

Matematica applicata

▼ 1.0 - Introduzione alla matematica applicata

▼ 1.1 - Definizione di probabilità e statistica

La **probabilità** e la **statistica** forniscono numeri che rappresentano l'incertezza di eventi che possono avere esiti diversi.

I numeri che vengono forniti dalla probabilità e dalla statistica sono compresi tra 0 e 1.

La **statistica descrittiva** si concentra sulla raccolta, l'organizzazione, la sintesi e la presentazione dei dati connessi ad un certo evento.

La **statistica inferenziale** utilizza i dati raccolti relativi ad un certo evento per fare una previsione.

▼ 1.2 - Richiami di calcolo combinatorio

Principio fondamentale del calcolo combinatorio (o di enumerazione)

Si considerano 2 esperimenti, si suppone che il primo abbia m esiti possibili e per ciascuno di questi esiti il secondo esperimento abbia n esiti possibili, allora le sequenze ordinate degli esiti dei 2 esperimenti saranno $m \cdot n$.

Sequenze distinte corrispondono a esiti finali distinti.

Esercizi

- ▼ Calcolare il numero di targhe automobilistiche italiane realizzabili.

Sapendo che una targa automobilistica ha la seguente struttura LLCCCLL, dove L rappresenta una lettera e C rappresenta una cifra, possiamo utilizzare il principio fondamentale del calcolo combinatorio per calcolare il numero di targhe realizzabili, sapendo che le lettere dell'alfabeto sono 26 e le cifre sono 10.

Il numero di targhe disponibili è dunque equivalente a: $26^4 \cdot 10^3$.

Disposizioni semplici

Si dicono **disposizioni semplici** di n oggetti distinti classe k le sequenze che si possono costruire di k oggetti presi tra n oggetti senza ripetizioni (l'ordine conta).

$$D_{n,k} = n(n-1) \dots (n-k+1) = \frac{n!}{(n-k)!}$$

Disposizioni con ripetizione

Si dicono **disposizioni con ripetizione** di n oggetti di classe k le sequenze che si possono costruire di k oggetti presi tra gli n con ripetizioni (l'ordine conta).

$$D_{n,k}^R = n \cdot n \dots n = n^k$$

Permutazioni semplici

Si dicono **permutazioni semplici** di n oggetti le sequenze diverse che si possono costruire di n oggetti presi tra gli n senza ripetizioni (l'ordine conta).

$$P_n = n(n-1) \dots 1 = n!$$

Combinazioni semplici

Si dicono **combinazioni semplici** di n oggetti di classe k gli insiemi che si possono costruire di k oggetti presi tra gli n senza ripetizioni (l'ordine non conta).

$$C_{n,k} = \frac{D_{n,k}}{P_k} = \frac{n!}{(n-k)!} \cdot \frac{1}{k!} = \binom{n}{k}$$

La formula è data dal fatto che le disposizioni semplici corrispondono alle combinazioni semplici in cui per ogni insieme di k oggetti diverso comprende anche tutte le sue permutazioni.

▼ 2.0 - Calcolo delle probabilità

▼ 2.1 - Introduzione alla probabilità e statistica

Definizione di elementi di probabilità e statistica

Gli **esiti** di un esperimento corrispondono ai possibili risultati di questo.

$$\text{esiti} = e_1, \dots, e_n$$

Lo **spazio campione** corrisponde all'insieme di tutti gli esiti di un esperimento.

$$\Omega \text{ oppure } S = e_1, \dots, e_n$$

Un **evento** corrisponde ad un sottoinsieme dello spazio campione.

$$E = \{e_1, \dots, e_n\}$$

L'**unione di eventi** $E \cup F$ corrisponde a tutti gli esiti che appartengono ad almeno 1 dei due eventi.

- Notazione: $E_1 \cup E_2 \cup \dots \cup E_n = \bigcup_{k=1}^n E_k$

L'**intersezione di eventi** $E \cap F$ corrisponde a tutti gli esiti comuni ad E e ad F .

- $E \cap F = \emptyset \implies$ i due eventi sono detti **disgiunti** o **mutuamente esclusivi** o **incompatibili**.

L'**evento complementare** $E^C = S \setminus E$ rappresenta l'insieme di tutti gli esiti di S che non stanno in E .

Esercizi

- ▼ Dato come esperimento il lancio di un dado cubico e dati i due eventi $E = \text{"esito multiplo di 3"}$ e $F = \text{"esito pari"}$, calcolare l'unione e l'intersezione dei due eventi.

Visto che il dado è cubico questo avrà 6 facce, dunque lo spazio campione è il seguente:

$$S = \{e_1, e_2, e_3, e_4, e_5, e_6\}$$

La struttura dei due eventi E ed F sarà dunque la seguente:

$$\begin{aligned}E &= \{e_3, e_6\} \\ F &= \{e_2, e_4, e_6\}\end{aligned}$$

Calcoliamo ora l'unione e l'intersezione di tali eventi:

$$\begin{aligned}E \cup F &= \{e_2, e_3, e_4, e_6\} \\ E \cap F &= \{e_6\}\end{aligned}$$

Definizione frequentista di probabilità

Dato uno spazio campione S e un evento $E \subset S$ si associa all'evento E un numero reale detto **probabilità** $P(E)$ nel seguente modo. Si ripete l'esperimento N volte e sia n_E il numero di volte in cui l'evento E si è verificato, allora:

$$P(E) = \frac{n_E}{N}$$

Questa definizione è valida se il numero di esiti contenuti in S è **finito** e se N è "**grande**".

Definizione classica di probabilità

Dato uno spazio campione S e un evento $E \subset S$ si associa all'evento E un numero reale detto **probabilità** $P(E)$ tale che:

$$P(E) = \frac{\text{n. esiti in } E}{\text{n. esiti in } S}$$

Questa definizione è valida se il numero di esiti contenuti in S è **finito** e se gli esiti sono tutti **equiprobabili**.

Si può arrivare a tale definizione anche utilizzando la definizione assiomatica di probabilità, definita in seguito, per uno spazio campione equiprobabile.

Supponiamo infatti di avere uno spazio campione $S = \{e_1, e_2, \dots, e_n\}$ che contiene n esiti equiprobabili, ovvero $p(e_1) = p(e_2) = \dots = p(e_n) = p$, abbiamo che:

$$\begin{aligned}1 = P(S) &= P(\cup_{j=1}^n e_j) = \sum_{j=1}^n P(e_j) = \sum_{j=1}^n p = np \\ \implies p &= \frac{1}{n}\end{aligned}$$

Sia $E = \{e_1, e_2, \dots, e_k\}$, calcoliamo $P(E)$:

$$P(E) = P(\cup_{j=1}^k e_j) = \sum_{j=1}^k P(e_j) = \sum_{j=1}^k \frac{1}{n} = \frac{k}{n} = \frac{\text{n. esiti in } E}{\text{n. esiti in } S}$$

Definizione assiomatica di probabilità (o assiomi di Kolmogorov)

Dato uno spazio campione S e un evento $E \subset S$ si associa all'evento E un numero reale detto **probabilità** $P(E)$ tale che:

- $P(E) \in [0, 1]$
- $P(S) = 1$
- Dati $E_1, E_2, \dots, E_n \in S$ disgiunti ($E_i \cap E_k = \emptyset$ se $i \neq k$), allora

$$P(\cap_{k=1}^n E_k) = \sum_{k=1}^n P(E_k)$$

Esercizi

▼ Data una scatola contenente 4 palline diverse, calcolare la probabilità che estraendo 2 palline con reimmissione si estrarrebbero 2 palline diverse:

L'evento di cui calcolare la probabilità è il seguente:

$$E = \{\text{estrazione di 2 palline diverse}\}$$

Possiamo calcolare la probabilità $P(E)$ di tale evento utilizzando la definizione classica di probabilità in quanto S è finito e gli esiti sono tutti equiprobabili, dunque:

$$P(E) = \frac{\text{n. esiti di } E}{\text{n. esiti di } S}$$

Il numero di esiti di E corrisponde al numero di esiti possibili estraendo senza reimmissione 2 palline dalla scatola, ovvero le disposizioni semplici di 4 oggetti classe 2.

Il numero di esiti di S corrisponde invece al numero di esiti possibili estraendo con reimmissione 2 palline dalla scatola, ovvero le disposizioni con reimmissione di 4 oggetti classe 2.

Possiamo dunque concludere che:

$$P(E) = \frac{\text{n. esiti di } E}{\text{n. esiti di } S} = \frac{D_{4,2} = \frac{4!}{2!} = 12}{4^2 = 16} = \frac{3}{4}$$

▼ Dati N individui nati in anni non bisestili, calcolare la probabilità dell'evento $E = \{\text{tutti gli individui hanno il compleanno in giorni diversi}\}$.

Utilizziamo la definizione classica di probabilità in quanto tutti gli esiti sono equiprobabili.

$$P(E) = \frac{\text{n. esiti di } E}{\text{n. esiti di } S} = \frac{D_{365,N} = \frac{365!}{(365-N)!}}{D_{365,N} = 365^N} = \frac{365!}{(365-N)! \cdot 365^N}$$

Conseguenze degli assiomi

Prima conseguenza

Dato uno spazio campione S e un evento $E \subseteq S$, allora:

$$P(E^C) = 1 - P(E)$$

Dimostrazione

Per definizione di evento complementare abbiamo che $E^C \cap E = \emptyset$ e $E^C \cup E = S$.

Utilizzando tali ipotesi e gli assiomi di Kolmogorov otteniamo:

$$1 \underset{A2}{=} P(S) = P(E^C \cup E) \underset{A3}{=} P(E^C) + P(E) \implies P(E^C) = 1 - P(E)$$

Q.E.D.

Seconda conseguenza

$$P(\emptyset) = 0$$

Dimostrazione

Sappiamo che $S^C = \emptyset$.

Dalla prima conseguenza degli assiomi sappiamo inoltre che $P(E^C) = 1 - P(E)$.

Utilizziamo dunque queste due ipotesi per calcolare $P(\emptyset)$:

$$P(\emptyset) = P(S^C) = 1 - P(S) \underset{A2}{=} 1 - 1 = 0$$

Q.E.D.

Terza conseguenza

Dato uno spazio campione S e gli eventi $E, F \subseteq S$, se $E \subset F$:

$$P(E) \leq P(F)$$

Dimostrazione

Siccome $E, F \subseteq S$ e $E \subset F$, allora $F = E \cup (E^C \cap F)$.

Calcoliamo dunque $P(F)$ utilizzando questa ipotesi:

$$P(F) = P(E \cup (E^C \cap F)) \underset{A2}{=} P(E) + P(E^C \cap F) \underset{A1}{\geq} P(E)$$

Q.E.D.

Quarta conseguenza

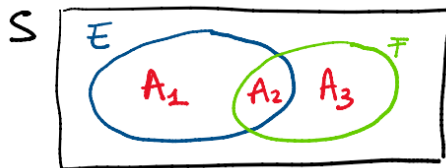
Dato uno spazio campione S e gli eventi $E, F \subseteq S$, si ha che:

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

Osserviamo che se $E \cap F = \emptyset$ (eventi disgiunti) $\implies P(E \cup F) = P(E) + P(F)$, in accordo con $A3$.

