Corso di Matematica Applicata

Simone Elia
Ingegneria Informatica
Università degli Studi di Bologna

1 - Definizioni di probabilità

@September 16, 2022

> **Definizione classica di probabilità**
>
> *Dato un esperimento con un **numero finito** di possibili esiti **equiprobabili**, un evento A associato a questo esperimento ha probabilità:*
>
> $$P(A) = \frac{\text{n}^\circ \text{ esiti favorevoli ad } A}{\text{n}^\circ \text{ esiti possibili}}$$
>
> *L'insieme di tutti gli esiti è definito **spazio campione**.*

Nonostante la sua semplicità, la definizione classica è comunque molto limitata:

- Gli esiti devono essere equiprobabili.
- Gli esiti devono essere in numero finito.
- A volte non è possibile "contare" gli esiti.

> **Definizione frequentista di probabilità**
>
> *Si ripete N volte un esperimento in maniera identica e indipendente, la probabilità dell'evento A si definisce come:*
>
> $$P(A) = \frac{\text{n}^\circ \text{ esperimenti con esito } A}{N}$$

Anche questa definizione, nonostante sia migliore, ha i suoi limiti:

Pro

- La definizione vale anche per eventi non equiprobabili.
- Non è richiesto di "contare" gli eventi.

Contro

- Esiti dell'esperimento devono essere in numero finito.
- A volte non è possibile ripetere l'esperimento molte volte → risultato impreciso.

2 - Calcolo combinatorio

> **Principio fondamentale del calcolo combinatorio - Principio di enumerazione**
>
> *Si supponga di realizzare 2 esperimenti e si supponga che il primo esperimento presenti n possibili esiti, mentre il secondo esperimento presenti m possibili esiti.*
> *Le coppie ordinate che contengono gli esiti del primo e del secondo esperimento saranno $n \times m$.*

▼ Esempio 1

✔ **Esempio**

> *Scatola con 6 palline rosse e 5 palline verdi. Estrazione* ***con reimmissione*** *3 palline. Quale è la probabilità di ottenere $2R + 1V$ (senza badare all'ordine di estrazione)?*

$A = 2R + 1V \text{senza ordine.}$

$\text{n}^\circ \text{ esiti possibili} = 11 \times 11 \times 11 = 11^3$

$\text{n}^\circ \text{ esiti favorevoli} = \underbrace{6 \times 6 \times 5}_{\text{RRV}} + \underbrace{6 \times 5 \times 6}_{\text{RVR}} + \underbrace{5 \times 6 \times 6}_{\text{VRR}} = 3 \times 5 \times 6^2$

$P(A) \underbrace{=}_{\text{DEF. CLASS.}} \frac{\text{n}^\circ \text{ esiti favorevoli}}{\text{n}^\circ \text{ esiti possibili}} = \frac{15 \times 6^2}{11^3}$

▼ Esempio 2

✔ **Esempio**

> *Scatola con 6 palline rosse e 5 palline verdi. Estrazione* ***senza reimmissione*** *3 palline. Quale è la probabilità di ottenere $2R + 1V$ (senza badare all'ordine di estrazione)?*

$A = 2R + 1V \text{senza ordine.}$

$\text{n}^\circ \text{ esiti possibili} = 11 \times 10 \times 9$

$\text{n}^\circ \text{ esiti favorevoli} = \underbrace{6 \times 5 \times 5}_{\text{RRV}} + \underbrace{6 \times 5 \times 5}_{\text{RVR}} + \underbrace{5 \times 6 \times 5}_{\text{VRR}} = 3 \times 5^2 \times 6$

$P(A) \underbrace{=}_{\text{DEF. CLASS.}} \frac{\text{n}^\circ \text{ esiti favorevoli}}{\text{n}^\circ \text{ esiti possibili}} = \frac{18 \times 5^2}{9 \times 10 \times 11}$

▼ Disposizione semplice

> **Disposizioni semplici di n elementi di classe k $(k \leq n)$**
>
> *Dati n oggetti distinti, si dicono disposizioni semplici di n elementi di classe k $(k \leq n)$ tutti gli allineamenti che si possono costruire prendendo k elementi tra gli n a disposizione senza ripetizioni.*
>
> $$D_{n,k} = n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$$
>
> ***N.B.*** Le disposizioni semplici sono associate alle estrazioni senza reimmissione.

▼ Disposizione con ripetizione

> **Disposizioni con ripetizione di n elementi di classe k**
>
> *Dati n oggetti distinti, si dicono disposizioni con ripetizione di n elementi di classe k, tutti gli allineamenti di k (anche ripetuti) presi dall'insieme degli n elementi dati.*

$$D^R_{n,k} = n^k$$

N.B. Le disposizioni con reimmissione sono associate alle estrazioni con reimmissione.

Permutazione semplice

Permutazioni semplici di n elementi

Dati n elementi distinti, si dicono permutazioni semplici degli n elementi, tutti gli allineamenti degli n elementi.

$$P_n = D_{n,n} = n!$$

N.B. Le disposizioni semplici sono associate alle estrazioni senza reimmissione.

Permutazione con ripetizione

Permutazioni con ripetizione di n elementi

Dati n elementi non necessariamente distinti, si dicono permutazioni con ripetizione degli elementi, tutti gli allineamenti di questi n elementi.

$$P^R_n = \frac{n!}{K_1!K_2!\ldots K_J!}$$

Esempio

✔️ **Anagrammi della parola "matematica"**

$n = 10$

- $A : 3$
- $M : 2$
- $T : 2$
- $E : 1$
- $I : 1$
- $C : 1$

$$P^R_n = \frac{10!}{3!2!2!1!1!1!} = \frac{10!}{24}$$

Combinazione semplice

Combinazioni semplici di n elementi di classe k $(k \leq n)$

Dati n elementi distinti, si dicono combinazioni semplici di n elementi di classe k $(k \leq n)$, tutti i gruppi di k elementi che si possono formare senza ripetizioni a partire dagli n dati.

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

@September 20, 2022

Paradosso dei compleanni

n persone nate in un anno non bisestile si trovano in un a stanza ($n \leq 365$). Qual è la probabilità che abbiano tutte date di compleanno diverse?

Ipotesi: tutte le date dell'anno sono equiprobabili

$C =$ *"Tutte le n persone hanno date di compleanno diverse"*

$P(C) = \frac{n^{\circ} \text{ esiti favorevili}}{n^{\circ} \text{ esiti totali}}$

$$P(C) = \frac{D_{365,n}}{D^R_{365,n}} = \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

Per trovare la probabilità che almeno due persone abbiano lo stesso compleanno considero l'evento complementare C^C

$$P(C^C) = 1 - P(C) = 1 - \frac{365 \times 364 \times \cdots \times (365 - n + 1)}{365^n}$$

Questo viene "*paradosso*" perchè $P(C^C)$ aumenta molto in fretta con l'aumentare di n ($n = 20 \Rightarrow P(C^C) \approx 41\%$, $n = 40 \Rightarrow P(C^C) \approx 89\%$, $n = 60 \Rightarrow P(C^C) \approx 99.4\%$)

3 - Notazione e definizioni preliminari

▼ Unione di eventi

> **Unione di eventi**
>
> Dati $A, B \subset S$, chiamiamo **unione di eventi** ($A \cup B$) l'evento di S che contiene tutti gli esiti contenuti in A e/o in B.

▼ Intersezione di eventi

> **Intersezione di eventi**
>
> Dati $A, B \subset S$, chiamiamo **intersezione di eventi** ($A \cap B$) l'evento di S che contiene tutti gli esiti contenuti sia in A che in B.

▼ Insieme vuoto

> **Insieme vuoto**
>
> Chiamiamo **insieme vuoto** ($\emptyset$) l'evento che non contiene esiti.

▼ Eventi mutuamente esclusivi

> **Eventi mutuamente esclusivi**
>
> Dati $A, B \subset S$, i due eventi si dicono mutuamente esclusivi se:
>
> $$A \cap B = \emptyset$$

▼ Evento complementare

> **Evento complementare**
>
> Dato $A \subset S$, definiamo **evento complementare** l'evento A^C tale che:
>
> $$A^C \equiv S - A$$

▼ Proprietà

Dati $A, B \subset S$

- $A^C \cap B^C = (A \cup B)^C$
- $A^C \cup B^C = (A \cap B)^C$

▼ Notazioni

Dati $A_1, A_2, \ldots, A_m \subset S$

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \cdots \cup A_m$$

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap \cdots \cap A_m$$

4 - Assiomi di Kolmogorov e proprietà

> **Assiomi di Kolmogorov**
>
> *Dato un esperimento che prevede più esiti possibili e a cui è associato uno spazio campione S, e dato un evento $E \subset S$, si definisce probabilità di E, $P(E) \in \mathbb{R}$, il numero tale che:*
>
> - **A1**. $0 \le P(E) \le 1$
> - **A2.** $P(S) = 1$
> - **A3.** Dati $E_1, E_2, \ldots, E_n \subset S$ mutuamente esclusivi, allora
>
> $$P(\bigcup_{k=1}^{m} E_k) = \sum_{k=1}^{m} P(E_k)$$

▼ Proprietà 1

$$\text{Dato } E \subset S$$
$$P(E^c) = 1 - P(E)$$

📎 **Dimostrazione**

$$1 \underbrace{=}_{A2} P(S) \underbrace{=}_{*} P(E \cup E^c) \underbrace{=}_{A3} P(E) + P(E^c)$$
$$P(E) + P(E^c) = 1$$
$$P(E^c) = 1 - P(E)$$

$* \to E \cup E^c = S$

▼ Proprietà 1bis

$$P(\emptyset) = 0$$

📎 **Dimostrazione**

$S^c = \emptyset$

$$P(\emptyset) = P(S^c) \underbrace{=}_{P1} 1 - P(S) \underbrace{=}_{A2} 1 - 1 = 0$$

▼ Proprietà 2

$$\text{Dati } A, B \subset S \text{ con } A \subset B$$
$$P(A) \le P(B)$$
$$B \cap A^c = B - A$$

📎 **Dimostrazione**

$\to B = A \cup (B \cap A^c)$

$\to A \cap (B \cap A^c) = \emptyset$

$$P(B) = P(A \cup (B \cap A^c)) \underbrace{=}_{A3} P(A) + \overbrace{P(B \cap A^c)}^{\ge 0} \underbrace{\ge}_{A1} P(A)$$
$$P(B) \ge P(A) \Rightarrow P(A) \le P(B)$$

▼ Proprietà 3

$$\text{Dati} \quad A, B \subset S$$
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

📎 **Dimostrazione**

→ $E_1 = A \backslash B$

→ $E_2 = A \cap B$

→ $E_3 = B \backslash A$

→ $A = E_1 \cup E_2$

→ $B = E_2 \cup E_3$

→ $A \cup B = E_1 \cup E_2 \cup E_3$

$$P(A \cup B) = P(E_1 \cup E_2 \cup E_3) \underbrace{=}_{A3}$$
$$= \overbrace{P(E_1) + P(E_2)}^{P(A)} + P(E_3) =$$
$$= P(A) + \overbrace{P(E_3) + P(E_2)}^{P(B)} - P(E_2) =$$
$$= P(A) + P(B) - P(A \cap B)$$

N.B. Se A e B sono mutualmente esclusivi:

$$A \cup B = \emptyset \Rightarrow P(A \cap B) = 0$$
$$P(A \cup B) = P(A) + P(B)$$

▼ Proprietà 4

$$\text{Dati} \quad A, B, C \subset S$$
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) -$$
$$-P(A \cap B) - P(B \cap C) - P(A \cap C) +$$
$$+P(A \cap B \cap C)$$

▼ Definizione classica di probabilità come conseguenza degli assiomi

Dato un esperimento che presenta un numero finito, N, di esiti equiprobabili:

$$S = \{e_1, e_2, \dots, e_N\} \qquad e_k = k\text{-esimo elemento di } S$$

$N \in \mathbb{N}$ con $N < +\infty$

$$P(e_1) = P(e_2) = \dots = P(e_N) = p$$

Quanto vale p?

$e_1, e_2, \dots, e_N$ sono mutualmente esclusivi in quanto esiti

$$1 \underbrace{=}_{A2} P(S) \overset{S=\bigcup_{k=1}^{N} e_k}{\overbrace{=}} P(\bigcup_{k=1}^{N} e_k) \underbrace{=}_{A3} \sum_{k=1}^{N} p(E_k) = \sum_{k=1}^{N} p = Np$$
$$1 = Np \Rightarrow p = \frac{1}{N}$$

Sia $A \subset S$ che contiene m esiti $1 \leq m \leq N$

$$A = \{e_1, e_2, \dots, e_m\}$$
$$A = \bigcup_{k=1}^{m} e_k$$

$$P(A) = P(\bigcup_{k=1}^{m} e_k) \underbrace{=}_{A3} \sum_{k=1}^{m} P(e_K) = \sum_{k=1}^{m} \frac{1}{N} = \frac{m}{N}$$
$$P(A) = \frac{m}{N} = \frac{\text{n}^\circ \text{ esiti contenuti in } A}{\text{n}^\circ \text{ esiti totali}}$$

5 - Probabilità condizionata e eventi indipendenti

@September 23, 2022

> **Probabilità condizionata**
>
> Dati 2 eventi $A, B \subset S$ con $P(B) \neq 0$, si definisce *probabilità di* A *condizionata da* B:
>
> $$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

> **Eventi indipendenti**
>
> Dati 2 eventi $A, B \subset S$, essi si dicono *indipendenti* se:
>
> $$P(A \cap B) = P(A)P(B)$$
>
> Dati 3 eventi $A, B, C \subset S$, essi si dicono indipendenti se:
>
> $$\begin{gathered} P(A \cap B \cap C) = P(A)P(B)P(C) \\ P(A \cap B) = P(A)P(B) \\ P(A \cap C) = P(A)P(C) \\ P(B \cap C) = P(B)P(C) \end{gathered}$$
>
> In generale Neventi sono indipendenti se tutte le loro intersezioni tra $N, N-1, N-2, \ldots, 2$ soddisfano condizioni simili a quelle scritte sopra.

N.B. Dati $A, B \subset S$ indipendenti con $P(B) \neq 0$:

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

> **Teorema**
>
> *Dati 2 eventi $A, B \subset S$, indipendenti, allora anche A e B^c sono indipendenti tra loro. (lo stesso vale per A^c e B e per A^c e B^c)*

Dimostrazione

→ $A = (A \cap B) \cup (A \cap B^c)$

$$P(A) = P((A \cap B) \cup (A \cap B^c)) \overbrace{=}^{A3} P(A \cap B) + P(A \cap B^c)$$

$$\begin{gathered} P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = \\ = P(A)(1 - P(B)) \underbrace{=}_{P1} P(A)P(B^c) \end{gathered}$$

In maniera simile si dimostra che A^c e B oppure A^c e B^c sono indipendenti.

