \leq x) = P(X \geq x) = \frac{1}{2}$
Noi scriviamo $Q_{\frac{1}{2}} := \{x \in \mathbb{R} : P(X \leq x) \geq \frac{1}{2} \wedge P(X \geq x) \geq \frac{1}{2}\}$

Proposizione

Un numero reale x è una mediana della variabile aleatoria X se e solo se:

- $F_X(x) \geq \frac{1}{2}$
- $\lim_{u \rightarrow x^-} F_X(u) \leq \frac{1}{2}$

Proposizione

$Q_{\frac{1}{2}}$ è sempre un insieme non vuoto (ovvero esiste sempre almeno una mediana per una variabile aleatoria).

Proposizione

Se una variabile aleatoria X è simmetrica rispetto al punto $x_0 \implies$
 $P(X \leq x_0) = P(X \geq x_0)$

Quantile di una variabile aleatoria

Dato un qualunque valore $q \in (0,1)$, chiamiamo quantile di ordine q (o q -quantile) della variabile aleatoria X un valore $x \in \mathbb{R}$ tale che $P(X \leq x) \geq q$ e $P(X \geq x) \geq 1 - q$.

In particolare, se $\int_{(-\infty, x_q]} f_X(x) d\mu_L(x) = q \implies x_q$ è un quantile di ordine q .

Proposizione

Se una variabile aleatoria ha funzione di distribuzione continua e strettamente crescente, allora il quantile (di ordine q) è unico e possiamo definire la **funzione quantile** Q_X come l'inversa della funzione di distribuzione stessa ($Q_X := F_X^{-1}$).

Valori critici di livello α

Sono due punti x_α^+, x_α^- tali che:

- $\int_{[x_\alpha^+, +\infty)} f_X(x) d\mu_L(x) = \alpha$
- $\int_{(-\infty, x_\alpha^-]} f_X(x) d\mu_L(x) = \alpha$

In particolare, $x_\alpha^- = \alpha$ -quantile e $x_\alpha^+ = (1 - \alpha)$ -quantile.

Media di una variabile aleatoria discreta finita

Sia X una variabile aleatoria discreta finita ($X(\Omega) = (x_k)_{k=1}^n$).

$$\mathbb{E}[X] = \sum_{k=1}^n p_k \cdot x_k, \text{ dove } p_k = P(X = x_k)$$

Media di una variabile aleatoria discreta infinita

Sia X una variabile aleatoria discreta infinita ($X(\Omega) = (x_n)_{n=1}^{+\infty}$).

X ammette media se $\sum_{n=1}^{+\infty} p_n \cdot |x_n| < +\infty$, dove $p_n = P(X = x_n)$.

In tal caso:

$$\mathbb{E}[X] = \sum_{n=1}^{+\infty} p_n \cdot x_n$$

Media di una variabile aleatoria continua

Sia X una variabile aleatoria continua ($X(\Omega) = \mathbb{R}$).

X ammette media se $\int_{\Omega} |X| dP < +\infty$

In tal caso:

$$\mathbb{E}[X] = \int_{\Omega} X^+ dP - \int_{\Omega} X^- dP$$

- $X^+(\omega) = \max\{X(\omega), 0\} = \frac{|X(\omega)| + X(\omega)}{2}$
- $X^-(\omega) = -\min\{X(\omega), 0\} = \frac{|X(\omega)| - X(\omega)}{2}$

In particolare, se $F_X(x)$ è assolutamente continua, la media di X esiste se $\int_{\mathbb{R}} |x| \cdot f_X(x) d\mu_L(x) < +\infty$, ed è data da:

$$\mathbb{E}[X] = \int_{\mathbb{R}} x \cdot f_X(x) d\mu_L(x)$$

Inoltre, se X ammette media (ovvero X ammette **momento di ordine 1** finito) $\implies X \in \mathcal{L}^1(\Omega; \mathbb{R})$

Proprietà di $\mathcal{L}^1(\Omega; \mathbb{R})$

- 1) $X \in \mathcal{L}^1(\Omega; \mathbb{R}) \implies |X| \in \mathcal{L}^1(\Omega; \mathbb{R})$
- 2) $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$
- 3) $X, Y \in \mathcal{L}^1(\Omega; \mathbb{R}) \implies \alpha X + \beta Y \in \mathcal{L}^1(\Omega; \mathbb{R}) \wedge \mathbb{E}[\alpha X + \beta Y] = \alpha \mathbb{E}[X] + \beta \mathbb{E}[Y]$
- 4) Se $X \leq Y \implies \mathbb{E}[X] \leq \mathbb{E}[Y]$
- 5) Se $X^n \in \mathcal{L}^1(\Omega; \mathbb{R}) \implies X \in \mathcal{L}^n(\Omega; \mathbb{R})$ (ovvero X ammette **momento di ordine n** finito)

Momento crudo di una variabile aleatoria

Sia $X \in \mathcal{L}^n(\Omega; \mathbb{R})$. Il momento crudo di ordine n si definisce come:

$$\mu'_n := \mathbb{E}[X^n]$$

Se $n = 1 \implies \mathbb{E}[X] \equiv \mu'_1 \equiv \mu$

Momento centrato di una variabile aleatoria

Sia $X \in \mathcal{L}^n(\Omega; \mathbb{R})$. Il momento centrato di ordine n si definisce come:

$$\mu_n := \mathbb{E}[(X - \mathbb{E}[X])^n]$$

Se $n = 1 \implies \mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[X] = 0 = \mu_1$

Se $n = 2 \implies \mathbb{E}[(X - \mathbb{E}[X])^2] \equiv \mathbb{D}^2[X] \equiv \text{Var}(X) \equiv \sigma_X^2 \equiv \mu_2 \equiv \text{Varianza di } X$

In particolare: $\mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X \cdot \mathbb{E}[X] + \mathbb{E}^2[X]] = \mathbb{E}[X^2] - 2\mathbb{E}[X \cdot \mathbb{E}[X]] + \mathbb{E}[\mathbb{E}^2[X]] = \mathbb{E}[X^2] - 2\mathbb{E}[X] \cdot \mathbb{E}[X] + \mathbb{E}^2[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X]$

Proprietà della varianza

- 1) $\mathbb{D}^2[X] \geq 0$
- 2) $\mathbb{D}^2[X] = 0 \iff \mathbb{E}[(X - \mathbb{E}[X])^2] = 0 \iff X \sim \text{Dirac}(\mathbb{E}[X])$

Momento standardizzato di una v.a.

Sia $X \in \mathcal{L}^n(\Omega; \mathbb{R})$, con $n \geq 2$. Il momento standardizzato di ordine n si definisce come:

$$\hat{\mu}_n := \mathbb{E} \left[\left(\frac{X - \mu}{\sigma_X} \right)^n \right]$$

Se $n = 2 \implies \mathbb{E}[(\frac{X-\mu}{\sigma_X})^2] = \mathbb{D}^2[\frac{X-\mu}{\sigma_X}] = \frac{1}{\sigma_X^2} \cdot \mathbb{D}^2[X - \mu] = \frac{1}{\sigma_X^2} \cdot \mathbb{D}^2[X] = 1 = \hat{\mu}_2$
 $\hat{\mu}_3$ è detto **skewness** (asimmetria) e misura quanto pesa la parte negativa della distribuzione di X rispetto a quella positiva.

$\hat{\mu}_4$ è detto **curtosi** e compara le code della distribuzione di X con quelle di una distribuzione normale. Per le variabili aleatorie normali o **mesocurtiche**, $\hat{\mu}_4 = 3$; per le variabili aleatorie **leptocurtiche** (più code meno spesse delle normali), $\hat{\mu}_4 < 3$; per le variabili aleatorie **platicurtiche** (con le code più spesse delle normali), $\hat{\mu}_4 > 3$.

Covarianza di due variabili aleatorie

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

Proprietà della covarianza

- 1) $Cov(X, X) = \mathbb{D}^2[X]$
- 2) $Cov(X, Y) = Cov(Y, X) \quad \forall X, Y \in \mathcal{L}^2(\Omega; \mathbb{R})$
- 3) $Cov(\alpha X + \beta Y, Z) = \alpha \cdot Cov(X, Z) + \beta \cdot Cov(Y, Z) \quad \forall X, Y, Z \in \mathcal{L}^2(\Omega; \mathbb{R})$
- 4) $|Cov(X, Y)| \leq \mathbb{D}[X] \mathbb{D}[Y]$

Prima disuguaglianza di Markov

Sia $X \geq 0$ una variabile aleatoria che ammette momento di ordine 1 finito e supponiamo che $\mathbb{E}[X] > 0$. Allora:

$$P(X \geq \lambda \mathbb{E}[X]) \leq \frac{1}{\lambda}$$

Corollario

$$P(X \geq k) \leq \frac{\mathbb{E}[X]}{k}$$

Seconda disuguaglianza di Markov

Sia X una variabile aleatoria e sia $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ una funzione boreliana non decrescente. Allora:

$$\varphi(x) \cdot P(X \geq x) \leq \mathbb{E}[\varphi(X)] \quad \forall x \in \mathbb{R}$$

Disuguaglianza di Tchebyshev

Sia X una qualunque variabile aleatoria e sia $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ una funzione boreliana, positiva e tale che:

- $\varphi(x) = \varphi(-x) \quad \forall x \in \mathbb{R}$
- $\varphi(x) \leq \varphi(y) \quad \forall x, y : 0 \leq x \leq y$
- $\varphi(X)$ sia una variabile aleatoria che ammette momento di ordine 1 finito.

