EPS
LEZIONE 1.

IL GIOCO DELLA ZARA (con 2 dadi onesti)

descrizione formale di un esperimento aleatorio e di numeri non deterministici.

$\mathcal{E}_2$ = esperimento → consiste nel lanciare due dadi onesti in maniera tale che si può parlare del primo e secondo dado

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

$\Omega \longrightarrow \mathbb{R}$

Trasformazione somma

attraverso la trasformazione dell'esito di $\mathcal{E}$ si ottiene $m \in \mathbb{R}$

1+2

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

6+6

Lo SPAZIO CAMPIONE $\Omega_2 = \{(i,j) : i,j = 1,2,\dots,6\} \to \{(1,1),(1,2),\dots,(6,6)\}$
→ insieme dei possibili risultati

Ne risulta che $|\Omega_2| = 6^2$ ovvero che in Ω_2 ci sono 36 coppie ordinate ognuna di quali è un PUNTO CAMPIONE

$S_2 = \{2,3,4,5,6,7,8,9,10,11,12\}$ → SPETTRO di $\mathcal{E}_2$
→ valori che l'esperimento può assumere tramite Z con probabilità non nulla

P (probabilità di un evento) = # CASI FAVOREVOLI / # CASI POSSIBILI

CASI POSSIBILI = SPAZIO CAMPIONE

$i,j = 1,2,\dots,6$ $\quad P_2(\{(i,j)\}) = \frac{1}{36}$ → è possibile utilizzare la definizione di probabilità introdotta da LAPLACE

Afferma che dato uno spazio di eventi finito, la probabilità è il rapporto tra il numero di casi favorevoli al verificarsi di un qualunque evento e il numero di casi possibili posto che gli eventi siano tutti equiprobabili

Considerando EVENTO ogni sottoinsieme di Ω_2 si ha:

$A \in \mathcal{P}(\Omega_2)$, $\quad P_2(A) = |A| / 36$

Raramente, però, si è interessati all'esito dell'esperimento, ma il valore numerico è una sua FUNZIONE

In questo caso ha senso considerare la somma dei punteggi dei due dadi:

$Z_2 : (i,j) \in \Omega_2 \longrightarrow (i+j) \in \mathbb{R}$ → da punto campione a numero reale

Assume con probabilità positiva i valori: 2, 3, ..., 11, 12
Per ognuno dei possibili valori è possibile determinare la sua CONTROIMMAGINE tramite Z_2

$P_2(Z_2=2) = P_2(Z_2=12) = 1/36$
$P_2(Z_2=4) = P_2(Z_2=10) = 3/36$
$P_2(Z_2=3) = P_2(Z_2=11) = 2/36$
$P_2(Z_2=5) = P_2(Z_2=9) = 4/36$
$P_2(Z_2=6) = P_2(Z_2=8) = 5/36$
$P_2(Z_2=7) = 6/36$

$$\sum_{s=2}^{12} P_2(Z_2=s) = \frac{2(1+2+3+4+5)+6}{36} = 1$$

$$(\Omega, \mathcal{P}(\Omega), P) \xrightarrow{Z} (\mathbb{R}, \mathcal{B}, P_Z)$$

$\mathcal{B}$ è una famiglia di sottoinsiemi di $\mathbb{R}$

1.28.40

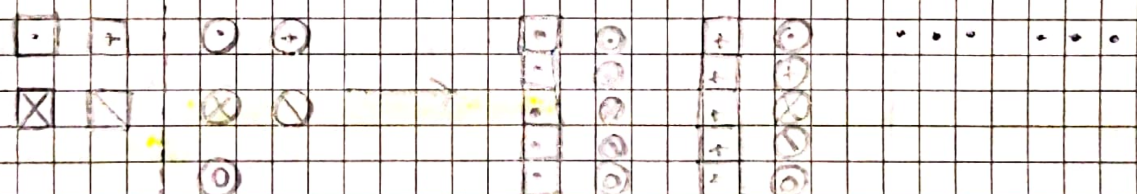
FORMALIZZAZIONE DI UN ESPERIMENTO ALEATORIO $\mathcal{E}$

- PROBLEMA DEL CONTARE
  → soluzione → Regola Moltiplicativa / principio fondamentale del calcolo combinatorio

principio di moltiplicazione

due insiemi

CASO $n = 2$



$4 \cdot 5 = 20$ MODI → procedura di scelta
→ - Come si classificano le scelte
- Come associare gli elementi nella procedura di scelta

$S = \{e_1, e_2, \dots, e_m\}$ $|S| = m$
alfabeto

R-selezione di S, $R \in N = \{1, 2, \dots, m\}$
→ in questo caso, costruzione di una parola dell'alfabeto S

- DISPOSIZIONI → R-selezioni in cui è importante l'ordine di presentazione degli elementi di S
↕
- COMBINAZIONI → R-selezioni in cui non è importante l'ordine in cui si presentano gli elementi di S

Sia combinazioni che disposizioni si suddividono in crea semplici e composizioni

LEZIONE 2

SPAZI DI ESITI EQUIPROBABILI → è uno spazio di esiti S
$S = \{1, 2, \dots, N\}$, in cui ogni esito ha la stessa probabilità di verificarsi

$P(\{1\}) = P(\{2\}) = \dots = P(\{N\})$

Principio di enumerazione ⟷ coefficiente binomiale
→ applicazione del principio di enumerazione nei problemi del contare

TOPIC DEL PROBLEMA → Contare quante sono le R-selezioni tra disposizioni e combinazioni

Si suppone di avere scelto l'alfabeto $S = \{e_1, \dots, e_m\}$ dove $m \in \mathbb{N}$ e $|S| = m$ e $R \in \mathbb{N}$

1) $D^{(r)}_{m,r}$ → numero delle disposizioni che ammettono ripetizioni lunghe R

1	2	...	R
•	•	.....	•
m	m	.....	m

$$\boxed{D^{(r)}_{m,R} = m^R}$$ ⇐

→ ogni sottoprocedura lavora su uno stesso insieme

2) $D_{m,R}$ → numero delle disposizioni lunghe R che NON ammettono ripetizioni

procedure di scelta in R sotto-procedure semplici

1	2	...	r
•	•	.....	•
m	(m-1)	.....	(m-R+1)

in questo caso le ripetizioni NON sono ammesse

- per la prima lettera, m possibili scelte
- per la seconda lettera, $m-1$ possibili scelte

E così via

→ **FATTORIALE DECRESCENTE** di ordine k

2.c) $m = k \Rightarrow P_m = D_{m,m}$ → disposizioni semplici in cui $k = m$

PERMUTAZIONI SEMPLICI

$= m \cdot (m-1) \cdot \ldots \cdot 2 \cdot 1 = m!$

→ $D_{m,k} = m \cdot (m-1) \cdot (m-k+1) \cdot \frac{\cancel{(m-k)}}{(m-k)} \cdot \frac{\cancel{(m-k+1)}}{(m-k-1)} \cdot \ldots \cdot \frac{\cancel{2}}{2} \cdot \frac{\cancel{1}}{1}$

$= \boxed{m! / (m-k)!}$

N.B.

$0! = 1$ → $m! / (m-m)!$

$\Rightarrow m! / 0! = m!$

ESEMPI DI APPLICAZIONE

- **Totocalcio** → AMMESSE RIPETIZIONI

pronosticare 14 simboli ⇒ 14 colonne

$S = \{1, X, 2\}$

$D^{(r)}_{3,14} = 3^{14} = 4\,782\,969$

14 partite

3 simboli

- **Corse TRIS**

cavalli

$S = \{1,2,3,4,5,6,7,8,9\}$

il giocatore deve indovinare il podio gioco.... 3 partorime dei 9 cavalli in cui non sono ammesse ripetizioni

senza ripetizioni

$D_{m,k} = D_{9,3} = m! / (m-k)! = 9! / 6!$

$= 9 \cdot 8 \cdot 7 \cdot \cancel{6!} / \cancel{6!} = 9 \cdot 8 \cdot 7 = 504$

3) $k \leq n$, $C_{n,k}$ → numero delle k-combinazioni semplici con gli elementi scelti da un insieme di cardinalità n, senza ripetizioni

$$= \boxed{D_{n,k} / P_k} = n! / k! \cdot (n-k)! = \binom{n}{k}$$ COEFFICIENTE BINOMIALE

numero delle k combinazioni semplici su un insieme di n cardinalità

4) $C^{(r)}_{n,k}$

Si suppone che $S = \{a, b, c, d, e, f\}$ con una k-selezione semplice lunga 4

$\{A, P\}$ ASSENZA PRESENZA

	a	b	c	d	e	f
abdf	P	P	A	P	A	P
abcd	P	P	P	P	A	A
ace	P	A	P	A	P	A
bcd	A	P	P	P	A	A

SENZA RIPETIZIONE COMBINAZIONI

questa colonna è fissa

	a	b	c	d	e	f				
aaaa	P	A	A	A	A	A	P	P	P	A
aaee	P	A	A	A	P	A	P	A	P	A
bbee	A	P	A	A	P	A	P	A	P	A
abcd	P	P	P	P	A	A	A	A	A	A
beef	A	P	A	A	P	P	A	P	A	A

$$\Rightarrow C^{(r)}_{n,k} = {}_{n-1,k}P^{(r)}_{n+k-1} = \binom{n+k-1}{k} = (n+k-1)! / k!(n-1)!$$

ESEMPIO 3.5.1

○○○○○○ 6
◉◉◉◉◉ 5

Estrazione a caso di due palline da un'urna che ne contiene 6 bianche e una che ne contiene 5 nere

Qui coppia di biglie ha la stessa probabilità di uscita

$$\frac{\binom{6}{1} \cdot \binom{5}{1}}{\binom{11}{2}}$$

← possibilità di prendere una bianca e una nera
← CASI POSSIBILI

$$\frac{\binom{6}{1}\cdot\binom{5}{1}}{\binom{11}{2}} = \frac{\frac{6!}{1!\cdot 5!}\cdot\frac{5!}{1!\cdot 4!}}{\frac{11!}{2!\cdot 9!}} = \frac{6\cdot 5}{\frac{110}{2}} = \frac{30}{55} = \frac{6}{11}$$

LEZIONE 3

RIEPILOGO

- Problema del contare → principio di enumerazione

$n, k \in \mathbb{N}$, $S = \{a_1, \dots, a_n\}$

- $D^r_{n,k} = n^k$
- $k \le n$, $D_{n,k} = n!/(n-k)!$
- $P_n = n!$
- $n_1 \dots n_s$ $P^r_n = \binom{n}{n_1,\dots,n_s}$
- $C^r_{n,k} = \binom{n+k-1}{n}$
- $k \le n$, $C_{n,k} = \binom{n}{k}$

VARIABILI ALEATORIO

↓

è un tipo di esperimento per il quale non si sa quale è l'esito ≠ CASUALE

è un caso particolare di aleatorio, più facile da interpretare ma più difficile da trattare: significa che nessuno degli esiti elementari è preferito rispetto agli altri, cioè hanno tutti la stessa importanza (probabilità)

$\Omega = \{(i,j) : i,j = 1,2,\dots,6\}$ SPAZIO CAMPIONE

↓

Tutti i sottoinsiemi di uno spazio campione sono eventi.