$\square$

6 - Partizione di uno spazio campione

Partizione di uno spazio campione

Dato uno spazio campione S, si dice ***partizione di*** *S una suddivisione in eventi $\{H_1, H_2, \ldots, H_m\} \quad (m \in \mathbb{N} \setminus \{0,1\})$ tali che $H_i \cap H_k = \emptyset$ se $i \neq k$ e $\bigcup_{k=1}^{m} H_k = S$.*

Gli elementi della partizione $H_1, H_2, \ldots, H_m$ sono detti **ipotesi.**

▼ Formula delle probabilità totali

Formula delle probabilità totali

Dato uno spazio campione S e una sua partizione$\{H_1, H_2, \ldots, H_m\}$e dato un evento $A \subset S$:

$$P(A) = \sum_{k=1}^{m} P(A|H_k)P(H_k)$$

📎 **Dimostrazione**

→

$$A = (A \cap H_1) \cup (A \cap H_2) \cup \cdots \cup (A \cap H_m) = \bigcup_{k=1}^{m} (A \cap H_k)$$

→ Gli H_k sono mutualmente esclusivi.

$$P(A) = P(\bigcup_{k=1}^{m} (A \cap H_k) \underbrace{=}_{A3} \sum_{k=1}^{m} P(A \cap H_k) =$$
$$= \sum_{k=1}^{m} P(A|H_k)P(H_k)$$

□

▼ Teorema di Bayes (o probabilità a posteriori)

Teorema di Bayes (o probabilità a posteriori)

Dato uno spazio campione S e una sua partizione $\{H_1, H_2, \ldots, H_m\}$ e considerato un evento $E \subset S$ con $P(E) \neq 0$, allora:

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

Dimostrazione

→ $P(E \cap H_j) = P(E|H_j)P(H_j)$

→ $P(E \cap H_j) = P(H_j|E)P(E)$

$$\frac{P(H_j|E)P(E)}{P(E)} = \frac{P(E|H_j)P(H_j)}{P(E)}$$

(con $P(E) \neq 0$)

$$P(H_j|E) = \frac{P(E|H_j)P(H_j)}{P(E)} \underbrace{=}_{\text{formula prob. tot.}} \\ = \frac{P(E|H_j)P(H_j)}{\sum_{k=1}^{m} P(E|H_k)P(H_k)}$$

$\square$

@September 27, 2022

Problema della rovina del giocatore

> *A e B giocano lanciando una moneta. se esce testa (T) B dà una moneta da 1€ ad A, se esce croce (C) A dà una moneta da 1€ a B. Il gioco continua fino a quando uno dei due giocatori rimane senza monete.*
>
> *Sia k (con $0 < k < n$) il numero di monete possedute inizialmente da A ed $(n-k)$ il numero di monete possedute da B. Qual è la possibilità che A vinca?*

A = "A vince"

$P(A) = p_k$

p_s = "Prob. di vincere possedendo s monete"

$P(T) = p$

$P(C) = 1 - P(T) = 1 - p = q$

Se $1 \leq k < n$ si dovrà sempre lanciare almeno una volta la moneta.

$$P(A) \underbrace{=}_{\text{prob. totali}} P(A|T)P(T) + P(A|C)P(C)$$

$$p_k = p_{k+1} \quad p \quad + \quad p_{k-1} \quad q$$

N.B.:

- $p_0 = 0, \; p_n = 1$
- $p + q = p + (1-p) = 1$

$$(p+q)p_k = (p_{k+1})p + (p_{k-1})q \\ q(p_k - p_{k-1}) = p(p_{k+1} - p_k) \\ p_{k+1} - p_k = (p_k - p_{k-1})\frac{q}{p}$$

Si arriva alla seguente conclusione:

$$\begin{array}{ll} k = 1 & p_2 - p_1 = (p_1 - \underbrace{p_0}_{0})\frac{q}{p} = p_1\frac{q}{p} \\ k = 2 & p_3 - p_2 = (\underbrace{p_2 - p_1}_{p_1\frac{q}{p}})\frac{q}{p} = p_1(\frac{q}{p})^2 \\ k = 3 & p_4 - p_3 = (\underbrace{p_3 - p_2}_{p_1(\frac{q}{p})^2})\frac{q}{p} = p_1(\frac{q}{p})^3 \\ & \dots \\ k = n-1 & p_n - p_{n-1} = (p_{n-1} - p_{n-2})\frac{q}{p} = p_1(\frac{q}{p})^{n-1} \end{array}$$

Sommo tutte le equazioni (si comportano come una serie telescopica):

$$p_n - p_1 = p_1 \frac{q}{p} + \cdots + p_1 \left(\frac{q}{p}\right)^{n-1}$$

Osservo che $p_n = 1$ e raggruppo p_1:

$$p_1 \left(1 + \frac{q}{p} + (\frac{q}{p})^2 + \cdots + (\frac{q}{p})^{n-1}\right) = 1$$

E quindi:

$$p_1 = \frac{1}{1 + \frac{q}{p} + (\frac{q}{p})^2 + \cdots + (\frac{q}{p})^{n-1}}$$

Da qui è possibile considerare prima il caso specifico della moneta equilibrata, poi il caso generico in cui $p \neq q$:

- **Moneta equilibrata**

  Se la moneta è equilibrata $p = q = \frac{1}{2}$ da ciò consegue che $\frac{q}{p} = 1$ e quindi

  $$p_1 = \frac{1}{1 + \frac{q}{p} + (\frac{q}{p})^2 + \cdots + (\frac{q}{p})^{n-1}} = \frac{1}{1 + 1 + \cdots + 1} = \frac{1}{N}$$

  Da qui è facile ottenere la formula più generale:

  $$\begin{gathered} p_{k+1} - p_k = p_k - p_{k-1} \\ p_2 - p_1 = p_1 - 0 \to p_2 = 2p_1 \\ p_3 - p_2 = p_2 - p_1 \to p_3 = 3p_1 \\ p_4 - p_3 = p_3 - p_2 \to p_4 = 4p_1 \end{gathered}$$

  Ovvero:

  $$p_j = jp_1 = \frac{j}{N}$$

- **Caso generale**

  Nel caso generale $\frac{q}{p} \neq 1$ è quindi possibile fare:

  $$\begin{aligned} p_1 &= \frac{1}{1 + \frac{q}{p} + (\frac{q}{p})^2 + \cdots + (\frac{q}{p})^{n-1}} \times \frac{1 - \frac{q}{p}}{1 - \frac{q}{p}} = \\ &= \frac{1 - \frac{q}{p}}{1 - \left(\frac{q}{p}\right)^n} \end{aligned}$$

  E si verifica quindi che

  $$p_j = \frac{1 - \left(\frac{q}{p}\right)^j}{1 - \left(\frac{q}{p}\right)^n}$$

Dispositivi in serie e parallelo

n dispositivi $D_1, D_2, \ldots, D_n$ indipendenti tra di loro hanno probabilità di funzionare P_k con $k = 1, 2, \ldots, n$.

Dispositivi in serie

Se i dispositivi formano un sistema in serie qual è la probabilità che il sistema funzioni?

F = "sistema funziona"

D_k = " il k-esimo dispositivo funziona"

$P(D_k) = P_k$

$$F = D_1 \cap D_2 \cap \cdots \cap D_n$$

$$\begin{aligned} P(F) = P(D_1 \cap D_2 \cap \cdots \cap D_n) = P(D_1) \cap P(D_2) \ldots P(D_n) = \\ = \prod_{k=1}^{n} P_k \end{aligned}$$

Dispositivi in parallelo

Se i dispositivi formano un sistema in parallelo qual è la probabilità che il sistema funzioni?

F = “sistema funziona”

D_k = “ il k-esimo dispositivo funziona”

$P(D_k) = P_k$

$$E = D_1 \cup D_2 \cup \cdots \cup D_n$$

$$\begin{aligned} P(E) &= 1 - P(E^C) \\ &= 1 - P(E^C) = \\ &= 1 - P((D_1 \cup D_2 \cup \cdots \cup D_n)^C) = \\ &= 1 - P(D_1^C \cap D_2^C \cap \cdots \cap D_n^C) = \\ &= 1 - P(D_1^C)P(D_2^C) \dots P(D_n^C) = \\ &= 1 - (1 - P_1)(1 - P_2) \dots (1 - P_n) = \\ &= 1 - \prod_{k=1}^{n} (1 - P_k) \end{aligned}$$

7 - Variabili casuali discrete

@September 30, 2022

> **Variabili casuali discrete**
>
> *Una **variabile casuale discreta** X è una corrispondenza tra gli eventi di Ω ed un insieme discreto (finito o numerabile) di numeri reali.*
>
> $$\begin{gathered} X \in \{a_1, a_2, \dots, a_n\} \\ \text{con} \quad a_k \in \mathbb{R} \qquad k = 1, 2, \dots, n \\ \text{e } n \text{ finito o numerabile.} \end{gathered}$$

Funzione di massa di probabilità

> **Funzione di massa di probabilità**
>
> *La **funzione di massa di probabilità** è una funzione di tipo $p : \mathbb{R} \to [0, 1]$ definita nel seguente modo:*
>
> $$p(b) = p(X = b) = \begin{cases} P(X = a_k) & \text{se } b = a_k \quad k = 1, \dots, n \\ 0 & \text{se } b \neq a_k \quad k = 1, \dots, n \end{cases}$$

Proprietà:

Proprietà 1

$$0 \leq p(b) \leq 1 \qquad \forall b \in \mathbb{R}$$

Proprietà 2

$$\sum_{k=1}^{n} p(a_k) = P(S) = 1$$

Funzione di ripartizione o distribuzione di probabilità

> **Funzione di ripartizione o distribuzione di probabilità**
>
> *La **funzione di ripartizione** è una funzione di tipo $F : \mathbb{R} \to [0, 1]$ definita come:*
>
> $$F(b) = P(X \leq b) \qquad \forall b \in \mathbb{R}$$

Proprietà

Proprietà 1

$$0 \leq F(b) \leq 1$$

Proprietà 2

$$\lim_{b \to +\infty} F(b) = P(X \leq +\infty) = 1$$

Proprietà 3

$$\lim_{b \to -\infty} F(b) = P(X \leq -\infty) = 0$$

Proprietà 4

$b, c \in \mathbb{R}$ tali che $b < c$

$$F(b) = P(X \leq b) \quad \leq \quad P(X \leq c) = F(c)$$
$$F(b) \leq F(c) \qquad \text{con } b < c$$

N.B. F è una funzione **non decrescente**.

▼ **Proprietà 5**

Se X è una variabile casuale discreta F è una funzione ‘a grumi’

8 - Variabili casuali continue

Variabili casuali continue

*Una **variabile casuale** si dice **continua** se ad essa è associata una funzione, detta **densità di probabilità**.*

$$f : \mathbb{R} \to \mathbb{R}^+ \cup \{0\} \quad \text{tale che } \forall \mathcal{B} \subseteq \mathbb{R}$$
$$P(X \in \mathcal{B}) = \int_{\mathcal{B}} f(s)ds$$

▼ Proprietà funzione di densità di probabilità

▼ Proprietà 1

$$f(x) \geq 0 \qquad \forall x \in \mathbb{R}$$

▼ Proprietà 2

$$P(X \in \mathbb{R}) = P(X \leq +\infty) = 1 \quad \Rightarrow \quad \int_{-\infty}^{+\infty} f(s)ds = 1$$

▼ Funzione di ripartizione o distribuzione di probabilità

Funzione di ripartizione o distribuzione di probabilità

*La **funzione di ripartizione** è una funzione di tipo $F : \mathbb{R} \to [0, 1]$ definita come:*

$$F(b) = P(X \leq b) \qquad \forall b \in \mathbb{R}$$

▼ Proprietà

▼ Proprietà 1

$$0 \leq F(b) \leq 1$$

▼ Proprietà 2

$$\lim_{b \to +\infty} F(b) = P(X \leq +\infty) = 1$$

▼ Proprietà 3

$$\lim_{b \to -\infty} F(b) = P(X \leq -\infty) = 0$$

▼ Proprietà 4

$b, c \in \mathbb{R}$ tali che $b < c$

$$F(b) = P(X \leq b) \quad \leq \quad P(X \leq c) = F(c)$$
$$F(b) \leq F(c) \qquad \text{con } b < c$$

N.B. F è una funzione **non decrescente**.

▼ Proprietà 5

$$F(b) = P(X \leq b) = \int_{-\infty}^{b} f(s)ds$$

$$f(b) = \frac{dF(b)}{db}$$

N.B. $f(b) \geq 0$ perchè $F(b)$ è non decrescente.

Considerazioni:

- Le variabili casuali discrete hanno una funzione di ripartizione discontinua e quindi non derivabile → non è possibile associare una funzione di densità.
- Per una variabile casuale continua non è significativo associare una funzione di massa:

$$p(b) = P(X = B) = \int_{X=b} f(s)ds = \int_b^b f(s)ds = 0$$
$$\forall b \in \mathbb{R}$$

- La funzione di ripartizione F è comoda per calcolare la probabilità di n intervalli:

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

9 - Coppie di variabili casuali

▼ Coppie di variabili casuali discrete

> **Funzione di massa di probabilità congiunta**
>
> La funzione di **massa di probabilità congiunta** è una funzione di tipo $p : \mathbb{R}^2 \to [0, 1]$ definita come:
>
> $$p(a, b) = P(X = a \cap Y = b) = P(X = a, Y = b)$$

▼ Proprietà

▼ Proprietà 1

$$p(a, b) \in [0, 1]$$

▼ Proprietà 2

$$\sum_{k=1}^{n}\sum_{j=1}^{m} p(x_k, y_j) = 1$$

> **Funzioni di massa marginali**
>
> $$p_X(a) = P(X = a) = P(X = a, Y = \text{qualsiasi}) = \\ = \sum_{j=1}^{m} P(X = a, Y = y_j)$$
>
> $$p_Y(b) = P(Y = b) = P(X = \text{qualsiasi}, Y = b) = \\ = \sum_{k=1}^{n} P(X = x_k, Y = b)$$

▼ Coppie di variabili casuali congiuntamente continue

@October 11, 2022

> **Coppie di variabili casuali congiuntamente continue**
>
> (X, Y), **variabili casuali**, si dicono **congiuntamente continue** se ad esse è assoaciata una funzione di tipo $f : \mathbb{R}^2 \to \mathbb{R}^+ \cup \{0\}$ tale che $\forall \mathcal{H} \subset \mathbb{R}^2$:
>
> $$P((X, Y) \in \mathcal{H}) = \iint f(s, t) ds dt$$
>
> f è detta **funzione di densità di probabilità conginuta.**

▼ Proprietà

▼ Proprietà 1

$$f(s, t) \geq 0$$

▼ Proprietà 2

$$\iint_{\mathbb{R}^2} f(s, t) ds dt = 1$$

> **Funzione di ripartizione di probabilità**
>
> Si definisce anche in questo caso una funzione $F : \mathbb{R}^2 \to [0, 1]$:

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

Proprietà

Le prime 4 proprietà sono le stesse del caso discreto

Proprietà 5

$$F(a,b) = \int_{-\infty}^{a} (\int_{-\infty}^{b} f(s,t)dt)ds$$

$$f(a,b) = \frac{\partial^2 F}{\partial a \partial b}$$

Funzioni di ripartizione di probabilità marginali

$$F_X(a) = P(X \le a) = P(X \le a, Y \le +\infty) = \\ = F(a, +\infty) = \int_{-\infty}^{a} (\int_{-\infty}^{+\infty} f(s,t)dt)ds$$

$$F_Y(b) = P(Y \le b) = P(X \le +\infty, Y \le b) = \\ = F(+\infty, b) = \int_{-\infty}^{b} (\int_{-\infty}^{+\infty} f(s,t)ds)dt$$

Funzioni di densità di probabilità marginali

$$f_X(a) = \frac{dF_X(a)}{da} = \frac{d}{da} \int_{-\infty}^{a} (\int_{-\infty}^{+\infty} f(s,t)dt)ds = \\ = \int_{-\infty}^{+\infty} f(a,t)dt$$

$$f_Y(b) = \frac{dF_Y(b)}{db} = \frac{d}{db} \int_{-\infty}^{b} (\int_{-\infty}^{+\infty} f(s,t)ds)dt = \\ = \int_{-\infty}^{+\infty} f(s,b)ds$$

Coppie variabili casuali indipendenti

Variabili casuali indipendenti

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

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

N.B. La definizione è semplice ma non operativa, non è sempre possibile verificarla $\forall \mathcal{A}, \mathcal{B} \subset \mathbb{R}$. Quindi si ricorre ai tre teoremi seguenti:

- **Teorema 1:**

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

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

- **Teorema 2**

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

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

- **Teorema 3**

  Condizione necessaria e sufficiente affinché due variabili casuali congiuntamente continue siano indipendenti è che:

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

10 - Valor medio o valore atteso o media teorica o speranza matematica

> **Valor medio**
>
> *Data una variabile casuale X si definisce, se esiste, il suo **valore medio** o valore atteso o media teorica o speranza matematica la seguente quantità:*
>
> $$E[X] = \begin{cases} \sum_{k=1}^{n} x_k p(x_k) & \text{se } X \text{ v.c. discreta} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{se } X \text{ v.c. continua} \end{cases}$$

▼ Proprietà

▼ Proprietà 1

Se $Y = h(X)$ con X variabile casuale nota:

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

se esistono.