Dimostrazione

Per effettuare la dimostrazione utilizziamo il seguente schema:



Notiamo dunque che:

$$\left. \begin{array}{l} A_1 = E \cap F^C \\ A_2 = E \cap F \\ A_3 = E^C \cap F \end{array} \right\} A_i \cap A_k = \emptyset \quad \text{se } i \neq k$$

Calcoliamo $P(E)$ e $P(F)$ in funzione di A_1, A_2 e A_3 :

$$E = A_1 \cup A_2 \implies P(E) = P(A_1 \cup A_2) \underset{A_3}{=} P(A_1) + P(A_2)$$

$$F = A_2 \cup A_3 \implies P(F) = P(A_2 \cup A_3) \underset{A_3}{=} P(A_2) + P(A_3)$$

Calcoliamo ora $P(E \cup F)$ in funzione di A_1, A_2 e A_3 :

$$E \cup F = A_1 \cup A_2 \cup A_3 \implies P(E \cup F) \underset{A_3}{=} P(A_1) + P(A_2) + P(A_3)$$

Infine calcoliamo lato sinistro dell'espressione da verificare $P(E) + P(F) + P(E \cap F)$ in funzione di A_1, A_2 e A_3 :

$$P(E) + P(F) - P(E \cap F) = P(A_1) + P(A_2) + P(A_2) + P(A_3) - P(A_2)$$

$$= P(A_1) + P(A_2) + P(A_3)$$

Notiamo dunque che $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.

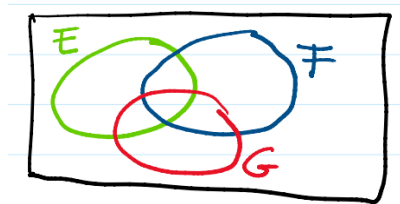
Q.E.D.

Quinta conseguenza

Dato uno spazio campione S e gli eventi $E, F, G \subseteq S$, si ha che:

$$P(E \cup F \cup G) =$$

$$P(E) + P(F) + P(G) - P(E \cap F) - P(E \cap G) - P(F \cap G) + P(E \cap F \cap G)$$



▼ 2.2 - Probabilità condizionata

Probabilità condizionata

Dato uno spazio campione S e due eventi $E, F \subset S$ con $P(F) \neq 0$ si dice probabilità di E **condizionata** da F e si scrive $P(E|F)$ la probabilità di E sapendo che si è verificato F , e si calcola nel seguente modo:

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

Proposizioni

- $P(E|F)P(F) = P(E \cap F) = P(F|E)P(E)$

Eventi indipendenti

Dato uno spazio campione S e n eventi $E_1, E_2, \dots, E_n \subset S$, essi si dicono **indipendenti** se:

$$P\left(\bigcap_{k=1}^n E_k\right) = \prod_{k=1}^n P(E_k)$$

Proposizioni

- Se $P(E \cap F) \neq P(E)P(F)$, E ed F sono dipendenti.
- Se E ed F sono indipendenti e $P(F) \neq 0$, $P(E|F) = P(E)$.

Dimostrazione

Utilizzando la definizione di probabilità condizionata e eventi indipendenti otteniamo che:

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

Eventi indipendenti complementari

Dato uno spazio campione S e due eventi $E, F \subset S$ indipendenti, allora E^C, F^C sono **indipendenti**.

□ Dimostrazione

Proprietà

- $(E_1 \cup E_2 \cup \dots \cup E_n)^C = E_1^C \cap E_2^C \cap \dots \cap E_n^C$
- $(E_1 \cap E_2 \cap \dots \cap E_n)^C = E_1^C \cup E_2^C \cup \dots \cup E_n^C$

Partizione dello spazio campione

Dato uno spazio campione S si dice **partizione dello spazio campione** S l'insieme degli eventi $\{H_1, H_2, \dots, H_n\}$ con $n \in \mathbb{N} \setminus \{0\}$ tale che:

$$H_i \cap H_k = \emptyset \quad i \neq k$$

$$\bigcup_{j=1}^n H_j = S$$

Gli eventi H_1, H_2, \dots, H_n sono detti **ipotesi**.

Teorema delle probabilità totali

Dato uno spazio campione S , un evento $E \subseteq S$ e una partizione di S $\{H_1, H_2, \dots, H_n\}$, allora:

$$P(E) = \sum_{k=1}^n P(E|H_k)P(H_k)$$

☐ Dimostrazione

Teorema di Bayes

Dato uno spazio campione S , un evento $E \subseteq S$ e una partizione di S $\{H_1, H_2, \dots, H_n\}$, allora:

$$P(H_j|E) = \frac{P(E|H_j)P(H_j)}{\sum_{k=1}^n P(E|H_k)P(H_k)}$$

☐ Dimostrazione

Problema della rovina del giocatore

Il seguente problema analizza le probabilità di vittoria di un giocatore che scommette in un gioco nel corso del tempo e al variare delle probabilità di vittoria.

[RovinaGiocatore2023.pdf](#)

▼ 3.0 - Variabili aleatorie

Variabile aleatoria

Una **variabile aleatoria** (o **casuale**) è una variabile che può assumere valori reali associati agli esiti di un esperimento.

Esempi

- $X = \begin{cases} 0 & \text{se esce testa} \\ 1 & \text{se esce croce} \end{cases}$

▼ 3.1 - Variabili aleatorie discrete

Variabile aleatoria discreta

Per definire una variabile aleatoria discreta occorre innanzitutto considerare gli esiti, finiti o numerabili, di un esperimento come dei valori reali discreti, ognuno differente dall'altro.

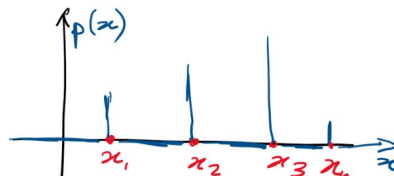
Definiamo quindi una **variabile aleatoria discreta** una variabile aleatoria che può assumere i valori discreti di tali esiti descritti in precedenza. Dunque:

$$X \in \{x_1, x_2, \dots, x_n\} \quad x_k \in \mathbb{R}$$

Funzione di massa di probabilità

Data una variabile aleatoria discreta, definiamo una **funzione di massa di probabilità** come una funzione del tipo $p : \mathbb{R} \rightarrow [0, 1]$ tale che:

$$p(x) \equiv P(X = x)$$



Intuitivamente la funzione di massa di probabilità associa ad ogni possibile risultato di un esperimento la sua probabilità di accadere.

Esempi

- Definita la variabile aleatoria $X = \begin{cases} 0 & \text{se esce testa} \\ 1 & \text{se esce croce} \end{cases}$, possiamo definire la funzione di massa di probabilità del lancio di una moneta nel seguente modo:

$$p(x) = \begin{cases} \frac{1}{2} & x = 0, 1 \\ 0 & x \neq 0, 1 \end{cases}$$

Proposizioni

- $\sum_{j=1}^n p(x_j) = 1$ perchè la funzione di massa di probabilità ha come dominio tutto lo spazio campione.

Funzione di distribuzione o ripartizione di probabilità

Data una variabile aleatoria discreta X , definiamo una **funzione di distribuzione o ripartizione di probabilità** come una funzione del tipo $F : \mathbb{R} \rightarrow [0, 1]$ tale che:

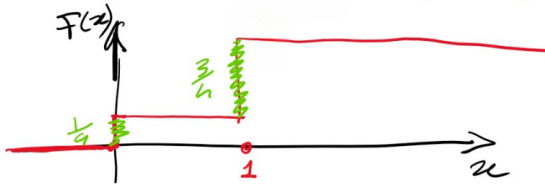
$$F(x) = P(X \leq x)$$

Intuitivamente la funzione di distribuzione o ripartizione di probabilità mostra come le probabilità si accumulano spostandosi lungo l'asse dei risultati di un esperimento.

Esempi

- Definita la variabile aleatoria $X = \begin{cases} 0 & \text{se esce testa} \\ 1 & \text{se esce croce} \end{cases}$ e la funzione di massa di probabilità $p(x) = \begin{cases} \frac{1}{4} & x = 0 \\ \frac{3}{4} & x = 1 \\ 0 & x \neq 0, 1 \end{cases}$ per il lancio di una moneta truccata, possiamo definire la funzione di distribuzione o ripartizione di probabilità nel seguente modo:

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{4} & x = 0 \\ \frac{1}{4} & 0 < x < 1 \\ 1 & x = 1 \\ 1 & x > 1 \end{cases}$$



Proposizioni

- $\lim_{x \rightarrow -\infty} F(x) = P(X \leq -\infty) = 0$
- $\lim_{x \rightarrow +\infty} F(x) = P(X \leq +\infty) = 1$
- Se $a < b$ con $a, b \in \mathbb{R}$, allora $F(a) \leq F(b)$, in quanto F è una funzione mai decrescente.
- Per una variabile aleatoria discreta, F è "a gradini", dunque non derivabile.

▼ 3.2 - Variabili aleatorie continue

Variabile aleatoria continua

Una variabile aleatoria si dice **continua** se ad essa è associata una funzione $f : \mathbb{R} \rightarrow \mathbb{R}^+ \cup \{0\}$ detta **funzione di densità di probabilità** tale che:

$$\forall A \in \mathbb{R} \quad P(X \in A) = \int_A f(x) dx$$

Intuitivamente una variabile aleatoria continua può assumere un numero infinito di valori all'interno di un intervallo (es. misura della lunghezza, temperatura, tempo ecc.). La funzione di densità di probabilità $f : \mathbb{R} \rightarrow \mathbb{R}^+ \cup \{0\}$ è una curva che descrive la distribuzione di probabilità di variabili continue all'interno di un intervallo, ma a differenza della funzione di massa di probabilità non fornisce direttamente la probabilità di un certo valore della variabile continua, ma ciò viene fornito dalla sua integrazione.

Ad esempio, immaginiamo di avere una variabile casuale continua X che rappresenta l'altezza delle persone in un certo gruppo. La funzione di densità di probabilità di X descrive la distribuzione di altezze all'interno di questo gruppo. Se vogliamo calcolare la probabilità che una persona abbia un'altezza tra 160 cm e 170 cm, possiamo prendere la funzione di densità di probabilità di X , che è la curva che rappresenta quanto è probabile ogni possibile valore di altezza, e quindi integriamo questa curva nell'intervallo da 160 cm a 170 cm. Questa integrazione ci darà la probabilità che una persona scelta casualmente in quel gruppo abbia un'altezza compresa tra 160 cm e 170 cm.

Funzione di distribuzione o ripartizione di probabilità con variabile aleatoria continua

Data una variabile aleatoria continua X , definiamo una **funzione di distribuzione o ripartizione di probabilità** come una funzione del tipo $F : \mathbb{R} \rightarrow [0, 1]$ tale che:

$$F(x) = P(X \leq x) = P(X \in]-\infty, x]) = \int_{-\infty}^x f(x)dx$$

Notiamo dunque che una funzione di distribuzione di probabilità con variabile aleatoria continua è continua e derivabile e la sua derivata è:

$$\frac{dF(x)}{dx} = \frac{d}{dx} \left(\int_{-\infty}^x f(x) dx \right) = f(x)$$

Proposizioni

- $f(x) \geq 0$

Dimostrazione

Siccome $F(x)$ è una funzione mai decrescente, allora $f(x) = \frac{dF(x)}{dx} \geq 0$.

Q.E.D.