↓

$P(\Omega)$ = FAMIGLIA DEGLI EVENTI

$\forall A \in P(\Omega)$, $\mathbb{P}(A) = |A|/|\Omega|$ ← definizione di LAPLACE

Descrizione dell'esperimento $\longrightarrow$ $(\Omega, F = P(\Omega), \mathbb{P})$

l'esito elementare è un numero reale

funzione di trasformazione

$$\Omega \longrightarrow \mathbb{R}$$
$$(i,j) \longrightarrow Z_2(i,j) = i+j$$

Con la funzione di trasformazione si può passare da un qualcosa di NON quantificabile a un numero reale quantificabile

$(\mathbb{R}, \mathcal{B}, \mathbb{P}_{Z_2})$

famiglia di insiemi di Borel $\longrightarrow$ La famiglia dei boreliani è ottenuta dagli intervalli con le operazioni di $\cup, \cap, ^c$. Operando in tutti i modi possibili con queste operazioni fino al numerabile si ottiene questa famiglia di sottoinsiemi di $\mathbb{R}$ che viene chiamata "FAMIGLIA DI BOREL"

SPETTRO $\longrightarrow$ $S_{Z_2} = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

$\mathbb{P}(Z_2 = 2) = \mathbb{P}(\{2\}) = 1/36$
$\mathbb{P}(Z_2 = 3) = \mathbb{P}(\{3\}) = 2/36$
$\mathbb{P}(Z_2 = 5) = \mathbb{P}(\{5\}) = 4/36$
$\mathbb{P}(Z_2 = 7) = \mathbb{P}(\{7\}) = 6/36$

Assegnare a ognuno dei valori dello spettro la sua probabilità in $\mathbb{R}$

Espansione da Z_2 a Z_3 $\longrightarrow$ al caso di tre dadi

$\mathcal{E}_3$ lancio di tre dadi onesti

$|\Omega_3| = 6^3 \Leftarrow \Omega_3 = \{(i,j,k) : i,j,k = 1,2,\dots,6\}$

ci sono 216 coppie ordinate ognuna delle quali è un punto campione $\longrightarrow$ ogni singoletto ha probabilità $1/216$

$$Z_3 : \Omega \longrightarrow \mathbb{R}$$
$$(i,j,k) \longrightarrow i+j+k \in \mathbb{R}$$

SPETTRO $= \{3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18\}$

la terna $(\Omega_3, F = P(\Omega_3), \mathbb{P}_3)$ descrive in maniera formale $\mathcal{E}_3$

$S_{Z_3} \cap]-10, 10]$

$\mathbb{P}_{Z_3}(\{3,4,5,6,7,8,9,10\}) = \mathbb{P}_{Z_3}\left(\bigcup_{s=3}^{10}\{s\}\right) = \sum_{s=3}^{10} \mathbb{P}_{Z_3}(\{s\})$

l'unione disgiunta è una SOMMA

Operazioni Algebriche → op. insiemistiche → determinazione di eventi che si verificano

Linguaggio logico degli eventi → linguaggio algebrico

$A \cap B \cup A \cap B^c \cup A^c \cap B \cup A^c \cap B^c = A \cup B$

$\Rightarrow A \cap B \cup A \cap B^c \cup A^c \cap B \cup A^c \cap B^c = \Omega$

→ 4 singole forme che creano, uniti disgiuntamente, lo spazio campione

UNIONE = l'unione di due eventi A e B è l'evento $E = A \cup B$ che si verifica se almeno uno dei due eventi A o B si verificano

INTERSEZIONE: l'intersezione di due eventi A e B è l'evento $E = A \cap B$ che si verifica se si verificano entrambi gli eventi A e B.

NEGAZIONE - la negazione di un evento A è l'evento A^c che si verifica se non si verifica A

Sfruttando le operazioni sugli insiemi è possibile definire i seguenti eventi:

→ EVENTO IMPOSSIBILE = è l'evento che non può mai verificarsi → $\emptyset = \Omega^c$

→ EVENTO CERTO = evento che si verifica sempre in quanto comprende tutti i possibili risultati dell'esperimento → $A \cup A^c = B \cup B^c = \Omega$

LEZIONE 4

$(?,?,?)$ $(\Omega, \quad , \quad)$

Determinare la terna della probabilità

$\mathcal{E}$ esperimento NON deterministico ALEATORIO → NON è prevedibile il risultato prima dell'effettuazione

↓

Ω (insieme di possibili esiti): SPAZIO CAMPIONE

→ quando Ω è $\mathbb{R}$, allora il secondo elemento nella descrizione di $\mathcal{E}$ è $P(\mathbb{R})$ e questo compromette il terzo elemento nella descrizione di $\mathcal{E}$

→ $P(\mathbb{R})$ sostituito dall'insieme $\mathcal{B}$ $(\mathcal{B} \subset P(\mathbb{R}))$

↓

FAMIGLIA DEGLI INSIEMI DI BOREL

$(\Omega, \cdot, \cdot)$

↑

CHI È? Si consideri $\mathcal{E}$: lancio di un dado onesto

- A: esce punteggio pari: 2, 4, 6
- B: esce punteggio alto: 5, 6

$A \cap B = \{6\}$ È UN EVENTO

esce pari e alto

$P(A) = 3/6 = 1/2$
$P(B) = 2/6 = 1/3$
$P(A \cap B) = 1/6$

A^c è un EVENTO ⇒ complementare dell'evento A: esce punteggio dispari: $\{1,3,5\}$

B^c è un EVENTO ⇒ complementare dell'evento B: esce punteggio BASSO: $\{1,2,3,4\}$

$A \cup B = \{2,4,6,5\}$ È UN EVENTO

METODO DEGLI ATOMI → atomizzare gli eventi iniziali A e B → frazionare gli eventi A e B

$A \cap B$
$A \cap B^c$
$A^c \cap B$
$A^c \cap B^c$
→ 4 atomi →
$A \cap B = \{6\}$
$A \cap B^c = \{2,4\}$
$A^c \cap B = \{5\}$
$A^c \cap B^c = \{1,3\}$

$\{1\}, \{2\}, \{3\}, \{4\}$ NON sono EVENTI

$G := \{A, B\}$ EVENTI GENERATORI

$\{\{6\}, \{5\}, \{1,3\}, \{2,4\}, \{5,6\}, \{1,3,5\}, \{2,4,5\}, \{6,1,3\}, \{2,4,6\}, \{1,2,3,4\},$
$\{1,3,5,6\}, \{2,4,5,6\}, \{1,2,3,4,6\}, \{1,3,2,4,5\}, \Omega, \emptyset\}$

Facendo l'unione tra gli atomi, si ottiene la famiglia di tutti i possibili eventi, tramite l'operazione σ

=> $\sigma(G) = \{\{5\}, \{6\}, \{1,3\}, \{2,4\}, \{5,6\}, \{1,3,5\}, \{2,4,5\}, \{1,3,6\}, \{2,4,6\}, \{1,2,3,4\}, \{1,3,5,6\}, \{2,4,5,6\}, \{1,2,3,4,6\}, \{1,2,3,4,5\}, \Omega, \emptyset\}$

sigma-operatore

EVENTO CERTO E EVENTO IMPOSSIBILE

=> il secondo elemento descrittivo dell'esperimento, è proprio l'insieme che si ottiene da tutti i generatori di eventi tramite l'op. σ

$\longrightarrow (\Omega, F = \sigma(G), ?)$

DEF.

Sia a una famiglia di sottoinsiemi di Ω. Si dice che a è un'algebra $\Leftrightarrow$
- $\Omega \in a$ (a1)
- $\forall A \in a,\ A^c \in a$ (a2)
- $\forall A_1, A_2 \in a,\ A_1 \cup A_2 \in a$ (a3)

→ ASSIOMI KOLMOGOROV

Sia F una famiglia di sottoinsiemi di Ω.
Si dice che F è una σ-algebra $\Leftrightarrow$

- Σ_1: $\Omega \in F$
- Σ_2: $\forall A \in F,\ A^c \in F$
- Σ_3: $(A_n)_{n \in \mathbb{N}} \subseteq F,\ \bigcup_{n \in \mathbb{N}} A_n \in F$

- $A \in a$ si può esprimere come A è un evento
- Ω è detto EVENTO CERTO
- Se A è un evento, A^c si chiama evento contrario di A
- $\emptyset = \Omega^c \longrightarrow$ evento impossibile
- A è stabile rispetto all'unione finita $\longrightarrow$ A stabile rispetto all'unione NUMERABILE
- Dalle leggi di De Morgan discende facilmente la stabilità di A rispetto all'intersezione numerabile

$\longrightarrow$ UNIONE NUMERABILE $\xleftrightarrow{\text{DE MORGAN}}$ INTERSEZIONE NUMERABILE

Sia a una σ-algebra, siano $A_1, A_2 \in a$ =>

$A_1 \cap A_2 \in a$: $(\underbrace{A_1^c}_{\in a} \cup \underbrace{A_2^c}_{\in a})^c \in a$ per (a2)