Allora:

$$\varphi(x) \cdot P(|X| \geq x) \leq \mathbb{E}[\varphi(X)] \quad \forall x \geq 0$$

Corollario

$$P(|X - \mathbb{E}[X]| \geq x) \leq \frac{\mathbb{D}^2[X]}{x^2}$$

Corollario

Sia X una variabile aleatoria che ammette momento di ordine 2 finito, e siano $a, b \in \mathbb{R}$ t.c. $a < \mathbb{E}[X] < b$. Allora:

$$P(a < X < b) \geq 1 - \frac{\mathbb{D}^2[X]}{k^2}$$

dove $k = \min\{\mathbb{E}[X] - a, b - \mathbb{E}[X]\}$

Disuguaglianza di Hölder

Siano p, q due esponenti coniugati (ovvero tali che $\frac{1}{p} + \frac{1}{q} = 1$) e siano $X \in \mathcal{L}^p(\Omega; \mathbb{R}), Y \in \mathcal{L}^q(\Omega; \mathbb{R})$. Allora:

$$\mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{\frac{1}{p}} \cdot \mathbb{E}[|Y|^q]^{\frac{1}{q}}$$

Proposizione

Due variabili aleatorie X, Y sono dette **ortogonali** se:

$$\mathbb{E}[XY] = \int_{\Omega} XY \, dP = 0$$

\mathcal{E} -vettore aleatorio

È una funzione $X : \Omega \rightarrow \mathbb{R}^N$ definita su uno spazio di probabilità (Ω, \mathcal{E}, P) . Dato lo spazio degli stati $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, per un \mathcal{E} -vettore aleatorio deve valere:

$$\forall B \in \mathcal{B}(\mathbb{R}^N) \quad \{X \in B\} = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{E}$$

In particolare, $X = (X_1, \dots, X_N)$ è un \mathcal{E} -vettore aleatorio $\iff X_1, \dots, X_N$ sono \mathcal{E} -variabili aleatorie.

Funz. di distribuz. di un \mathcal{E} -vettore aleatorio

È una funzione $F_X : \mathbb{R}^N \rightarrow \mathbb{R}$ definita come:

$$F_X(x_1, \dots, x_N) = P(X_1 \leq x_1, \dots, X_N \leq x_N)$$

Proprietà della funz. di distribuz. di un \mathcal{E} -vettore aleatorio

- 1) $\lim_{x_1 \rightarrow +\infty} \dots \lim_{x_N \rightarrow +\infty} F_X(x_1, \dots, x_N) = 1$
- 2) $\lim_{x_k \rightarrow -\infty} F_X(x_1, \dots, x_k, \dots, x_N) = 0$
- 3) Se mandiamo a $+\infty$ tutte le variabili di F_X tranne x_k , otteniamo la funzione di distribuzione (monodimensionale) della \mathcal{E} -variabile aleatoria X_k .

Densità di un \mathcal{E} -vettore aleatorio

La funzione di distribuzione F_X di un \mathcal{E} -vettore aleatorio X è **assolutamente continua** se:

$$F_X(x_1, \dots, x_N) = \int_{\prod_{k=1}^N (-\infty, x_k]} \frac{\partial^N F_X(u_1, \dots, u_N)}{\partial u_1, \dots, \partial u_N} d\mu_L^N(u_1, \dots, u_N)$$

Quando ciò è verificato, possiamo definire la funzione di densità $f_X : \mathbb{R}^N \rightarrow \mathbb{R}$ come:

$$f_X(x_1, \dots, x_N) := \frac{\partial^N F_X(x_1, \dots, x_N)}{\partial x_1, \dots, \partial x_N}$$

Media di un \mathcal{E} -vettore aleatorio

Un \mathcal{E} -vettore aleatorio $X = (X_1, \dots, X_N)$ ammette momento di ordine 1 finito se:

$$\int_{\Omega} \|X\|_2 dP < +\infty, \text{ dove } \|X\|_2 = \sqrt{\sum_{k=1}^N X_k^2}$$

In tal caso abbiamo:

$$\mathbb{E}[X] = \mathbb{E}[(X_1, \dots, X_N)] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_N])$$

Proposizione

$X = (X_1, \dots, X_N)$ ammette momento di ordine 1 finito se X_1, \dots, X_N ammettono tutte momento di ordine 1 finito.

Varianza di un \mathcal{E} -vettore aleatorio

Un \mathcal{E} -vettore aleatorio $X = (X_1, \dots, X_N)$ ammette momento di ordine 2 finito se:

$$\int_{\Omega} \|X\|_2^2 dP < +\infty, \text{ dove } \|X\|_2 = \sqrt{\sum_{k=1}^N X_k^2}$$

In tal caso possiamo definire la varianza di X come:

$$\mathbb{D}^2[X] = \begin{bmatrix} \mathbb{D}^2[X_1] & Cov(X_1, X_2) & \cdots & Cov(X_1, X_N) \\ Cov(X_2, X_1) & \mathbb{D}^2[X_2] & \cdots & Cov(X_2, X_N) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_N, X_1) & Cov(X_N, X_2) & \cdots & \mathbb{D}^2[X_N] \end{bmatrix}$$

Si tratta di una matrice simmetrica e semi-definita positiva (cioè ha tutti gli autovalori positivi).

Proposizione

$X = (X_1, \dots, X_N)$ ammette momento di ordine 2 finito se tutti i possibili prodotti $X_j \cdot X_k$ ammettono momento di ordine 1 finito $\forall j, k = 1, \dots, N$.

Indipendenza di variabili aleatorie

Due variabili aleatorie X, Y si dicono indipendenti se:

$$P(X \leq x, Y \leq y) = P(X \leq x) \cdot P(Y \leq y)$$

Inoltre, una variabile aleatoria X è indipendente da una famiglia \mathcal{F} di eventi se:

$$\forall E \in \sigma(X) \quad \forall F \in \mathcal{F} \quad P(E \cap F) = P(E) \cdot P(F)$$

dove $\sigma(X)$ è l'informazione (o meglio la σ -algebra degli eventi) generata da X .

Proposizione

Due variabili aleatorie X, Y sono indipendenti $\iff \sigma(X), \sigma(Y)$ sono indipendenti, ovvero:

$$\forall E \in \sigma(X) \quad \forall F \in \sigma(Y) \quad P(E \cap F) = P(E) \cdot P(F)$$

Proposizione

Siano X, Y due variabili aleatorie indipendenti, e siano $g : \mathbb{R} \rightarrow \mathbb{R}$, $h : \mathbb{R} \rightarrow \mathbb{R}$ due funzioni boreliane. Allora $g(X), h(Y)$ sono a loro volta variabili aleatorie indipendenti.

Insieme di variabili aleatorie indipendenti

Supponiamo di avere tante variabili aleatorie $(X_j)_{j \in J}$. Queste sono:

- Indipendenti pairwise se X_{j_1}, X_{j_2} sono indipendenti $\forall j_1, j_2 \in J : j_1 \neq j_2$
- Indipendenti totalmente se
$$P(X_{j_1} \leq x_{j_1}, \dots, X_{j_n} \leq x_{j_n}) = \prod_{k=1}^n P(X_{j_k} \leq x_{j_k}) \quad \forall \{j_1, \dots, j_n\} \subseteq J$$

Funzione di distribuzione congiunta

È una funzione così definita:

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

Inoltre, abbiamo che:

$$F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y) \iff X, Y \text{ sono indipendenti}$$

Proposizione

Consideriamo una funzione di distribuzione congiunta $F_{X,Y}(x, y)$ assolutamente continua. Allora:

- X è una variabile aleatoria assolutamente continua e ha densità $f_X(x)$.
- Y è una variabile aleatoria assolutamente continua e ha densità $f_Y(y)$.
- È possibile definire la **densità congiunta** $f_{X,Y}(x, y)$.

Non vale il viceversa: se due variabili aleatorie X, Y sono assolutamente continue, non è detto che anche $F_{X,Y}(x, y)$ sia assolutamente continua (a meno che X, Y sono indipendenti).