- $1 = P(X \in \mathbb{R}) = \int_{-\infty}^{+\infty} f(x)dx$

Osservazioni

- $p(a) = P(X = a) = \int_a^a f(x)dx = 0 \quad \forall a \in \mathbb{R}$.
- Non è possibile definire la funzione di densità di probabilità $f(x)$ per variabili discrete in quanto $f(x) = \frac{dF(x)}{dx}$ e per variabili discrete $F(x)$ è "a gradini", dunque non derivabile.
- Per calcolare la probabilità in un certo intervallo di una variabile casuale continua:

1. $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$
2. $P(a \leq X \leq b) = P(X = a) + P(a < X \leq b) = F(b) - F(a)$
3. $P(a < X < b) = P(a < X \leq b) - P(X = b) = F(b) - F(a)$

▼ 3.3 - Coppie di variabili aleatorie

Coppie di variabili aleatorie discrete

Una **coppia di variabili aleatorie discrete** è formata da due variabili casuali discrete che valutano due tipologie di esiti diversi dell'esperimento.

$$\begin{aligned}X &= \{x_1, x_2, \dots, x_n\} \\ Y &= \{y_1, y_2, \dots, y_n\}\end{aligned}$$

Funzione di massa di probabilità

La **funzione di massa di probabilità congiunta** è definita nel seguente modo:

$$p(a, b) = P(X = a, Y = b)$$

La **funzione di massa di probabilità marginale** è una funzione del tipo $\mathbb{R} \rightarrow [0, 1]$ tale che:

- per X

$$p_X(a) = P(X = a) = P(X = a, Y \leq +\infty) = \sum_{k=1}^n p(a, y_k)$$

- per Y

$$p_Y(b) = P(Y = b) = P(X \leq +\infty, Y = b) = \sum_{k=1}^n p(a_k, y)$$

Intuitivamente la funzione di massa di probabilità marginale consente di calcolare il valore della funzione di massa di probabilità per una delle variabili aleatorie in un sistema composto da due o più variabili aleatorie, indipendentemente dalle altre.

Funzione di ripartizione o distribuzione di probabilità

La **funzione di ripartizione o distribuzione di probabilità congiunta** è definita nel seguente modo:

$$F(a, b) = P(X \leq a, Y \leq b)$$

La **funzione di ripartizione o distribuzione di probabilità marginale** è una funzione del tipo $\mathbb{R} \rightarrow [0, 1]$ tale che:

- per X

$$F_X(a) = P(X \leq a) = P(X \leq a, Y \leq +\infty) = F(a, +\infty)$$

- per Y

$$F_Y(b) = P(Y \leq b) = P(X \leq +\infty, Y \leq b) = F(+\infty, b)$$

Intuitivamente la funzione di ripartizione o distribuzione di probabilità marginale consente di calcolare il valore della funzione di ripartizione o distribuzione di probabilità per una delle variabili aleatorie in un sistema composto da due o più variabili aleatorie, indipendentemente dalle altre.

Coppie di variabili aleatorie continue

Due variabili si dicono **congiuntamente continue** se ad esse è associata una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+ \cup \{0\}$ detta **funzione di densità di probabilità congiunta** tale che:

$$\forall C \in \mathbb{R}^2 \quad P((X, Y) \in C) = \iint_C f(x, y) \, dx \, dy$$

Proposizioni

- $f(x, y) \geq 0$
- $P((X, Y) \in \mathbb{R}^2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \, dx \, dy = 1$

Funzione di ripartizione o distribuzione di probabilità

La **funzione di ripartizione o distribuzione di probabilità congiunta** per una coppia di variabili aleatorie continue è definita nel seguente modo:

$$F(a, b) = P(X \leq a, Y \leq b) \quad \forall a, b \in \mathbb{R}$$

La **funzione di ripartizione o distribuzione di probabilità marginale** per una coppia di variabili aleatorie continua è una funzione del tipo $\mathbb{R} \rightarrow [0, 1]$ tale che:

- per X

$$p_X(a) = P(X \leq a) = P(X \leq a, Y \leq +\infty) = F(a, +\infty)$$

- per Y

$$p_Y(b) = P(Y \leq b) = P(X \leq +\infty, Y \leq b) = F(+\infty, b)$$

Proposizioni

- $F : \mathbb{R}^2 \rightarrow [0, 1]$
- $F(a, b) = P(X \leq a, Y \leq b) = \int_{-\infty}^a (\int_{-\infty}^b f(x, y) dx) dy$
Si dimostra che $\frac{\partial^2 F}{\partial a \partial b} = f(a, b)$.
- $F(-\infty, -\infty) = 0$
 $F(+\infty, +\infty) = 1$

Funzione di densità marginale di probabilità

La funzione di densità marginale di probabilità è una funzione del tipo $\mathbb{R} \rightarrow [0, 1]$ tale che:

- per X

$$f_X(a) = \int_{-\infty}^{+\infty} f(a, y) dy$$

- per Y

$$f_Y(b) = \int_{-\infty}^{+\infty} f(x, b) dx$$

□ Dimostrazione

Indipendenza tra variabili aleatorie

Due variabili casuali X e Y si dicono indipendenti se $\forall A, B \in \mathbb{R}$:

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

Tale definizione è semplice ma praticamente impossibile da applicare in quanto occorrerebbe verificarla per qualsiasi valore di A e B in \mathbb{R} . Per facilitare la verifica si introducono dunque i seguenti teoremi.

Teorema 1

Condizione necessaria e sufficiente affinché due **variabili aleatorie** siano **indipendenti** è che:

$$F(a, b) = F_X(a)F_Y(b) \quad \forall a, b \in \mathbb{R}$$

Teorema 2

Condizione necessaria e sufficiente affinché due **variabili aleatorie discrete** siano **indipendenti** è che:

$$p(a, b) = p_X(a)p_Y(b) \quad \forall a, b \in \mathbb{R}$$

Teorema 3

Condizione necessaria e sufficiente affinché due **variabili aleatorie continue** siano **indipendenti** è che:

$$f(a, b) = f_X(a)f_Y(b) \quad \forall a, b \in \mathbb{R}$$

▼ 3.4 - Caratteristiche numeriche delle variabili aleatorie

Valore atteso

Il valore atteso è anche detto valore medio, media teorica, media o speranza matematica.

Data una variabile casuale X si definisce, se esiste, **valore atteso** di X :

$$E[X] = \begin{cases} \sum_{k=1}^n x_k p(x_k) & \text{se } X \text{ variabile discreta} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{se } X \text{ variabile continua} \end{cases}$$

Intuitivamente il valore atteso di una variabile casuale rappresenta la media ponderata dei possibili risultati della variabile, dove i pesi sono dati dalle probabilità associate a ciascun risultato.

Si può notare che il valore atteso può non essere coincidente con uno dei valori assunti da X , ma sicuramente appartiene a $[X_{min}, X_{max}]$.

Proprietà

- Data una variabile casuale X e una funzione $g(X)$:

$$E[g(X)] = \begin{cases} \sum_{k=1}^n g(x_k) p(x_k) & \text{se } X \text{ variabile discreta} \\ \int_{-\infty}^{+\infty} g(x) f(x) dx & \text{se } X \text{ variabile continua} \end{cases}$$

- Data una variabile casuale continua X e $\alpha, \beta \in \mathbb{R}$:

$$E[\alpha X + \beta] = \alpha E[X] + \beta$$

□ Dimostrazione

- Date due variabili casuali continue X e Y e una funzione $h(X, Y)$:

$$E[h(X, Y)] = \begin{cases} \sum_{k=1}^n \sum_{j=1}^m h(x_k, y_j) p(x_k, y_j) & \text{se } X \text{ variabile discreta} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) f(x, y) dx dy & \text{se } X \text{ variabile continua} \end{cases}$$

- Date due variabili casuali continue X e Y con valori attesi $E[X]$ e $E[Y]$:

$$E[X + Y] = E[X] + E[Y]$$

Questa proprietà si estende ad N variabili aleatorie $(E[\sum_{i=1}^N X_i] = \sum_{i=1}^N E[X_i])$.

☐ Dimostrazione

Varianza

Data una variabile casuale X con valor medio $E[X] = \mu$, si definisce **varianza** di x :

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

Intuitivamente la varianza fornisce una misura di quanto i valori di una variabile casuale si discostano dal valore atteso.

Osservazione

- Sia X una variabile casuale discreta con $X = \{x_1, \dots, x_n\}$, allora:

$$Var(X) = \sum_{k=1}^n (x_k - \mu)^2 p(x_k) \geq 0$$

Sia X una variabile casuale continua con $f(x)$ funzione di densità di probabilità, allora:

$$Var(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx \geq 0$$

Proprietà

- $Var(X) \geq 0$
- $Var(X) = \underbrace{E[X^2]}_{\text{momento di ordine 2}} - \left(\underbrace{E[X]}_{\text{momento di ordine 1}} \right)^2$

☐ Dimostrazione

- $Var(\alpha X + \beta) = \alpha^2 Var(X)$

☐ Dimostrazione

Scarto quadratico medio o deviazione standard

Data una variabile casuale X con varianza $Var(X)$, si definisce **scarto quadratico medio** o **deviazione standard**:

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

Intuitivamente lo scarto quadratico medio di una variabile aleatoria rappresenta la dispersione media dei valori della variabile rispetto al valore atteso. Mettendolo a confronto con la varianza, quest'ultima è espressa in unità di misura al quadrato della variabile originale, mentre lo scarto quadratico medio è espressa nella stessa unità di misura della variabile originale. Quindi, se la variabile casuale misura una grandezza in metri, la varianza sarà in metri quadri, mentre lo scarto quadratico medio sarà in metri.

Funzione generatrice dei momenti

Data una variabile casuale X si definisce **funzione generatrice dei momenti** di X la seguente funzione:

$$\phi(t) = E[e^{tX}] \quad t \in \mathbb{R}$$

Proprietà

- $\phi(0) = E[e^{0x}] = E[1] = 1$
- Sia X una variabile casuale discreta $X = \{x_1, \dots, x_n\}$, allora:

$$\phi(t) = \sum_{k=1}^n e^{tx_k} p(x_k)$$

Sia X una variabile casuale continua con funzione di densità di probabilità $f(x)$:

$$\phi(t) = E[e^{tX}] = \int_{-\infty}^{+\infty} e^{tx} f(x) dx$$

- $\lim_{t \rightarrow 0} \frac{d\phi}{dt} = E[X]$

Si può estendere a $\lim_{t \rightarrow 0} \frac{d^n \phi}{dt^n} = E[X^n]$

Dimostrazione

$$\lim_{t \rightarrow 0} \frac{d\phi}{dt} = \lim_{t \rightarrow 0} \frac{d}{dt} (E[e^{tX}]) = \lim_{t \rightarrow 0} E\left[\frac{de^{tX}}{dt}\right] = E[Xe^{tX}] = E[X]$$

- Siano X e Y due variabili aleatorie indipendenti, allora:

$$\phi_{X+Y}(t) = \phi_X(t) + \phi_Y(t)$$

- Siano X e Y due variabili aleatorie, se hanno la stessa funzione generatrice allora sono **identicamente distribuite**, ovvero hanno la stessa funzione di ripartizione o

distribuzione di probabilità, la stessa funzione di densità di probabilità se sono continue o la stessa funzione di massa di probabilità se sono discrete.