$\longrightarrow \in a$

DIMOSTRAZIONI DELLA STABILITÀ DELL'INTERSEZIONE NUMERABILE IN UNA σ-ALGEBRA

Sia $\mathcal{A}$ una σ-algebra
$m \in \mathbb{N} : m \geq 2$ e

- Siano $A_1, A_2, \dots, A_m \in \mathcal{A}$

$$\bigcap_{i=1}^{m} A_i = \Big(\bigcup_{i=1}^{m} \underbrace{A_i^c}_{\in \mathcal{A}} \Big)^c \in \mathcal{A}$$

$\in \mathcal{A}$

Sia F una σ-algebra e $(A_m)_{m \in \mathbb{N}} \subseteq F$ e =>

$$\Rightarrow \bigcap_{m \in \mathbb{N}} A_m = \Big(\underbrace{\bigcup_{m \in \mathbb{N}} A_m^c}_{\in F} \Big)^c \longrightarrow \in F$$

$\in F$

- ESEMPIO

$\mathcal{E}$: Si lancia una moneta un numero indefinito di volte
∞ → numero indefinito di volte

$\Omega = \{T, C\}$

$\underline{\omega} = (\omega_1, \omega_2, \omega_3, \dots) \longrightarrow$ l'elemento generico è una successione di Testa o Croce

↳ $(\omega_m)_{m \in \mathbb{N}}$

$A =$ " due T consecutive prima di due C consecutive "

↳ $A_1 : T_1 T_2$
$A_2 : C_1 T_2 T_3$
$A_3 : T_1 C_2 T_3 T_4$
$A_4 : C_1 T_2 C_3 T_4 T_5$

$\longrightarrow A = A_1 \cup A_2 \cup A_3 \cup A_4 \dots$

↓

A è unione di una successione di eventi

$$\Rightarrow \mathcal{E} = (\Omega, \underbrace{F = \sigma(G)}, ?)$$

$\longrightarrow G = (T_m)_{m \in \mathbb{N}}$

il secondo elemento che descrive formalmente un esperimento aleatorio è una σ-algebra F

ORA

Come si DEFINISCE la probabilità di un evento?

Modi di definire la probabilità di un evento

1) LAPLACE : $A \in F$, $\mathbb{P}(A) = |A| / |\Omega| \geq 0$ come è solo nel caso in cui si riscontra perfetta simmetria per i punti campioni che sono un numero finito

2) **FREQUENTISTA** (o STATISTICA) $\longrightarrow$ la probabilità di un esito è considerata una proprietà dell'esito stesso

si ripete l'esperimento un numero $m \in \mathbb{N}$ di volte e si considera il rapporto:

$$\mathbb{P}_f(A) = \lim_{m \to \infty} \frac{m_A}{m} \geq 0$$

$m_A \longrightarrow$ numero delle volte in cui A si è verificato

$\frac{m_A}{m} \longrightarrow$ frequenza relativa dell'evento A

$\downarrow$

la probabilità di due eventi disgiunti sarà uguale alla somma delle probabilità

3) **SOGGETTIVO**

La probabilità di un evento è il prezzo che un individuo equo e coerente paga per ricevere ~~uno~~ 1 nel caso che A si verifichi

$\longrightarrow$ è il grado di fiducia che ciascun individuo attribuisce al verificarsi dell'evento, in base alle informazioni a sua disposizione

$\mathbb{P}_c(\Omega) = \mathbb{P}_f(\Omega) = \mathbb{P}_s(\Omega) = 1$

$A \in F$, $\mathbb{P}_c(A) \geq 0$, $\mathbb{P}_f(A) \geq 0$, $\mathbb{P}_s(A) \geq 0$

FINITA ADDITIVITÀ

$A, B \in F$, $A \cap B = \emptyset \longrightarrow$
$\mathbb{P}_c(A \cup B) = \mathbb{P}_c(A) + \mathbb{P}_c(B)$
$\mathbb{P}_f(A \cup B) = \mathbb{P}_f(A) + \mathbb{P}_f(B)$
$\mathbb{P}_s(A \cup B) = \mathbb{P}_s(A) + \mathbb{P}_s(B)$

Da cui deriviamo gli assiomi di Kolmogorov

$\downarrow$

$(\Omega, F = \sigma(G), \boxed{\mathbb{P}})$

$\downarrow$

deve rispettare i tre assiomi $\Sigma_1, \Sigma_2, \Sigma_3$

$\longrightarrow$ Approccio ASSIOMATICO:
1 - $\mathbb{P}(A) \geq 0$
2 - $\mathbb{P}(\Omega) = 1$
3 - $A \cap B \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

FINITA ADDITIVITÀ

- $0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\emptyset) = 0$
- $B \subset A \longrightarrow \mathbb{P}(B) \leq \mathbb{P}(A)$
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(B) = 1 \longrightarrow \mathbb{P}(B \cap A) = \mathbb{P}(A)$
- $\mathbb{P}(B \cup A) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$\mathbb{P}: F \longrightarrow \mathbb{R}\,[0,1]$

RISULTATO FINALE: $\mathcal{E} = (\Omega, F = \sigma(G), \mathbb{P}) \longrightarrow$ TERNA DI PROBABILITÀ

LEZIONE 5

calcolo delle probabilità ⟶ NUMERO ALEATORIO (non deterministico)

$\mathcal{E} = (\Omega, F = \sigma(G), P)$ ⟶ è una misura di probabilità

ASSIOMI DI KOLMOGOROV

$H_1: A \in F,\ P(A) \geq 0$

$H_2: P(\Omega) = 1$

$H_3: (A_m)_{m \in \mathbb{N}} \subseteq F,\ A_i \cap A_j = \emptyset \Rightarrow P(\bigcup A_m) = \sum P(A_m)$ ⟶ NUMERABILE ADDITIVITÀ

TH. 1

$\emptyset \in F$ (il vuoto è un evento)

DIM ⟶ vale Σ_1 ovvero Ω è un evento $\Rightarrow$ per Σ_2 allora $\Omega^c = \emptyset$ il quale è un evento $\square$

TH. 2

$(\forall m \in \mathbb{N},\ m > 1)\ (\forall A_1, A_2, \dots, A_m \in F)$
Allora, $\bigcup_{i=1}^{m} A_i \in F$

DIM ⟶ $\forall i \mid i = 1, 2, \dots, m,\ B_i = A_i \in F$
con $i > m,\ B_i = \emptyset$ ⟶ $(B_m)_{m \in \mathbb{N}}$

per Σ_3 allora si ha $\bigcup_{m \in \mathbb{N}} B_m \in F \Leftrightarrow \bigcup_{i=1}^{m} \in F$ ⟶ l'unione finita di eventi è un evento $\square$

TH. 3

$(\forall m \in \mathbb{N} \mid m > 1)\ (\forall A_1, A_2, \dots, A_m) \in F$
Allora, $\bigcap_{i=1}^{m} A_i \in F$

DIM

$\bigcap_{i=1}^{m} A_i = \left(\bigcup_{i=1}^{m} A_i^c\right)^c \in F$] RELAZIONE DI DE MORGAN $\square$

TH. 4

$\forall A \in F,\ (A^c)^c = A$ | RISULTATO BANALE

TH. 5

$\forall A, B \in F$, (i) $A \cup B = A \cup (B \cap A^c)$
(ii) $A = A \cap B \cup A \cap B^c$

DIM (ii)

$A = A \cap \Omega$

$A = A \cap (B \cup B^c)$] Ω in funzione di B

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

→ prop distributiva

DIM (i)

→ non contribuisce all'unione

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

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

TH. 6

$\mathbb{P}(\emptyset) = 0$

DIM

$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c)$

$1 = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ ← NON POSSIBILE PER TH. 6 in quanto si sta parlando di Ω insieme

→ Sia $m \in \mathbb{N}$, $A_m = \emptyset$

$\Rightarrow A_i \cap A_j$ con $i \neq j = \emptyset$

→ $\mathbb{P}\left(\bigcup_{m \in \mathbb{N}} A_m\right) \overset{P_3}{=} \sum_{m \in \mathbb{N}} \mathbb{P}(A_m) \Leftrightarrow \mathbb{P}(\emptyset) = \sum \mathbb{P}(\emptyset)$

$\mathbb{P}(\emptyset) = \mathbb{P}(\emptyset) \cdot \sum_{m \in \mathbb{N}} 1$

questa serie diverge e ciò non può verificarsi in quanto andrebbe contro assioma $\Sigma 1$

→ $\boxed{\mathbb{P}(\emptyset) = \sum_{m=1}^{\infty} = 0 = 0}$ $\square$

TH. 7

$\forall m \in \mathbb{N}$, $\forall A_1, A_2, \dots, A_m \in \mathcal{F}$ con $A_i \cap A_j = \emptyset$

Allora $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$

La prob l'unione finita di eventi a due a due disgiunti è uguale alla somma delle probabilità

DIM

Ponendo una successione ausiliaria $i = 1, 2, \dots, m$ tali che $B_i = A_i$ e $i > m$, $B_i = \emptyset$

$\mathbb{P}(\bigcup_{i\in\mathbb{N}} B_i) = \sum_{i\in\mathbb{N}} \mathbb{P}(B_i) = \sum_{i=1}^{m} \mathbb{P}(B_i)$

$= \mathbb{P}(\bigcup_{i=1}^{m} A_i) \longleftrightarrow = \sum_{i=1}^{n} \mathbb{P}(A_i)$ $\square$

=> la finita additività è una conseguenza degli assiomi

COROLLARIO TH7 (i)

$\forall A_1, A_2 \in F \mid A_1 \cap A_2 = \emptyset$

$\Rightarrow \mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$

COROLLARIO TH7 (ii)

$\forall A \in F$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ (COMPLEMENTO A UNO)

DIM

$1 = \mathbb{P}(\Omega) = \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) \Rightarrow 1 - \mathbb{P}(A) = \mathbb{P}(A^c)$ $\square$

COROLLARIO TH7 (iii)

$\forall A \in F$, $\mathbb{P}(A) \leq 1$

DIM

$\mathbb{P}(A) = \underbrace{1 - \underbrace{\mathbb{P}(A^c)}_{\geq 0 \text{ per } M_1}}_{\leq 1}$ $\square$

TH8 (Proprietà di monotonia)