▼ Proprietà 2

Caso particolare: $Y = \alpha X + \beta$ con $\alpha \in \mathbb{R}, \beta \in \mathbb{R}$:

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

> 📎 **Dimostrazione - caso discreto**
>
> → X v.c. discreta con $X \in \{x_1, x_2, \ldots, x_m\}$
>
> $$\begin{aligned} E[Y] &= \sum_{k=1}^{m} h(x_k) p(x_k) = \\ &= \sum_{k=1}^{m} (\alpha x_k + \beta) p(x_k) = \\ &= \alpha \sum_{k=1}^{m} x_k p(x_k) + \beta \sum_{k=1}^{m} p(x_k) = \\ &= \alpha E[X] + \beta \end{aligned}$$
>
> *In maniera analoga è dimostrabile il caso continuo*
>
> □

▼ Proprietà 3

se $Z = g(X, Y)$ con X e Y v.c. note:

$$E[Z] = E[g(X, Y)] =$$
$$= \begin{cases} \sum_{k=1}^{m} \sum_{j=1}^{n} g(x_k, y_j) p(x_k, y_j) & \text{se } X, Y \text{ v.c. discrete} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) & \text{se } X, Y \text{ v.c. continue} \end{cases}$$

se esistono.

▼ Proprietà 4

Date X, Y v.c. e $Z = X + Y$:

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

@October 18, 2022

Dimostrazione - caso continuo

$$
\begin{aligned}
E[Z] = E[X+Y] = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y)f(x,y)dxdy = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}xf(x,y)dxdy + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}yf(x,y)dxdy = \\
&= \int_{-\infty}^{+\infty}x(\int_{-\infty}^{+\infty}f(x,y)dy)dx + \int_{-\infty}^{+\infty}y(\int_{-\infty}^{+\infty}f(x,y)dx)dy = \\
&= \int_{-\infty}^{+\infty}xf_x(x)dx + \int_{-\infty}^{+\infty}yf_y(y)dy = \\
&= E[X]+E[Y]
\end{aligned}
$$

In maniera analoga si dimostra il caso discreto.

□

▼ **Proprietà 4-bis**

La proprietà 4 può essere estesa:

Dati $X_1, X_2, \ldots, X_N$ v.c. con valor medio definito:

$$E[X_1, X_2, \ldots, X_N] = E[X_1] + E[X_2] + \cdots + E[X_N]$$

Momento n-esimo di una variabile casuale

Il **momento n-esimo**, se esiste della v.c. X è:

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

con $n \in \mathbb{N}$

Teorema

Se X e Y sono variabili casuali con $E[X] = \mu_X$, $E[Y] = \mu_Y$ e sono indipendenti, allora:

$$E[XY] = E[X]E[Y]$$

N.B. Dal teorema si ricava che se X e Y sono indipendenti

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

Attenzione! La covarianza nulla non implica che i due eventi siano indipendenti, ma se gli eventi sono indipendenti hanno covarianza nulla

📎 **Dimostrazione - caso continuo**

$$
\begin{aligned}
E[XY] &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf(x,y)dxdy \underbrace{=}_{\text{Th 3 sulle var. indip.}} \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xyf_X(x)f_Y(y)dxdy = \\
&= \int_{-\infty}^{+\infty} yf_Y(y)(\underbrace{\int_{-\infty}^{+\infty} xf_X(x)dx}_{E[X]})dy = \\
&= \int_{-\infty}^{+\infty} yf_Y(y)E[X]dy = \\
&= E[X]\underbrace{\int_{-\infty}^{+\infty} yf_Y(y)dy}_{E[Y]} \\
&= E[X]E[Y]
\end{aligned}
$$

□

11 - Varianza di una variabile casuale

> **Varianza di una variabile casuale**
>
> *Data una variabile casuale X con valor medio $E[X] = \mu$ definito, si dice, se esiste, **varianza di** X:*
>
> $$Var(X) = \begin{cases} \sum_i (x_i - \mu)^2 P(x_i) & \text{caso discreto} \\ \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx & \text{caso continuo} \end{cases}$$
>
> $$Var(X) = E[(X - E[X])^2]$$

N.B. la varianza è **sempre positiva**.

Proprietà

Proprietà 1

La varianza è il quadrato del momento di ordine 1 sottratto al quadrato del momento di ordine 1.

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

> **Dimostrazione**
>
> $$\begin{aligned} Var(X) &= E[(X - E[X])^2] = \\ &= E[X^2 + E[X]^2 - 2E[X]X] = \\ &= E[X^2] + E[E[X]^2] + E[-2E[X]X] = \\ &= E[X^2] + E[X]^2 + (-2E[X]E[X]) = \\ &= E[X^2] - E[X]^2 \end{aligned}$$
>
> □

Proprietà 2

Caso particolare:

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

> **Dimostrazione**
>
> → $Y = \alpha X + \beta$
>
> → $E[Y] = E[\alpha X + \beta] = \alpha E[X] + \beta$
>
> ---
>
> $$\begin{aligned} Var(Y) &= E[(Y - E[Y])^2] = \\ &= E[(\alpha X + \beta - (\alpha E[X] + \beta))^2] = \\ &= E[(\alpha X - \alpha E[X])^2] = \\ &= E[\alpha^2 (X - E[X])^2] = \\ &= \alpha^2 E[(X - E[X])^2] = \\ &= \alpha^2 Var(X) \end{aligned}$$
>
> □

Proprietà 3

Date X e Y variabili casuali con valore atteso e varianza definiti:

$$Var(X, Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

📎 **Dimostrazione**

La dimostrazione è la medesima della proprietà ***6*** *della covarianza.*

Proprietà 3-bis

Date N *variabili casuali* $X_1, X_2 \ldots, X_N$

$$Var(\sum_{k=1}^{N} X_k) =$$
$$= \sum_{k=1}^{N} Var(X_k) + \sum_{k=1}^{N}\sum_{j=1}^{N} Cov(X_k, X_j) \qquad \text{con } k \neq j$$

12 - Covarianza

> **Covarianza**
>
> *Date X e Y, variabili casuali con $E[X] = \mu_X$ ed $E[Y] = \mu_Y$ si definisce covarianza:*
>
> $$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$
>
> **Caso discreto:**
>
> $$Cov(X, Y) = \sum_{k=1}^{m} \sum_{j=1}^{n} (x_k - \mu_X)(y_j - \mu_Y)p(x_k, y_j)$$
>
> **Caso continuo:**
>
> $$Cov(X, Y) = \int_{-\infty}^{+\infty} \int_{\infty}^{+\infty} (x - \mu_X)(y - \mu_Y)f(x, y)dxdy$$

Proprietà

Proprietà 1

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

Proprietà 2

$$Cov(X, X) = Var(X)$$

> 📎 **Dimostrazione**
>
> $$\begin{aligned} Cov(X, X) &= E[(X - \mu_X)(X - \mu_X)] = \\ &= E[(X - \mu_X)^2] = \\ &= E[(X - E[X])^2] = \\ &= Var(X) \end{aligned}$$
>
> □

Proprietà 3

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

N.B. Se due variabili casuali sono indipendenti $E[XY] = E[X]E[Y]$ quindi la covarianza è nulla.

Proprietà 4

$$Cov(\alpha X, Y) = Cov(X, \alpha Y) = \alpha Cov(X, Y)$$

Proprietà 5

Siano $X_1, X_2, ..., X_N$ e $Y_1, Y_2, ..., Y_M$ variabili casuali, allora

$$Cov(\sum_{k=1}^{N} X_k, \sum_{j=1}^{M} Y_j) = \sum_{k=1}^{N} \sum_{j=1}^{M} Cov(X_k, Y_j)$$

Proprietà 6

$$Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$$

Dimostrazione

→ $Z = X + Y$

→ $E[X] = \mu_X; \qquad E[Y] = \mu_Y$

→ $E[Z] = E[X + Z] = E[X] + E[Y] = \mu_X + \mu_Y$

$$\begin{aligned} Var(Z) &= E[(Z - E[Z])^2] = \\ &= E[(X + Y - (\mu_X + \mu_Y))^2] = \\ &= E[\{(X - \mu_X) + (Y - \mu_Y)\}^2] = \\ &= E[(X - \mu_X)^2 + (Y - \mu_Y)^2 + 2(X - \mu_X)(Y - \mu_Y)] = \\ &= E[(X - \mu_X)^2] + E[(Y - \mu_Y)^2] + 2E[(X - \mu_X)(Y - \mu_Y)] = \\ &= Var(x) + Var(Y) + 2Cov(X, Y) \end{aligned}$$

□

Coefficiente di correlazione

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

13 - Funzioni generatrici dei momenti

@October 25, 2022

Variabili casuali identicamente distribuite

*Due o più variabili sono **identicamente distribuite** se hanno la stessa funzione di ripartizione (distribuzione) di probabilità:*

- *Variabili casuali discrete hanno la stessa funzione di massa di probabilità*
- *Variabili casuali continue hanno la stessa funzione di densità di probabilità*

Funzione generatrice dei momenti

Data una variabile casuale X, si definisce funzione generatrice dei momenti di X (se esiste):

$$\phi(t) = E[e^{tX}] \qquad \phi : \mathbb{R} \to \mathbb{R}$$

Proprietà funzione generatrice dei momenti

Proprietà 0

$$\phi(0) = E[1] = 1 \qquad \forall X$$

Proprietà 1

$$\frac{d\phi(t)}{dt}\bigg|_{t=0} = \frac{d}{dt}E[e^{tX}]\bigg|_{t=0} = E[Xe^{tX}]\bigg|_{t=0} = E[X]$$

Proprietà 2

$$\frac{d^k\phi(t)}{dt^k}\bigg|_{t=0} = E[X^k] \qquad \text{per } K \geq 1$$

Proprietà 3

Se X e Y variabili casuali indipendenti

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

📎 **Dimostrazione - caso continuo**

$$\begin{aligned}
\phi_{X+Y}(t) &= E[e^{t(X+Y)}] = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{t(x+y)} f(x,y)dxdy = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{t(x+y)} f_X(x)f_Y(y)dxdy = \\
&= (\int_{-\infty}^{+\infty} e^{tx} f_X(x)dx)(\int_{-\infty}^{+\infty} e^{ty} f_Y(y)dy) = \\
&= E[e^{tX}]E[e^{tY}] = \\
&= \phi_X(t)\phi_Y(t)
\end{aligned}$$

Analogamente si dimostra il caso discreto

$\square$

Proprietà 4

*Se due o più variabili casuali hanno la stessa funzione generatrice dei momenti allora sono **identicamente distribuite**.*

Proprietà di riproducibilità delle v.c. binomiali

Date X, T, variabili casuali indipendenti binomiali:

$$X \sim B(n, p)$$
$$T \sim B(m, p)$$

Allora:

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

Dimostrazione

$$\phi_X(t) = (e^t p + q)^n$$
$$\phi_T(t) = (e^t p + q)^m$$
$$\phi_{X+T}(t) = \phi_X(t)\phi_T(t) = (e^t p + q)^{n+m}$$

14 - Disuguaglianze di Markov e Čebyčev

> **Disuguaglianza di Markov**
>
> *Data una variabile casuale $X \geq 0$ a valori non negativi, dato $a \in \mathbb{R}^+$:*
>
> $$P(X \geq a) \leq \frac{E[X]}{a}$$
>
> ***N.B.*** Si riesce a stimare $P(X \geq a)$ usando solo $E[X]$, affinchè la disuguaglianza dia una stima utile si deve avere $\frac{E[X]}{a} < 1$

📎 **Dimostrazione - caso continuo**

$$\begin{aligned}
E[X] &= \int_{-\infty}^{+\infty} xf(x)dx \underbrace{=}_{x \geq 0} \\
&= \int_{0}^{+\infty} xf(x)dx = \\
&= \int_{0}^{a} xf(x)dx + \int_{a}^{+\infty} xf(x)dx \geq \\
&\geq \int_{a}^{+\infty} xf(x)dx \underbrace{\geq}_{x \geq a} \\
&\geq \int_{a}^{+\infty} af(x)dx = \\
&= a\int_{a}^{+\infty} f(x)dx = \\
&= aP(X \geq a)
\end{aligned}$$

$$\begin{aligned}
E[X] &\geq aP(X \geq a) \\
P(X \geq a) &\leq \frac{E[X]}{a}
\end{aligned}$$

La dimostrazione è analoga per il caso discreto.

□

> **Deviazione standard**
>
> *Definiamo scarto quadratico medio o deviazione standard:*
>
> $$\sigma = \sqrt{Var(X_k)}$$

> **Disuguaglianza di Čebyčev**
>
> *Data una variabile casuale X (con media $E[X] = \mu$ e varianza $Var(X) = \sigma^2$) e dato $r \in \mathbb{R}^+$:*
>
> $$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

N.B. La disuguaglianza è utile se $\frac{\sigma^2}{r^2} < 1$

Dimostrazione

→ $P(|X-\mu| \geq r) = P((X-\mu)^2 \geq r^2)$

→ $Y = (X-\mu)^2 \Rightarrow Y \geq 0$

→ $E[Y] = E[(X-\mu)^2] = Var(X)$

$$P(|X-\mu| \geq r) = P((X-\mu)^2 \geq r^2) = P(Y \geq r^2)$$
$$P(Y \geq r^2) \leq \frac{E[Y]}{r^2} = \frac{Var(X)}{r^2}$$

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

□

15 - Legge dei grandi numeri

▼ Media campionaria

Immaginiamo di associare ad ogni esperimento una variabile casuale che rappresenti l'esito dell'esperimento.