Covarianza

Date due variabili casuali X e Y si definisce **covarianza** di X e Y :

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

Intuitivamente la covarianza fornisce un'indicazione di come le variazioni di una variabile sono associate alle variazioni dell'altra. Ad esempio, se si ha una covarianza positiva significa che le due variabili vanno insieme nella stessa direzione (es. temperatura esterna e vendita di gelati), covarianza negativa per direzioni opposte (es. quantità di esercizio fisico e peso corporeo) e covarianza vicina a zero indica che non c'è una relazione lineare evidente tra le variazioni delle due variabili.

Proprietà

- $Cov(X, Y) > 0 \rightarrow$ se X cresce anche Y cresce.
 $Cov(X, Y) < 0 \rightarrow$ se X cresce, Y decresce e viceversa.
- $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$
□ Dimostrazione
- $Cov(X, Y) = Cov(Y, X)$
- $Cov(X, X) = Var(X)$
- $Cov(\alpha X, Y) = Cov(X, \alpha Y) = \alpha Cov(X, Y)$
- $Cov(X, Y) = E[XY] - E[X]E[Y]$

Coefficiente di correlazione

Date due variabili casuali X e Y si definisce **coefficiente di correlazione** di X e Y :

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} \in [-1, 1]$$

Intuitivamente il coefficiente di correlazione fornisce le stesse indicazioni della covarianza fornendo una misura standardizzata che non cambia a seconda del range dei valori delle due variabili, infatti tale coefficiente assume valori che vanno da -1 a 1.

Proprietà

- Se $Y = \alpha X + \beta$, con $\alpha, \beta \in \mathbb{R}$ e $\alpha \neq 0$, allora:

$$\text{Corr}(X, Y) = \begin{cases} +1 & \text{se } \alpha > 0 \\ -1 & \text{se } \alpha < 0 \end{cases}$$

Teorema

Date due variabili casuali indipendenti X e Y con valori attesi $E[X]$ e $E[Y]$, allora:

$$E[XY] = E[X]E[Y] \quad \text{e quindi} \quad \text{Cov}(X, Y) = 0$$

È importante notare che tutte le variabili indipendenti hanno covarianza nulla, ma non è vero il contrario, in quanto non è detto che tutte le variabili con covarianza nulla siano indipendenti.

☐ Dimostrazione

▼ 3.5 - Modelli di variabili casuali discrete

Variabili casuali di Bernoulli

Le **variabili casuali di Bernoulli** ($Be(p)$) rappresentano un modello di variabili casuali discrete le quali comprendono solo i valori 0 e 1:

$$X \sim Be(p) \implies X \in \{0, 1\} \rightarrow \begin{cases} X = 0 \rightarrow \text{non si verifica l'evento A/insuccesso} \\ X = 1 \rightarrow \text{si verifica l'evento A/successo} \end{cases}$$

- **Funzione generatrice dei momenti:** $\phi(t) = q + e^t \cdot p$

☐ Dimostrazione

- **Valore atteso:** $E[X] = p$

☐ Dimostrazione

- **Varianza:** $\text{Var}(X) = p \cdot q$

☐ Dimostrazione

Variabili casuali binomiali

Le **variabili casuali binomiali** ($B(n, p)$) rappresentano un modello di variabili casuali discrete le quali, per un esperimento che si ripete n volte in maniera identica e indipendente, comprendono i valori $\{0, 1, \dots, n\}$ ed il loro valore indica il numero di volte in cui un certo evento A con probabilità p si è verificato:

$$X \sim B(n, p) \implies \begin{cases} X = 0 \rightarrow \text{l'evento } A \text{ non si è mai verificato} \\ X = 1 \rightarrow \text{l'evento } A \text{ si è verificato 1 volta} \\ \dots \\ X = n \rightarrow \text{l'evento } A \text{ si è verificato } n \text{ volte} \end{cases}$$

Notiamo che se $n = 1$, allora la variabile casuale binomiale diventa una variabile casuale di Bernoulli.

- **Funzione di massa:** $p(a) = C_{n,a} \cdot p^a \cdot q^{n-a}$

▼ Dimostrazione

Considerando un valore $k \in \{0, 1, \dots, n\}$, calcolare $P(X = k)$ equivale a calcolare la probabilità che l'evento A si verifichi k volte, ovvero negli n esperimenti sono stati ottenuti k successi e $n - k$ insuccessi. Tale valore si calcola moltiplicando:

- il numero totale di possibili combinazioni in cui si hanno k successi, ovvero $C_{n,k}$.
- la probabilità che A si verifichi k volte, ovvero p^k .
- la probabilità che A non si verifichi $n - k$ volte, ovvero q^{n-k} .

- **Funzione generatrice dei momenti:** $\phi_X(t) = (q + e^t \cdot p)^n$

□ Dimostrazione

- **Valore atteso:** $E[X] = n \cdot p$

▼ Dimostrazione

Il calcolo del valore atteso di una v. c. binomiale tramite definizione è abbastanza ostico:

$$E[X] = \sum_{k=0}^n k \cdot P(X = k) = \sum_{k=0}^n k \cdot C_{n,k} \cdot p^k \cdot q^{n-k}$$

Per questo è possibile facilitarlo considerando la v. c. binomiale come la somma di n v. c. di Bernoulli che indicano se l' i -esimo esperimento si è verificato, ovvero:

$$X = \sum_{i=1}^n X_i \quad \text{con } X_i \sim Be(p)$$

Considerato ciò, si ha che:

$$\begin{aligned} E[X] &= E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = n \cdot E[X_i] \\ &= n \cdot p \end{aligned}$$

- **Varianza:** $Var[X] = n \cdot p \cdot q$

▼ **Dimostrazione**

Anche la varianza può essere calcolata con la stessa tecnica del valore atteso, dunque:

$$\begin{aligned} Var[X] &= Var\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n Var[X_i] = n \cdot Var[X_i] \\ &= n \cdot p \cdot q \end{aligned}$$

Proprietà di riproducibilità di variabili casuali binomiali

Considerando due variabili casuali binomiali indipendenti X e Y , tali che $X \sim B(n, p)$ e $Y \sim B(m, p)$ (ovvero con parametro p uguale), si ha che:

$$X + Y \sim B(n + m, p)$$

□ Dimostrazione

Variabili casuali geometriche

Le **variabili casuali geometriche** ($G(p)$) rappresentano un modello di variabili casuali discrete le quali, per un esperimento che si ripete più volte in maniera identica e indipendente, comprendono i valori $\{1, \dots, +\infty\}$ ed il loro valore indica il numero di volte in cui si è dovuto ripetere l'esperimento al fine di far verificare un certo evento A avente probabilità p :

$$X \sim G(p) \implies \begin{cases} X = 1 \rightarrow \text{l'esperimento è stato ripetuto 1 volta} \\ \text{affinchè } A \text{ si sia verificato} \\ X = 2 \rightarrow \text{l'esperimento è stato ripetuto 2 volte} \\ \text{affinchè } A \text{ si sia verificato} \\ \dots \end{cases}$$

- **Funzione di massa:** $p(a) = q^{a-1} \cdot p$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \frac{1}{p}$

☐ Dimostrazione

- **Varianza:** $Var(X) = \frac{q}{p^2}$

☐ Dimostrazione

Variabili casuali binomiali negative

Le **variabili casuali binomiali negative** ($NB(r, p)$) rappresentano un modello di variabili casuali discrete le quali, per un esperimento che si ripete più volte in maniera identica e indipendente, comprendono i valori $\{r, \dots, +\infty\}$ ed il loro valore indica il numero di volte in cui si è dovuto ripetere l'esperimento al fine di far verificare un numero r di volte un certo evento A avente probabilità p :

$$X \sim NB(r, p) \implies \begin{cases} X = r \rightarrow \text{l'esperimento è stato ripetuto } r \text{ volte} \\ \text{affinchè } A \text{ si sia verificato } r \text{ volte} \\ X = r + 1 \rightarrow \text{l'esperimento è stato ripetuto } r + 1 \text{ volte} \\ \text{affinchè } A \text{ si sia verificato } r \text{ volte} \\ \dots \end{cases}$$

- **Funzione di massa:** $p(a) = C_{a-1, r-1} \cdot p^r \cdot q^{a-r}$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \frac{r}{p}$

☐ Dimostrazione

- **Varianza:** $Var(X) = \frac{r \cdot q}{p^2}$

☐ Dimostrazione

Variabili casuali di Poisson (o degli elementi rari)

Le **variabili casuali di Poisson** ($Po(\lambda)$) rappresentano un modello di variabili casuali discrete le quali, per un esito A che in un certo intervallo di tempo si verifica in media λ volte, comprendono i valori $\{0, 1, \dots, +\infty\}$ ed il loro valore indica il numero di volte in cui l'evento A si è verificato durante tale intervallo di tempo:

$$X \sim Po(\lambda) \implies \begin{cases} X = 0 \rightarrow \text{l'evento non si è mai verificato} \\ X = 1 \rightarrow \text{l'evento si è verificato 1 volta} \\ \dots \end{cases}$$

- **Funzione di massa:** $p(a) = \frac{\lambda^a}{a!} \cdot e^{-\lambda}$
- **Funzione generatrice dei momenti:** $\phi(t) = e^{\lambda \cdot (e^t - 1)}$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \lambda$

☐ Dimostrazione

- **Varianza:** $Var(X) = \lambda$

☐ Dimostrazione

Proprietà di riproducibilità di variabili casuali di Poisson

Considerando due variabili casuali di Poisson indipendenti X e Y , tali che $X \sim Po(\lambda_1)$ e $Y \sim Po(\lambda_2)$, si ha che:

$$X + Y \sim Po(\lambda_1 + \lambda_2)$$

☐ Dimostrazione

Approssimare una variabile casuale di Poisson ad una variabile casuale binomiale

La v. c. di Poisson di parametro λ può essere vista come il limite di una v. c. binomiale di parametri n e p con $n \gg 1$ e $np = \lambda$.

☐ Dimostrazione

▼ 3.6 - Modelli di variabili casuali continue

Variabili casuali uniformi

Le **variabili casuali uniformi** ($U(\alpha, \beta)$) rappresentano un modello di variabili casuali continue le quali all'interno di un certo intervallo $[\alpha, \beta]$ assumono lo stesso valore di probabilità $\frac{1}{\beta - \alpha}$ (in modo che $\int_{-\infty}^{+\infty} f(x)dx = 1$), e 0 altrove:

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{se } x \in [\alpha, \beta] \\ 0 & \text{altrove} \end{cases}$$

Si dicono casuali in quanto tutti gli esiti presenti all'interno dell'intervallo $[\alpha, \beta]$ hanno la stessa probabilità di accadere.

Considerando un intervallo $[c, d]$ sottoinsieme di $[\alpha, \beta]$, si ha che la probabilità di avere $X \in [c, d]$ sarà uguale a $\frac{d - c}{\beta - \alpha}$.