$\forall A, B \in F \mid A \subseteq B$ → ogni volta si verifica A, si verifica pure B, non vale il viceversa
$\Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$

DIM

$B = A \cup (B \cap A^c)$ ← Finita add

$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \Rightarrow \mathbb{P}(B) \geq \mathbb{P}(A)$ $\square$

TH9

$\forall A, B \in F$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ ← formula di inclusione ed esclusione

DIM

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

$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap A^c) \Rightarrow \mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \cap A^c)$
$\Rightarrow \mathbb{P}(B \cap A^c) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$

$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ $\square$

LEZIONE 6 22.10

$\forall A, B \in F$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ ← Formula di unione ed esclusione

DIM:

$$\boxed{\begin{array}{l} A \cup B = A \cup (B \cap A^c) \\ A = A \cap B \cup A \cap B^c \\ B = B \cap A \cup B \cap A^c \end{array}} \Rightarrow P(A \cup B) = P(A) + P(B \cap A^c)$$

$$= P(A) + P(B) - P(A \cap B)$$

$P(B) = P(A \cap B) + P(A^c \cap B)$

$P(B \cap A^c) = P(B) - P(A \cap B)$

ESERCIZIO

1) Una scatola contiene una biglia rossa, una verde e una blu

1 - Descrivere lo spazio degli esiti dell'esperimento che consiste nell'estrarre una biglia, rimetterla nella scatola ed estrarne una seconda volta

2 - Ripeti l'esercizio senza la rimessa della prima biglia

1- $\Omega_1 = \{(R,R), (R,V), (R,B), (V,V), (V,B), (V,R), (B,B), (B,R), (B,V)\}$

$\hookrightarrow |\Omega| = 3^2 = 9$

Disposizioni con ripetizione lunghe 2

2- $\Omega_2 = \{(R,V), (R,B), (V,R), (V,B), (B,R), (B,V)\}$

$\hookrightarrow |\Omega_2| = 6$ ← Disposizioni semplici lunghe 2

2) Si tira tre volte una moneta. Qual è lo spazio degli esiti di questo esperimento casuale? Scrivi esplicitamente l'evento "si ottengono più teste che croci"

- $\Omega = \{(T,T,T), (T,T,C), (T,C,T), (C,T,T), (C,C,C), (C,C,T), (C,T,C), (T,C,C)\}$

Disposizioni lunghe 3 con $|\text{valori}| = 2$

- "Si ottengono più teste che croci"

$E = \{(T,T,T), (T,T,C), (T,C,T), (C,T,T)\}$

3) Siano $S = \{1,2,3,4,5,6,7\}$
$E = \{1,3,5,7\}$
$F = \{7,4,6\}$
$G = \{1,4\}$

Scrivere gli elementi dei seguenti eventi

a) $E \cap F = \{1,3,5,7\}$ NO
a) $E \cap F = \{7\}$
b) $E \cap G^c = \{3,5,7\}$
$G^c = \{2,3,5,6,7\}$

PROPRIETA' DISTRIBUTIVA

c) $E^c \cap (F \cup G) = E^c \cap F \cup E^c \cap G = \{4,6\} \cup \{4\} = \{4,6\}$
$E^c = \{2,4,6\}$
d) $E \cup (F \cap G) = E \cup \{4\} = \{1,3,4,5,7\}$
e) $(E \cap F^c) \cup G = \{1,3,5\} \cup G = \{1,3,4,5\}$
$F^c = \{1,2,3,5\}$
f) $(E \cap G) \cup (F \cap G) = \{1\} \cup \{4\} = \{1,4\} = G$

4) Si tirano 2 dadi. Sia E l'evento che la somma dei punteggi sia pari, F che il primo dado realizzi 1, e G che la somma sia 5. Si descrivano gli eventi

a) $E \cap F$; b) $E \cup F$; c) $F \cap G$; d) $E \cap F^c$; e) $E \cap F \cap G$

$E = \{\Sigma = 2k : k = 1,2,3,4,5,6\}$
$F = \{(1,x) \mid x = 1,2,3,4,5,6\}$
$G = \{\Sigma = 5\} = \{(1,4),(2,3),(3,2),(4,1)\}$

a) $E \cap F = \{(1,1),(1,3),(1,5)\}$
b) $E \cup F = E \cup \{(1,2),(1,4),(1,6)\}$
c) $F \cap G = \{(1,4)\}$
d) $E \cap F^c =$
e) $E \cap F \cap G = \emptyset$

11) Vale la proprietà della subadditività di P:

Se $E_1, E_2, \dots, E_m$ eventi qualsiasi,

$$P\left(\bigcup_{i=1}^{m} E_i\right) \leq \sum_{i=1}^{m} P(E_i)$$

- $m=2$ (BASE INDUTTIVA)

$\forall\ E_1, E_2 \in F$

$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
$\Rightarrow P(E_1 \cup E_2) \geq P(E_1) + P(E_2)$

18) Un gruppo di 5 bambini e 10 bambine è in fila in ordine casuale, nel senso che tutte le 15! possibili permutazioni si suppongono equiprobabili.

a) Qual è la probabilità che il quarto della fila sia un bambino M

PERMUTAZIONI SEMPLICI

E $(m/f, m/f, m/f, m, m/f, m/f, \dots, m/f)$

$P(E_1) = 14!/15! = \frac{14!}{15 \cdot 14!} = \frac{1}{15}$

$\downarrow$

$H = 11$

per prima

IL PROBLEMA DELLE CONCORDANZE

Sia n un numero intero, e si abbia a disposizione un mazzetto di n cartoncini.
Essi vengono numerati da 1 a n sul fronte inizialmente bianco. Sul dorso i cartoncini presentano la stessa immagine.

$\mathcal{E}$: MISCHIARE I CARTONCINI, APPOGGIARE SU UN TAVOLO IL MAZZETTO MOSTRANDONE IL DORSO, MOSTRARE IL FRONTE DEL CARTONCINO PIÙ ALTO E RIPORRE ACCANTO AL MAZZETTO IL CARTONCINO MOSTRANDONE IL FRONTE

$\rightarrow$ Si dice concordanza l'evento che si verifica quando si alza la carta i-esima e si legge $i = 1, 2, \dots, n$

C_i (concordanza)

1°. Determinare la probabilità di avere almeno una concordanza

2°. Determinare la probabilità di avere 0 concordanze

3°. Determinare la probabilità di avere 1 concordanza

oss. ci sono $n!$ possibili mischiate

$P(C_i) = (n-1)!/n!$

1) $P\left(\bigcup_{i=1}^{n} C_i\right) = \sum_{i=1}^{n} P(C_i) - \sum_{i,j=1}^{n} P(C_i \cap C_j) + \sum_{\substack{i,j,k=1 \\ i<j<k}}^{n} P(C_i \cap C_j \cap C_k) \dots$

$+ \dots + (-1)^{n-1} \cdot P(C_1 \cap \dots \cap C_n)$

$\forall i = 1, \dots, n \quad P(C_i) = (n-1)!/n!$

$= \frac{(n-1)!}{n!} \sum_{i=1}^{n} 1 - \frac{(n-2)!}{n!} \sum_{i,j=1}^{n} 1 \dots$

$P(C_1 \cap C_2) = (n-2)!/n! \Rightarrow i,j = 1, \dots n \wedge i < j \quad P(C_i \cap C_j) = (n-2)!/n!$

LEZIONE 7

$$\mathbb{P}\left(\bigcup_{i=1}^{n} e_i\right) = \frac{(n-1)!}{n!}\sum_{i=1}^{n} 1 - \frac{(n-2)!}{n!}\sum_{i=1}^{n}\sum_{j=i+1}^{n} 1 + \frac{(n-3)!}{n!}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sum_{k=j+1}^{n} 1 + \dots + (-1)^{n-1}\frac{1}{n!}$$

$$= \frac{(n-1)!}{n!}\binom{n}{1} - \frac{(n-2)!}{n!}\binom{n}{2} + \frac{(n-3)!}{n!}\binom{n}{3} + \dots + (-1)^{n-1}\frac{1}{n!}$$

$$= \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} + \dots + (-1)^{n-1}\frac{1}{n!}$$

2- Probabilità di osservare zero concordanze nelle n chiamate

Si indichi con $E_{0,n}$ l'evento che si presenta quando in una esecuzione dell'esperimento non si osserva alcuna concordanza

$\Rightarrow E_{0,n} = \left(\bigcup_{i=1}^{n} e_i\right)^c \Rightarrow \mathbb{P}(E_{0,n}) = 1 - \mathbb{P}\left(\bigcup_{i=1}^{n} e_i\right)$

$\Downarrow$

$$\mathbb{P}(E_{0,n}) = \sum_{i=0}^{n} (-1)^i \frac{1}{i!}$$

$\mathbb{P}(E_{0,n}) =$ n° delle permutazioni con 0 concordanze $/ n!$

3- Probabilità di osservare solo una concordanza nelle n chiamate

$F_{1,1,n}$ l'evento che si presenta quando si ottiene nell'esp. esattamente una concordanza e questa avviene alla prima chiamata

$N_{1,1,n}$ il numero delle mescolate che fanno verificare solo la concordanza nella prima chiamata e zero concordanze nelle successive

$\rightarrow N_{1,1,n} = N_{0,n-1}$

$$\mathbb{P}(F_{1,1,n}) = N_{1,1,n}/n! = N_{0,n-1}/n! = (n-1)!\,\mathbb{P}(E_{0,n-1})/n!$$

$$= \frac{1}{n}\mathbb{P}(E_{0,n-1}) \quad \square$$

- Assiomi di Kolmogorov $\longrightarrow$ sistema assiomatico nato come base di partenza per il calcolo delle probabilità

DEF. LAPLACE

$\forall A, B \in F$: $\frac{n(A\cap B)}{n}$ num. casi favorevoli / numero casi

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

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

$P(B) > 0$

$$P(A|B) = n\,A\cap B / n_B$$

probabilità di A una volta che l'evento B si è verificato

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

probabilità di un evento condizionata da un altro evento

evento condizionante

$\boxed{P(A\cap B) = P(B)\cdot P(A|B)}$ LEGGE PROBABILITÀ COMPOSTE

ESEMPIO

$\Omega = \{(i,j) : i,j = 1,\dots,6\}$

$$P(Z_2 = 8 \mid Z_1 = 3) = \frac{P(Z_2 = 8, Z_1 = 3)}{P(Z_1 = 3)} = \frac{P(Z_2 = 8 \cap Z_1 = 3)}{P(Z_1 = 3)}$$

$\{Z_2 = 8\} = \{(2,6),(3,5),(4,4),(5,3),(6,2)\}$

$\{Z_1 = 3\} = \{(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)\}$

$P(Z_2 = 8, Z_1 = 3) = \{(3,5)\} \Rightarrow P\{(3,5)\} = 1/36$

$$P(Z_2 = 8 \mid Z_1 = 3) = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{36}\cdot 6 = \frac{1}{6}$$

Con la scrittura $P(A|B)$ non si rappresenta la probabilità dell'evento "A|B", ma solo un modo per indicare il rapporto che compare tra due σ-algebra, A e B ∈ F

ESEMPIO 3.6.2

La organizzazione per cui lavora il signor Jones organizza una cena tra uomini per i dipendenti e i loro figli.
Sono invitati i dipendenti padri di figli maschi, assieme al maggiore fra i loro figli maschi.
Jones ha due figli, ed è invitato alla cena.
Qual è la probabilità condizionata che entrambi i suoi figli siano maschi?