Proposizione

X, Y sono variabili aleatorie indipendenti $\iff f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$.

Proposizione

Se X, Y sono variabili aleatorie indipendenti $\implies \mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y] \implies \text{Cov}(X, Y) = 0$

Vettori gaussiani

Siano X_1, \dots, X_N variabili aleatorie reali e sia $X := (X_1, \dots, X_N)^T$. X si dice gaussiano se:

$$\forall (c_1, \dots, c_N) \in \mathbb{R}^N \quad \sum_{k=1}^N c_k X_k \sim N(\mu, \sigma^2)$$

È possibile anche avere $\sigma^2 = 0$ (ovvero delle distribuzioni di Dirac, che possiamo considerare come un caso degenerare delle distribuzioni gaussiane).

Criteri

1) Basta trovare una sola combinazione lineare che non dia luogo a una distribuzione gaussiana per stabilire che X non è un vettore gaussiano.

2) X è un vettore gaussiano se tutte le sue componenti X_1, \dots, X_N sono variabili aleatorie gaussiane indipendenti tra loro (ma il viceversa non è vero).

3) Sia $(X_1, \dots, X_N)^T$ un vettore gaussiano, sia $A \equiv (a_{m,n})_{m=1, n=1}^{M, N}$ una matrice con M righe e N colonne, e sia $b \in \mathbb{R}^M$ un vettore. Allora il vettore $(Y_1, \dots, Y_M)^T$ dato da:

$$(Y_1, \dots, Y_M)^T := b + A \cdot (X_1, \dots, X_N)^T$$

è a sua volta un vettore gaussiano se A ha rango massimo (ovvero se $\text{rank}(A) = \min\{M, N\}$).

4) Sia $(X_1, \dots, X_N)^T$ un vettore gaussiano (con X_1, \dots, X_N variabili aleatorie gaussiane standard indipendenti), sia $A \equiv (a_{m,n})_{m=1, n=1}^{M, N}$ una matrice con M righe e N colonne, e sia $\mu \in \mathbb{R}^M$ un vettore. Allora il vettore $(Y_1, \dots, Y_M)^T$ dato da:

$$(Y_1, \dots, Y_M)^T := \mu + A \cdot (X_1, \dots, X_N)^T$$

è a sua volta un vettore gaussiano.

- Vettore delle medie: $\mu \equiv (\mu_1, \dots, \mu_M)^T : \mu_k = \mathbb{E}[Y_k] \quad \forall k = 1, \dots, M$
- Matrice delle covarianze: $AA^T \equiv (\sigma_{m,n})_{m,n=1}^M \equiv \Sigma^2 \sigma_{m,n} \equiv \text{Cov}(Y_m, Y_n)$

Proposizione

Se le variabili aleatorie X_1, \dots, X_N che costituiscono il vettore gaussiano X sono scorrelate, allora sono anche indipendenti.

Proposizione

Se il vettore gaussiano $X = (X_1, \dots, X_N)^T$ è non degenerare (ovvero nessuna combinazione lineare delle sue componenti dà luogo a una distribuzione di Dirac), allora:

- è assolutamente continuo;
- $f_{X_1, \dots, X_N}(x_1, \dots, x_N) = \frac{1}{\sqrt{(2\pi)^N \cdot \det(\Sigma^2)}} \cdot e^{-\frac{1}{2}(x-\mu)^T (\Sigma^2)^{-1} (x-\mu)}$

Condizionamento di variabili aleatorie

Sia $X \in \mathcal{L}^1(\Omega, \mathbb{R})$ una variabile aleatoria e sia $F \in \mathcal{E}$ un evento. Definiamo il condizionamento di X rispetto a F ($\mathbb{E}[X|F]$) nel seguente modo:

$$\begin{cases} 0 & \text{se } P(F) = 0 \\ \frac{1}{P(F)} \int_F X dP & \text{altrimenti} \end{cases}$$

In particolare, se F è un insieme discreto ($F = \{\omega_1, \dots, \omega_n\}$) tale che $P(F) > 0$:

$$\mathbb{E}[X|F] = \frac{1}{P(F)} \sum_{k=1}^n X(\omega_k) P(\omega_k)$$

Teorema di Radon Nykodim

Sia $P_{\mathcal{F}}^X : \mathcal{F} \rightarrow \mathbb{R}$ una funzione così definita:

$$P_{\mathcal{F}}^X(F) = \int_F X dP, \quad \text{dove } X \in \mathcal{L}^1(\Omega, \mathbb{R})$$

Allora:

- $P_{\mathcal{F}}^X(\bigcup_{n=1}^{+\infty} F_n) = \int_{\bigcup_{n=1}^{+\infty} F_n} X dP = \sum_{n=1}^{+\infty} \int_{F_n} X dP = \sum_{n=1}^{+\infty} P_{\mathcal{F}}^X(F_n)$
(con $F_{n1} \cap F_{n2} = \emptyset \quad \forall n_1 \neq n_2$)
- Se X è una variabile aleatoria assolutamente continua \implies
 $P_{\mathcal{F}}^X(F) = 0 \quad \forall F : P(F) = 0$
- $\exists \hat{X}$ tale che $\int_F X dP = \int_F \hat{X} dP_{\mathcal{F}}^X$, dove \hat{X} è una \mathcal{F} -variabile aleatoria.

Condizionamento rispetto a una famiglia di eventi

Sia X una \mathcal{E} -variabile aleatoria tale che $X \in \mathcal{L}^1(\Omega, \mathbb{R})$ e sia $\mathcal{F} \subseteq \mathcal{E}$. Allora $\mathbb{E}[X|\mathcal{F}]$ è una \mathcal{F} -variabile aleatoria tale che $\int_F \mathbb{E}[X|\mathcal{F}] dP_{\mathcal{F}} = \int_F X dP$

Se inoltre $\mathcal{F} = \sigma(Y) \implies \mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X|\sigma(Y)] \equiv \mathbb{E}[X|Y]$

Proposizione

Data una famiglia di eventi \mathcal{F} , $\exists (F_n)_{n \in \mathbb{N}} \quad (N \subseteq \mathbb{N})$ tale che $\mathcal{F} = \sigma((F_n)_{n \in N})$.
Dunque:

$$\mathbb{E}[X|\mathcal{F}] = \sum_{n \in N} \mathbb{E}[X|F_n] \cdot \mathbf{1}_{F_n}$$

Proposizione

Sia Y una variabile aleatoria discreta tale che $Y(\Omega) = (y_n)_{n \in N}$ ($N \subseteq \mathbb{N}$) $\implies \sigma(Y) = \sigma((Y = y_n)_{n \in N})$. Dunque:

$$\mathbb{E}[X|Y] = \sum_{n \in N} \mathbb{E}[X|Y = y_n] \cdot \mathbf{1}_{\{Y=y_n\}}$$

Proprietà della speranza condizionata

- 1) Se X è una \mathcal{F} -variabile aleatoria $\implies \mathbb{E}[X|\mathcal{F}] = X$
- 2) $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]] = \mathbb{E}[X] \quad \forall X \in \mathcal{L}^1(\Omega, \mathbb{R})$
- 3) Siano \mathcal{G}, \mathcal{F} due σ -algebre tali che $\mathcal{G} \subseteq \mathcal{F}$. Allora:

$$\mathbb{E}[\mathbb{E}[X|\mathcal{F}] | \mathcal{G}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] | \mathcal{F}] = \mathbb{E}[X|\mathcal{G}]$$

- 4) Se X è una variabile aleatoria indipendente dalla famiglia $\mathcal{F} \implies \mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$
- 5) Se X, Y sono due variabili aleatorie indipendenti $\implies \mathbb{E}[X|Y] = \mathbb{E}[X]$
- 6) Sia $\phi : \mathbb{R} \rightarrow \mathbb{R}$ una funzione convessa, con $\phi \circ X \in \mathcal{L}^1(\Omega, \mathbb{R})$. Allora:

$$\phi(\mathbb{E}[X|\mathcal{F}]) \leq \mathbb{E}[\phi(X)|\mathcal{F}]$$

Caso particolare: $|\mathbb{E}[X|\mathcal{F}]|^p \leq \mathbb{E}[|X|^p|\mathcal{F}]$

- 7) $\mathbb{E}[|\mathbb{E}[X|\mathcal{F}]|^p] \leq \mathbb{E}[|X|^p] \quad \forall X \in \mathcal{L}^p(\Omega, \mathbb{R})$

Questa è una conseguenza delle proprietà 2 e 6.