- **Funzione di ripartizione:**

$$F(x) = \begin{cases} 0 & \text{se } x < \alpha \\ \frac{x - \alpha}{\beta - \alpha} & \text{se } \alpha \leq x \leq \beta \\ 1 & \text{se } x > \beta \end{cases}$$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \frac{\beta + \alpha}{2}$

☐ Dimostrazione

- **Varianza:** $Var(X) = \frac{(\beta - \alpha)^2}{12}$

☐ Dimostrazione

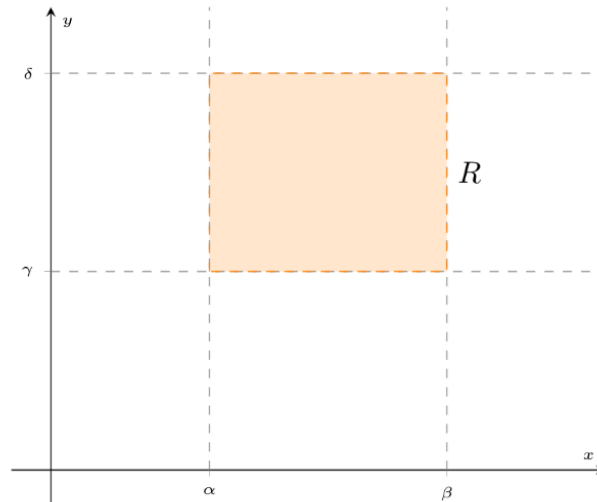
☐ Problemi sulla probabilità che un sistema in serie e parallelo funzioni fino ad un istante a utilizzando v. c. uniformi

Coppia di variabili casuali uniformi

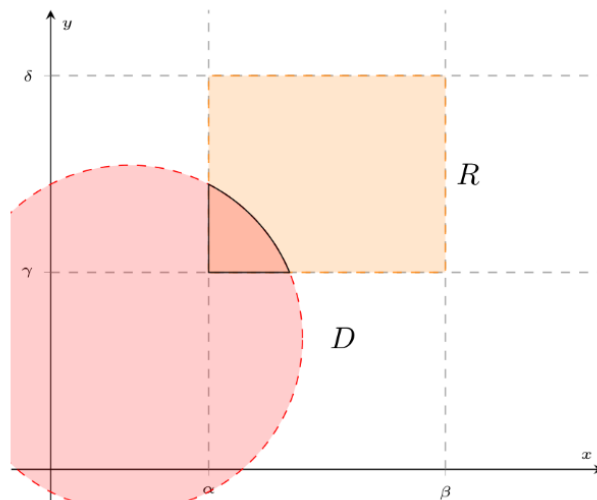
Considerando due variabili casuali continue uniformi indipendenti tali che

$$X \sim U(\alpha, \beta) \quad Y \sim U(\gamma, \delta)$$

è possibile definire l'insieme R , detto anche **rettangolo di definizione di (X, Y)** :



Può capitare di domandarsi la probabilità che la coppia (X, Y) appartenga ad un certo insieme D , graficamente:



Tale probabilità equivale a:

$$P((X, Y) \in D) = \frac{\text{Area di } D \cap R}{\text{Area di } R}$$

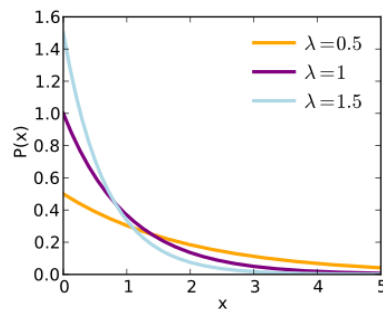
□ **Dimostrazione**

Variabili casuali esponenziali

Le **variabili casuali esponenziali** ($E(\lambda)$) rappresentano un modello di variabili casuali continue le quali possiedono la seguente funzione di densità di probabilità:

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda \cdot x} & \text{se } x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

Il valore della variabile rappresenta il tempo intercorso tra due "eventi rari" che avvengono in maniera indipendente e continua ad un tasso medio di λ .



- **Funzione di ripartizione:**

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda \cdot x} & \text{se } x \geq 0 \end{cases}$$

☐ Dimostrazione

- **Funzione generatrice dei momenti:** $\phi(t) = \frac{\lambda}{\lambda - t}$ per $t \leq \lambda$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \frac{1}{\lambda}$

☐ Dimostrazione

- **Varianza:** $Var(X) = \frac{1}{\lambda^2}$

☐ Dimostrazione

☐ Problemi sulla probabilità che un sistema in serie e parallelo funzioni fino ad un istante a utilizzando v. c. esponenziali

Variabile casuale esponenziale moltiplicata per uno scalare

Considerando una v. c. esponenziale $X \sim E(\lambda)$, se $Y = c \cdot X$ (con $x \in \mathbb{R}^+$) si ha che $Y \sim E(\frac{\lambda}{c})$.

☐ Dimostrazione

Mancanza di memoria per variabili casuali esponenziali

Considerando una v. c. esponenziale $X \sim E(\lambda)$, si ha che è caratterizzata dalla **mancanza della memoria**, ovvero si ha che considerando $a, b \in \mathbb{R}^+$ vale:

$$P(X > a + b | X > a) = P(X > b)$$

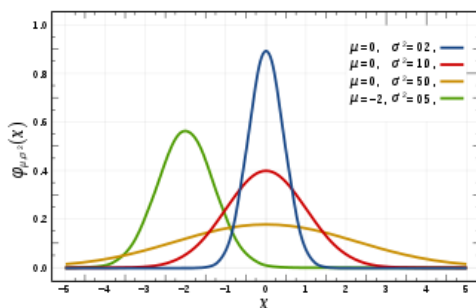
Intuitivamente questo teorema spiega come una volta che si è atteso un certo periodo di tempo senza che l'evento di interesse si sia verificato, la distribuzione di probabilità del tempo residuo rimanente non cambia, indipendentemente da quanto tempo si è già atteso.

☐ Dimostrazione

Variabili casuali gaussiane

Le **variabili casuali gaussiane** ($N(\mu, \sigma)$) rappresentano un modello di variabili casuali continue le quali possiedono la seguente funzione di densità di probabilità:

$$f(x) = \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



- **Funzione di ripartizione:** $F(a) = \int_{-\infty}^a \frac{1}{\sigma \cdot \sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$

La funzione di ripartizione è un integrale convergente, ma che tuttavia non presenta una primitiva esprimibile analiticamente. Per effettuare i calcoli, è quindi necessario utilizzare apposite tabelle.

☐ Dimostrazione

- **Funzione generatrice dei momenti:** $\phi(t) = e^{t\mu} + \frac{t^2 \cdot \sigma^2}{2}$

☐ Dimostrazione

- **Valore atteso:** $E[X] = \mu$

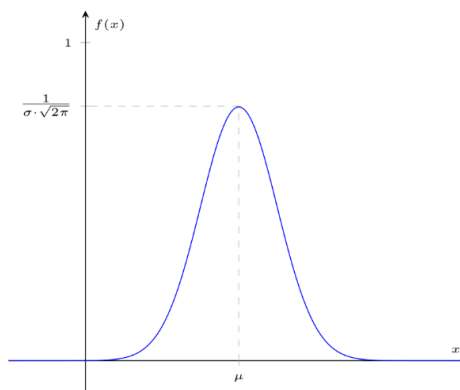
☐ Dimostrazione

- **Varianza:** $Var(X) = \sigma^2$

☐ Dimostrazione

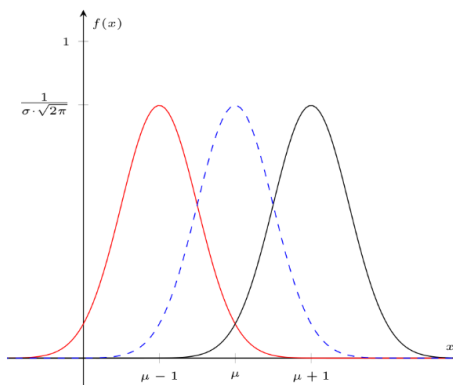
Effetto della variazione dei parametri sulla funzione di densità

Considerando la funzione di densità:

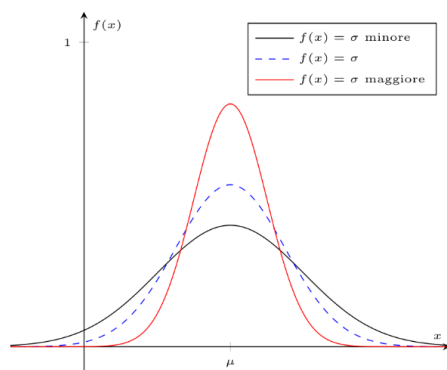


è possibile effettuare le seguenti considerazioni sui parametri:

- Il massimo della funzione e l'asse di simmetria della funzione si trovano entrambi in corrispondenza del valore atteso μ .
- Al variare di μ si ottiene una traslazione orizzontale della curva.



- Al variare di σ si ottiene un cambiamento sul valore del picco (al crescere di σ il valore del picco diminuisce e viceversa).



Trasformazione lineare di una v. c. gaussiana

Considerando una variabile casuale gaussiana $X \sim N(\mu, \sigma)$ e considerando la variabile casuale $Y = \alpha X + \beta$ si avrà che:

$$Y \sim N(\alpha\mu + \beta, |\alpha|\sigma)$$

☐ Dimostrazione

Variabile casuale gaussiana standard

Una variabile casuale gaussiana X si definisce **standard** se:

$$X \sim N(0, 1)$$

Ricondurre una qualsiasi v. c. gaussiana a una v. c. gaussiana standard

È possibile ricondurre una qualsiasi v. c. gaussiana normale $X \sim N(\mu, \sigma)$ a una v. c. gaussiana standard $Z = \frac{X - \mu}{\sigma}$ ($Z \sim (0, 1)$).

☐ Dimostrazione

Calcolo della funzione di ripartizione di una gaussiana utilizzando la standard

È importante notare che $F_X(a) = F_Z\left(\frac{a - \mu}{\sigma}\right)$. Tramite questa osservazione è possibile calcolare il valore della funzione di ripartizione in un qualsiasi punto per una qualsiasi variabile casuale gaussiana avendo tuttavia sottomano solamente la tabella di valori della funzione di ripartizione della gaussiana standard.

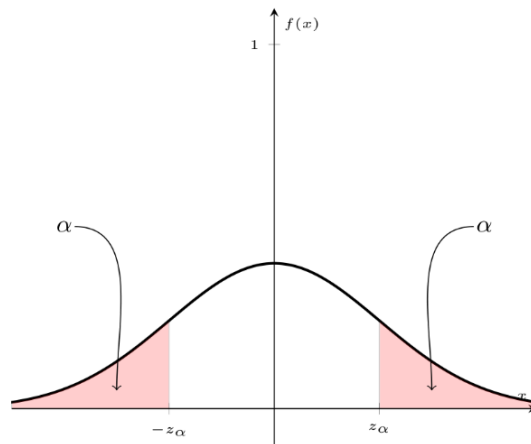
□ Dimostrazione

Valore critico di una v. c. gaussiana standard

Considerando una variabile casuale gaussiana standard $X \sim N(0, 1)$ si definisce valore critico z_α il valore:

$$z_\alpha : P(Z > z_\alpha) = P(Z < -z_\alpha) = \alpha$$

ovvero il valore z_α per cui l'area sottesa alla curva negli intervalli $(-\infty, -z_\alpha]$ e $[z_\alpha, +\infty)$ è α .



Si sa per definizione che l'area sottesa alla funzione di densità di probabilità di una variabile casuale è 1, dunque $z_{\frac{1}{2}} = 0$, in quanto le due metà della funzione di densità hanno un'area sottesa pari a $\frac{1}{2}$.