$S = \{(m,m),(m,f),(f,m),(f,f)\}$ ⟶ Tutti gli esiti sono equiprobabili

A: "almeno uno dei due è maschio"
B: "entrambi sono maschi"

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

- $\mathbb{P}(A) = \{(m,f),(f,m),(m,m)\}$
  $\mathbb{P}(B) = \{(m,m)\}$

$$= \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{4} \cdot \frac{4}{3} = \frac{1}{3}$$

ESEMPIO 3.6.3 (Per legge probabilità composta)

Il Signor Perez è convinto che vi sia il 30% di probabilità che la sua azienda apra un nuovo ufficio a Phoenix.
Nel caso si verifica, egli stima di avere 60% di probabilità di assumere il ruolo di manager della nuova filiale.
Che Probabilità vi è che egli diventi il manager della nuova sede di Phoenix?

- U, M e F

U := "viene aperto un nuovo ufficio a Phoenix"
M := "Perez viene promosso a Phoenix come manager"

$$\mathbb{P}(U \cap M) = \mathbb{P}(U) \cdot \mathbb{P}(M|U) = 0{,}3 \times 0{,}6 = 0{,}18$$

ESEMPIO 3.7.1

Una società di assicurazione ritiene che la popolazione possa essere divisa in due categorie: - persone inclini a provocare incidenti.
- persone non incline.
Una persona incline a incidenti ha un incidente in un anno con probabilità 0,4, mentre 0,2 per persone non incline.
Assumendo che il 30% della popolazione sia incline agli incidenti, quanto vale la probabilità che un nuovo assicurato abbia un incidente entro un anno della stipula del contratto assicurativo?

H: persone inclini a provocare incidenti.
A_1: prob che un nuovo assicurato abbia un incidente

$$\mathbb{P}(A_1) = \mathbb{P}(A_1 \cap H) + \mathbb{P}(A_1 \cap H^c) = \mathbb{P}(A_1|H)\mathbb{P}(H) + \mathbb{P}(A_1|H^c)\mathbb{P}(H^c)$$

Se si applica la legge delle probabilità composte si ottiene la FORMULA DELLE ALTERNATIVE

$$= 0{,}4 \cdot 0{,}3 + 0{,}2 \cdot 0{,}7 = 0{,}26 = 26\%$$

$$\mathbb{P}(H|A_1) = \frac{\mathbb{P}(H \cap A_1)}{\mathbb{P}(A_1)} = \boxed{\frac{\mathbb{P}(A_1) \cdot \mathbb{P}(H|A_1)}{\mathbb{P}(A_1)}}$$

FORMULA DI BAYES

$$P(H|A_1) = \frac{P(H)\cdot P(A_1|H)}{P(A_1)} = \boxed{P(H)\cdot P(A_1|H) \Big/ P(H)\cdot P(A_1|H) + P(H^c)\cdot P(A_1|H^c)}$$

TEOREMA DI BAYES

Per il nostro esempio

$$0{,}3 \cdot 0{,}4 \Big/ 0{,}26 = \frac{0{,}12}{0{,}26} = 0{,}465 = 46{,}15$$

LEZIONE 8

13) Dimostra le due equazioni seguenti

a) $P(E \cap F^c) = P(E) - P(E \cap F)$

DIM (formula di decomposizione)

$E = (E \cap F) \cup (E \cap F^c) \Rightarrow P(E) = P(E \cap F) + P(E \cap F^c)$

$\Rightarrow P(E \cap F^c) = P(E) - P(E \cap F)$

b) $P(E^c \cap F^c) = 1 - P(E) - P(F) + P(E \cap F)$

DIM

$E^c \cap F^c = (E \cup F)^c \Rightarrow P(E^c \cap F^c) = 1 - P(E \cup F) =$

$= 1 - P(E) - P(F) + P(E \cup F)$

14) Dimostra che la probabilità che si realizzi uno e uno solo degli eventi E, F è pari a

$P(E) + P(F) - 2P(E \cap F)$

$E \cup F = (E \cup F^c) \cup (E^c \cup F) \cup (E \cap F) \Rightarrow$

$\Rightarrow P(E \cup F) = P(E) + P(F) - P(E \cap F)$ (somma, esclusione)

7) Semplifica, dove possibile, le espressioni che seguono

- $E \cup E^c = \Omega$
- $E \cap E^c = \emptyset$
- $(E \cup F) \cap (E \cup F^c) = ((E \cup F) \cap E) \cup ((E \cup F) \cap F^c)$

$(E \cap E) \cup (F \cap E) \cup (E \cap F^c) \cup (F \cap F^c)$

$E \cup (F \cap E) \cup (E \cap F^c) \cup \emptyset$

$= E \cup (E \cap F) \cup (E \cap F^c)$

$= E \cup E = E$

ESEMPIO 3.7.3

In una prova a risposta multipla, nel rispondere ad una domanda uno studente può conoscere la risposta, oppure provare a indovinare. Sia p la probabilità che conosca la risposta e $1-p$ la probabilità che tiri a indovinare.
Si assume che, se prova ad indovinare, risponde correttamente con probabilità $1/m$, dove m è il numero di alternative nella scelta multipla.
Qual è la probabilità condizionata che egli conoscesse la risposta a una domanda alla quale ha risposto correttamente?

$C, R \in \mathcal{F}$

C = "sceglie la risposta giusta"
R = "conosce la risposta giusta"

$\longrightarrow \mathbb{P}(R|C) = \frac{\mathbb{P}(R \cap C)}{\mathbb{P}(C)} \Rightarrow \mathbb{P}(C) > 0$

$= \mathbb{P}(R|C) = \frac{\mathbb{P}(R|C)\,\mathbb{P}(C)}{\mathbb{P}(R)}$

$\mathbb{P}(R) = p$, $\mathbb{P}(R^c) = 1-p$ con $p \in (0,1)$

$\mathbb{P}(C|R) = 1$, $\mathbb{P}(C|R^c) = 1/m$ con $m \in \mathbb{N}$

$\mathbb{P}(C) = \mathbb{P}(C|R)\,\mathbb{P}(R) = 1 \times p = p + \mathbb{P}(C|R^c)\,\mathbb{P}(R^c) = p + (1/m)(1-p)$

$$\mathbb{P}(R|C) = \frac{p}{p + (1/m)(1-p)} = \frac{mp}{1 + (m-1)p}$$

ESEMPIO 3.7.4

Una particolare analisi del sangue è efficace al 99% nell'individuare una certa malattia quando essa è presente.
Si possono, però, verificare dei falsi positivi con probabilità 1%.
Se l'incidenza di questo male sulla popolazione è dello 0,5%, qual è la probabilità che un soggetto sia malato condizionata dal fatto che le analisi abbiano dato esito positivo?

$M, E \in \mathcal{F}$

M = "il soggetto è malato"
E = "il risultato delle analisi è positivo"

$\mathbb{P}(M|E) = \frac{\mathbb{P}(M \cap E)}{\mathbb{P}(E)}$ con $\mathbb{P}(E) > 0$

Verosimiglianza

$=$ formula delle alternative $\longrightarrow$ $= \frac{\mathbb{P}(E|M)\,\mathbb{P}(M)}{\mathbb{P}(E|M)\,\mathbb{P}(M) + \mathbb{P}(E|M^c)\,\mathbb{P}(M^c)}$

$= \frac{0{,}99 \times 0{,}005}{0{,}99 \times 0{,}005 + 0{,}01 \times 0{,}995} = 0{,}3322 \Rightarrow$ il 33% delle persone che risultano positive alle analisi sono realmente affette dalla malattia

ESEMPIO 3.4.5

Ad un certo stadio delle indagini su un crimine, l'investigatore capo è convinto al 60% della colpevolezza di un certo sospetto. Si suppone che si scopra un nuovo indizio che mostra che il colpevole deve possedere una certa caratteristica distintiva; inoltre anche il sospettato la possiede. Se tale particolarità interessa il 20% della popolazione, quanto sicuro deve essere l'investigatore della colpevolezza del sospettato?