Se l'esperimento viene ripetuto in maniera identica e indipendente N volte → N variabili casuali indipendenti e identicamente distribuite $X_1, X_2, \dots, X_N$:

Chiamiamo **media campionaria**:

$$\bar{X} = \frac{X_1, X_2, \dots, X_N}{N}$$

$X_1, X_2, \dots, X_N$ sono identicamente distribuite: hanno tutte $E[X_k] = \mu$ e $Var(X_k) = \sigma^2$

$$\begin{aligned} E[\bar{X}] &= E[\frac{1}{N}\sum_{k=1}^{N} X_k] = \\ &= \frac{1}{N}E[\sum_{k=1}^{N} X_k] = \\ &= \frac{1}{N}\sum_{k=1}^{N} E[X_k] = \\ &= \frac{1}{N}\sum_{k=1}^{N} \mu = \\ &= \frac{N}{N}\mu = \\ &= \mu \end{aligned}$$

$$\begin{aligned} Var(\bar{X}) &= Var(\frac{1}{N}\sum_{k=1}^{N} X_k) = \\ &= \frac{1}{N^2}Var(\sum_{k=1}^{N} X_k) = \\ &= \frac{1}{N^2}\sum_{k=1}^{N} Var(X_k) = \\ &= \frac{1}{N^2}\sum_{k=1}^{N} \sigma^2 = \\ &= \frac{N\sigma^2}{N^2} = \\ &= \frac{\sigma^2}{N} \\ &\Downarrow \\ \sqrt{Var(\bar{X})} &= \frac{\sigma}{\sqrt{N}} \end{aligned}$$

> **Legge dei grandi numeri**
>
> *Data una successione di variabili casuali indipendenti identicamente distribuite $X_1, X_2, \dots, X_N$ (con $E[X_k] = \mu$ e $Var(X_k) = \sigma^2$ con $k = 1, 2, \dots, N$) allora $\forall \epsilon > 0$ (piccolo a piacere)*
>
> $$P(|\bar{X} - \mu| \geq \epsilon) \underset{N \to +\infty}{\longrightarrow} 0$$
>
> *ovvero la media campionaria converge in probabilità alla media teorica.*

Dimostrazione

→ $E[\bar{X}] = \mu$

→ $Var(\bar{X}) = \frac{\sigma^2}{N}$

$$P(|\bar{X} - \mu| \geq \epsilon) = P(|\bar{X} - E[\bar{X}]| \geq \epsilon) \underbrace{\leq}_{\text{Dis. di Čeb.}}$$

$$\leq \frac{Var(\bar{X})}{\epsilon^2} = \frac{\sigma^2}{N\epsilon^2}$$

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

$$\lim_{N \to +\infty} P(|\bar{X} - \mu| \geq \epsilon) \leq 0$$

Per l'assioma 1 di Kologorov:

$$\lim_{N \to +\infty} P(|\bar{X} - \mu| \geq \epsilon) \geq 0$$

$$\lim_{N \to +\infty} P(|\bar{X} - \mu| \geq \epsilon) = 0$$

▼ Corollario di Bernoulli

> **Corollario di Bernoulli**
>
> *Data una successione di variabili casuali di Bernoulli indipendenti identicamente distribuite* $X_1, X_2, \ldots, X_N$ *(con* $X_k \sim Be(p)$ *con* $k = 1, 2, \ldots, N$*) e* $\forall \epsilon > 0$ *(piccolo a piacere)*
>
> $$P(|\frac{1}{N}\sum_{k=1}^{N} X_k - p| \geq \epsilon) \underset{N \to +\infty}{\longrightarrow} 0$$

Dimostrazione

→ $E[X_k] = p \qquad (X_k \sim Be(p))$

→ $Var(X_k) = pq$

Per la legge dei grandi numeri:

$$\lim_{N \to +\infty} P(|\frac{1}{N}\sum_{k=1}^{N} X_k - E[X_j]| \geq \epsilon) = 0$$

N.B. Esiste un enunciato alternativo per il corollario:

Al crescere indefinito del numero di prove la frequenza relativa di un evento A converge in probabilità alla probabilità teorica di A

16 - Modelli di variabili casuali discrete

▼ Variabile casuale di Bernoulli

> **Variabile casuale di Bernoulli**
>
> *In un esperimento l'evento A (successo) si verifica con* $P(A) = p$
>
> $$X \sim Be(p) \qquad X = \begin{cases} 1 & \text{se } A \text{ si verifica} \\ 0 & \text{altrimenti} \end{cases}$$
>
> $$p(a) = p; \quad p(0) = 1 - p = q$$

▼ Valor medio

$$E[X] = p$$

> 📎 **Dimostrazione**
>
> $$\begin{aligned} E[X] &= \sum x_k p(x_k) = \\ &= 0 \cdot q + 1 \cdot p = \\ &= p \end{aligned}$$

▼ Varianza

$$Var(X) = pq$$

▼ Funzione generatrice dei momenti

$$\phi(t) = pe^t + q$$

▼ Variabile casuale binomiale

> **Variabile casuale binomiale**
>
> *Si ripete un esperimento in maniera identica e indipendente n volte. L'evento A si verifica in una prova con probabilità* $P(A) = p$
>
> $X =$ 'numero di esperimenti in cui si verifica A'
>
> $X \in \{0, 1, 2, \dots, n\}$
>
> $$X \sim B(n, p)$$
>
> $0 \leq k \leq n$ con $(k \in \mathbb{N})$
>
> $$p(k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}$$
>
> Dati n esperimenti indipendenti:
>
> $$Y_1, Y_2, \dots, Y_n \sim Be(p)$$
>
> Allora X è la somma di n variabili casuali indipendenti identicamente distribuite:

> $$X = \sum_{j=1}^{n} Y_j$$

- **Valor medio**

$$E[X] = np$$

- **Varianza**

$$Var(X) = npq$$

- **Funzione generatrice dei momenti**

$$\phi(t) = (pe^t + q)^n$$

N.B. La binomiale ha la seguente proprietà, detto **principio di riproducibilità**:
Dati $X \sim B(n, p)$ e $Z \sim B(m, p)$, indipendenti

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

Variabile casuale di Poisson

> **Variabile casuale di Poisson (degli eventi rari)**
>
> *Una v. c. si Poisson rappresenta in genere un numero di eventi che si osservano in un certo intervallo temporale/spaziale a patto che il singolo evento abbia una probabilità bassa di verificarsi (evento raro)*
>
> $$X \sim Po(\lambda) \qquad \text{con } \lambda \in \mathbb{R}^+$$
> $$X \in \mathbb{N}$$
> $$p(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

- **Funzione generatrice dei momenti:**

$$\phi(t) = e^{\lambda(e^t - 1)}$$

📎 **Dimostrazione**

→ $\sum_{i=0}^{+\infty} \frac{\alpha^i}{i!} = e^{\alpha}$ per $\alpha > 0$

$$\begin{aligned}
\phi(t) &= E[e^{tX}] \\
&= \sum_{i=0}^{+\infty} (e^{ti} \frac{\lambda^i}{i!} e^{-\lambda}) = \\
&= e^{-\lambda} \sum_{i=0}^{+\infty} \frac{(e^t \lambda)^i}{i!} \underbrace{=}_{e^t \lambda = \alpha} \\
&= e^{-\lambda} \sum_{i=0}^{+\infty} \frac{\alpha^i}{i!} = \\
&= e^{-\lambda} e^{\alpha} = \\
&= e^{\lambda(e^t - 1)}
\end{aligned}$$

$\square$

- **Valor medio**

$$E[X] = \lambda$$

📎 **Dimostrazione**

$$\begin{aligned} E[X] &= \left.\frac{d\phi(t)}{dt}\right|_{t=0} = \\ &= \left.\frac{d}{dt}[e^{\lambda(e^t-1)}]\right|_{t=0} = \\ &= \left.e^{\lambda(e^t-1)}\lambda e^t\right|_{t=0} = \\ &= \lambda \end{aligned}$$

□

▼ Varianza

$$Var(X) = \lambda$$

📎 **Dimostrazione**

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

$$\begin{aligned} E[X^2] &= \left.\frac{d^2\phi(t)}{dt^2}\right|_{t=0} = \\ &= \left.\frac{d}{dt}[e^{\lambda(e^t-1)}\lambda e^t]\right|_{t=0} = \\ &= \left.[\lambda e^t e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}\lambda e^t]\right|_{t=0} = \\ &= \lambda + \lambda^2 \end{aligned}$$

$$Var(X) = \lambda + \lambda^2 - (\lambda)^2 = \lambda$$

□

N.B. Una variabile casuale di Poisson si comporta come il limite di una binomiale $X \sim B(n, p)$ con $n \gg 1$ e $p \ll 1$. Sostituendo $\lambda = np$.

Rimane valido il **principio di riproducibilità**:

$$\begin{gathered} X \sim Po(\lambda) \qquad Y \sim Po(\mu) \\ X + Y \sim Po(\lambda + \mu) \end{gathered}$$

▼ Esempi

▼ Esempio 1

Una compagnia di assicurazioni riceve in media 5 richieste di rimborso al giorno.

1. *Che probabilità c'è che in un giorno arrivino meno di 3 chiamate?*
2. *Con che probabilità in 5 giorni arriveranno 20 richieste?*

Primo punto:

X = *"n° di richieste in un giorno"*

$X \sim Po(\lambda)$ con $E[X] = 5 = \lambda$

$$\begin{aligned} P(X < 3) &= P(X = 0) + P(X = 1) + P(X = 2) = \\ &= \frac{5^0}{0!}e^{-5} + \frac{5^1}{1!}e^{-5} + \frac{5^2}{2!}e^{-5} = \\ &= e^{-5}[1 + 5 + \frac{25}{4}] = \\ &= \frac{49}{4}e^{-5} \end{aligned}$$

Secondo punto:

X_k = *"n° di richieste il k-esimo giorno"*

$X_k \sim Po(5)$ le X_k sono indipendenti

Y = *"n° di richieste in 5 giorni"*

$Y = \sum_{k=1}^{5} \sim Po(\lambda + \lambda + \lambda + \lambda + \lambda) = Po(5\lambda) = Po(25)$

$$P(Y = 20) = \frac{25^{20}}{20!} e^{-25}$$

▼ Variabile casuale geometrica

> **Variabile casuale geometrica**
>
> *Si ripete l'esperimento in maniera identica e indipendente fino ad osservare un successo (l'evento A con $P(A) = p$)*
>
> $X =$ 'numero di prove per osservare un successo'
>
> $$X \sim G(p)$$
>
> $X \in \mathbb{N} \setminus \{0\}$
>
> $k = 1, 2, \dots, +\infty$
>
> $$p(k) = (1-p)^{k-1} p = q^{k-1} p$$

La probabilità di massa può essere anche interpretata come "*probabilità che non succeda A nei primi $k-1$ eventi moltiplicata per la probabilità che succeda A nel k-esimo*"

▼ Valor medio

$$E[X] = \frac{1}{p}$$

> 📎 **Dimostrazione**
>
> → $\sum_{i=0}^{+\infty} \alpha^i = \frac{1}{1-\alpha}$
>
> $$\begin{aligned} E[X] &= \sum_{k=1}^{+\infty} k p(k) = \\ &= \sum_{k=1}^{+\infty} k q^{k-1} p = \\ &= p \sum_{k=1}^{+\infty} k q^{k-1} = \\ &= p \sum_{k=0}^{+\infty} k q^{k-1} = \\ &= p \frac{d}{dq} \sum_{k=0}^{+\infty} q^k \underbrace{=}_{*} \\ &= p \frac{d}{dq} \left(\frac{1}{1-q}\right) = \\ &= \frac{p}{(1-q)^2} = \\ &= \frac{1}{p} \end{aligned}$$
>
> $\square$

▼ Varianza

$$Var(X) = \frac{q}{p^2}$$

Dimostrazione

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

$$\begin{aligned} E[X^2] &= \sum_{k=1}^{+\infty} k^2 q^{k-1} p = \\ &= p\sum_{k=1}^{+\infty}(k^2 - k + k)q^{k-1} = \\ &= p[\sum_{k=1}^{+\infty}(k^2 - k)q^{k-1} + \sum_{k=1}^{+\infty} kq^{k-1}] = \\ &= p(\frac{1}{p^2}) + p\sum_{k=0}^{+\infty} k(k-1)q^{k-1} = \\ &= \frac{1}{p} + qp\sum_{k=0}^{+\infty} k(k+1)q^{k-2} = \\ &= \frac{1}{p} + pq\frac{d^2}{dq^2}\sum_{k=0}^{+\infty} q^k = \\ &= \frac{1}{p} + pq\frac{2}{(1-(1-p))^3} = \\ &= \frac{1}{p} + 2\frac{q}{p^2} = \\ &= \frac{p + 2(1-p)}{p^2} = \\ &= \frac{2-p}{p^2} \end{aligned}$$

$$\begin{aligned} Var(X) = \frac{2-p}{p^2} - (\frac{1}{p})^2 &= \frac{2-p-1}{p^2} = \\ = \frac{1-p}{p^2} &= \frac{q}{p^2} \end{aligned}$$

$\square$

@November 4, 2022

▼ Variabile casuale binomiale negativa

Variabile casuale binomiale negativa

Si ripete un esperimento in maniera identica e indipendente fino ad osservare l'evento A (successo) r volte (con $r \in \mathbb{N} \setminus \{0\}$)

$X =$ 'numero di esperimenti necessari per osservare A r volte'

$p = P(A)$ → probabilità del verificarsi di A in un singolo esperimento

$$X \sim NB(r, p)$$

$X \in \{r, r+1, r+2, \dots, +\infty\}$

$k = r, r+1, \dots$

$$p(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

📎 Dimostrazione

$$\begin{aligned} p(k) = p(X = k) = \\ &= P(\text{esce } A\ (r-1) \text{ volte in } (k-1) \text{ prove}) \cdot \\ &\cdot P(\text{esce } A \text{ nella } k\text{-esima prova}) = \\ &= \binom{k-1}{r-1} p^{r-1} (1-p)^{k-1-(r-1)} \cdot p = \\ &= \binom{k-1}{r-1} p^r (1-p)^{k-r} \end{aligned}$$

N.B. $X \sim NB(r, p)$ è descrivibile come:

$$X = \sum_{j=1}^{r} Y_j$$

con $Y_j \sim G(p)$ e $Y_1, Y_2, \ldots, Y_r$

▼ Valor medio

$$E[X] = \frac{r}{p}$$

📎 **Dimostrazione**

→ $E[Y_j] = \frac{1}{p}$

$$\begin{aligned} E[X] = E[\sum_{j=1}^{r} Y_j] = \\ &= \sum_{j=1}^{r} E[Y_j] = \\ &= \sum_{j=1}^{r} \frac{1}{p} = \\ &= \frac{r}{p} \end{aligned}$$

▼ Varianza

$$Var(X) = \frac{rq}{p^2}$$

📎 **Dimostrazione**

→ $Var(Y_j) = \frac{q}{p^2}$

$$\begin{aligned} Var(X) = Var(\sum_{j=1}^{r} Y_j) = \\ &= \sum_{j=1}^{r} Var(Y_j) = \\ &= \sum_{j=1}^{r} \frac{q}{p^2} = \\ &= \frac{rq}{p^2} \end{aligned}$$

17 - Modelli di variabili casuali continue

▼ Variabile casuale uniforme

> *Una variabile casuale continua si dice **uniforme** se:*
>
> $$X \sim U(\alpha, \beta) \qquad \alpha, \beta \in \mathbb{R} \qquad \alpha < \beta$$
>
> tale che:
>
> $$f(x) = \begin{cases} k & \text{se } x \in [\alpha, \beta] \\ 0 & \text{altrimenti} \end{cases}$$

Il valore di k è calcolabile nel seguente modo:

$$1 = \int_{-\infty}^{+\infty} f(x)dx = \int_{\alpha}^{\beta} kdx = k[x]_{\alpha}^{\beta} = k(\beta - \alpha)$$
$$\Downarrow$$
$$k = \frac{1}{\beta - \alpha}$$

quindi:

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

▼ Funzione di distribuzione di probabilità

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

> 📎 **Dimostrazione**
>
> $$\begin{aligned} F(a) &= P(X \leq a) = \\ &= \int_{-\infty}^{a} f(x)dx = \\ &= \begin{cases} \int_{-\infty}^{a} 0dx & \text{se } a < \alpha \\ \int_{-\infty}^{\alpha} 0dx + \int_{\alpha}^{a} \frac{1}{\beta-\alpha}dx & \text{se } \alpha \leq a \leq \beta \\ \int_{-\infty}^{\alpha} 0dx + \int_{\alpha}^{\beta} \frac{1}{\beta-\alpha}dx + \int_{\beta}^{a} 0dx & \text{se } a > \beta \end{cases} = \\ &= \begin{cases} 0 & \text{se } a < \alpha \\ \frac{1-\alpha}{\beta-\alpha} & \text{se } \alpha \leq a \leq \beta \\ 1 & \text{se } a > \beta \end{cases} \end{aligned}$$