- 8) $\mathbb{D}^2[\mathbb{E}[X|\mathcal{F}]] \leq \mathbb{D}^2[X] \quad \forall X \in \mathcal{L}^2(\Omega, \mathbb{R})$
- 9) Se X, \mathcal{F} sono indipendenti $\implies \mathbb{D}^2[\mathbb{E}[X|\mathcal{F}]] = \mathbb{D}^2[\mathbb{E}[X]] = 0$
- 10) Se $X \in \mathcal{L}^2(\Omega, \mathbb{R}) \implies \mathbb{E}[X|\mathcal{F}] = \min_{Y \in \mathcal{L}^2(\Omega_{\mathcal{F}}, \mathbb{R})} (\mathbb{E}[(X - Y)^2])$, dove $\mathcal{L}^2(\Omega_{\mathcal{F}}, \mathbb{R}) \subseteq \mathcal{L}^2(\Omega, \mathbb{R})$ e $\mathcal{L}^2(\Omega_{\mathcal{F}}, \mathbb{R})$ è la proiezione ortogonale di $\mathcal{L}^2(\Omega, \mathbb{R})$.

Proposizione

Se $(X, Y) \sim N(\mu, \Sigma^2) \implies$

$$\mathbb{E}[X|Y] = \mathbb{E}[X] + \text{Corr}(X, Y) \cdot \frac{\mathbb{D}[X]}{\mathbb{D}[Y]} \cdot (Y - \mathbb{E}[Y])$$

$$\mathbb{E}[X^2|Y] = \mathbb{D}^2[X] - \frac{\text{Cov}(X, Y)}{\mathbb{D}^2[Y]} + \left[\mathbb{E}[X] + \frac{\text{Cov}(X, Y)}{\mathbb{D}^2[Y]} \cdot (Y - \mathbb{E}[Y]) \right]^2$$

dove:

$$\text{Corr}(X, Y) = \begin{cases} \frac{\text{Cov}(X, Y)}{\mathbb{D}[X]\mathbb{D}[Y]} & \text{se } \mathbb{D}[X]\mathbb{D}[Y] > 0 \\ 0 & \text{se } \mathbb{D}[X]\mathbb{D}[Y] = 0 \end{cases}$$

Proposizione

Sia \mathcal{N} una σ -algebra degli eventi, sia $N \in \mathcal{N}$ un evento e sia $P_Y : \mathcal{N} \rightarrow \mathbb{R}$ una funzione di probabilità. Allora, per il Teorema di Radon Nykodim, esistono una funzione $P_Y^X : \mathcal{N} \rightarrow \mathbb{R}$ e una funzione dP_Y^X/dP_Y tali che:

$$P_Y^X(N) = \int_N dP_Y^X/dP_Y dP_Y$$

e possiamo definire $\mathbb{E}[X|Y = y] := dP_Y^X/dP_Y$.

Proposizione

Siano X, Y due variabili aleatorie congiuntamente assolutamente continue. Allora $\mathbb{E}[X|Y = y]$ può essere definita nel seguente modo:

$$\mathbb{E}[X|Y = y] = \int_{\mathbb{R}} x f_{X|Y}(x, y) d\mu_L(x) \quad , \quad \text{dove} \quad f_{X|Y}(x, y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

D'altra parte, $\mathbb{E}[h(X)|Y = y]$ è uguale a :

$$\mathbb{E}[h(X)|Y = y] = \int_{\mathbb{R}} h(x) f_{X|Y}(x, y) d\mu_L(x)$$

Successioni di variabili aleatorie

Convergenza quasi certa

Sia $(X_n)_{n \geq 1}$ (con $X_n : \Omega \rightarrow \mathbb{R}$) una successione di variabili aleatorie. Allora:
 $X_n \xrightarrow{a.s.} X$ se $\exists E \in \mathcal{E} \text{ t.c. } P(E) = 0 \wedge \forall \omega \in \Omega - E \quad \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)$

Proposizione

Se $X_n \xrightarrow{a.s.} X \wedge X_n \xrightarrow{a.s.} Y \implies X = Y \text{ a.s.}$, ovvero:

$$\exists E \in \mathcal{E} \text{ t.c. } P(E) = 0 \wedge \forall \omega \in \Omega - E \quad X(\omega) = Y(\omega)$$

Proposizione

Se $X_n \xrightarrow{a.s.} X \wedge g : \mathbb{R} \rightarrow \mathbb{R}$ è una funzione **continua** $\implies g(X_n) \xrightarrow{a.s.} g(X)$

Proposizione

Se $X_n \xrightarrow{a.s.} X \wedge Y_n \xrightarrow{a.s.} Y \implies$

- $\alpha X_n + \beta Y_n \xrightarrow{a.s.} \alpha X + \beta Y$
- $X_n Y_n \xrightarrow{a.s.} XY$
- $\frac{X_n}{Y_n} \xrightarrow{a.s.} \frac{X}{Y}$ (ma solo se $P(Y_n = 0) = 0 \wedge P(Y = 0) = 0$)

Teorema

$$X_n \xrightarrow{a.s.} X \iff$$

- $\lim_{m \rightarrow +\infty} P(\bigcap_{n \geq m} \{|X_n - X| < \epsilon\}) = 1 \quad \forall \epsilon \in \mathbb{R}^+$
- $\lim_{m \rightarrow +\infty} P(\bigcup_{n \geq m} \{|X_n - X| \geq \epsilon\}) = 0 \quad \forall \epsilon \in \mathbb{R}^+$

Convergenza in probabilità

Sia $(X_n)_{n \geq 1}$ (con $X_n : \Omega \rightarrow \mathbb{R}$) una successione di variabili aleatorie. Allora:

$$X_n \xrightarrow{P} X \quad \text{se } \forall \epsilon > 0:$$

- $\lim_{n \rightarrow +\infty} P(\{|X_n - X| < \epsilon\}) = 1$
- $\lim_{n \rightarrow +\infty} P(\{|X_n - X| \geq \epsilon\}) = 0$

Proposizione

$$X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X$$

Proposizione

$$\text{Se } X_n \xrightarrow{P} X \implies \exists \text{ sottosuccessione } (X_{n_k})_{k \geq 1} \text{ t.c. } X_{n_k} \xrightarrow{a.s.} X$$

Teorema di Slutsky

Assumiamo che $X_n \xrightarrow{P} X$. Allora, per ogni funzione continua $g : \mathbb{R} \rightarrow \mathbb{R}$, abbiamo $g(X_n) \xrightarrow{P} g(X)$.

Convergenza debole

Sia $(X_n)_{n \geq 1}$ (con $X_n : \Omega_n \rightarrow \mathbb{R}$) una successione di variabili aleatorie, e siano $F_{X_n} : \mathbb{R} \rightarrow \mathbb{R}$ le corrispettive funzioni di distribuzione. Allora:

$$X_n \xrightarrow{W} X \quad \text{se } \lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x) \quad \forall x : F_X \text{ è continua in } x.$$

Proposizione

$$X_n \xrightarrow{P} X \implies X_n \xrightarrow{W} X$$

Proposizione

$$\text{Se } X_n \xrightarrow{W} X \wedge X_n \xrightarrow{W} Y \text{ allora non è detto che } X = Y \text{ a.s.}$$

Proposizione

$$\text{Se } X_n \xrightarrow{W} X \wedge g : \mathbb{R} \rightarrow \mathbb{R} \text{ è una funzione continua } \implies g(X_n) \xrightarrow{W} g(X)$$

Proposizione

Se $X_n \xrightarrow{W} X \wedge Y_n \xrightarrow{W} Y \wedge X_n, Y_n$ sono variabili aleatorie **scorrelate** \implies

$$\alpha X_n + \beta Y_n \xrightarrow{W} \alpha X + \beta Y$$

Proposizione

Se le variabili aleatorie X_n di una successione sono definite tutte sullo stesso spazio di probabilità $\Omega \wedge X_n \xrightarrow{W} X \wedge X \sim \text{Dirac}(x_0) \implies X_n \xrightarrow{P} X$

Convergenza in media p-esima

Sia $(X_n)_{n \geq 1}$ (con $X_n : \Omega \rightarrow \mathbb{R}$) una successione di variabili aleatorie tali che $X_n \in \mathcal{L}^p(\Omega; \mathbb{R})$, con $p \geq 1$. Allora:

$$X_n \xrightarrow{\mathcal{L}^p} X \quad \text{se} \quad \mathbb{E}[|X_n - X|^p]^{\frac{1}{p}} \rightarrow 0$$

Proposizione

Se $X_n \xrightarrow{\mathcal{L}^p} X \wedge X_n \xrightarrow{\mathcal{L}^p} Y \implies X = Y \text{ a.s.}$

Proposizione

Se $X_n \xrightarrow{P} X \wedge \exists Y \text{ t.c. } |X_n| \leq Y, Y \in \mathcal{L}^p(\Omega; \mathbb{R}) \implies$

- $X \in \mathcal{L}^p(\Omega; \mathbb{R})$
- $X_n \xrightarrow{\mathcal{L}^p} X$