$G, C \in F$

G: "il sospetto è colpevole"
C: "il sospetto possiede il tratto distintivo del colpevole"

$$P(G|C) = \frac{P(G\cap C)}{P(C)} = \boxed{\frac{P(C|G)\cdot P(G)}{P(C|G)\cdot P(G) + P(C|G^c)\cdot P(G^c)}}$$

evento certo

$$= \frac{1 \cdot 0{,}6}{1\cdot 0{,}6 + 0{,}2\cdot 0{,}4} = \frac{0{,}6}{0{,}68} = 0{,}88 = 88\%$$

Si è stabilito che la probabilità che il sospetto abbia la caratteristica nel caso in cui non è colpevole sia quella generale della popolazione $\longrightarrow P(G|C) = 88\%$

FORMULA DI ALTERNATIVE DI UNA SUCCESSIONE

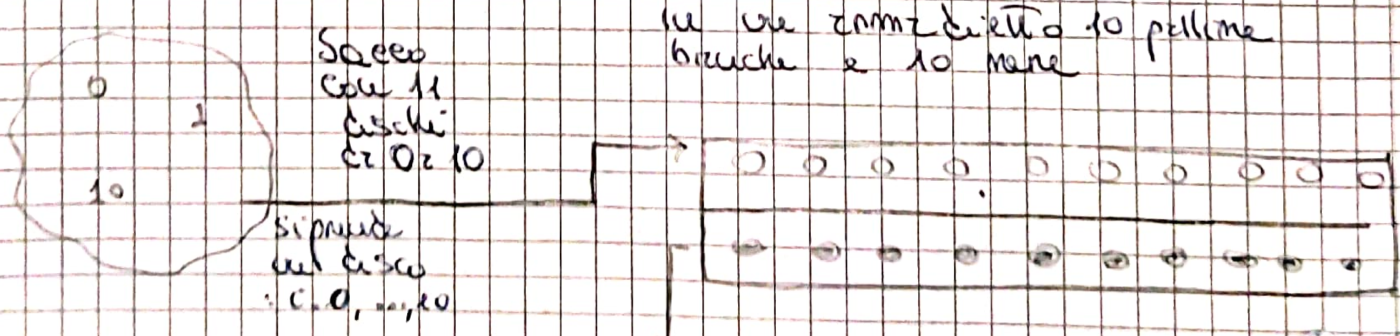
$\forall\, (H_n)_{n\in \mathbb{N}} \subseteq F$

1) $\forall n \in \mathbb{N},\ P(H_n) > 0$
2) $i \neq j \in \mathbb{N},\ H_i \cap H_j = \emptyset$
3) $\bigcup H_n = \Omega$

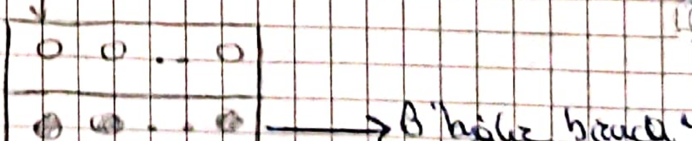
Con queste premesse, questa successione costituisce un sistema completo di alternative

$$P(A) = P(A \cap \Omega) = (A \cap \bigcup H_n)$$
$$= P[\bigcup (A\cap H_n)] = \sum_{n\in\mathbb{N}} P(A \cap H_n) = \sum_{n\in\mathbb{N}} [P(H_n)\cdot P(A|H_n)]$$

ESEMPIO



Dopo aver visto i sul disco si mette in un'urna i palline bianche e n-i palline nere



H_i "esce il disco i"

$$\mathbb{P}(B) = \sum_{i=0}^{10} \mathbb{P}(B|H_i) \cdot \mathbb{P}(H_i)$$

$$= \sum_{i=0}^{10} \frac{i}{10} \cdot \frac{1}{11} = \frac{1}{110} \cdot \sum_{i=1}^{100} i = \frac{110/2}{110} = \frac{110}{2} \cdot \frac{1}{110} = \frac{1}{2}$$

LEZIONE 9

RIEPILOGO

formula di Bayes per due eventi:

$\forall A, E \in F$

$$\mathbb{P}(A) = \mathbb{P}(A|E) \cdot \mathbb{P}(E) + \mathbb{P}(A|E^c) \cdot \mathbb{P}(E^c) \quad \leftarrow \text{formula delle alternative}$$

$$\mathbb{P}(E|A) = \frac{\mathbb{P}(A|E) \cdot \mathbb{P}(E)}{\mathbb{P}(A|E) \cdot \mathbb{P}(E) + \mathbb{P}(A|E^c) \cdot \mathbb{P}(E^c)}$$

- $\forall m \in \mathbb{N}$, $m \geq 2$

Sia $(H_m)_{m \in \mathbb{N}}$, successione finita di eventi.

→ è un sistema completo di alternative

si ha che: $\mathbb{P}(E) > 0$

$\mathbb{P}(A) = \sum_m \mathbb{P}(A|H_m) \cdot \mathbb{P}(H_m)$

↓

dato che A è verificato, si vuole aggiornare la probabilità dell'i-esima alternativa

→ $\mathbb{P}(H_i|A) = \dfrac{\mathbb{P}(A|H_i) \cdot \mathbb{P}(H_i)}{\sum_{m \in \mathbb{N}} \mathbb{P}(A|H_m) \cdot \mathbb{P}(H_m)}$

↓

formula di Bayes nel caso generico

ESEMPIO 3.7.7

Un aereo è scomparso, e si suppone che possa essere caduto in una qualsiasi di tre regioni, con uguale probabilità.
Per $i = 1,2,3$, sia $1-\alpha_i$ la probabilità di ritrovare un velivolo che cada nella i-esima regione.
Qual è la probabilità che l'aereo si trovi in ciascuna delle tre regioni se una ricerca della regione 1 ha dato esito negativo?

$\forall i = 1,2,3$, si denota con R_i l'evento "il velivolo si trova nella regione i-esima"

Sia E l'evento "la ricerca nella regione 1 non ha successo"

10.00

Dalla formula di Bayes si ottiene

$$\mathbb{P}(R_i|E) = \frac{\mathbb{P}(E|R_i)\cdot\mathbb{P}(R_i)}{\sum_{i=1}\mathbb{P}(E|R_i)\,\mathbb{P}(R_i)}$$

$$= \frac{\alpha_1/3}{\alpha_1/3+1/3+1/3} = \frac{\alpha_1}{\alpha_1+2}$$

$\mathbb{P}(H_1) = \mathbb{P}(H_2) = \mathbb{P}(H_3) = \frac{1}{3}$

$\downarrow$

le alternative sono equiprobabili

ESEMPIO

E_1: Giovanni
E_2: Enrico

O: "E_2 è onesto", $\mathbb{P}(O) = 0,9$
$B = O^c$ $\quad \mathbb{P}(B) = 0,10 = \mathbb{P}(O^c)$
V_1: "Enrico vince l'eposituo"

$\mathbb{P}(V_1|B) = 1$
$\mathbb{P}(V_1|O) = 1/2$

$$\mathbb{P}(B|V_1) = \frac{\mathbb{P}(V_1|B)\,\mathbb{P}(B)}{\mathbb{P}(V_1|B)\,\mathbb{P}(B)+\mathbb{P}(V_1|O)\cdot\mathbb{P}(O)} = \frac{\frac{1}{10}}{\frac{1}{10}+\frac{9}{10}\left(\frac{1}{2}\right)}$$

$$= \frac{1}{\frac{11}{2}} = 1\cdot\frac{2}{11} = \frac{2}{11} \cong 0,18$$

V_2: "E_2 vince il primo e il secondo eposituo"

$\mathbb{P}(V_2|B) = 1$
$\mathbb{P}(V_2|O) = (1/2)^2 = 1/4$ *

$$\mathbb{P}(B|V_2) = \frac{\frac{1}{10}}{\frac{1}{10}+\frac{1}{4}\cdot\frac{9}{10}} = \frac{1}{\frac{4}{13}} = 1\cdot\frac{13}{4} = \frac{13}{4} \cong 0,3$$

V_6: "E_2 ne vince 6"

$$\mathbb{P}(B|V_6) = \frac{\frac{1}{10}}{\frac{1}{10}+\frac{1}{2^6}\cdot\frac{9}{10}} = \frac{1}{\frac{1}{64}+9} = \frac{64}{73} \cong 0,9$$

Si può notare che la probabilità che è un baro dato che vince AUMENTA

$\downarrow$

INFERENZA DIRETTA

oppure $\longrightarrow$ INFERENZA ITERATIVA $\longrightarrow$ Calcolando le probabilità a priori, trasformando quelle a posteriori in quelle a priori dopo il primo evento

* Sia $\mathbb{P}(B) > 0$, $A \in \mathcal{F}$

$$\mathbb{P}(A|B) \begin{cases} > \mathbb{P}(A) \\ = \mathbb{P}(A) \\ < \mathbb{P}(A) \end{cases}$$

1) se $\mathbb{P}(A|B) = \mathbb{P}(A)$ allora A è indipendente da B

$\Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

Se l'evento A è indipendente dall'evento B allora la probabilità risultante $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

quando A è indipendente da B allora per SIMMETRIA, B è indipendente da A

3) $\mathbb{P}(A) > 0$, $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

DEF.

A e B sono INDIPENDENTI $\Leftrightarrow \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

PROPOSIZIONE 3.P.1

Se A e B sono eventi indipendenti, lo sono anche A e B^c

DIM

Bisogna dimostrare che $\mathbb{P}(A \cap B^c) = \mathbb{P}(A)\mathbb{P}(B^c)$

$A = A \cap B \cup A \cap B^c$

$$\begin{aligned} \mathbb{P}(A \cap B^c) &= \mathbb{P}(A) - \mathbb{P}(A \cap B) \\ &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)[1 - \mathbb{P}(B)] = \mathbb{P}(A)\mathbb{P}(B^c) \end{aligned}$$

$\square$

ESEMPIO 3.P.3

Si tirano due dadi non truccati. Sia E_7 l'evento "la somma dei due punteggi è pari a 7", sia F l'evento "il primo dado totalizza un 4" e sia G l'evento "il secondo dado totalizza un 3"

Si può dimostrare che E_7 è indipendente da F come pure da G

$E_7 = \{Z_2 = 7\}$
$F = \{Z^1 = 3\}$
$G = \{Z^2 = 4\}$

$\mathbb{P}(\{Z_2 = 7\} \mid \{Z^1 = 3\} \cap \{Z^2 = 4\}) = 1$
$\mathbb{P}(Z_2 = 7 \mid \{Z^1 = 3\}) = 1/6$
$\mathbb{P}(Z_2 = 7 \mid \{Z^1 = 4\}) = 1/6$

$\mathbb{P}(Z_2 = 7) = \frac{6}{36} = 1/6$

DEFINIZIONE 3.8.2

Siano $A, B, C \in F$

A, B, C sono indipendenti tra loro $\iff$ $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$
$\mathbb{P}(A \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(C)$
$\mathbb{P}(B \cap C) = \mathbb{P}(B) \cdot \mathbb{P}(C)$
$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \cdot \mathbb{P}(B) \cdot \mathbb{P}(C)$