▼ Valor medio

$$E[X] = \frac{\beta + \alpha}{2}$$

Dimostrazione

$$E[X] = \int_{-\infty}^{+\infty} xf(x)dx =$$
$$= \int_{-infty}^{\alpha} x \cdot 0dx + \int_{\alpha}^{\beta} \frac{x}{\beta-\alpha}dx + \int_{\beta}^{+\infty} x \cdot 0dx =$$
$$= \int_{\alpha}^{\beta} \frac{x}{\beta-\alpha}dx =$$
$$= \frac{1}{\beta-\alpha}[\frac{x^2}{2}]_{\alpha}^{\beta} =$$
$$= \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} =$$
$$= \frac{\beta-\alpha}{2}$$

□

Varianza

$$Var(X) = \frac{\alpha^2+\beta^2+\alpha\beta}{3} - (\frac{\beta+\alpha}{2})^2$$

Dimostrazione

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

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 f(x)dx =$$
$$= \int_{-infty}^{\alpha} x^2 \cdot 0dx + \int_{\alpha}^{\beta} \frac{x^2}{\beta-\alpha}dx + \int_{\beta}^{+\infty} x^2 \cdot 0dx =$$
$$= \int_{\alpha}^{\beta} \frac{x^2}{\beta-\alpha}dx =$$
$$= \frac{1}{\beta-\alpha}[\frac{x^3}{3}]_{\alpha}^{\beta} =$$
$$= \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} =$$
$$= \frac{\beta^2+\alpha^2+\alpha\beta}{2}$$

$$Var(X) = \frac{\alpha^2+\beta^2+\alpha\beta}{3} - (\frac{\beta+\alpha}{2})^2$$

□

Coppie di variabili casuali uniformi indipendenti

Date due variabili casuali uniformi e indipendenti:

$$\begin{array}{rl} X \sim U(\alpha, \beta) & \alpha, \beta \in \mathbb{R} \quad \alpha < \beta \\ Y \sim U(\gamma, \delta) & \gamma, \delta \in \mathbb{R} \quad \gamma < \delta \end{array}$$

$$f(x,y) = f_X(x)f_Y(y) =$$
$$= \begin{cases} \frac{1}{\beta-\alpha} \cdot \frac{1}{\delta-\gamma} & \text{se} \quad x \in [\alpha, \beta], y \in [\gamma, \delta] \\ 0 & \text{altrove} \end{cases}$$

N.B. $(\beta-\alpha) \cdot (\delta-\gamma)$ può essere interpretata come l'area di un rettangolo (R) *.*

Dato $\mathcal{B} \subset \mathbb{R}^2$*, la probabilità che la coppia di variabili casuali* $X, Y \in \mathcal{B}$ *è pari a:*

$$P((X,Y) \in \mathcal{B}) = \frac{Area(\mathcal{B} \cap R)}{Area(R)}$$

📎 **Dimostrazione**

$$
\begin{aligned}
P((X,Y)\in\mathcal{B}) = P((X,Y)\in\mathcal{B}\cap R) = \\
&= \iint_{\mathcal{B}\cap R} f(x,y)dxdy = \\
&= \iint_{\mathcal{B}\cap R} \frac{1}{Area(R)}dxdy = \\
&= \frac{1}{Area(R)}\iint_{\mathcal{B}\cap R} 1dxdy = \\
&= \frac{Area(\mathcal{B}\cap R)}{Area(R)}
\end{aligned}
$$

▼ Metodi di Monte Carlo

> *Si definiscono **metodi di Monte Carlo** i metodi che sfruttano la casualità e la probabilità per calcolare una quantità deterministica*

▼ Esempi

> **Misura della superficie di un lago attraverso il lancio di palle di cannone**
>
> Si lanciano delle palle di cannone casualmente all'interno del rettangolo di lati l_1 e l_2 che contiene il lago

→ $X \sim U(0,l_1)$
→ $Y \sim U(0,l_2)$
→ (X,Y) = "*Punto di arrivo della palla di cannone*"

$$P(\text{palla di cannone nel lago}) = P((X,Y)\in L) = \frac{Area(L)}{Area(R)}$$

Se su n lanci n_L finiscono nel lago, per il corollario di Bernoulli, se $n \gg 1$:

$$\frac{n_L}{n} \approx P(\text{palla di cannone finisce nel lago}) = \frac{Area(L)}{Area(R)}$$
$$\Downarrow$$
$$Area(R) \approx Area(R)\frac{n_L}{n}$$

@November 8, 2022

▼ Problema degli spilli di Buffon

Correlato ai metodi di Monte Carlo vi è il problema del 1777 detto "problema degli spilli" proposto da Buffon. In esso si propone di stimare il π lanciando degli spilli di lunghezza $2a$ su un piano diviso da rette parallele a distanza $2b$ l'una dall'altra (con $2a < 2b$).

Si inizia considerando due variabili casuali uniformi, chiamando M il punto medio dello spillo si ha:

X = "Distanza di M dalla retta più vicina"

$X \sim U(0,b)$

θ = "Angolo tra la retta di direzione comune passante per M e lo spillo"

$\theta \sim U(0,\pi)$

Si ha intersezione tra la retta e lo spillo se:

$$X \le a\sin(\theta)$$

dunque:

$$P(\text{spillo interseca la retta } r) = P(X \le a\sin(\theta))$$

Ipotizzando X e θ variabili casuali continue e indipendenti:

$$f(x,y) = f_X(x)f_\theta(y)$$

$$f_X(x) = \\ = \begin{cases} \frac{1}{b} & \text{se } x \in [0, b] \\ 0 & \text{altrove} \end{cases}$$

$$f_\theta(y) = \\ = \begin{cases} \frac{1}{\pi} & \text{se } y \in [0, \pi] \\ 0 & \text{altrove} \end{cases}$$

$$\begin{aligned} P(X \leq a\sin(\theta)) &= \int_0^\pi \int_0^{a\sin(y)} f(x, y)dxdy = \\ &= \int_0^\pi \int_0^{a\sin(y)} \frac{1}{b\pi}dxdy = \\ &= \int_0^\pi \frac{1}{b\pi}[x]_0^{a\sin(y)}dy = \\ &= \frac{1}{b\pi}\int_0^\pi a\sin(y)dy = \\ &= \frac{a}{b\pi}[-\cos(y)]_0^\pi dy = \\ &= \frac{2a}{b\pi} \end{aligned}$$

Ripetendo n volte il lancio dello spillo ($n \gg 1$) per il corollario di Bernoulli:

$$f_A = \frac{n_A}{n} \\ \Downarrow \\ \frac{2a}{b\pi} \approx \frac{n_A}{n} \Rightarrow \pi \approx \frac{2an}{bn_A}$$

Variabile casuale esponenziale

> *Una variabile casuale si dice **esponenziale** se*
>
> $$X \sim E(\lambda) \qquad \lambda \in \mathbb{R}^+$$
>
> tale che:
>
> $$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x \geq 0 \\ 0 & \text{altrove} \end{cases}$$

Funzione di distribuzione di probabilità

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

> **Dimostrazione**
>
> $$\begin{aligned} F(a) &= P(X \leq a) = \\ &= \int_{-\infty}^a f(x)dx = \\ &= \begin{cases} \int_{-\infty}^a 0dx & \text{se } a < 0 \\ \int_{-\infty}^0 0dx + \int_0^a \lambda e^{-\lambda x}dx & \text{se } a \geq 0 \end{cases} = \\ &= \begin{cases} 0 & \text{se } a < 0 \\ 1 - e^{-\lambda a} & \text{se } a \geq 0 \end{cases} \end{aligned}$$

Funzione generatrice dei momenti

$$\phi(t) = \begin{cases} \text{diverge} & \text{se } \lambda - t \leq 0 \\ \frac{\lambda}{\lambda - t} & \text{se } \lambda - t > 0 \end{cases}$$

📎 **Dimostrazione**

$$
\begin{aligned}
\phi(t) = E[e^{tX}] = \\
&= \int_{-\infty}^{+\infty} e^{tx} f(x) dx = \\
&= \int_{-\infty}^{0} e^{tx} dx + \int_{0}^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \\
&= \lambda \int_{0}^{+\infty} e^{-(\lambda - t)x} dx = \\
&= \begin{cases} \text{diverge} & \text{se } \lambda - t \leq 0 \\ \lambda[-\frac{e^{-(\lambda-t)x}}{\lambda - t}]_0^{+\infty} & \text{se } \lambda - t > 0 \end{cases} = \\
&= \begin{cases} \text{diverge} & \text{se } \lambda - t \leq 0 \\ \frac{\lambda}{\lambda - t} & \text{se } \lambda - t > 0 \end{cases}
\end{aligned}
$$

▼ **Valor medio**

$$E[X] = \frac{1}{\lambda}$$

📎 **Dimostrazione**

$$
\begin{aligned}
E[X] &= \frac{d\phi(t)}{dt}\Big|_{t=0} = \\
&= \frac{\lambda}{(\lambda - t)^2}\Big|_{t=0} = \\
&= \frac{1}{\lambda}
\end{aligned}
$$

▼ **Varianza**

$$Var(X) = \frac{1}{\lambda^2}$$

📎 **Dimostrazione**

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

$$
\begin{aligned}
E[X^2] &= \frac{d^2\phi(t)}{dt^2}\Big|_{t=0} = \\
&= \frac{2\lambda}{(\lambda - t)^3}\Big|_{t=0} = \\
&= \frac{2\lambda}{\lambda^3} = \\
&= \frac{2}{\lambda^2}
\end{aligned}
$$

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

N.B. Le variabili casuali esponenziali non sono riproducili

▼ **Proprietà**

▼ **Proprietà 1**

Data una variabile casuale $X \sim E(\lambda)$ e una costante c $\in \mathbb{R}^+$

$$Y = cX \sim E(\frac{\lambda}{c})$$

📎 **Dimostrazione**

$$\begin{aligned}\phi_Y(t) &= E[e^{tY}] = \\ &= E[e^{tcX}] = \\ &= \phi_X(tc) = \\ &= \frac{\lambda}{\lambda - tc} = \\ &= \frac{\frac{\lambda}{c}}{\frac{\lambda}{c} - t}\end{aligned}$$

▼ **Proprietà 2**

*Dati n **dispositivi in serie** $D_1, D_2, \ldots, D_n$ ciascuno con un funzionamento di tipo esponenziale.*
D_k ha tempo di funzionamento $X_k \sim E(\lambda_k)$ con $\lambda_k \in \mathbb{R}^+$ e $k = 1, \ldots, n$, e $X_1, X_2, \ldots, X_n$ sono indipendenti.
Allora il tempo di funzionamento del sistema T è rappresentabile come:

$$T \sim E(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

📎 **Dimostrazione**

$$\begin{aligned}F_T(t) &= P(T \leq t) = \\ &= P(\min(X_1, X_2, \ldots, X_n) \leq t) = \\ &= 1 - P(\min(X_1, X_2, \ldots, X_n) > t) = \\ &= 1 - P(X_1 > t, X_2 > t, \ldots, X_n > t) = \\ &= 1 - P(X_1 > t)P(X_2 > t)\ldots P(X_n > t) = \\ &= 1 - \begin{cases} 1 & \text{se } t < 0 \\ e^{-\lambda_1 t} & \text{se } t \geq 0 \end{cases} \cdots \begin{cases} 1 & \text{se } t < 0 \\ e^{-\lambda_n t} & \text{se } t \geq 0 \end{cases} = \\ &= \begin{cases} 0 & \text{se } t < 0 \\ 1 - e^{-(\lambda_1 + \lambda_2 + \cdots + \lambda_n)t} & \text{se } t \geq 0 \end{cases} \\ &\Downarrow \\ T &\sim E(\lambda_1 + \lambda_2 + \cdots + \lambda_n)\end{aligned}$$

$\square$

▼ **Proprietà 3**

*Dati n **dispositivi in parallelo**, ognuno si comporta come una variabile casuale $X_k \sim E(\lambda_k)$ con $k = 1, 2, \ldots, n$. Chiamando S il tempo di funzionamento del sistema, allora la funzione di distribuzione di probabilità di S sarà:*

$$F_S(a) = \begin{cases} 0 & \text{se } a < 0 \\ \prod_{k=1}^{n}(1 - e^{-\lambda_k a}) & \text{se } a \geq 0 \end{cases}$$

📎 **Dimostrazione**

$$\begin{aligned}F_S(a) &= P(\max(X_1, X_2, \ldots, X_n) \leq a) = \\ &= P(X_1 \leq a, X_2 \leq a, \ldots, X_n \leq a) = \\ &= P(X_1 \leq a)P(X_2 \leq a)\ldots P(X_n \leq a) = \\ &= F_{X_1}(a)F_{X_2}(a)\ldots F_{X_n}(a) = \\ &= \begin{cases} 0 & \text{se } a < 0 \\ 1 - e^{-\lambda_1 a} & \text{se } a \geq 0 \end{cases} \cdots \begin{cases} 0 & \text{se } a < 0 \\ 1 - e^{-\lambda_n a} & \text{se } a \geq 0 \end{cases} = \\ &= \begin{cases} 0 & \text{se } a < 0 \\ \prod_{k=1}^{n}(1 - e^{-\lambda_1 a}) & \text{se } a \geq 0 \end{cases}\end{aligned}$$

$\square$

▼ **Proprietà 4**

Assenza di memoria

Data una variabile esponenziale $X \sim E(\lambda)$ *con* X *= "tempo di funzionamento di un dispositivo"*

$$P(X > s+t | X > t) = P(X > s) \qquad \forall s,t \in \mathbb{R}$$

📎 **Dimostrazione**

$$\begin{aligned} P(X > s+t|X>t) &= \frac{P(X > s+t \cap X > t)}{P(X>t)} = \\ &= \frac{P(X>s+t)}{P(X>t)} = \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = \\ &= e^{-\lambda s} = \\ &= P(X>s) \end{aligned}$$

$\square$

@November 11, 2022

▼ Variabile casuale gaussiana (o normale)

> *Una variabile si dice* ***gaussiana*** *o* ***normale*** *se*
>
> $$X \sim N(\mu, \sigma^2) \qquad \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+$$
>
> Tale che:
>
> $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

▼ Funzione di distribuzione di probabilità

$$F(a) = \int_{-\infty}^{a} f(x)dx = \int_{-\infty}^{a} \frac{1}{\sigma\sqrt{s\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

▼ Funzione generatrice dei momenti

$$\phi(t) = e^{t\mu + \frac{t^2}{2}\sigma}$$

▼ **Corollario per la dimostrazione**

$$I = \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx$$

I = ?