Proposizione

Se $X_n \xrightarrow{\mathcal{L}^p} X \implies \exists$ sottosuccessione $(X_{n_k})_{k \geq 1} \text{ t.c. } X_{n_k} \xrightarrow{\text{a.s.}} X$

Proposizione

$$X_n \xrightarrow{\mathcal{L}^p} X \implies X_n \xrightarrow{P} X$$

$$\begin{array}{ccccc}
X_n \xrightarrow{\text{a.s.}} X & \not\Rightarrow & X_n \xrightarrow{L^p} X \\
\Downarrow & \Leftrightarrow & \Downarrow \\
X_n \xrightarrow{\mathbf{P}} X & X_n \xrightarrow{\mathbf{P}} X & X_n \xrightarrow{\mathbf{P}} X \\
\Downarrow & \Uparrow & \Downarrow \\
X_n \xrightarrow{\mathbf{W}} X & (\Omega_n = \Omega \wedge X = x_0 1_\Omega) &
\end{array}$$

Media e varianza campionaria

Sia X una variabile aleatoria e siano X_1, X_2, \dots, X_n alcuni suoi campioni (e.g. X può rappresentare l'altezza e X_1, X_2, \dots, X_n sono le altezze rispettivamente delle persone 1, 2, ..., n). Siano inoltre $\mu_X = \mathbb{E}[X]$ e $\sigma_X^2 = \mathbb{D}^2[X]$.

- **Media campionaria:** $\bar{X}_n := \frac{1}{n} \sum_{k=1}^n X_k$
 $\mathbb{E}[\bar{X}_n] = \mu_X$
 $\mathbb{D}^2[\bar{X}_n] = \frac{\sigma_X^2}{n}$
- **Varianza campionaria non distorta:** $S_n^2 := \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2$
 $\mathbb{E}[S_n^2] = \sigma_X^2$
 Se la curtosi $\hat{\mu}_4$ di X esiste finita $\implies \mathbb{D}^2[S_n^2] = \frac{\sigma_X^4}{n} \left(\hat{\mu}_4 - \frac{n-3}{n-1} \right)$
- **Varianza campionaria distorta:** $\tilde{S}_n^2 := \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^2$
 $\mathbb{E}[\tilde{S}_n^2] = \frac{n-1}{n} \sigma_X^2$
- **Varianza campionaria sapendo μ_X :** $S_n^2(\mu_X) := \frac{1}{n} \sum_{k=1}^n (X_k - \mu_X)^2$
 $\mathbb{E}[S_n^2(\mu_X)] = \sigma_X^2$
 Se la curtosi $\hat{\mu}_4$ di X esiste finita $\implies \mathbb{D}^2[S_n^2(\mu_X)] = \frac{1}{n} (\hat{\mu}_4 - \sigma_X^4)$

Leggi dei grandi numeri

1° legge

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie indipendenti e identicamente distribuite t.c. $X_n \sim \text{Ber}(p)$. Allora:

$$\bar{X}_n \xrightarrow{P} p$$

2° legge

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti.

- Ammettono momento di ordine 2 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.
- Hanno una varianza pari a $\mathbb{D}^2[X_n] = \sigma_n^2$.
- Sono tali che $\lim_{n \rightarrow +\infty} \frac{1}{n^2} \sum_{k=1}^n \sigma_k^2 = 0$.

Allora:

$$\overline{X}_n \xrightarrow{P} \mu$$

3° legge (Khinchine)

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti e identicamente distribuite.
- Ammettono momento di ordine 1 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.

Allora:

$$\overline{X}_n \xrightarrow{P} \mu$$

4° legge

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti.
- Ammettono momento di ordine 2 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.
- Hanno una varianza pari a $\mathbb{D}^2[X_n] = \sigma_n^2$.
- Sono tali che $\sum_{n=1}^{+\infty} \frac{\sigma_n^2}{n^2} < +\infty$.

Allora:

$$\overline{X}_n \xrightarrow{a.s.} \mu$$

5° legge (Kolmogorov)

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti e identicamente distribuite.
- Ammettono momento di ordine 1 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.

Allora:

$$\overline{X}_n \xrightarrow{a.s.} \mu$$

6° legge

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti e identicamente distribuite.
- Ammettono momento di ordine 2 finito.
- Hanno tutte media nulla.
- Hanno tutte la stessa varianza $\mathbb{D}^2[X_n] = \sigma^2$.

Allora:

$$\tilde{S}_n^2 \xrightarrow{a.s.} \sigma^2 \quad ; \quad S_n^2 \xrightarrow{a.s.} \sigma^2$$

7° legge

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono scorrelate.
- Ammettono momento di ordine 2 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.
- Sono tali che $\mathbb{D}^2[\bar{X}_n] \leq \Sigma$.

Allora:

$$\bar{X}_n \xrightarrow{\mathcal{L}^2} \mu$$

Teorema del limite centrale

Sia $(X_n)_{n \geq 1}$ una successione di variabili aleatorie che:

- Sono indipendenti e identicamente distribuite.
- Ammettono momento di ordine 2 finito.
- Hanno tutte la stessa media $\mathbb{E}[X_n] = \mu$.
- Hanno tutte la stessa varianza $\mathbb{D}^2[X_n] = \sigma^2$.

Sia inoltre $Z_n := \sum_{k=1}^n X_k$ il **random walk**, e sia $\tilde{Z}_n := \frac{Z_n - n\mu}{\sigma\sqrt{n}}$ la sua standardizzazione. Allora:

$$\tilde{Z}_n \xrightarrow{W} N(0, 1)$$

Simple random sample

È una successione $(X_k)_{k=1}^n$ di variabili aleatorie indipendenti, estratte da una stessa variabile aleatoria X e, quindi, equamente distribuite ($X_k \sim X \ \forall k = 1, \dots, n$).

Statistica

È una qualunque funzione di Borel di un simple random sample ed è definita come:

$$G_n(\omega) := g(X_1(\omega), \dots, X_n(\omega)) \quad \forall \omega \in \Omega$$

Esempi di statistiche

- Somma di variabili aleatorie (Z_n)
 $Se X \sim Ber(p) \implies Z_n \sim Bin(n, p)$
 $Se X \sim Poiss(\lambda) \implies Z_n \sim Poiss(n\lambda)$
 $Se X \sim N(\mu, \sigma^2) \implies Z_n \sim N(n\mu, n\sigma^2)$
 $Se X \sim Exp(\lambda) \implies Z_n \sim \Gamma(n, \lambda) \implies Z_n \sim \frac{1}{2\lambda} \chi_{2n}^2$
 $Se X \sim \chi_1^2 \implies Z_n \sim \chi_n^2$
- Media campionaria (\bar{X}_n)
 $Se X \sim N(\mu, \sigma^2) \implies \bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$
- Massimo tra variabili aleatorie (\check{X}_n)
- Minimo tra variabili aleatorie (\hat{X}_n)
 $Se X \sim Exp(\lambda) \implies \hat{X}_n \sim Exp(n\lambda)$
- Somma dei quadrati di variabili aleatorie (Q_n)
 $Se X \sim N(0, 1) \implies Q_n \sim \chi_n^2$
- Varianza campionaria conoscendo la media ($S_n^2(\mu)$)
- Varianza campionaria non distorta ($S_n^2(X)$)
- Varianza campionaria distorta ($\tilde{S}_n^2(X)$)

Teorema

Se $X \sim N(\mu, \sigma^2) \implies$ le statistiche $\bar{X}_n, S_n^2(X)$ sono indipendenti.

Teorema

Se $X \sim N(\mu, \sigma^2) \implies (n-1) \frac{S_n^2(X)}{\sigma^2} \sim \chi_{n-1}^2$

Teorema

Se $X \sim N(\mu, \sigma^2) \implies \frac{\bar{X}_n - \mu}{\frac{S_n}{\sqrt{n}}} \sim t_{n-1}$

dove t_{n-1} rappresenta una distribuzione di Student con n-1 gradi di libertà.

Teorema

Se $X \sim N(\mu, \sigma^2) \implies \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$

Stima puntuale

Una statistica $G_n : \Omega \rightarrow \mathbb{R}^n$ è detta **stimatore puntuale** (che indicheremo anche con $\hat{\theta}_n$) del parametro ignoto θ se possiamo sfruttare le realizzazioni $G_n(\omega)$ di G_n per stimare il valore vero di θ .

Data una qualunque realizzazione x_1, \dots, x_n del campione X_1, \dots, X_n (dove $x_k = X_k(\omega)$), per qualche $\omega \in \Omega$ e per ogni $k = 1, \dots, n$, il numero reale

$$G_n(\omega) := g(X_1(\omega), \dots, X_n(\omega)) \equiv g(x_1, \dots, x_n)$$

è chiamato **stima puntuale** di θ (che indicheremo anche con $\hat{\theta}_n(\omega)$).