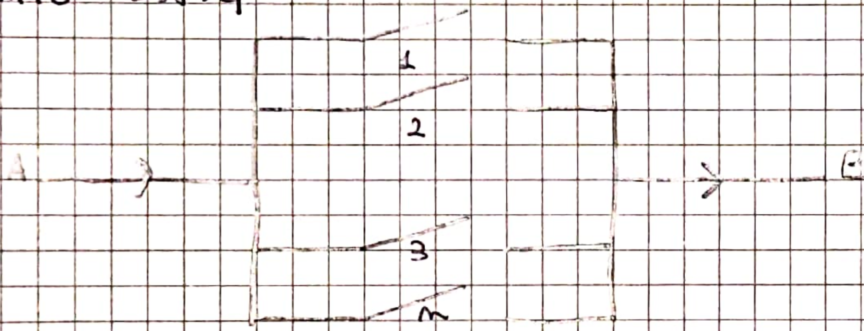
Si noti che se tre eventi A, B, C sono indipendenti, allora ciascuno di essi è indipendente da qualunque evento si possa costruire con gli altri due

Siano A, B, C indipendenti tra loro

$\mathbb{P}([A \cap (B \cup C)]) = \mathbb{P}([A \cap B] \cup [A \cap C])$
$= \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}(A \cap B \cap C)$ ← UNIONE ESCLUSIONE
$= \mathbb{P}(A) \cdot \mathbb{P}(B) + \mathbb{P}(A) \cdot \mathbb{P}(C) - \mathbb{P}(A) \cdot \mathbb{P}(B \cap C)$
$= \mathbb{P}(A) \cdot [\mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C)]$
$= \mathbb{P}(A) \cdot \mathbb{P}(B \cup C)$

LEZIONE 10

ESEMPIO 3.8.4



Un sistema in parallelo, funziona se la corrente passa da A a B

Un sistema composto da n componenti distinti si dice in parallelo se funziona fino a che almeno uno dei componenti funziona
Sia dato un sistema per il quale, per $i = 1, 2, \dots, n$ il componente i-esimo funziona indipendentemente da tutti gli altri, con probabilità p_i. Qual è la probabilità che l'intero sistema funzioni?

$A_i \in F$
A_i = "il componente i funziona"

$\mathbb{P}(\text{il sistema funziona}) = 1 - \mathbb{P}(\text{il sistema NON funziona})$
$= 1 - \mathbb{P}(\text{nessuna componente funziona})$
$\mathbb{P}(\cup E_i) = \mathbb{P}[(\cap E_i^c)^c]$ ← $= 1 - \mathbb{P}(A_1^c \cap A_2^c \cap \dots \cap A_n^c)$
$= 1 - \prod_{i=1}^{n} (1 - p_i)$

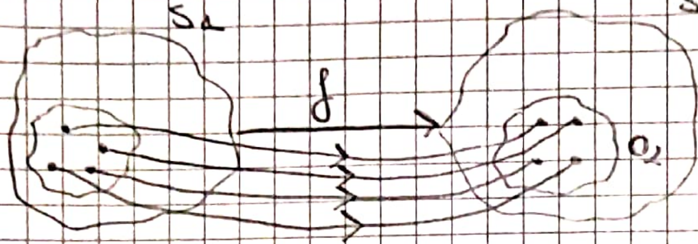
- La continuità (funzioni) riguarda la topologia
- La misurabilità riguarda la famiglia dei misurabili.

(S_1, τ_1) , (S_2, τ_2)

$f : S_1 \longrightarrow S_2$

f è continua $\iff \forall\, O_2 \in \tau_2,\ \{f(s_1) \in O_2\} \in \tau_1$

elementi di S_1 la cui immagine sta in O_2



(S_1, F_1) , (S_2, F_2)

$F_1, F_2 = \sigma$-algebra

$\forall\, A_1, A_m \in F_1 \wedge \forall\, A_1, A_m \in F_2 \longrightarrow$ elementi MISURABILI

$f : S_1 \longrightarrow S_2$

$s_1 \rightsquigarrow s_2 = f(s_1)$

f è misurabile $\iff \forall\, B_2 \in F_2,\ \{f(s_1) \in B_2\} \in F_1$

elementi di S_1 la cui immagine sta in B_2

ESEMPIO

(Ω, F) $(\mathbb{R}, \mathcal{B})$

$X : \Omega \xrightarrow[F\text{-mis}]{} \mathbb{R}$

la controimmagine di un Boreliano è un sottoinsieme dello spazio campione

$\mathcal{E} = (\Omega, F, \mathbb{P})$ con la misurabilità si ottiene $\mathbb{P}_X \Rightarrow (\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$

$B \in \mathcal{B},\ \mathbb{P}_X(B) = \mathbb{P}(\{X \in B\}) = \mathbb{P}(X \in B)$

$B \in \mathcal{B},\ \mathbb{P}_X(B) = \mathbb{P}(X \in B) \geq 0$

$\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}) = \mathbb{P}(\Omega) = 1$

$(B_n)_{n \in \mathbb{N}} \subseteq \mathcal{B},\ B_i \cap B_j = \emptyset,\ \mathbb{P}_X(\cup B_n) = \mathbb{P}(X \in \cup B_n)$

$= \mathbb{P}(\cup \{X \in B_n\}) = \sum_{n \in \mathbb{N}} \mathbb{P}(X \in B_n) = \sum_{n \in \mathbb{N}} \mathbb{P}_X(B_n)$

$\mathcal{E}$: "lancio una moneta truccata un numero indefinito di volte"

$\Omega = \{T, C\}^{\infty}$, $\omega \in \Omega = (\omega_1, \omega_2, \dots, \omega_m) : \omega_m = \begin{cases} T \\ C \end{cases}$

$F = \sigma(G)$, $G = (T_m)_{m \in \mathbb{N}}$

$T_1 = \{(T, \omega_1, \omega_2, \omega_3, \dots, \omega_m, \dots)\}$
$T_2 = \{(\omega_1, T, \omega_3, \dots, \omega_m, \dots)\}$
$T_m = \{(\omega_1, \omega_2, \omega_3, \dots, T, \dots)\}$

$X_1 : \Omega \longrightarrow \mathbb{R}$

$\omega \rightsquigarrow X_1(\omega) = \begin{cases} 1 & \text{se } \omega_1 = T \\ 0 & \text{se } \omega_1 = C \end{cases}$

$\forall B \in \mathcal{B}$

- se B non contiene 0 e non contiene 1 $\{X_1 \in B\} = \emptyset \in F$
- se B contiene 1 e non contiene 0 $\{X_1 \in B\} = T \in F$
- se B contiene 0 e non contiene 1 $\{X_1 \in B\} = T^c \in F$
- se B contiene 0 e contiene 1 $\{X_1 \in B\} = \Omega \in F$

$S_{X_1} = \{0, 1\}$, supponendo che p è la probabilità che il è uscita testa, $p \in]0,1[$

$\mathbb{P}_{X_1}(\{1\}) = \mathbb{P}(X_1 = 1) = p$
$\mathbb{P}_{X_1}(\{0\}) = \mathbb{P}(X_1 = 0) = 1 - p = q$

DEF 5.1.1. "Si supponga che venga realizzata una prova il cui esito può essere un "successo" o un "fallimento". Si definisce la variabile aleatoria X in modo che sia $X = 1$ nel primo caso e $X = 0$ nel secondo, la funzione di MASSA di probabilità è data da

$\mathbb{P}(X=0) = 1 - p$
$\mathbb{P}(X=1) = p$ $\longrightarrow 0 \leq p \leq 1$

in questo caso si dice che la variabile aleatoria X è una var. di Bernoulli o bernoulliana, per un'opportuna scelta del parametro p."

$X_m \sim B(1, p)$

In altri termini, una variabile aleatoria di Bernoulli può assumere solo i valori 0 e 1 con val. atteso:

$E(X) := 1 \cdot \mathbb{P}(X=1) + 0 \cdot \mathbb{P}(X=0) = p$

LEZIONE 11

$(\{T,C\}^{\infty}, \mathcal{F} = \sigma[(T_m)_{m\in\mathbb{N}}], \mathbb{P})$

$p := \mathbb{P}(\text{"}T\text{"}) \in (0,1)$

$m \in \mathbb{N}$, $X_m := \begin{cases} 1, & T_m = T \\ 0 & \text{altrimenti} \end{cases}$ $\Rightarrow$ $X_m : \{T,C\}^{\infty} \longrightarrow \mathbb{R}$

$\omega \longrightarrow X_m(\omega)$

[Successione numerabile di numeri aleatori]

$\omega_1 = (C, T, T, T, C, T, \dots) \longrightarrow X_3 = 1 \longrightarrow X_6(\omega_1) = 1$

$\omega_2 = (C, C, T, C, C, C, \dots) \longrightarrow X_3(\omega_2) = 1 \longrightarrow X_6(\omega_2) = 0$

$\omega_3 = (C, T, T, T, C, T, \dots) \longrightarrow X_3(\omega_3) = 1 \rightarrow X_6(\omega_3) = 1$

$\omega_4 = (T, C, C, C, C, T, \dots) \longrightarrow X_3(\omega_4) = 0 \rightarrow X_6(\omega_4) = 1$

X_m variabile aleatoria $\longrightarrow$ $X_m(\omega_m) = m$ REALIZZAZIONE

$\longrightarrow X_m : \{T,C\}^{\infty} \longrightarrow \mathbb{R}$

$\downarrow$

$\mathcal{F}$-misurabile

$S_2 := X_1 + X_2$ $= \mathcal{S}_{S_2} = \{0,1,2\}$ "spettro"

somma delle prime due prove

Ciò mostra che la somma di due funzioni misurabili è una FUNZIONE MISURABILE

$B \in \mathcal{B}$, $\{S_2 \in B\} =$

$\mathcal{S}_{S_2} \cap B = \emptyset \Rightarrow \emptyset$

$\mathcal{S}_{S_2} \cap B = \{0\} \Rightarrow C_1 \cap C_2$

$\mathcal{S}_{S_2} \cap B = \{1\}$, $\left| \begin{array}{l} T_1 \cap C_2 \\ \cup \\ C_1 \cap T_2 \end{array} \right.$