Considero I^2

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \cdot \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{(x^2+y^2)}{2}} dxdy \end{aligned}$$

Passo a coordinate polari:

- $x = r\cos(\theta)$
- $y = r\sin(\theta)$
- $dxdy = rdrd\theta$

- $x^2 + y^2 = r^2$

$$\begin{aligned} I^2 &= \int_0^{2\pi} \left(\int_0^{+\infty} e^{-\frac{r^2}{2}} r dr \right) d\theta = \\ &= \int_0^{2\pi} \left[-e^{-\frac{r^2}{2}} \right]_0^{+\infty} d\theta = \\ &= \int_0^{2\pi} 1 d\theta = \\ &= 2\pi \end{aligned}$$

$$\Downarrow$$

$$I = \sqrt{2\pi}$$

📎 **Dimostrazione**

$$\begin{aligned} \phi(t) &= E[e^{tX}] = \\ &= \int_{-\infty}^{+\infty} \frac{e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx \end{aligned}$$

Applico un cambio di variabile $y = \frac{x-\mu}{\sigma}$ *e* $dy = \frac{dx}{\sigma}$

$$\begin{aligned} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t(\sigma y + \mu)} e^{-\frac{y^2}{2}} dy = \\ &= \frac{e^{t\mu}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{t\sigma y - \frac{y^2}{2} + (-\frac{t^2\sigma^2}{2} + \frac{t^2\sigma^2}{2})} dy = \\ &= \frac{e^{t\mu + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(y-t\sigma)^2}{2}} dy \end{aligned}$$

Applico un altro cambio di variabile $z = y - t\sigma$ *e* $dz = dy$

$$\begin{aligned} &= \frac{e^{t\mu + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \\ &= \frac{e^{t\mu + \frac{t^2\sigma^2}{2}}}{\sqrt{2\pi}} \sqrt{2\pi} = \\ &= e^{t\mu + \frac{t^2\sigma^2}{2}} \end{aligned}$$

▼ Valor medio

$$E[X] = \mu$$

📎 **Dimostrazione**

$$\begin{aligned} E[X] &= \left. \frac{d\phi(t)}{dt} \right|_{t=0} = \\ &= \left. [e^{t\mu + \frac{t^2}{2}\sigma^2} (\mu + t\sigma^2)] \right|_{t=0} = \\ &= \mu \end{aligned}$$

▼ Varianza

$$Var(X) = \sigma^2$$

📎 **Dimostrazione**

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

$$\begin{aligned} E[X^2] &= \frac{d^2\phi}{dt^2}\bigg|_{t=0} = \\ &= [e^{t\mu+\frac{t^2}{2}\sigma^2}(\mu+t\sigma^2)^2 + e^{t\mu+\frac{t^2}{2}\sigma^2}\sigma^2]\bigg|_{t=0} = \\ &= \mu^2+\sigma^2 \end{aligned}$$

$$Var(X) = \mu^2+\sigma^2-\mu^2 = \sigma^2$$

□

▼ Proprietà

▼ Proprietà 1

Data una variabile casuale normale $X \sim N(\mu, \sigma^2)$ *e* $Y = \alpha X + \beta$ *(con* $\alpha \in \mathbb{R} \setminus \{0\}$ *e* $\beta \in \mathbb{R}$*)*

$$Y \sim N(\alpha\mu+\beta, \alpha^2\sigma^2)$$

📎 **Dimostrazione**

$$\begin{aligned} \phi_Y(t) &= E[e^{tY}] = \\ &= E[e^{t(\alpha X+\beta)}] = \\ &= E[e^{t\alpha X}e^{t\beta}] = \\ &= e^{t\beta}E[e^{(t\alpha)X}] = \\ &= e^{t\beta}\phi_X(t\alpha) = \\ &= e^{t\beta}(e^{(t\alpha)\mu+\frac{(t\alpha)^2}{2}\sigma^2}) = \\ &= e^{t(\alpha\mu+\beta)+\frac{t^2\alpha^2\sigma^2}{2}} \end{aligned}$$

$$\Downarrow$$

$$Y \sim (\alpha\mu+\beta, \alpha^2\sigma^2)$$

□

▼ Proprietà 2

Date due variabili casuali normali e indipendenti $X \sim N(\mu_1, \sigma_1^2)$ *e* $Y \sim N(\mu_2, \sigma_2^2)$ *allora:*

$$T = X + Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$$

📎 **Dimostrazione**

$$\phi_X(t) = e^{t\mu_1+\frac{t^2}{2}\sigma_1^2}$$
$$\phi_Y(t) = e^{t\mu_2+\frac{t^2}{2}\sigma_2^2}$$

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) = e^{t(\mu_1+\mu_2)+\frac{t^2}{2}(\sigma_1^2\sigma_2^2)}$$
$$\Downarrow$$
$$T \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$$

□

▼ Variabile casuale lognormale

N.B. *Per affrontare questa sezione è necessario aver compreso la parte sulle funzioni di variabili casuali.*

Data una variabile casuale normale $X \sim N(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ e $\sigma^2 \in \mathbb{R}^+$, definiamo una variabile casuale continua Y ***lognormale***:

$$Y = e^X \sim Lognormale(\mu, \sigma^2)$$

quindi:

$$\begin{aligned} Y = g(x) = e^X \geq 0 \\ a = e^X \Rightarrow X = g^{-1}(a) = \ln(a) \qquad \text{con } a > 0 \end{aligned}$$

con:

$$f_Y(a) = \left|\frac{dg^{-1}(a)}{da}\right| f_X(g^{-1}(a))$$

$$\frac{dg^{-1}(a)}{da} = \frac{1}{a}$$

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

▼ Funzione di distribuzione di probabilità

$$F_Y(a) = \begin{cases} P(\frac{X-\mu}{\sigma} \leq \frac{\ln(a)-\mu}{\sigma}) & \text{se } a > 0 \\ 0 & \text{se } a \leq 0 \end{cases}$$

Dimostrazione

$$\begin{aligned} F_Y(a) &= P(Y \leq a) = \\ &= P(e^X \leq a) = \\ &= \begin{cases} P(X \leq \ln(a)) & \text{se } a > 0 \\ 0 & \text{se } a \leq 0 \end{cases} = \\ &= \begin{cases} P(\frac{X-\mu}{\sigma} \leq \frac{\ln(a)-\mu}{\sigma}) & \text{se } a > 0 \\ 0 & \text{se } a \leq 0 \end{cases} \end{aligned}$$

Chiamando $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$F_Y(a) = \begin{cases} F_Z(\frac{\ln(a)-\mu}{\sigma}) & \text{se } a > 0 \\ 0 & \text{se } a \leq 0 \end{cases}$$

▼ Valor medio

$$E[Y] = e^{\mu + \frac{\sigma^2}{2}}$$

Dimostrazione

$$\begin{aligned} E[Y] &= E[e^X] = \\ &= \phi_X(1) = \\ &= e^{t\mu + \frac{t^2}{2}\sigma^2}\Big|_{t=1} = \\ &= e^{\mu + \frac{\sigma^2}{2}} \end{aligned}$$

□

▼ Varianza

$$Var(Y) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

📎 **Dimostrazione**

$$Var(Y) = E[Y^2] - E[Y]^2$$

$$\begin{aligned} E[Y^2] &= E[e^{2X}] = \\ &= \phi_X(2) = \\ &= e^{t\mu + \frac{t^2}{2}\sigma^2}\Big|_{t=2} = \\ &= e^{2\mu + 2\sigma^2} \end{aligned}$$

$$Var(Y) = e^{2\mu+2\sigma^2} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

$\square$

▼ Proprietà

▼ Proprietà 1

Date due variabili casuali lognormali, il loro prodotto è lognormale.

$$Y_1 \sim Lognormal(\mu_1, \sigma_1^2) \qquad Y_2 \sim Lognormal(\mu_2, \sigma_2^2)$$

$$T \sim Lognormal(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

📎 **Dimostrazione**

$Y_1 \sim Lognormal(\mu_1, \sigma_1^2)$ $\qquad$ $Y_2 \sim Lognormal(\mu_2, \sigma_2^2)$

$$Y_2 = e^{X_2}$$

con $X_1 \sim N(\mu_1, \sigma_1^2)$ $\qquad$ con $X_1 \sim N(\mu_2, \sigma_2^2)$

$$T = Y_1 Y_2 = e^{X_1} e^{X_2} = e^{X_1 + X_2} = e^W$$

con $W = X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$, per la proprietà di riproducibilità della gaussiana.

$$T \sim Lognormal(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$\square$

18 - Processo stocastico di Poisson

> **Processo stocastico**
>
> Un processo stocastico è una famiglia di variabili casuali che dipendono da un parametro reale (spesso il tempo)

> **Processo stocastico di Poisson**
>
> Un processo stocastico si dice *di Poisson* se rappresenta una famiglia di variabili casuali
>
> $N(t)$ = "n° eventi che si verificano in $]0,t]$ (dove t è il tempo)"
>
> $N(t)$ variabili casuali discrete.
>
> con le ipotesi:
>
> 1. $N(0)=0$
> 2. Il numero di eventi che si verificano in un certo intervallo di tempo dipende dalla lunghezza dell'intervallo, ma non dipende dalla posizione sull'asse reale (gli intervalli $]0,t]$ e $]s,s+t]$ si comportano nello stesso modo)
> 3. Il numero di eventi che si verificano in un dato intervallo è indipendente dal numero di eventi che si verificano in un intervallo disgiunto dal primo
> 4. $\lim_{h\to 0}\frac{P(N(h)=1)}{h}=\lambda$ con $\lambda \in \mathbb{R}^+$
> 5. $\lim_{h\to 0}\frac{P(N(h)\geq 2)}{h}=0$

N.B. $P(N(t)=k)$ si comporta come una binomiale $B(m,\frac{\lambda t}{m})$, ma per $m\to +\infty$:

$$\begin{gathered} B(m,\frac{\lambda t}{m}) \longrightarrow Po(m\frac{\lambda t}{m}) \\ \Downarrow \\ N(t)\sim Po(\lambda t) \end{gathered}$$

@November 15, 2022

Ora definiamo gli **intertempi** $X_1, X_2, \dots$ come:

X_1 = "Tempo che intercorre tra l'istante iniziale e il verificarsi del primo evento"

X_1 variabile casuale continua

Se $s\geq 0$:

$$P(X_1>s) \underbrace{=}_{\text{il numero di eventi è nullo}} P(N(s)=0) \underbrace{=}_{N(s)\sim Po(\lambda s)} \\ = \frac{(\lambda s)^0}{0!}e^{-\lambda s} = e^{-\lambda s}$$

$$F_{X_1}(s) = P(X_1\leq s) = 1-P(X_1>s) = \begin{cases} 1-e^{-\lambda s} & \text{se } s\geq 0 \\ 0 & \text{se } s<0 \end{cases}$$

Considero dunque un nuovo intertempo:

X_2= "tempo che intercorre tra il primo e il secondo evento"

X_2 variabile casuale continua

Considero $w\geq 0$:

$$\begin{aligned} P(X_2>w) &= P(\text{ci siano 0 eventi in }]s_1, s_1+w]) = \\ &= P(\text{ci siano 0 eventi in }]0,w]) = \\ &= P(N(w)=0) = \\ &= \frac{(\lambda w)^0}{0!}e^{-\lambda w} = \\ &= e^{-\lambda w} \end{aligned}$$

$$F_{X_2}(w) = P(X_2 \leq w) = 1 - P(X_2 > w) = \begin{cases} 1 - e^{-\lambda w} & \text{se } s \geq 0 \\ 0 & \text{se } s < 0 \end{cases}$$
$$\Downarrow$$
$$X_2 \sim E(\lambda)$$

N.B. In modo simile si dimostra che, per $k > 1$:

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

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

19 - Funzione di una variabile casuale nota

> **Funzione di una variabile casuale nota**
>
> *Data una variabile casuale nota X, definiamo una sua funzione*
>
> $$Y = g(X)$$
>
> *con valor medio*
>
> $$E[Y] = E[g(X)] = \begin{cases} \sum_{j=1}^{n} g(x_j)p(x_j) & X \text{ v.c. discreta} \\ \int_{-\infty}^{+\infty} g(x)f(x)dx & X \text{ v.c. continua} \end{cases}$$
>
> *e*
>
> $$E[Y^2] = E[g^2(X)] = \begin{cases} \sum_{j=1}^{n} g^2(x_j)p(x_j) & X \text{ v.c. discreta} \\ \int_{-\infty}^{+\infty} g^2(x)f(x)dx & X \text{ v.c. continua} \end{cases}$$

▼ Funzione di ripartizione di Y per le variabili continue

Nel caso continuo è possibile determinare la funzione di densità di una funzione di una variabile casuale:

▼ Funzione monotona crescente

Data una variabile casuale continua X nota, con funzione di densità di probabilità $f_X(x)$:

$$\begin{aligned} F_Y(a) &= P(Y \le a) = \\ &= P(g(X) \le a) = \\ &= P(X \le x^*) = \\ &= \int_{-\infty}^{x^*} f_X(x)dx = \\ &= \int_{-\infty}^{g^{-1}(a)} f_X(x)dx \end{aligned}$$

Dove $g(x^*) = a$ e quindi $x^* = g^{-1}(a)$

$$\begin{aligned} f_Y(a) &= \frac{d}{da}F_Y(a) = \\ &= \frac{d}{da}\int_{-\infty}^{g^{-1}(a)} f_X(x)dx = \\ &= f_X(g^{-1}(a))\frac{d}{da}g^{-1}(a) \end{aligned}$$

▼ Funzione monotona decrescente

$$\begin{aligned} F_Y(a) &= P(Y \le a) = \\ &= P(g(X) \le a) = \\ &= P(X \ge x^*) = \\ &= \int_{x^*}^{+\infty} f_X(x)dx = \\ &= \int_{g^{-1}(a)}^{+\infty} f_X(x)dx \end{aligned}$$

Dove $g(x^*) = a$ e quindi $x^* = g^{-1}(a)$

$$
\begin{aligned}
f_Y(a) &= \frac{d}{da}F_Y(a) = \\
&= \frac{d}{da}\int_{g^{-1}(a)}^{+\infty} f_X(x)dx = \\
&= -f_X(g^{-1}(a))\frac{d}{da}g^{-1}(a)
\end{aligned}
$$

@November 17, 2022

▼ Funzione non monotona

Se la funzione $g(X)$ non è monotona, è necessario individuare tutti i punti in cui si ha $g(x_k) = a$ e considerare gli intervalli contenuti in questi punti.

$$
\begin{aligned}
F_Y(a) &= P(Y \le a) = \\
&= P(g(X) \le a) = \\
&= P(X \le x_1^* \cup X \in [x_2^*, x_3^*] \cup X \in [x_4^*, x_5^*] \cup \dots) = \\
&= \int_{-\infty}^{x_1^*} f_X(x)dx + \int x_2^{*x_3^*} f_X(x)dx + \int x_4^{*x_5^*} f_X(x)dx + \dots
\end{aligned}
$$

E di conseguenza

$$
f_Y(a) = \frac{d}{da}F_Y(a) = \dots
$$

Si noti che non è possibile scrivere una formula generale.

20 - Funzione di due variabili casuali continue

Date due variabili casuali continue X, Y, considero la funzione nota $f(x,y)$ e la variabile casuale $Z = h(X,Y)$. Allora:

$$\begin{aligned} F_Z &= P(z \le a) \\ &= P(h(X,Y) \le a) \\ &= \iint_{\mathcal{D}} f(x,y)dxdy \end{aligned}$$

Dove $\mathcal{D}$ è l'area della proiezione del grafico con $z \le a$.

▼ Caso particolare: $Z = X + Y$

Considero una variabile casuale continua come la somma di due variabili casuali continue:

$$Z = X + Y$$

▼ Funzione di distribuzione di probabilità

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

📎 **Dimostrazione**

$$\begin{aligned} F_Z(a) = P(Z \le a) = \\ = P(Y \le -X + a) = \\ = \int_{-\infty}^{+\infty} (\int_{-\infty}^{-x+a} f(x,y)dy)dx \end{aligned}$$

Analogamente si dimostra la seconda uguaglianza

□

▼ Funzione di densità di probabilità

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

📎 **Dimostrazione**

$$\begin{aligned} f_Z(a) = \frac{dF_Z(a)}{da} = \\ = \frac{d}{da} \int_{-\infty}^{+\infty} (\int_{-\infty}^{-x+a} f(x,y)dy)dx = \\ = \int_{-\infty}^{+\infty} (\frac{d}{da} \int_{-\infty}^{-x+a} f(x,y)dy)dx = \\ = \int_{-\infty}^{+\infty} f(x, -x+a)dx \end{aligned}$$

Analogamente si dimostra la formula in dy.