Errore quadratico medio

$$\text{MSE}(\hat{\theta}_n) := \mathbb{E}[(\hat{\theta}_n - \theta)^2]$$

Distorsione

$$\text{Bias}(\hat{\theta}_n) := \mathbb{E}[\hat{\theta}_n] - \theta$$

Proposizione

$$\text{MSE}(\hat{\theta}_n) = \mathbb{D}^2[\hat{\theta}_n] + \text{Bias}^2(\hat{\theta}_n)$$

Errore standard

$$\text{SE}(\hat{\theta}_n) := \sqrt{\mathbb{D}^2[\hat{\theta}_n]}$$

Consistenza di uno stimatore puntuale

Definizione

Uno stimatore puntuale $\hat{\theta}_n$ di θ è (asintoticamente) **consistente in probabilità** se $\hat{\theta}_n \xrightarrow{P} \theta$ per $n \rightarrow +\infty$.

Definizione

Uno stimatore puntuale $\hat{\theta}_n$ di θ è (asintoticamente) **consistente in media quadratica** se $\hat{\theta}_n \xrightarrow{\mathcal{L}^2} \theta$ per $n \rightarrow +\infty$.

Proposizione

Se uno stimatore $\hat{\theta}_n$ corretto (\equiv non distorto) di θ è consistente in media quadratica, allora è consistente anche in probabilità.

Valore critico inferiore

Dato $\alpha \in (0, 1)$, chiamiamo valore critico inferiore di livello α della variabile aleatoria X e denotiamo con x_{α}^{-} il minimo α -quantile di X . In simboli: $x_{\alpha}^{-} := \tilde{x}_{\alpha}$

Valore critico superiore

Dato $\alpha \in (0, 1)$, chiamiamo valore critico superiore di livello α della variabile aleatoria X e denotiamo con x_{α}^{+} il massimo $(1 - \alpha)$ -quantile di X . In simboli: $x_{\alpha}^{+} := \hat{x}_{1-\alpha}$

Limite inferiore di confidenza

Dato $\alpha \in (0, 1)$, diciamo che una statistica $\underline{\theta} : \Omega \rightarrow \mathbb{R}$ è un limite inferiore di confidenza al livello di confidenza $1 - \alpha$ per il parametro vero θ se abbiamo:

$$P(\underline{\theta} \leq \theta) \geq 1 - \alpha \quad \forall \theta \in \Theta$$

Inoltre, qualunque valore $\underline{\theta}(\omega) \in \mathbb{R}$ preso dalla statistica $\underline{\theta}$ all'occorrenza di un sample point $\omega \in \Omega$ è detto **realizzazione del limite inferiore di confidenza** $\underline{\theta}$.

Limite superiore di confidenza

Dato $\alpha \in (0, 1)$, diciamo che una statistica $\bar{\theta} : \Omega \rightarrow \mathbb{R}$ è un limite superiore di confidenza al livello di confidenza $1 - \alpha$ per il parametro vero θ se abbiamo:

$$P(\bar{\theta} \geq \theta) \geq 1 - \alpha \quad \forall \theta \in \Theta$$

Inoltre, qualunque valore $\bar{\theta}(\omega) \in \mathbb{R}$ preso dalla statistica $\bar{\theta}$ all'occorrenza di un sample point $\omega \in \Omega$ è detto **realizzazione del limite superiore di confidenza** $\bar{\theta}$.

Intervallo di confidenza

Dato $\alpha \in (0, 1)$, diciamo che due statistiche $\phi : \Omega \rightarrow \mathbb{R}$ e $\psi : \Omega \rightarrow \mathbb{R}$ costituiscono un intervallo di confidenza (ϕ, ψ) al livello di confidenza $1 - \alpha$ per il parametro vero θ se abbiamo:

$$P(\phi \leq \theta \leq \psi) \geq 1 - \alpha \quad \forall \theta \in \Theta$$

Inoltre, qualunque intervallo $(\phi(\omega), \psi(\omega)) \subseteq \mathbb{R}$, dove $\phi(\omega), \psi(\omega)$ sono i valori presi da ϕ e ψ all'occorrenza di un sample point $\omega \in \Omega$, è detto **realizzazione dell'intervallo di confidenza** per θ .

Proposizione

Dato $\alpha \in (0, 1)$, siano $\underline{\theta} : \Omega \rightarrow \mathbb{R}$ e $\bar{\theta} : \Omega \rightarrow \mathbb{R}$ rispettivamente il limite inferiore di confidenza e il limite superiore di confidenza al livello di confidenza $1 - \frac{\alpha}{2}$ per il parametro vero θ . Allora la coppia $(\underline{\theta}, \bar{\theta})$ costituisce un intervallo di confidenza al livello di confidenza $1 - \alpha$ per il parametro vero θ .

Ampiezza di (ϕ, ψ)

È la statistica $\psi - \phi$.

Precisione di (ϕ, ψ)

È il numero reale $\frac{1}{\mathbb{E}[\psi - \phi]}$.

Intervalli di confidenza per la media di una popolazione

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 nota. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza $1 - \alpha$ per il parametro μ è dato dalle seguenti statistiche:

$$\bar{X}_n - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad ; \quad \bar{X}_n + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 ignota. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza $1 - \alpha$ per il parametro μ è dato dalle seguenti statistiche:

$$\bar{X}_n - t_{\frac{\alpha}{2}, n-1} \frac{S_n}{\sqrt{n}} \quad ; \quad \bar{X}_n + t_{\frac{\alpha}{2}, n-1} \frac{S_n}{\sqrt{n}}$$

dove $t_{\frac{\alpha}{2}, n-1} \equiv t_{\frac{\alpha}{2}, n-1}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria di Student con $n-1$ gradi di libertà.

Proposizione

Sia X una variabile aleatoria con distribuzione ignota ma con momento di ordine 4 finito. Assumiamo inoltre che la dimensione n del campione X_1, \dots, X_n sia

grande ($n > 40$). Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza approssimativamente $1 - \alpha$ per il parametro μ è dato dalle seguenti statistiche:

$$\bar{X}_n - z_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}} \quad ; \quad \bar{X}_n + z_{\frac{\alpha}{2}} \frac{S_n}{\sqrt{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Proposizione

Sia X una variabile aleatoria bernoulliana con parametro di successo p ignoto. Assumiamo inoltre che la dimensione n del campione X_1, \dots, X_n sia grande ($n > 40$). Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza approssimativamente $1 - \alpha$ per il parametro p è dato dalle seguenti statistiche:

$$\frac{\bar{X}_n + \frac{z_{\frac{\alpha}{2}}^2}{2n} - z_{\frac{\alpha}{2}}^+ \sqrt{\frac{1}{n} \bar{X}_n (1 - \bar{X}_n) + \frac{z_{\frac{\alpha}{2}}^2}{4n^2}}}{1 + \frac{z_{\frac{\alpha}{2}}^2}{n}} \quad ; \quad \frac{\bar{X}_n + \frac{z_{\frac{\alpha}{2}}^2}{2n} + z_{\frac{\alpha}{2}}^+ \sqrt{\frac{1}{n} \bar{X}_n (1 - \bar{X}_n) + \frac{z_{\frac{\alpha}{2}}^2}{4n^2}}}{1 + \frac{z_{\frac{\alpha}{2}}^2}{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Intervalli di confidenza per la varianza di una popolazione

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 ignota. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza $1 - \alpha$ per il parametro σ^2 è dato dalle seguenti statistiche:

$$\frac{(n-1)S_{X,n}^2}{\chi_{n-1, \frac{\alpha}{2}}^{2,+}} \quad ; \quad \frac{(n-1)S_{X,n}^2}{\chi_{n-1, \frac{\alpha}{2}}^{2,-}}$$

dove $\chi_{n-1, \frac{\alpha}{2}}^{2,-}$ è il valore critico inferiore di livello $\frac{\alpha}{2}$ di χ_{n-1}^2 , mentre $\chi_{n-1, \frac{\alpha}{2}}^{2,+}$ è il valore critico superiore di livello $\frac{\alpha}{2}$ di χ_{n-1}^2 .

Proposizione

Sia X una variabile aleatoria qualsiasi con momento di ordine 4 finito. Assumiamo inoltre che la dimensione n del campione X_1, \dots, X_n sia grande ($n > 40$). Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza approssimativamente $1 - \alpha$ per il parametro σ^2 è dato dalle seguenti statistiche:

$$\frac{S_{X,n}^2}{1 - z_{\frac{\alpha}{2}} \sqrt{\frac{1}{n} (\widetilde{Kurt}_{X,n} - 1)}} \quad ; \quad \frac{S_{X,n}^2}{1 + z_{\frac{\alpha}{2}} \sqrt{\frac{1}{n} (\widetilde{Kurt}_{X,n} - 1)}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ di $N(0, 1)$ e $\widetilde{Kurt}_{X,n}$ è lo stimatore distorto della curtosi standardizzata di X.