Proprietà di riproducibilità di una v. c. gaussiana

Considerando due variabili casuali gaussiane indipendenti $X \sim N(\mu_1, \sigma_1)$ e $Y \sim N(\mu_2, \sigma_2)$ si ha che:

$$X + Y \sim N(\mu_1 + \mu_2, \sqrt{\sigma_1^2 + \sigma_2^2})$$

□ Dimostrazione

▼ 3.7 - Funzioni di variabili casuali continue

Funzione di una variabile casuale continua

Considerando una variabile casuale continua X la cui funzione di densità è $f_X(x)$ è possibile definire una **funzione di tale variabile casuale continua** come:

$$Y = g(X)$$

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

Momento di ordine 2: $E[Y^2] = E[(g(X))^2] = \int_{-\infty}^{+\infty} (g(x))^2 \cdot f_X(x) dx$

- **Varianza:**

$$\begin{aligned} \text{Var}(X) &= E[Y^2] - (E[Y])^2 \\ &= \int_{-\infty}^{+\infty} (g(x))^2 \cdot f_X(x) dx - \left(\int_{-\infty}^{+\infty} g(x) \cdot f_X(x) dx \right)^2 \end{aligned}$$

- **Funzione di densità e di ripartizione:**

Per calcolare la funzione di densità e di ripartizione di una v. c. continua funzione di un'altra v. c. continua occorre prendere in considerazione 3 diversi casi:

1. $g(X)$ è una funzione monotona crescente:

$$\begin{aligned} P(Y \leq a) &= \int_{-\infty}^{g^{-1}(a)} f_X(x) dx \\ f_Y(a) &= \left(\frac{d}{da} g^{-1}(a) \right) \cdot f_X(g^{-1}(a)) \end{aligned}$$

☐ Dimostrazione

2. $g(X)$ è una funzione monotona decrescente:

$$\begin{aligned} P(Y \leq a) &= \int_{g^{-1}(a)}^{+\infty} f_X(x) dx \\ f_Y(a) &= -\left(\frac{d}{da} g^{-1}(a) \right) \cdot f_X(g^{-1}(a)) \end{aligned}$$

☐ Dimostrazione

3. $g(X)$ è una funzione non monotona:

non esiste una formula generale, vedi dimostrazione

☐ Dimostrazione

Funzione di due variabili casuali continue

Considerando una coppia di variabili casuali continue (X, Y) la cui funzione di densità congiunta è $f(x, y)$, è possibile definire la **funzione di tali variabili casuali continue** come:

$$Z = h(X, Y)$$

- **Valore atteso:** $E[Z] = E[h(X, Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) \cdot f(x, y) \, dx dy$

Momento di ordine 2:

$$E[Z^2] = E[(h(x, y))^2] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (h(x, y))^2 \cdot f(x, y) \, dx dy$$

- **Varianza:**

$$\begin{aligned} \text{Var}(Z) &= E[Z^2] - (E[Z])^2 \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (h(x, y))^2 \cdot f(x, y) \, dx dy - \left(\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) \cdot f(x, y) \, dx dy \right)^2 \end{aligned}$$

- **Funzione di densità e di ripartizione:**

Non esistono formule generali, dunque analizziamo solo il caso particolare in cui:

$$Z = X + Y \implies \begin{cases} F_Z(a) = \int_{-\infty}^{-x+a} \left(\int_{-\infty}^{+\infty} f(x, y) \, dx \right) dy \\ f_Z(a) = \int_{-\infty}^{+\infty} f(x, -x+a) \, dx \end{cases}$$

□ Dimostrazione

Variabile casuale lognormale

Considerando una variabile casuale gaussiana $W \sim N(\mu, \sigma)$, è possibile definire la **variabile casuale lognormale** $X \sim \text{lognormale}(\mu, \sigma)$ come:

$$X = e^W$$

- **Valore atteso:** $E[X] = e^{\mu + \frac{\sigma^2}{2}}$

□ Dimostrazione

- **Varianza:** $\text{Var}(X) = e^{2\mu + \sigma^2} \cdot (e^{\sigma^2} - 1)$

☐ Dimostrazione

- **Funzione di ripartizione:**

$$F_X(a) = \begin{cases} 0 & \text{se } a < 0 \\ F_Z\left(\frac{\ln a - \mu}{\sigma}\right) & \text{se } a \geq 0 \end{cases} \quad \text{con } Z \sim N(0, 1)$$

☐ Dimostrazione

- **Funzione di densità:**

$$f_X(a) = \begin{cases} 0 & \text{se } a < 0 \\ \frac{1}{\sigma a \sqrt{2\pi}} e^{-\frac{(\frac{\ln a - \mu}{\sigma})^2}{2}} & \text{se } a \geq 0 \end{cases}$$

☐ Dimostrazione

Proprietà di riproducibilità di variabili casuali lognormali

Considerando due variabili casuali lognormali indipendenti X_1 e X_2 tali che $X_1 \sim \text{lognormale}(\mu_1, \sigma_1)$ e $X_2 \sim \text{lognormale}(\mu_2, \sigma_2)$, si ha che:

$$X_1 \cdot X_2 \sim \text{lognormale}(\mu_1 + \mu_2, \sqrt{(\sigma_1)^2 + (\sigma_2)^2})$$

☐ Dimostrazione

▼ 3.8 - Legge dei grandi numeri

Disuguaglianza di Markov

Data una variabile casuale X a valori non negativi ($X \geq 0$) e un numero $a \in \mathbb{R}^+$:

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

☐ Dimostrazione

Disuguaglianza di Čebyšëv

Data una variabile casuale X (con $E[X] = \mu$ e $Var(X) = \sigma^2$) e dato un numero $r \in \mathbb{R}^+$:

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

□ Dimostrazione

Legge dei grandi numeri

Consideriamo un esperimento ripetuto n volte nelle stesse condizioni e in maniera indipendente. Sia $\{X_1, X_2, \dots, X_n\}$ l'insieme delle variabili aleatorie delle singole esecuzioni. Supponendo che X_1, \dots, X_n siano identicamente distribuite e indipendenti con $E[X_k] = \mu$ e $Var(X_k) = \sigma^2$, è possibile affermare che la probabilità della media aritmetica di tali variabili converge al valore atteso μ , ovvero:

$$\lim_{n \rightarrow +\infty} P(|\bar{X} - \mu| \geq \epsilon) = 0 \quad \begin{array}{l} \forall \epsilon \in \mathbb{R}^+ \\ (\epsilon \text{ piccolo a piacere}) \end{array}$$

$$\text{con } \bar{X} = \sum_{k=1}^n \frac{X_k}{n}$$

Intuitivamente ciò significa che:

$$n \rightarrow +\infty \implies P(\bar{X} = \mu) = 1 \text{ e } P(\bar{X} \neq \mu) = 0$$

□ Dimostrazione

Corollario di Bernoulli

Al tendere ad infinito del numero di prove, la definizione frequentista dell'evento A :

$$f_A = \frac{n. \text{ tentativi in cui } A \text{ si verifica}}{n. \text{ tentativi svolti}}$$

si ha che converge alla probabilità teorica di A ($p(A)$), ovvero:

$$\lim_{n \rightarrow +\infty} P(|f_A - p(A)| \geq \epsilon) = 0 \quad \begin{array}{l} \forall \epsilon \in \mathbb{R}^+ \\ (\epsilon \text{ piccolo a piacere}) \end{array}$$

☐ Dimostrazione

Metodi Monte Carlo

I **metodi Monte Carlo** sono una classe di metodologie che utilizzano eventi casuali e la teoria della probabilità per determinare quantità non casuali.

☐ Problema degli spilli di Buffon

▼ 3.9 - Teorema del limite centrale

Teorema del limite centrale

Considerando una successione di variabili casuali identicamente distribuite e indipendenti X_1, \dots, X_n di valore atteso μ e varianza σ^2 , si ha che la funzione di distribuzione della variabile casuale

$$T_n = \frac{X_1 + \dots + X_n - n \cdot \mu}{\sigma \cdot \sqrt{n}}$$

tende, per $n \rightarrow +\infty$, a quella di una variabile casuale gaussiana standard:

$$F_{T_n}(a) \xrightarrow{n \rightarrow +\infty} F_Z(a) \quad \forall a \in \mathbb{R} \text{ e } Z \sim N(0, 1)$$

☐ Dimostrazione

- **Valore atteso:** $E[T_n] = 0$

☐ Dimostrazione

- **Varianza:** $Var(T_n) = 1$

☐ Dimostrazione

Conseguenze del limite centrale

1. Approssimazione di v. c. binomiali a v. c. gaussiane

Considerando una variabile casuale binomiale $Y \sim B(n, p)$ si ha che è possibile dire che per n abbastanza grandi è valida l'approssimazione:

$$Y \sim B(n, p) \sim N(n \cdot p, \sqrt{n \cdot p \cdot q})$$

□ Dimostrazione

Da notare che nella pratica (esercizi) “ n abbastanza grandi” corrisponde a $n \geq 30$.

Approssimazione alla continuità

Consideriamo le v. c. $D \sim B(n, p)$ e $M \sim N(n \cdot p, \sqrt{n \cdot p \cdot q})$ con n abbastanza grande, ovvero è possibile approssimare D con M , e consideriamo la probabilità che assumano un certo valore uguale, minore o maggiore di k . Siccome D è per natura una variabile discreta e M una variabile continua, le modalità per calcolare tale probabilità sono diverse per le due variabili:

$$P(D = k) = C_{n,k} \cdot p^k \cdot q^{n-k}$$
$$P(M = k) = \int_k^k f_M(x) dx = 0$$

Un'idea per recuperare questa differenza può essere seguire le seguenti regole di approssimazione:

- $P(D = k) = P(k - \frac{1}{2} < M \leq k + \frac{1}{2})$
- $P(D \leq k) = P(M \leq k + \frac{1}{2})$
- $P(D < k) = P(M \leq k - \frac{1}{2})$
- $P(D \geq k) = P(M \geq k - \frac{1}{2})$
- $P(D > k) = P(M \geq k + \frac{1}{2})$

2. Approssimazione della somma di variabili casuali

Considerando la somma di n variabili casuali indipendenti di valore atteso μ e varianza σ^2 , $S_n = X_1 + \dots + X_n$ (il cui valore atteso è $n \cdot \mu$ e la deviazione standard è $\sqrt{n} \cdot \sigma$), si ha che è possibile dire che per n abbastanza grandi è valida l'approssimazione:

$$S_n \sim N(n \cdot \mu, \sigma \cdot \sqrt{n})$$

□ Dimostrazione

3. Approssimazione della media campionaria di variabili casuali

Considerando la media di n variabili casuali indipendenti di valore atteso μ e varianza σ^2 , $\overline{X} = \frac{X_1 + \dots + X_n}{n}$ (il cui valore atteso è μ e la deviazione standard è $\frac{\sigma}{\sqrt{n}}$), si ha che è possibile dire che per n abbastanza grandi è valida l'approssimazione:

$$\overline{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Intuitivamente, se X_1, \dots, X_n rappresentano i risultati di n esperimenti, la loro media si comporta come una gaussiana se $n \geq 30$, indipendentemente dalla distribuzione probabilistica di X_k .