□

▼ Caso particolare: v.c. indipendenti

Se due variabili casuali sono indipendenti posso considerare la funzione:

$$f(x,y) = f_X(x) f_Y(x)$$

di conseguenza:

$$f_z(a) = \begin{cases} \int_{-\infty}^{+\infty} f_X(-y+a) f_Y(y) dy \\ \text{oppure} \\ \int_{-\infty}^{+\infty} f_X(x) f_Y(-x+a) dx \end{cases}$$

21 - Teorema del limite centrale

@November 22, 2022

> **Teorema del limite centrale**
>
> *Data una successione di v.c. indipendenti e identicamente distribuite* $X_1, X_2, \ldots, X_n$ *con* $E[X_k] = \mu$ *e* $Var(X_k) = \sigma^2$ *con* $\mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+ \quad \forall k = 1, \ldots, n$. *Definita la v.c.* Y *nel seguente modo:*
>
> $$Y_n = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n\sigma^2}}$$
>
> *allora* $\forall a \in \mathbb{R}$*:*
>
> $$\lim_{n \to +\infty} P(Y_n \leq a) = F_Z(a)$$
>
> dove $Z \sim N(0, 1)$
>
> Ovvero:
> Y_n ***converge in distribuzione ad una v.c. normale standard per*** $n \to +\infty$

▼ Dimostrazione

📎 **Dimostrazione**

$S_n = \sum_{k=1}^{n} X_k$ → variabile casuale

$$E[S_n] = E[\sum_{k=1}^{n} X_k] = \sum_{k=1}^{n} E[X_k] = n\mu$$

$$Var(S_n) = Var(\sum_{k=1}^{n} X_k) = \sum_{k=1}^{n} Var(X_k) = n\sigma^2$$

$$Y_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$$
$$\Downarrow$$
$$\begin{aligned} E[Y_n] &= E[\frac{S_n}{\sigma\sqrt{n}} - \frac{n\mu}{\sigma\sqrt{n}}] = \\ &= \frac{1}{\sigma\sqrt{n}} E[S_n] - \frac{n\mu}{\sigma\sqrt{n}} = \\ &= \frac{n\mu}{\sigma\sqrt{n}} - \frac{n\mu}{\sigma\sqrt{n}} = \\ &= 0 \end{aligned}$$

$$\begin{aligned} Var(Y_n) &= Var(\frac{S_n}{\sigma\sqrt{n}} - \frac{n\mu}{\sigma\sqrt{n}}) = \\ &= \frac{1}{(\sigma\sqrt{n})^2} Var(S_n) = \\ &= \frac{1}{\sigma^2 n} n\sigma^2 = \\ &= 1 \end{aligned}$$

$\square$

▼ Ipotesi aggiuntive

Supponiamo che $\mu = 0$ e $\sigma^2 = 1$.

(***N.B.*** se tale condizione non è verificata per E_k è sempre possibile definire nuove variabili $T_k = \frac{X_k - \mu}{\sigma}$ in modo tale che $E[T_k] = 0$ e $Var(T_k) = 1$)

Allora:

$$Y_n = \frac{\sum_{k=1}^{n} X_k}{\sqrt{n}}$$

Lo scopo della seguente dimostrazione è quello di verificare:

$$\lim_{n \to +\infty} P(Y_n \leq a) = \lim_{n \to +\infty} F_{Y_n}(a) = F_Z(a)$$

Con $Z \sim N(0,1), \forall a \in \mathbb{R}$. Ovvero:

$$\lim_{n \to +\infty} \phi_{Y_n}(t) = \phi_Z(t) = e^{\frac{t^2}{2}}$$

Ovvero:

$$\lim_{n \to +\infty} \ln(\phi_{Y_n}(t)) = \frac{t^2}{2}$$

Assegno a $L(t)$ il logaritmo precedente

$$L(t) = \ln(\phi_{X_k}(t)) \Rightarrow \phi_{X_k} = e^{L(t)}$$

$\forall k$ le variabili casuali sono identicamente distribuite.

$$L(0) = \ln(\phi_{X_k}(0)) = \ln(1) = 0$$

$$\begin{aligned} L'(0) &= \left.\frac{dL}{dt}\right|_{t=0} = \\ &= \left.\frac{d}{dt}\ln(\phi_{X_k}(t))\right|_{t=0} = \\ &= \left.\left[\frac{1}{\phi_{X_k}(t)}\frac{d}{dt}\phi_{X_k}(t)\right]\right|_{t=0} = \\ &= \frac{1}{1}\underbrace{E[X_k]}_{=0} = 0 \end{aligned}$$

$$\begin{aligned} L''(0) &= \left.\frac{d^2L}{dt^2}\right|_{t=0} = \\ &= \left.\frac{d}{dt}\left(\frac{1}{\phi_{X_k}(t)}\frac{d\phi_{X_k}(t)}{dt}\right)\right|_{t=0} = \\ &= \left.\left(\frac{d}{dt}(\frac{1}{\phi_{X_k}(t)})\frac{d\phi_{X_k}(t)}{dt} + \frac{1}{\phi_{X_k}(t)}\frac{d^2\phi_{X_k}(t)}{dt^2}\right)\right|_{t=0} = \\ &= \left.\left((-\frac{1}{\phi_{X_k}(t)^2}\frac{d\phi_{X_k}(t)}{dt})\frac{d\phi_{X_k}(t)}{dt}\right)\right|_{t=0} + \frac{1}{1}E[X^2] \\ &= -E[X_k]^2 + E[X^2] = \\ &= \sigma^2 = \\ &= 1 \end{aligned}$$

Riassumendo:

→ $\phi_{X_k}(t) = e^{L(t)} \quad \forall k = 1, 2, \ldots, n$

→ $L(0) = 0; L'(0) = 0; L''(0) = 1$

$$\begin{aligned} \phi_{Y_n}(t) &= \phi_{\sum_k \frac{X_k}{\sqrt{n}}}(t) = \\ &= \phi_{\frac{X_1}{\sqrt{n}}}(t)\phi_{\frac{X_2}{\sqrt{n}}}(t)\phi_{\frac{X_3}{\sqrt{n}}}(t)\ldots\phi_{\frac{X_n}{\sqrt{n}}}(t) = \\ &= \left(\phi_{\frac{X_1}{\sqrt{n}}}(t)\right)^n \end{aligned}$$

$$\phi_{\frac{X_1}{\sqrt{n}}}(t) = E[e^{t\frac{X_1}{\sqrt{n}}}] = E[e^{(\frac{t}{\sqrt{n}}X_1)}] = \phi_{X_1}(\frac{t}{\sqrt{n}})$$

$$\ln(\phi_{Y_n}(t)) = \ln\left[\left(\phi_{X_1}(\frac{t}{\sqrt{n}})\right)^n\right] =$$
$$= n\ln\left(\phi_{X_1}(\frac{t}{\sqrt{n}})\right) =$$
$$= nL(\frac{t}{\sqrt{n}})$$

$$\lim_{n\to+\infty} \ln(\phi_{Y_n}(t)) = \lim_{n\to+\infty} nL(\frac{t}{\sqrt{n}}) =$$
$$= \lim_{n\to+\infty} \frac{L(\frac{t}{\sqrt{n}})}{n^{-1}} =$$
$$= \lim_{n\to+\infty} \frac{L'(\frac{t}{\sqrt{n}})(-\frac{1}{2}tn^{-\frac{3}{2}})}{-n^{-2}} =$$
$$= \lim_{n\to+\infty} \frac{L'(\frac{t}{\sqrt{n}})\frac{1}{2}t}{n^{-\frac{1}{2}}} =$$
$$= \lim_{n\to+\infty} \frac{\frac{1}{2}tL''(\frac{t}{\sqrt{n}})(-\frac{t}{2}n^{-\frac{3}{2}})}{-\frac{1}{2}n^{-\frac{3}{2}}} =$$
$$= \lim_{n\to+\infty} \frac{1}{2}t^2L''(\frac{t}{\sqrt{n}}) =$$
$$= \frac{1}{2}t^2L''(0) = \frac{1}{2}t^1$$

Da cui si ricava:

$$\lim_{n\to+\infty} \phi_{Y_n}(t) = e^{\frac{t^2}{2}} = \phi_Z(t)$$

$\square$

Applicazioni

Applicazione 1

$X \sim B(n,p)$

$$X = \sum_{k=1}^{n} V_k$$

con $V_k \sim Be(p)$ indipendenti e identicamente distribuite.

Se $n \geq 30$ (valore considerato di solito) allora:

$$X \dot{\sim} N(np, npq)$$

Regole di approssimazione

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

Applicazione 2

Date $X_1, X_2, \ldots, X_n$ variabili casuali indipendenti e identicamente distribuite con $\mu = E[X_k]$ e $\sigma^2 = Var(X_k)$:

$$\bar{X}) \sum_{k=1}^{n} \frac{X_k}{n} \dot{\sim} N(\mu, \frac{\sigma^2}{n})$$

22 - Inferenza statistica

@November 25, 2022

> **Statistica descrittiva**
>
> *La **statistica descrittiva** è la disciplina che si occupa di studiare e presentare dati che si riferiscono a tutti gli individui di una popolazione relativamente ad una o più caratteristiche misurabili*

> **Inferenza statistica**
>
> *L'**inferenza statistica** è la disciplina che parte dallo studio di un sottoinsieme della popolazione (detto campione) ed analizzando le caratteristiche degli individui del campione cerca di inferire qualcosa sulla popolazione.*

N.B. Il campione deve essere scelto in maniera del tutto casuale e deve essere sufficientemente numeroso

> **Popolazione**
>
> *Definiamo **popolazione** un insieme di individui molto numeroso (o infinito)*

Considerando la seguente **ipotesi fondamentale**:

"Le caratteristiche degli individui di una stessa popolazione si suppongono variabili casuali indipendenti e identicamente distribuite"

Di conseguenza il **campione** è:

"Insieme di variabili casuali indipendenti e identicamente distribuite $X_1, X_2, \dots, X_N$, dove Nè la numerosità del campione"

e definiamo la **statistica:**

"Qualsiasi funzione degli elementi del campione $g(X_1, X_2, \dots, X_N)$"

L'**inferenza statistica**, a partire dal campione, cerca di determinare le caratteristiche incognite della funzione di distribuzione di probabilità della popolazione. Essa è definita **inferenza non parametrica** se la forma della funzione di distribuzione è incognita, **inferenza parametrica** se è nota la funzione di distribuzione a meno di una o più parametri.

Nel caso di **inferenza parametrica** chiamiamo F_θ la funzione e θ il parametro incognito.

Definiamo:

> **Stimatore**
>
> *Lo stimatore $\hat{\theta}$ una statistica, ovvero una funzione delle variabili casuali del campione che viene impiegata per stimare θ.*

> **Stima di θ**
>
> *La stima di θ è il valore dello stimatore per un campione fissato.*

▼ Proprietà di un buono stimatore

▼ **Stimatore corretto**

$$E[\hat{\theta}] = \theta$$

▼ **Stimatore efficiente**

Tra tutti gli stimatori viene scelto quello con varianza minima

▼ **Stimatore consistente**

$$\lim_{n \to \infty} E[(\hat{\theta} - \theta)^2] = 0$$

N.B. Se lo stimatore è corretto:

$$\lim_{N \to \infty} Var(\hat{\theta}) = 0$$

▼ Stimatori per una popolazione gaussiana

Si dimostra che per una popolazione gaussiana:

- $\bar{X}$ (media campionaria) è lo stimatore corretto, efficiente e consistente per μ.

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

- S^2 (varianza campionaria) è lo stimatore corretto, efficiente e consistente per σ^2.

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

▼ Media campionaria

Per la legge dei grandi numeri abbiamo già dimostrato che la media campionaria è uno stimatore corretto e consistente:

$$E[\bar{X}] = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{N}$$

Caso 1 - $X_1, \ldots, X_N$ gaussiane

$$\bar{X} = \sum_{k=1}^{N} \frac{X_k}{N} \sim N(\mu, \frac{\sigma^2}{N})$$

Caso 2- $X_1, \ldots, X_N$ non gaussiane

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

Se $N < 30$ la distribuzione di $\bar{X}$ dipende dalla popolazione di partenza.

Se $N \geq 30$, per il *Teorema del Limite Centrale*:

$$\bar{X} \dot{\sim} N(\mu, \frac{\sigma^2}{N})$$

▼ Varianza campionaria

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

Caso 1 - Un solo individuo

Se compare un solo individuo $\bar{X} = X_1$ e $N - 1 = 0$, si ottiene quindi una forma indeterminata $\frac{0}{0}$

Caso 2 - Più individui

→ $E[X_k] = \mu$
→ $Var(X_k) = \sigma^2$
→ $E[X_k^2] = \sigma^2 + \mu^2 *$

→ $E[\bar{X}] = \mu$
→ $Var(\bar{X}) = \frac{\sigma^2}{N}$
→ $E[\bar{X}^2] = \frac{\sigma^2}{N} + \mu^2 *$

$* \, E[X^2] = Var(X) + E[X]^2$

$$\begin{aligned} E[S^2] &= E[\frac{1}{N-1} \sum_{k=1}^{N} (X_k - \bar{X})^2] = \\ &= \frac{1}{N-1} \left(\sum_{k=1}^{N} E[X_k^2] + \sum_{k=1}^{N} E[\bar{X}^2] - 2E[\sum_{k=1}^{N} \bar{X} X_k] \right) = \\ &= \frac{1}{N-1} \left(N(\sigma^2 + \mu^2) + N(\frac{\sigma^2}{N} + \mu^2) - 2NE[\bar{X}^2] \right) = \\ &= \frac{1}{N-1} (N-1)\sigma^2 = \\ &= \sigma^2 \end{aligned}$$

S^2 è uno stimatore corretto per σ^2

23 - Distribuzione χ quadro a n gradi di libertà

> *Date* $Z_1, Z_2, ..., Z_n \sim N(0,1)$ *indipendenti:*
> → $E[Z_k] = 0$
> → $Var(Z_k) = 1$
> → $E[Z_k^2] = Var(Z_k) + E^2[Z_k] = 1$
>
> $$C_n = \sum_{k=1}^{n} \sim \chi_n^2$$

$$E[C_n] = E[\sum_{k=1}^{n} Z_k^2] = \sum_{k=1}^{n} E[Z_k^2] = n \cdot 1$$

> Definiamo il valore critico $\chi^2_{\beta,n}$ come:
>
> $$P(C_n \geq \chi^2_{\beta,n}) = \beta$$

Se la popolazione è Gaussiana N è la numerosità del campione si dimostra che:

$$\frac{(N-1)S^2}{\sigma^2} \sim \chi^2_{N-1}$$

N.B. Se $C_n \sim \chi_n^2$ e $n \gg 1$, per il teorema del limite centrale:

$$C_n \dot{\sim} N(n, ?)$$

▼ Varianza

$$Var(C_n) = 2n$$

Dimostrazione

→ $Z_k \sim N(0,1)$

→ $E[Z_1^2] = Var(Z_1) + E[Z_1]^2 = 1 + 0^2 = 1$

$$Var(C_n) = Var(\sum_{k=1}^{n} Z_k^2) = \sum_{k=1}^{n} Var(Z_k^2)$$

$$Var(Z_1^2) = E[Z_1^4] - E[Z_1^2]^2$$

Ricordiamo la funzione generatrice dei momenti di una normale:

$$\phi_{Z_1}(t) = e^{\frac{t^2}{2}}$$

$$\begin{aligned} E[Z_1^4] &= \frac{d^4}{dt^4}\phi_{Z_1}(t)\Big|_{t=0} = \\ &= \frac{d^3}{dt^3}\frac{2t}{2}e^{\frac{t^2}{2}}\Big|_{t=0} = \\ &= \frac{d^2}{dt^2}(e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}})\Big|_{t=0} = \\ &= \frac{d}{dt}(te^{\frac{t^2}{2}} + 2te^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}})\Big|_{t=0} = \\ &= \left[e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} + 2e^{\frac{t^2}{2}} + 2t^2 e^{\frac{t^2}{2}} + 3t^2 e^{\frac{t^2}{2}} + t^4 e^{\frac{t^2}{2}}\right]\Big|_{t=0} = \\ &= 3 \end{aligned}$$

$$Var(Z_1^2) = 3 - 1 = 2$$

$$Var(C_n) = \sum_{k=1}^{n} Var(Z_k^2) = 2n$$

24 - Intervalli di confidenza

> **t di Student con n gradi di libertà**
> → $Z \sim N(0,1)$
> → $C_n \sim \chi_n^2$
>
> $$T_n \sim t_n$$
>
> $$T_n = \frac{Z}{\sqrt{\frac{C_n}{n}}} \in \mathbb{R}$$
>
> $t_{\beta,n}$ è il **valore critico** della χ_n^2
>
> $$P(T_n \geq t_{\beta,n}) = \beta$$
> $$P(T_n \leq -t_{\beta,n}) = \beta$$

N.B. Il grafico f_{T_n} è simmetrico rispetto all'asse y.