Intervalli di confidenza per la differenza tra le medie di due popolazioni

Proposizione

Siano X,Y due variabili aleatorie gaussiane con varianze σ_X^2, σ_Y^2 note. Assumiamo inoltre che i campioni X_1, \dots, X_m e Y_1, \dots, Y_n siano indipendenti. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza $1 - \alpha$ per la differenza $\mu_X - \mu_Y$ è dato dalle seguenti statistiche:

$$\bar{X}_m - \bar{Y}_n - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}} \quad ; \quad \bar{X}_m - \bar{Y}_n + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{m} + \frac{\sigma_Y^2}{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Proposizione

Siano X,Y due variabili aleatorie gaussiane con la stessa varianza σ^2 ignota. Assumiamo inoltre che i campioni X_1, \dots, X_m e Y_1, \dots, Y_n siano indipendenti. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza $1 - \alpha$ per la differenza $\mu_X - \mu_Y$ è dato dalle seguenti statistiche:

$$\bar{X}_m - \bar{Y}_n - t_{\frac{\alpha}{2}, m+n-2} \sqrt{S_p^2 \left(\frac{1}{m} + \frac{1}{n} \right)} \quad ; \quad \bar{X}_m - \bar{Y}_n + t_{\frac{\alpha}{2}, m+n-2} \sqrt{S_p^2 \left(\frac{1}{m} + \frac{1}{n} \right)}$$

dove $t_{\frac{\alpha}{2}, m+n-2} \equiv t_{\frac{\alpha}{2}, m+n-2}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria di Student con m+n-2 gradi di libertà e S_p^2 è la **pooled sample variance** che è data da:

$$S_p^2 = \frac{(m-1)S_{X,m}^2 + (n-1)S_{Y,n}^2}{m+n-2}$$

Proposizione

Siano X,Y due variabili aleatorie gaussiane con varianze σ_X^2, σ_Y^2 differenti e ignote. Assumiamo inoltre che i campioni X_1, \dots, X_m e Y_1, \dots, Y_n siano indipendenti. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza approssimativamente $1 - \alpha$ per la differenza $\mu_X - \mu_Y$ è dato dalle seguenti statistiche:

$$\bar{X}_m - \bar{Y}_n - t_{\frac{\alpha}{2}, \hat{\nu}} \sqrt{\frac{S_{X,m}^2}{m} + \frac{S_{Y,n}^2}{n}} \quad ; \quad \bar{X}_m - \bar{Y}_n + t_{\frac{\alpha}{2}, \hat{\nu}} \sqrt{\frac{S_{X,m}^2}{m} + \frac{S_{Y,n}^2}{n}}$$

dove $t_{\frac{\alpha}{2}, \hat{\nu}} \equiv t_{\frac{\alpha}{2}, \hat{\nu}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria di Student con $\hat{\nu}$ gradi di libertà, dove:

$$\hat{\nu} = \left\lfloor \frac{\left(\frac{s_{X,m}^2}{m} + \frac{s_{Y,n}^2}{n} \right)^2}{\frac{s_{X,m}^4}{(m-1)m^2} + \frac{s_{Y,n}^4}{(n-1)n^2}} \right\rfloor$$

Proposizione

Siano X, Y due variabili aleatorie qualsiasi con momenti di ordine 2 finiti. Assumiamo inoltre che i campioni X_1, \dots, X_m e Y_1, \dots, Y_n siano indipendenti e che le dimensioni m, n di entrambi i campioni siano grandi ($m > 40, n > 40$). Allora, fissato $\alpha \in (0, 1)$, l'intervallo di confidenza con livello di confidenza approssimativamente $1 - \alpha$ per la differenza $\mu_X - \mu_Y$ è dato dalle seguenti statistiche:

$$\bar{X}_m - \bar{Y}_n - z_{\frac{\alpha}{2}} \sqrt{\frac{S_{X,m}^2}{m} + \frac{S_{Y,n}^2}{n}} \quad ; \quad \bar{X}_m - \bar{Y}_n + z_{\frac{\alpha}{2}} \sqrt{\frac{S_{X,m}^2}{m} + \frac{S_{Y,n}^2}{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Intervalli di predizione

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 nota. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di predizione con livello di confidenza $1 - \alpha$ per il campione X_{n+1} è dato dalle seguenti statistiche:

$$\bar{X}_n - z_{\frac{\alpha}{2}} \sigma \sqrt{1 + \frac{1}{n}} \quad ; \quad \bar{X}_n + z_{\frac{\alpha}{2}} \sigma \sqrt{1 + \frac{1}{n}}$$

dove $z_{\frac{\alpha}{2}} \equiv z_{\frac{\alpha}{2}}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria gaussiana.

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 ignota. Allora, fissato $\alpha \in (0, 1)$, l'intervallo di predizione con livello di confidenza $1 - \alpha$ per il campione X_{n+1} è dato dalle seguenti statistiche:

$$\bar{X}_n - t_{\frac{\alpha}{2}, n-1} S_{X,n} \sqrt{1 + \frac{1}{n}} \quad ; \quad \bar{X}_n + t_{\frac{\alpha}{2}, n-1} S_{X,n} \sqrt{1 + \frac{1}{n}}$$

dove $t_{\frac{\alpha}{2}, n-1} \equiv t_{\frac{\alpha}{2}, n-1}^+$ è il valore critico superiore di livello $\frac{\alpha}{2}$ della variabile aleatoria di Student con $n-1$ gradi di libertà.

Test d'ipotesi

Definizione

L'**ipotesi nulla**, comunemente denotata con H_0 , è l'affermazione riguardante un parametro θ di una variabile aleatoria X (come $\theta = \theta_0$) che inizialmente è assunta essere vera.

L'**ipotesi alternativa**, che contraddice l'ipotesi nulla, è solitamente denotata con H_1 oppure H_α ; può essere espressa in uno dei seguenti tre modi:

- $\theta \neq \theta_0$
- $\theta > \theta_0$
- $\theta < \theta_0$

Definizione

Si ha un **errore del I tipo** se viene rigettata l'ipotesi nulla H_0 quando essa in realtà è vera, e si ha che:

$$\alpha := P(\text{rigetto } H_0 \mid H_0 \text{ vera})$$

Definizione

Si ha un **errore del II tipo** se viene accettata l'ipotesi nulla H_0 quando essa in realtà è falsa, e si ha che:

$$\beta := P(\text{accetto } H_0 \mid H_0 \text{ falsa})$$

Test d'ipotesi per la media di una popolazione

Posposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 nota.

Effettuiamo l'ipotesi nulla $H_0 : \mu = \mu_0$ e consideriamo l'ipotesi alternativa H_1 .

La statistica che interviene è:

$$Z_0 = \frac{\bar{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Nell'ipotesi che H_0 sia vera, tale statistica ha una distribuzione gaussiana con media 0 e varianza 1.

Allora, fissato $\alpha \in (0, 1)$, abbiamo che:

$$\alpha = P(\text{rigetto } H_0 \mid H_0 \text{ vera}) = \begin{cases} P(Z_0 < z_\alpha^-) & \text{se } H_1 : \mu < \mu_0 \\ P(Z_0 > z_\alpha^+) & \text{se } H_1 : \mu > \mu_0 \\ P(Z_0 < z_{\frac{\alpha}{2}}^-) + P(Z_0 > z_{\frac{\alpha}{2}}^+) & \text{se } H_1 : \mu \neq \mu_0 \end{cases}$$

Dopodiché, fissato $\mu_1 \neq \mu_0$, la probabilità di commettere un errore del II tipo è in funzione di μ_1 ed è data da:

$$\beta(\mu_1) = P(\text{accetto } H_0 \mid \mu = \mu_1) = \begin{cases} 1 - P\left(Z_1 \leq z_\alpha^- - \frac{\mu_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) & \text{se } H_1 : \mu < \mu_0 \\ P\left(Z_1 \leq z_\alpha^+ - \frac{\mu_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) & \text{se } H_1 : \mu > \mu_0 \\ P\left(Z_1 \leq z_{\frac{\alpha}{2}}^+ - \frac{\mu_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) - P\left(Z_1 \leq z_{\frac{\alpha}{2}}^- - \frac{\mu_1 - \mu_0}{\frac{\sigma}{\sqrt{n}}}\right) & \text{se } H_1 : \mu \neq \mu_0 \end{cases}$$

Qui la statistica che è intervenuta è:

$$Z_1 = \frac{\bar{X}_n - \mu_1}{\frac{\sigma}{\sqrt{n}}}$$

Nell'ipotesi che $\mu = \mu_1$, tale statistica ha una distribuzione gaussiana con media 0 e varianza 1.

Proposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 ignota.