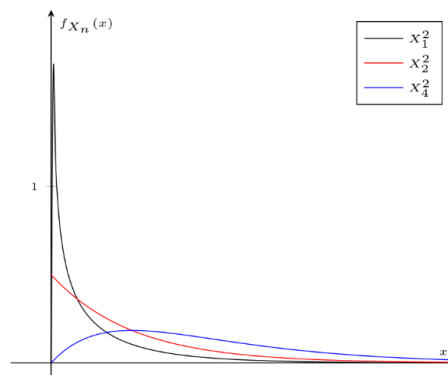
□ Dimostrazione

Variabile casuale chi-quadro a n gradi di libertà

Una variabile casuale continua X si definisce **chi-quadro a n gradi di libertà**, e si scrive $X \sim X_n^2$, se è la somma di n variabili casuali gaussiane standard elevate al quadrato, ovvero:

$$X = \sum_{k=1}^n (Z_k)^2 \quad \text{con } Z_k \sim N(0, 1) \text{ indipendenti}$$

- **Funzione di densità:** particolarmente complicata. Sappiamo però che è ≥ 0 e che è caratterizzata dal seguente grafico:



- **Valore atteso:** $E[X] = n$

□ Dimostrazione

- **Varianza:** $Var(X) = 2n$

Proprietà di riproducibilità

Considerando due variabili chi-quadro indipendenti X_1 e X_2 , tali che $X_1 \sim X_n^2$ e $X_2 \sim X_m^2$, si ha che:

$$X_1 + X_2 \sim X_{n+m}^2$$

Approssimazione ad una variabile casuale gaussiana standard

Considerando $n \rightarrow +\infty$, è possibile approssimare una v. c. chi-quadro ad una v. c. gaussiana standard grazie al teorema del limite centrale:

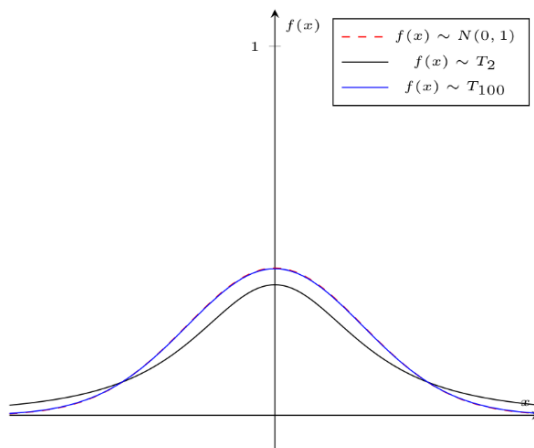
$$X \sim X_n^2 \xrightarrow{n \rightarrow +\infty} X \sim N(0, 1)$$

Variabile casuale t di Student

Una variabile casuale continua X si definisce **t di Student a n gradi di libertà**, e si scrive $X \sim T_n$, se corrisponde a:

$$X = \frac{Z}{\sqrt{\frac{C_n}{n}}} \quad \text{con } Z \sim N(0, 1) \text{ e } C_n \sim X_n^2 \text{ indipendenti}$$

- Funzione di densità: particolarmente complicata. Sappiamo però che è ≥ 0 e che è caratterizzata dal seguente grafico:



Approssimazione ad una variabile casuale gaussiana standard

Considerando $n \rightarrow +\infty$, è possibile approssimare una variabile casuale t di Student ad una variabile casuale gaussiana standard grazie al teorema del limite centrale:

$$X \sim T_n \xrightarrow{n \rightarrow +\infty} X \sim N(0, 1)$$

▼ 3.10 - Processi stocastici

Processo stocastico

Un **processo stocastico** è una famiglia di variabili casuali che dipendono da un parametro (solitamente il tempo).

Processo stocastico di Poisson

Un processo stocastico si definisce di **Poisson** se le variabili casuali rappresentano il numero di eventi che accadono in un certo intervallo di tempo, ovvero

$$N(t) = \text{"numero di eventi in }]0, t]\text{" con } t \in \mathbb{R}^+$$

sotto le seguenti ipotesi:

- In $t = 0$ si ha che il numero di eventi è **nullo**, ovvero $N(0) = 0$.
- Gli incrementi sono **indipendenti**, ovvero il numero di eventi che avvengono in intervalli di tempo disgiunti è indipendente.
- Gli incrementi sono **stazionari**, ovvero il numero di eventi che accadono in un certo intervallo di tempo dipende dall'ampiezza dell'intervallo, ma non dalla sua posizione nella retta reale. Si ha quindi che nell'intervallo $]s, s + t]$ il numero di eventi è $N(t)$ esattamente come nell'intervallo $]0, t]$.
- Considerando $\lambda \in \mathbb{R}^+$, si ha che

$$\lim_{h \rightarrow 0} \frac{P(N(h) = 1)}{h} = \lambda$$

ovvero

$$P(N(h) = 1) \simeq h \cdot \lambda \quad \text{se } h < 1$$

λ è detta **costante del processo stocastico**.

- Si ha che

$$\lim_{h \rightarrow 0} \frac{P(N(h) \geq 2)}{h} = 0$$

ovvero

$$P(N(h) \geq 2) \simeq 0 \quad \text{se } h < 1$$

Processo stocastico di Poisson come variabile casuale di Poisson

Considerando un processo stocastico di Poisson si ha che ha una distribuzione di Poisson di parametro $\lambda \cdot t$, ovvero

$$N(t) \sim Po(\lambda \cdot t)$$

☐ Dimostrazione

Intertempi del processo stocastico di Poisson

Considerando un processo stocastico di Poisson, è possibile definire l'intertempo

$X_k =$ "tempo che intercorre tra il $(k - 1)$ -esimo e il k -esimo evento"

ovvero una variabile casuale continua.

Intertempi nel processo stocastico di Poisson come v. c. esponenziali

Considerando un processo stocastico di Poisson $N(t)$ con un generico intertempo X_k , si ha che ogni variabile X_k è assimilabile ad una variabile casuale esponenziale di parametro λ , ovvero

$$X_k \sim E(\lambda)$$

☐ Dimostrazione

▼ 4.0 - Inferenza statistica

▼ 4.1 - Concetti introduttivi

Popolazione

Si definisce **popolazione** un insieme infinito (o molto grande) di individui (o oggetti) che possiedono una certa caratteristica misurabile.

Campione

Si definisce **campione** un sottoinsieme finito di individui (o oggetti) della popolazione scelto in maniera casuale.

Campione come insieme di variabili casuali

Data una popolazione ed un certo campione, è possibile rappresentare una determinata caratteristica tramite una variabile casuale X . Per rendere più semplice l'analisi, ipotizzeremo che ogni variabile casuale sia indipendente e identicamente distribuita (ovvero con stessa funzione di ripartizione F).

È quindi possibile definire un campione come un insieme di variabili casuali indipendenti e identicamente distribuite, ovvero:

$$\text{Campione} = (X_1, \dots, X_n)$$

Inferenza statistica

L'**inferenza statistica** è il processo per cui, date delle caratteristiche su un campione, si inducono su una popolazione.

In base alla conoscenza o meno della funzione di distribuzione F , è possibile considerare due diversi tipi di inferenza, ovvero:

- **Inferenza parametrica**, se la funzione F è conosciuta eccetto che per un parametro. In questo caso lo scopo dell'inferenza è determinare tale parametro.

Un tipico problema di inferenza parametrica consiste nel, dopo aver riconosciuto il tipo di distribuzione, ad esempio una binomiale di parametri n e p , dover stimare uno o entrambi i parametri n e p .

- **Inferenza non parametrica**, se la funzione F non è nota. Lo scopo dell'inferenza in questo caso è conoscere il tipo di distribuzione.

Statistica

Si definisce statistica Y una funzione delle variabili casuali che compongono il campione, ovvero:

$$Y = g(X_1, \dots, X_n)$$

Stimatore

Considerando una funzione di ripartizione F_σ , ovvero conosciuta a meno di un parametro σ , si ha che per stimare il valore di σ , detto **stima**, è possibile utilizzare uno **stimatore** $\hat{\sigma}$, ovvero una statistica che associa ad ogni campione la sua stima:

$$\sigma = \hat{\sigma}(X_1, \dots, X_n)$$

Affinchè tale stimatore sia valido, deve essere:

- **Corretto**, ovvero $E[\hat{\sigma}] = \sigma$.
- **Efficiente**, ovvero si ha che deve avere una varianza più piccola possibile.
- **Consistente**, ovvero $Var(\hat{\sigma}) \xrightarrow{n \rightarrow +\infty} 0$.

Esempi di stimatori

- **Stimatore media campionaria**

Considerando il campione (X_1, \dots, X_n) , dove il valor medio e la varianza di ogni variabile casuale valgono rispettivamente μ e σ^2 , è possibile definire lo stimatore media campionaria come:

$$\bar{X} = \frac{\sum_{k=1}^n X_k}{n}$$

- **Valore atteso:** $E[\bar{X}] = \mu$

☐ Dimostrazione

- **Varianza:** $Var(X) = \frac{\sigma^2}{n}$

☐ Dimostrazione

È possibile verificare che \bar{X} è corretto per il parametro μ in quanto $E[\bar{X}] = \mu$ e consistente in quanto $\lim_{n \rightarrow +\infty} Var(\bar{X}) = \lim_{n \rightarrow +\infty} \frac{\sigma^2}{n} = 0$.

La distribuzione dello stimatore media campionaria è gaussiana se:

- la popolazione da cui proviene il campione è gaussiana, per via della proprietà di riproducibilità.
 - la popolazione da cui proviene il campione non è gaussiana ma $n \geq 30$, per via del teorema del limite centrale.
- **Stimatore varianza campionaria**

Considerando il campione (X_1, \dots, X_n) , dove il valor medio e la varianza di ogni variabile casuale valgono rispettivamente μ e σ^2 , è possibile definire lo stimatore varianza campionaria come:

$$S^2 = \frac{1}{n-1} \cdot \sum_{k=1}^n (X_k - \bar{X})^2$$

(viene utilizzato $n - 1$ poichè lo stimatore deve essere corretto)

- **Valore atteso:** $E[S^2] = \sigma^2$

□ Dimostrazione

Si dimostra che se la popolazione è gaussiana è possibile ricondurre uno stimatore varianza campionaria ad una variabile chi-quadro ad $n - 1$ gradi di libertà nel seguente modo:

$$\frac{(n-1)S^2}{\sigma^2} \sim X_{n-1}^2$$

È possibile inoltre ricondurre uno stimatore varianza campionaria ad una variabile casuale t di Student nel seguente modo:

$$\frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim T_{n-1}$$

□ Dimostrazione

▼ 4.2 - Intervalli di confidenza

Intervallo di confidenza

Per la stima del parametro σ si può optare per l'utilizzo di uno stimatore in un dato set di esperimenti. Tale stima non è però puntuale in quanto, ripetendo gli esperimenti, si possono ottenere valori diversi per lo stimatore. Si preferisce quindi considerare un intervallo di valori possibili, detto intervallo di confidenza.

L'intervallo di confidenza è un intervallo di valori in cui si ha fiducia $1 - \beta$ (con $\beta \in [0, 1]$) che il parametro da stimare abbia valore compreso in tale intervallo.

Tipologie di intervalli di confidenza

- **Intervallo di confidenza bilaterale per la media di una popolazione gaussiana con varianza nota**

Considerando il campione (X_1, \dots, X_n) , dove $X_k \sim N(\mu, \sigma)$ con σ nota e μ incognita, si vuole stimare la media μ .