> *La stima puntuale di θ è il valore di $\hat{\theta}$ per un certo insieme di misurazioni (per un certo campione), ma questa stima è troppo sensibile al variare dei risultati sperimentali. Si preferisce dunque utilizzare un intervallo di valori in cui si ha **confidenza** che possa cadere il reale valore di θ. Questo intervallo è detto **Intervallo di confidenza**.*

Seguono esempi di costruzione di intervalli di confidenza, tutti accomunati dall'ipotesi: **Popolazione Gaussiana con almeno un parametro incognito.**

▼ Intervallo di confidenza per μ incognito e σ^2 noto

$$X_1, X_2, \ldots, X_n \sim N(\mu, \sigma^2)$$

utilizziamo

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

come stimatore di μ.

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{N}) \Rightarrow \underbrace{\frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{N}}}}_{Z} \sim N(0,1)$$

chimando z_β il valore critico di $N(0,1)$:

$$P(Z \geq z_\beta) = \beta$$

$$P(Z \leq -z_\beta) = \beta \Rightarrow P(-z_\beta \leq Z \leq z_\beta) = 1 - 2\beta$$

$$P(-z_\beta \leq \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{N}}} \leq z_\beta) = 1 - 2\beta$$
$$P(-z_\beta \frac{\sigma}{\sqrt{N}} \leq \bar{X} - \mu \leq z_\beta \frac{\sigma}{\sqrt{N}}) = 1 - 2\beta$$
$$P(-\bar{X} - z_\beta \frac{\sigma}{\sqrt{N}} \leq -\mu \leq -\bar{X} + z_\beta \frac{\sigma}{\sqrt{N}}) = 1 - 2\beta$$
$$P(\bar{X} + z_\beta \frac{\sigma}{\sqrt{N}} \geq \mu \geq \bar{X} - z_\beta \frac{\sigma}{\sqrt{N}}) = 1 - 2\beta$$
$$P(\bar{X} - z_\beta \frac{\sigma}{\sqrt{N}} \leq \mu \leq \bar{X} + z_\beta \frac{\sigma}{\sqrt{N}}) = 1 - 2\beta$$

$1 - 2\beta$ è detto **confidenza dell'intervallo**.

$$\mu \in [\bar{X} - z_\beta \frac{\sigma}{\sqrt{N}}, \bar{X} + z_\beta \frac{\sigma}{\sqrt{N}}]$$

l'intervallo ha ampiezza $2z_\beta \frac{\sigma}{\sqrt{N}}$, per aumentare l'ampiezza è possibile ridurre z_β (tuttavia questo porterebbe all'aumento di β e quindi alla riduzione della confidenza) oppure aumentare, se possibile, N.

N.B. SI parla di confidenza e non di probabilità perchè una volta che sono stati condotti gli esperimenti $\bar{X}$ assume un valore numerico ed è possibile determinare un intervallo numerico, μ non è una variabile casuale ma una quantità deterministica.

▼ Intervallo di confidenza per μ con σ^2 incognito

Considero N variabili $X_1, X_2, \ldots, X_N \sim N(\mu, \sigma^2)$ con media campionaria:

$$\bar{X} = \sum_{k=1}^{N} \sim N(\mu, \frac{\sigma^2}{N})$$

considero poi:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\frac{\sigma^2}{N}}} \sim N(0,1)$$

e

$$C_{N-1} = \frac{(N-1)S^2}{\sigma^2} \sim \chi^2_{N-1}$$

tali che:

$$T_{N-1} = \frac{Z}{\sqrt{\frac{C_{N-1}}{N-1}}} = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}} \sim t_{N-1}$$

Allora:

$$P(-t_{\beta,N-1} \leq T_{N-1} \leq t_{\beta,N-1}) = 1 - 2\beta$$
$$P(-t_{\beta,N-1} \leq \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{N}}} \leq t_{\beta,N-1}) = 1 - 2\beta$$
$$P(\bar{X} - t_{\beta,N-1}\frac{S}{\sqrt{N}} \leq \mu \leq \bar{X} + t_{\beta,N-1}\frac{S}{\sqrt{N}}) = 1 - 2\beta$$

L'intervallo di confidenza è quindi:

$$\mu \in [\bar{X} - t_{\beta,N-1}\frac{S}{\sqrt{N}}, \bar{X} + t_{\beta,N-1}\frac{S}{\sqrt{N}}]$$

N.B. Come nel caso precedente per ridurre l'ampiezza è possibile ridurre $t_{\beta,N-1}$ (ovvero la confidenza) oppure aumentare N.

▼ Intervallo di confidenza per σ^2

Considero:

$$C_{N-1} = \frac{(N-1)S^2}{\sigma^2} \sim \chi^2_{N-1}$$

con popolazione gaussiana.

S^2 è lo stimatore di σ^2 con $S^2 = \sum_i \frac{(X_i - \bar{X})^2}{N-1}$.

$$P(\chi^2_{1-\beta,N-1} \leq C_{N-1} \leq \chi_{\beta,N-1}) = 1 - 2\beta$$
$$P\left(\chi^2_{1-\beta,N-1} \leq \frac{(N-1)S^2}{\sigma^2} \leq \chi_{\beta,N-1}\right) = 1 - 2\beta$$
$$P\left(\frac{\chi^2_{1-\beta,N-1}}{(N-1)S^2} \leq \frac{1}{\sigma^2} \leq \frac{\chi_{\beta,N-1}}{(N-1)S}\right) = 1 - 2\beta$$
$$P\left(\frac{(N-1)S^2}{\chi^2_{1-\beta,N-1}} \geq \sigma^2 \geq \frac{(N-1)S}{\chi_{\beta,N-1}}\right) = 1 - 2\beta$$
$$P\left(\frac{(N-1)S^2}{\chi^2_{\beta,N-1}} \leq \sigma^2 \leq \frac{(N-1)S}{\chi_{1-\beta,N-1}}\right) = 1 - 2\beta$$

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

25 - Regressione

Avendo $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ dati in entrata e un uscita Y, associati da una funzione

Regressione: determinare f oppure i coefficienti in essa contenuti.

Regressione lineare: si ha per f funzione lineare delle variabili d'ingresso.

Regressione lineare semplice: una sola variabile in ingresso ed f lineare.

Caso regressione lineare semplice:

è ora necessario stimare α e β.

Ciò che viene osservato in uscita è:

$$Y = \beta X + \alpha + \text{errore casuale}$$

siano A e B gli stimatori di α e β:

$$SS^2 = \sum_{k=1}^{M} (Y_k - (BX_k + A))^2$$

essi devono essere scelti in modo da minimizzare SS^2 (si usa il metodo dei minimi quadrati)

$$\begin{cases} \frac{\partial SS^2}{\partial A} = 0 \\ \frac{\partial SS^2}{\partial B} = 0 \end{cases} \Rightarrow \begin{cases} \sum_k 2(Y_k - (BX_k + A)) \cdot (-1) = 0 \\ \sum_k 2(Y_k - (BX_k + A)) \cdot (-X_k) = 0 \end{cases}$$

Si ottiene quindi:

1.

$$\sum_{k=1}^{M} Y_k - B \sum_{k=1}^{M} X_k - \sum_{k=1}^{M} A = 0$$

chiamando $\bar{X}$ e $\bar{Y}$ le rispettive medie aritmetiche è possibile scrivere l'equazione nel seguente modo:

$$\begin{gathered} M\bar{Y} - BM\bar{Y} - MA = 0 \\ \bar{Y} - B\bar{X} - A = 0 \\ A = \bar{Y} - B\bar{X} \end{gathered}$$

2.

$$\begin{gathered} \sum_{k=1}^{M} (Y_k X_k - BX_k^2 - AX_k) = 0 \\ \sum_{k=1}^{M} Y_k X_k - \sum_{k=1}^{M} BX_k^2 - \sum_{k=1}^{M} AX_k = 0 \end{gathered}$$

Sostituisco con i valori di **1.**

$$\begin{gathered} \sum_{k=1}^{M} Y_k X_k - \sum_{k=1}^{M} BX_k^2 - \sum_{k=1}^{M} AX_k = 0 \\ \sum_{k=1}^{M} Y_k X_k - \sum_{k=1}^{M} BX_k^2 - \bar{Y} \sum_{k=1}^{M} X_k + B\bar{X} \sum_{k=1}^{M} X_k = 0 \\ \sum_{k=1}^{M} Y_k X_k - \sum_{k=1}^{M} BX_k^2 - \bar{Y} M \bar{X} + BM\bar{X}^2 = 0 \\ B(\sum_{k=1}^{M} X_k^2 - M\bar{X}^2) = \sum_{k=1}^{M} X_k Y_k + M\bar{X}\bar{Y} \\ B = \frac{\sum_{k=1}^{M} Y_k X_k - M\bar{X}\bar{Y}}{\sum_{k=1}^{M} X_k^2 - M\bar{X}} = \frac{\sum_{k=1}^{M} Y_k (X_k - \bar{X})}{\sum_{k=1}^{M} X_k^2 - M\bar{X}} \end{gathered}$$

RIASSUMENDO:

$$B=\frac{\sum_{k=1}^{M}Y_k(X_k-\bar{X})}{\sum_{k=1}^{M}X_k^2-M\bar{X}}$$

$$A=\bar{Y}-B\bar{X}$$

▼ Ipotesi aggiuntive

1 - ***Le quantità** $X_1,X_2,\ldots,X_M$ **sono quantità deterministiche prive di errore:***

$$X_k=x_k\Rightarrow\bar{X}=\bar{x}=\sum_{k=1}^{M}\frac{x_k}{M}$$

(non si può parlare di media campionaria dato che X_k non sono variabili casuali appartenenti allo stesso campione)

2 - $Y_k\sim N(\beta x_k+\alpha,\sigma^2)$ ***con*** $k=1,2,\ldots,M$ ***e le*** Y_k ***sono indipendenti:***

$$Y_k=\beta X_k+\alpha+\text{errore}$$

Per il teorema del limite centrale $\dot\sim N(0,\sigma^2)$.

$\bar{Y}=\frac{\sum Y_k}{M}$ è una media aritmetica di variabili casuali ma **non** è una media campionaria dato che le Y_k presentano $E[Y_k]$ diverse (Y_k non sono identicamente distribuite), non appartengono allo stesso campione.

Assunte queste due ipotesi verifico che gli stimatori siano corretti.

Stimatore B

$$E[B]=E\left[\frac{\sum_k Y_k(x_k-\bar{x}_k)}{\sum_k x_k^2-M\bar{x}^2}\right]$$

Per l'ipotesi **1** Y_k sono le uniche variabili casuali

$$=\frac{\sum_k(x_k-\bar{x})E[Y_k]}{\sum_k x_k^2-M\bar{x}^2}=$$

Per l'ipotesi **2**:

$$=\frac{\sum_k(x_k-\bar{x})(\beta x_k+\alpha)}{\sum_k x_k^2-M\bar{x}^2}=$$

$$=\beta\frac{\sum_k x_k(x_k-\bar{x})}{\sum_k x_k^2-M\bar{x}^2}+\alpha\frac{\overbrace{\sum_k(x_k-\bar{x})}^{=0}}{\sum_k x_k^2-M\bar{x}^2}=$$
$$=\beta\frac{\sum_k x_k^2-M\bar{x}^2}{\sum_k x_k^2-M\bar{x}^2}=$$
$$=\beta$$

Stimatore A

$$
\begin{aligned}
E[A] &= E[\bar{Y} - B\bar{X}] = \\
&= E[\bar{Y}] - E[B]\bar{x} = \\
&= E[\sum_{k=1}^{M} \frac{Y_k}{M}] - \beta\bar{x} = \\
&= \frac{1}{M}\sum_{k=1}^{M} E[Y_k] - \beta\bar{x} = \\
&= \frac{1}{M}\sum_{k=1}^{M} (\beta x_k + \alpha) - \beta\bar{x} = \\
&= \beta\sum_{k=1}^{M} \frac{x_k}{M} + \alpha\frac{M}{M} - \beta\bar{x} = \\
&= \beta\bar{x} + \alpha - \beta\bar{x} = \\
&= \alpha
\end{aligned}
$$

Calcolo la varianza di B

$$
\begin{aligned}
Var(B) &= Var\left(\sum_{k=1}^{M} \frac{(x_k - \bar{x})Y_k}{\sum_k x_k^2 - M\bar{x}}\right) = \\
&= \sum_{k=1}^{M} \frac{(x_k - \bar{x})^2}{(\sum_k x_k^2 - M\bar{x}^2)^2} \underbrace{Var(Y_k)}_{=\sigma^2} = \\
&= \sigma^2 \frac{\sum_k (x_k - \bar{x})^2}{(\sum_k x_k^2 - M\bar{x}^2)^2} = \\
&= \sigma^2 \frac{\sum_k (x_k - \bar{x})^2}{(\sum_k (x_k - \bar{x})^2)^2} = \\
&= \frac{\sigma^2}{\sum_k (x_k - \bar{x})^2}
\end{aligned}
$$

Immaginando che σ^2 sia assegnato, per minimizzare la varianza si può intervenire solo sul denominatore, considerando dei valori “molto grandi” e distanti tra loro.

Note dell’autore

Gennaio 2023

Carissimi compagni di corso, presenti o futuri. Ho creato questa piccola-grande dispensa per permettere di facilitare tutti con lo studio e la comprensione degli argomenti. Ovviamente pubblico tutto in maniera gratuita e incondizionata, ma qualora qualcuno trovasse veramente utile questo libretto, lascio la possibilità di regalarmi qualche spicciolo per un caffè :).
Auguro a tutti voi buona fortuna e un buon proseguimento degli studi,
vi ringrazio,

Simone Elia

Pay Simone Elia using PayPal.Me

Go to paypal.me/simoneeelia and type in the amount. Since it's PayPal, it's easy and secure. Don't have a PayPal account? No worries.

http://paypal.me/simoneeelia