Effettuiamo l'ipotesi nulla $H_0 : \mu = \mu_0$ e consideriamo l'ipotesi alternativa H_1 .

La statistica che interviene è:

$$T_{n-1}^{(0)} = \frac{\bar{X}_n - \mu_0}{\frac{S_n}{\sqrt{n}}}$$

Nell'ipotesi che H_0 sia vera, tale statistica ha una distribuzione di Student con n-1 gradi di libertà.

Allora, fissato $\alpha \in (0, 1)$, abbiamo che:

$$\alpha = P(\text{rigetto } H_0 \mid H_0 \text{ vera}) = \begin{cases} P(T_{n-1}^{(0)} < t_{n-1, \alpha}^-) & \text{se } H_1 : \mu < \mu_0 \\ P(T_{n-1}^{(0)} > t_{n-1, \alpha}^+) & \text{se } H_1 : \mu > \mu_0 \\ P(T_{n-1}^{(0)} < t_{n-1, \frac{\alpha}{2}}^-) + P(T_{n-1}^{(0)} > t_{n-1, \frac{\alpha}{2}}^+) & \text{se } H_1 : \mu \neq \mu_0 \end{cases}$$

Dopodiché, fissato $\mu_1 \neq \mu_0$, la probabilità di commettere un errore del II tipo è in funzione di n e di d (dove $d := \frac{\mu_1 - \mu_0}{\sigma}$) ed è data da:

$$\beta(n, d) = P(\text{accetto } H_0 \mid \mu = \mu_1) =$$

$$\begin{cases} 1 - P(T_{n-1}^{(0)} \leq t_{n-1, \alpha}^- \mid \mu = \mu_1) & \text{se } H_1 : \mu < \mu_0 \\ P(T_{n-1}^{(0)} \leq t_{n-1, \alpha}^+ \mid \mu = \mu_1) & \text{se } H_1 : \mu > \mu_0 \\ P(T_{n-1}^{(0)} \leq t_{n-1, \frac{\alpha}{2}}^+ \mid \mu = \mu_1) - P(T_{n-1}^{(0)} \leq t_{n-1, \frac{\alpha}{2}}^- \mid \mu = \mu_1) & \text{se } H_1 : \mu \neq \mu_0 \end{cases}$$

Proposizione

Sia X una variabile aleatoria con distribuzione e varianza σ^2 ignote, e supponiamo di avere un campione X_1, \dots, X_n di dimensioni elevate.

Effettuiamo l'ipotesi nulla $H_0 : \mu = \mu_0$ e consideriamo l'ipotesi alternativa H_1 . La statistica che interviene è:

$$\tilde{Z}_0 = \frac{\bar{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Nelle ipotesi che H_0 sia vera e che n sia sufficientemente elevato, tale statistica ha una distribuzione approssimativamente gaussiana con media 0 e varianza 1. Di conseguenza, si applicano approssimativamente i risultati della prima proposizione.

Proposizione

Sia X una variabile aleatoria bernoulliana con parametro di successo p ignoto, e supponiamo di avere un campione X_1, \dots, X_n di dimensioni elevate.

Effettuiamo l'ipotesi nulla $H_0 : p = p_0$ e consideriamo l'ipotesi alternativa H_1 .

Nell'ipotesi che n sia sufficientemente elevato, la sample sum Z_n (che è per natura una variabile aleatoria binomiale) ha una distribuzione approssimativamente gaussiana con media $\mu_{Z_n} = np$ e varianza $\sigma_{Z_n}^2 = np(1-p)$.

Perciò, la sample mean $\bar{X}_n = \frac{1}{n} Z_n$ ha una distribuzione approssimativamente gaussiana con media $\mu_{\bar{X}_n} = p$ e varianza $\sigma_{\bar{X}_n}^2 = \frac{p(1-p)}{n}$.

In definitiva, la statistica che interviene è:

$$Z_0 = \frac{\bar{X}_n - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Nell'ipotesi che H_0 sia vera, tale statistica ha una distribuzione approssimativamente gaussiana con media 0 e varianza 1. Di conseguenza, per quanto riguarda l'errore del I tipo, si applicano approssimativamente i risultati della prima proposizione.

Dopodiché, fissato $p_1 \neq p_0$, la probabilità di commettere un errore del II tipo è in funzione di p_1 ed è data da:

$$\begin{aligned} \beta(p_1) &= P(\text{accetto } H_0 \mid p = p_1) = \\ &\begin{cases} 1 - P\left(Z_1 \leq \frac{p_0 - p_1 + z_{\alpha}^- \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p_1(1-p_1)}{n}}}\right) & \text{se } H_1 : \mu < \mu_0 \\ P\left(Z_1 \leq \frac{p_0 - p_1 + z_{\alpha}^+ \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p_1(1-p_1)}{n}}}\right) & \text{se } H_1 : \mu > \mu_0 \\ P\left(Z_1 \leq \frac{p_0 - p_1 + z_{\frac{\alpha}{2}}^+ \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p_1(1-p_1)}{n}}}\right) - P\left(Z_1 \leq \frac{p_0 - p_1 + z_{\frac{\alpha}{2}}^- \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p_1(1-p_1)}{n}}}\right) & \text{se } H_1 : \mu \neq \mu_0 \end{cases} \end{aligned}$$

Qui la statistica che è intervenuta è:

$$Z_1 = \frac{\bar{X}_n - p_1}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Nelle ipotesi che $p = p_1$ e che n sia sufficientemente elevato, tale statistica ha una distribuzione approssimativamente gaussiana con media 0 e varianza 1.

Test d'ipotesi per la varianza di una popolazione

Posposizione

Sia X una variabile aleatoria gaussiana con varianza σ^2 ignota.

Effettuiamo l'ipotesi nulla $H_0 : \sigma^2 = \sigma_0^2$ e consideriamo l'ipotesi alternativa H_1 .

La statistica che interviene è:

$$X_{n-1}^2 = \frac{(n-1)S_n^2}{\sigma_0^2}$$

Nell'ipotesi che H_0 sia vera, tale statistica ha una distribuzione di chi-quadro con $n-1$ gradi di libertà.

Allora, fissato $\alpha \in (0, 1)$, abbiamo che:

$$\alpha = P(\text{rigetto } H_0 \mid H_0 \text{ vera}) = \begin{cases} P(X_{n-1}^2 < \chi_{n-1, \alpha}^{2, -}) & \text{se } H_1 : \sigma^2 < \sigma_0^2 \\ P(X_{n-1}^2 > \chi_{n-1, \alpha}^{2, +}) & \text{se } H_1 : \sigma^2 > \sigma_0^2 \\ P(X_{n-1}^2 < \chi_{n-1, \frac{\alpha}{2}}^{2, -}) + P(X_{n-1}^2 > \chi_{n-1, \frac{\alpha}{2}}^{2, +}) & \text{se } H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

Proposizione

Sia X una variabile aleatoria qualunque con varianza σ^2 ignota e con momento di ordine 4 finito, e supponiamo di avere un campione X_1, \dots, X_n di dimensioni elevate.

Effettuiamo l'ipotesi nulla $H_0 : \sigma^2 = \sigma_0^2$ e consideriamo l'ipotesi alternativa H_1 .

La statistica che interviene è:

$$Z_0 = \frac{S_n^2 - \sigma_0^2}{\sqrt{\left(\frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^4 - \sigma_0^4\right) \frac{1}{n}}}$$

Nelle ipotesi che $\sigma^2 = \sigma_0^2$ e che n sia sufficientemente elevato, tale statistica ha una distribuzione approssimativamente gaussiana con media 0 e varianza 1.

Allora, fissato $\alpha \in (0, 1)$, abbiamo che:

$$\alpha = P(\text{rigetto } H_0 \mid H_0 \text{ vera}) = \begin{cases} P(Z_0 < z_{\alpha}^{-}) & \text{se } H_1 : \sigma^2 < \sigma_0^2 \\ P(Z_0 > z_{\alpha}^{+}) & \text{se } H_1 : \sigma^2 > \sigma_0^2 \\ P(Z_0 < z_{\frac{\alpha}{2}}^{-}) + P(Z_0 > z_{\frac{\alpha}{2}}^{+}) & \text{se } H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

Cose extra da ricordare

1) Funzione di distribuzione della variabile aleatoria gaussiana:

$$F_X(x) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma \sqrt{2}} \right) \right]$$

dove:

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

2)

$$\int_{-\infty}^{+\infty} e^{-\frac{(x-\mu)^2}{\sigma}} dx = \sqrt{\sigma\pi} \quad \text{dove } \sigma > 0$$

3)

$$\int_{-\infty}^{+\infty} e^{-bx^2+cx+d} dx = \sqrt{\frac{\pi}{b}} \cdot e^{\frac{c^2}{4b}+d} \quad \text{dove } b > 0$$