Si ha che:

$$\mu \in \left[\bar{X} - z_{\frac{\beta}{2}} \cdot \frac{\sigma}{\sqrt{n}}, z_{\frac{\beta}{2}} \cdot \frac{\sigma}{\sqrt{n}} + \bar{X} \right]$$

con confidenza $1 - \beta$, dove $[-z_{\frac{\beta}{2}}, z_{\frac{\beta}{2}}]$ è l'intervallo in cui si ha probabilità $1 - \beta$ che Z assuma valore all'interno di esso. $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ e $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

☐ Dimostrazione

• **Intervallo di confidenza bilaterale per la media di una popolazione gaussiana con media e varianza incognite**

Considerando il campione (X_1, \dots, X_n) , dove $X_k \sim N(\mu, \sigma)$ con μ e σ incognite, si vuole stimare la media μ .

Si ha che:

$$\mu \in \left[\bar{X} - t_{n-1, \frac{\beta}{2}} \cdot \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1, \frac{\beta}{2}} \cdot \frac{S}{\sqrt{n}} \right]$$

con confidenza $1 - \beta$, dove $[-t_{n-1, \frac{\beta}{2}}, t_{n-1, \frac{\beta}{2}}]$ è l'intervallo in cui si ha probabilità $1 - \beta$ che t_{n-1} assuma valore all'interno di esso. $t_{n-1} = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim T_{n-1}$ e $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

☐ Dimostrazione

• **Intervallo di confidenza bilaterale per la varianza di una popolazione gaussiana con media e varianza incognite**

Considerando il campione (X_1, \dots, X_n) , dove $X_k \sim N(\mu, \sigma)$ con μ e σ incognite, si vuole stimare la varianza σ .

Si ha che:

$$\sigma^2 \in \left[\frac{(n-1) \cdot S^2}{X_{n-1, \frac{\beta}{2}}^2}, \frac{(n-1) \cdot S^2}{X_{n-1, 1-\frac{\beta}{2}}^2} \right]$$

con confidenza $1 - \beta$, dove $[X_{n-1, 1-\frac{\beta}{2}}^2, X_{n-1, \frac{\beta}{2}}^2]$ è l'intervallo in cui sia ha probabilità $1 - \beta$ che C_{n-1} assuma valore all'interno di esso. $C_{n-1} = \frac{(n-1) \cdot S^2}{\sigma^2} \sim X_{n-1}^2$ e $S^2 = \frac{1}{n-1} \cdot \sum_{k=1}^n (X_k - \bar{X})^2$.

☐ Dimostrazione

▼ **4.3 - Regressione**

Regressione

La **regressione** è una tecnica che consiste nel determinare la relazione tra le variabili casuali in ingresso (Z_1, Z_2, \dots) note e la variabile Y in uscita.

Regressione lineare

Nel caso in cui tra le variabile in ingresso e la variabile in uscita vi sia una relazione lineare, ovvero del tipo

$$Y = \alpha + \beta_1 \cdot Z_1 + \beta_2 \cdot Z_2 + \dots$$

in cui sono incogniti i coefficienti, ovvero α, β_1, β_2 , si parla di **regressione lineare**.

Il problema, in questo caso, consiste nello stimare tali coefficienti introducendo appropriati stimatori.

Regressione lineare semplice

Nel caso in cui si abbia un'unica variabile casuale in ingresso e in uscita, ovvero del tipo

$$Y = \alpha + \beta \cdot Z$$

si parla di **regressione lineare semplice**.

Tuttavia, nella realtà, la misura Y è affetta da un errore di misura ϵ . Ciò che si osserva sarà quindi $Y = \alpha + \beta \cdot Z + \epsilon$. Dal punto di vista pratico ciò corrisponde ad avere dati sperimentali che non sono posti su una retta (come dovrebbe essere considerando che la relazione tra le variabili è lineare), ma che costituiscono una nuvola di punti. Il problema è quindi capire quale è la retta che interpola meglio questi punti.

Problemi di regressione lineare con il metodo dei minimi quadrati

In un problema di regressione lineare semplice del tipo $Y = \alpha + \beta \cdot Z$, è possibile stimare i coefficienti α e β utilizzando il metodo dei minimi quadrati.

Definiamo quindi gli stimatori A e B rispettivamente per α e β . Considerando quindi che la variabile casuale Z assumerà i valori x_1, x_2, \dots, x_n , si ha che i valori teorici di Y legati ad essi saranno $A + B \cdot x_k$, e occorrerà paragonarli ai valori misurati realmente, ovvero Y_1, Y_2, \dots, Y_n .

Il nostro scopo sarà dunque quello di minimizzare la distanza tra i valori teorici e quelli reali, ossia la seguente:

$$\text{distanza} = Y_k - (A + B \cdot x_k)$$

Tale obiettivo può essere raggiunto minimizzando il valore della somma SS di tutte le distanze tra i valori teorici e quelli reali. Per evitare però che la somma di tali distanze sia un valore nullo, in quanto la distanza potrebbe essere sia positiva che negativa, consideriamo la somma dei quadrati, ovvero:

$$SS = \sum_{k=1}^n (Y_k - A - B \cdot x_k)^2$$

A questo punto occorre dunque scegliere A e B in modo da minimizzare SS , risolvendo quindi il seguente sistema:

$$\begin{cases} \frac{\partial SS}{\partial A} = 0 \\ \frac{\partial SS}{\partial B} = 0 \end{cases}$$

Iniziamo a risolvere il sistema:

- Prima equazione:

$$\begin{aligned} \frac{\partial SS}{\partial A} &= 0 && \Longleftrightarrow \\ \frac{\partial}{\partial A} \left(\sum_{k=1}^n (Y_k - A - B \cdot x_k)^2 \right) &= 0 && \Longleftrightarrow \\ -2 \cdot \sum_{k=1}^n (Y_k - A - B \cdot x_k) &= 0 && \Longleftrightarrow \\ \sum_{k=1}^n (Y_k - A - B \cdot x_k) &= 0 \end{aligned}$$

- Seconda equazione:

...

$$\begin{aligned}
\frac{\partial SS}{\partial B} &= 0 && \Longleftrightarrow \\
\frac{\partial}{\partial B} \left(\sum_{k=1}^n (Y_k - A - B \cdot x_k)^2 \right) &= 0 && \Longleftrightarrow \\
-2 \cdot \sum_{k=1}^n x_k \cdot (Y_k - A - B \cdot x_k) &= 0 && \Longleftrightarrow \\
-2 \cdot \sum_{k=1}^n (Y_k \cdot x_k - A \cdot x_k - B \cdot (x_k)^2) &= 0 && \Longleftrightarrow \\
\sum_{k=1}^n (Y_k \cdot x_k - A \cdot x_k - B \cdot (x_k)^2) &= 0
\end{aligned}$$

Riscriviamo dunque il sistema iniziale inserendo ciò che abbiamo appena calcolato:

$$\begin{cases} \sum_{k=1}^n (Y_k - A - B \cdot x_k) = 0 \\ \sum_{k=1}^n (Y_k \cdot x_k - A \cdot x_k - B \cdot (x_k)^2) = 0 \end{cases} \Longleftrightarrow
\begin{cases} \sum_{k=1}^n Y_k - n \cdot A - B \cdot \sum_{k=1}^n x_k = 0 \\ \sum_{k=1}^n (Y_k \cdot x_k) - A \cdot \sum_{k=1}^n x_k - B \cdot \sum_{k=1}^n (x_k)^2 = 0 \end{cases}$$

Per risolvere tale sistema occorre considerare la media aritmetica \bar{Y} delle Y_k , ovvero

$$\bar{Y} = \frac{1}{n} \cdot \sum_{k=1}^n Y_k$$

e la media aritmetica \bar{x} delle x_k , ovvero

$$\bar{x} = \frac{1}{n} \cdot \sum_{k=1}^n x_k$$

Terminiamo dunque la risoluzione del sistema:

- Prima equazione:

$$\begin{aligned}
\sum_{k=1}^n Y_k - n \cdot A - B \cdot \sum_{k=1}^n x_k &= 0 && \Longleftrightarrow \\
n \cdot \bar{Y} - n \cdot A - B \cdot n \cdot \bar{x} &= 0 && \Longleftrightarrow \\
A &= \bar{Y} - B \cdot \bar{x}
\end{aligned}$$

- Seconda equazione:

$$\begin{aligned}\sum_{k=1}^n (Y_k \cdot x_k) - A \cdot \sum_{k=1}^n x_k - B \cdot \sum_{k=1}^n (x_k)^2 &= 0 \quad \Longleftrightarrow \\ \sum_{k=1}^n (Y_k \cdot x_k) - A \cdot n \cdot \bar{x} - B \cdot \sum_{k=1}^n (x_k)^2 &= 0 \quad \Longleftrightarrow\end{aligned}$$

e sostituendo ad A il valore ottenuto dalla prima equazione, si procede

$$\begin{aligned}\sum_{k=1}^n (Y_k \cdot x_k) - (\bar{Y} - B \cdot \bar{x}) \cdot n \cdot \bar{x} - B \cdot \sum_{k=1}^n (x_k)^2 &= 0 \quad \Longleftrightarrow \\ \sum_{k=1}^n (Y_k \cdot x_k) - \bar{Y} \cdot n \cdot \bar{x} + B \cdot n \cdot (\bar{x})^2 - B \cdot \sum_{k=1}^n (x_k)^2 &= 0 \quad \Longleftrightarrow \\ B \cdot n \cdot (\bar{x})^2 - B \cdot \sum_{k=1}^n (x_k)^2 &= - \sum_{k=1}^n (Y_k \cdot x_k) + \bar{Y} \cdot n \cdot \bar{x} \quad \Longleftrightarrow \\ B &= \frac{\bar{Y} \cdot n \cdot \bar{x} - \sum_{k=1}^n (Y_k \cdot x_k)}{n \cdot (\bar{x})^2 - \sum_{k=1}^n (x_k)^2} \quad \Longleftrightarrow \\ B &= \frac{\sum_{k=1}^n (Y_k \cdot (x_k - \bar{x}))}{\sum_{k=1}^n (x_k)^2 - n \cdot (\bar{x})^2}\end{aligned}$$

Tramite il metodo dei due quadrati abbiamo dunque stimato i valori di α e β ottenendo:

$$\begin{cases} A = \bar{Y} - B \cdot \bar{x} \\ B = \frac{\sum_{k=1}^n (Y_k \cdot (x_k - \bar{x}))}{\sum_{k=1}^n (x_k)^2 - n \cdot (\bar{x})^2} \end{cases}$$

Verifiche necessarie

È possibile verificare la correttezza dei due stimatori A e B ottenendo che $E[A] = \alpha$ e $E[B] = \beta$.

☐ Dimostrazione

È possibile inoltre verificare l'efficienza di B ottenendo che $Var(B) = \frac{\sigma^2}{\sum_{j=1}^n (x_j)^2 - n \cdot (\bar{x})^2}$, che verifica l'efficienza di B , in quanto i valori x_j al denominatore sono deterministici (permettendo quindi di essere grandi e distanti a piacere).

☐ Dimostrazione

Distribuzione di B

Considerando lo stimatore $B = \frac{\sum_{k=1}^n (Y_k \cdot (x_k - \bar{x}))}{\sum_{k=1}^n (x_k)^2 - n \cdot (\bar{x})^2}$, si ha che è combinazione lineare delle variabili Y_k (che sono variabili casuali gaussiane e indipendenti) ed ha quindi la seguente distribuzione:

$$B \sim N \left(\beta, \sqrt{\frac{\sigma^2}{\sum_{k=1}^n (x_k)^2 - n \cdot (\bar{x})^2}} \right)$$