$\mathcal{S}_{S_2} \cap B = \{2\} \Rightarrow T_1 \cap T_2$

$\mathcal{S}_{S_2} \cap B = \{0,1\} \Rightarrow \left[\begin{array}{l} C_1 \cap C_2 \\ \cup \\ T_1 \cap C_2 \\ \cup \\ C_1 \cap T_2 \end{array} \right.$

$\mathcal{S}_{S_2} \cap B = \{0,2\} \Rightarrow \left[\begin{array}{l} C_1 \cap C_2 \\ \cup \\ T_1 \cap T_2 \end{array} \right.$

$\mathcal{S}_{S_2} \cap B = \{1,2\} \Rightarrow \left[\begin{array}{l} T_1 \cap C_2 \\ \cup \\ C_1 \cap T_2 \\ \cup \\ T_1 \cap T_2 \end{array} \right.$

$\mathcal{S}_{S_2} \cap B = \{0,1,2\} \Rightarrow \left[\begin{array}{l} T_1 \cap C_2 \\ \cup \\ C_1 \cap T_2 \\ \cup \\ T_1 \cap T_2 \\ \cup \\ C_1 \cap C_2 \end{array} \right.$

$\longrightarrow$

- $p_0 := \mathbb{P}(S_{S_2} = 0) = \mathbb{P}(C_1 \cap C_2) = \mathbb{P}(C_1) \cdot \mathbb{P}(C_2) = (1-p)^2$
- $p_1 := \mathbb{P}(S_{S_2} = 1) = \mathbb{P}(C_1 \cap T_2 \cup T_1 \cap C_2) = \mathbb{P}(C_1 \cap T_2) + \mathbb{P}(T_1 \cap C_2)$
  $= \mathbb{P}(C_1) \cdot \mathbb{P}(T_2) + \mathbb{P}(T_1) \cdot \mathbb{P}(C_2) = (1-p)p + p(1-p)$
  $= 2p(1-p)$
- $p_2 := \mathbb{P}(S_2 = 2) = \mathbb{P}(T_1 \cap T_2) = \mathbb{P}(T_1) \cdot \mathbb{P}(T_2) = p \cdot p = p^2$

$S_{S_2} = \{0, 1, 2\} \longrightarrow ((1-p)^2, 2p(1-p), p^2)$

DISTRIBUZIONE DI PROBABILITA' di S_2

- $F_X(x)$
- $\int_A x\,dx = f_x$
- $F_X(x) = \mathbb{P}(X = x)$

è un modello che associa una probabilità a ogni modalità osservabile di una variabile aleatoria

questa misura può essere sia DISCRETA che CONTINUA

MEDIA di numeri aleatori ← la somma dei valori dello spettro moltiplicati per le rispettive probabilità

$\mathbb{E}(X_1) = 0 \cdot (1-p) + 1 \cdot p = p$

$\mathbb{E}(S_2) = 0(1-p)^2 + 1(2p(1-p)) + 2p^2$

$= 2p \cancel{-2p^2} \cancel{+2p^2} = 2p$

la media di variabili di bernoulli coincide con il prodotto dei parametri

Sia $S_3 = X_1 + X_2 + X_3$, con lo spettro $= S_{S_3} = \{0, 1, 2, 3\}$ $\quad q = 1-p$

$q^3, \quad 3 \cdot q^2 p, \quad 3 \cdot q p^2, \quad p^3$

$\mathbb{E}(S_3) = 0q^3 + 1(3q^2p) + 2(3qp^2) + 3p^3$

$= 3q^2p + 6qp^2 + 3p^3$

$= 3p(q^2 + 2qp + p^2) = 3p\underbrace{(q+p)^2}_{1} = 3p$

FORMULA 5.1.3

$S_m = X_1 + X_2 + \dots + X_m$

$S_{S_m} = \{0, 1, 2, \dots, m\}$

$k \in S_{S_m}, \quad p_k := \mathbb{P}(S_m = k) = \binom{m}{k} \cdot p^k \cdot q^{m-k}$

distribuzione di probabilità binomiale con parametri m e p

$\mathbb{P}(S_n \in \mathbb{R}) = 1$

$$\sum_{k=0}^{n} \left(\binom{n}{k} p^k q^{n-k} \right) = (p+q)^n = 1^n = 1$$

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$$

$$\mathbb{E}(S_n) = \sum_{k=0}^{n} \left(k \cdot \binom{n}{k} \cdot p^k \cdot q^{n-k} \right)$$

$$= \sum_{k=1}^{n} \left(k \cdot \binom{n}{k} \cdot p^k \cdot q^{n-k} \right)$$

$$= \sum_{k=1}^{n} \left(k \cdot \frac{n!}{k!\,(n-k)!} \cdot p^k \cdot q^{n-k} \right)$$

$$= \sum_{k=1}^{n} \left(\frac{n!}{(k-1)! \cdot (n-k)!} \cdot p^k \cdot q^{n-k} \right)$$

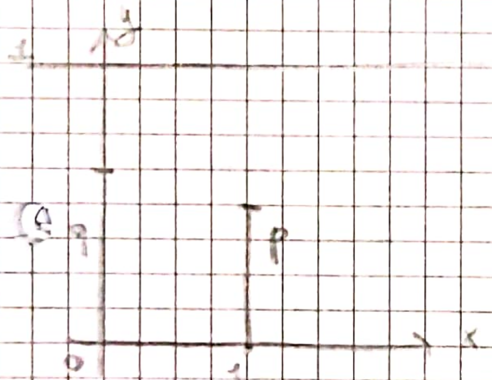
$$= \sum_{k=1}^{n} n \left(\frac{(n-1)!}{(k-1)! \cdot (n-k)!} \cdot p^k \cdot q^{n-k} \right)$$

$$= n \cdot \sum_{k=1}^{n} \left(\frac{(n-1)!}{(k-1)! \cdot (n-k)!} \cdot p^k \cdot q^{n-k} \right) \Rightarrow$$

$$\Rightarrow \mathbb{E}(S_n) = n \cdot p$$

$X_1, \quad S_{X_1} = \{0, 1\}$
(q, p)

Rappresentazione graficamente i valori dello spettro



Lezione 12

Un altro modo per scrivere $\mathbb{E}(S_n)$ è il seguente

$$\mathbb{E}(S_n) = \mathbb{E}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{E}(X_k)$$

Trasforma una variabile aleatoria in un numero deterministico

Tutte le prove hanno la stessa distribuzione di Bernoulli.

$$= \sum_{k=1}^{n} p = p \sum_{k=1}^{n} 1 = np$$

ESERCIZIO

Determinare la probabilità che nel prossimo concorso esca il numero 5 sulla ruota di Venezia

Sia E l'evento in questione

$i = 1, \dots, 5$, E_i: "E si verifica alla i-esima estrazione"

$E = \bigcup_{i=1}^{5} E_i \Rightarrow \mathbb{P}(E) = \sum \mathbb{P}(E_i)$

$\mathbb{P}(E_1) = 1/90$ *

$\mathbb{P}(E_2) = \mathbb{P}(\{5\}_1^c \cap \{5\}_2) = \mathbb{P}(\{5\}_1^c) \cdot \mathbb{P}(\{5\}_2 | \{5\}_1^c)$

$= \cancel{89}/90 \cdot 1/\cancel{89} = \frac{1}{90}$ *

gli eventi sono equiprobabili

$\Downarrow$

$\mathbb{P}(E_3) = \cancel{89}/90 \cdot \cancel{88}/\cancel{89} \cdot \frac{1}{\cancel{88}} = \frac{1}{90}$ *

$\mathbb{P}(E_i) = 5/90 = 1/18$

Un altro valore importante è il Tempo T, il quale:

$T \geq 1 : X_T = 1$ "per la prima volta"

$(X_n) n \in \mathbb{N}$, $p = \mathbb{P}(T)$ $X_n \sim B(1, p)$

$S_T = \{1, 2, 3, \dots, n\} = \mathbb{N}$

$n \in \mathbb{N}$, $\mathbb{P}(T = n) = \mathbb{P}(C_1 \cap C_2 \cap \dots \cap C_{n-1} \cap T_n)$

$= \mathbb{P}(C_1) \cdot \mathbb{P}(C_2) \cdot \dots \cdot \mathbb{P}(C_{n-1}) \cdot \underbrace{\mathbb{P}(T_n)}_{p}$

$= p \cdot q^{n-1}$

valore T

valori di uscita di croci

$T \sim G(p)$ → T ha distribuzione geometrica

descrive il numero di lanci necessari affinché si verifichi il primo successo in una successione di esperimenti bernoulliani indipendenti e identicamente distribuiti.

Ora bisogna verificare che: $\sum_{n=1}^{\infty} \mathbb{P}(T_n) = 1$

$$= \sum_{m=1}^{\infty} p \cdot q^{m-1} = p \sum_{m=1}^{\infty} q^{m-1} = p \sum_{m=0}^{\infty} q^m \to$$ serie geometrica di ragione q

$$= p \cdot \frac{1}{1-q} = p \cdot \frac{1}{1-(1-p)} = p/p = 1$$

$$E(T) = \sum_{m=1}^{\infty} m \cdot p \cdot q^{m-1} = p + 2pq + 3pq^2 + 4pq^3 + \dots$$

valore numerico della spesa

$$= p(1 + 2q + 3q^2 + 4q^3 + \dots)$$
$$= 1 + 2q + 3q^2 + 4q^3 + \dots$$
$$= 1 + q + q^2 + q^3 + \dots = 1/p$$
$$= q + q^2 + q^3 + \dots = q(1 + q + q^2 + \dots) = q/p$$
$$= q^2 + q^3 + \dots = q^2(1 + q + \dots) = q^2/p$$
$$= q^3 + \dots = q^3(1 + \dots) \; q^3/p$$

$$= \cancel{p}\left(\frac{1}{\cancel{p}} + \frac{q}{\cancel{p}} + \frac{q^2}{\cancel{p}} + \frac{q^3}{\cancel{p}} + \dots\right) = 1/p$$

$W_1 = T \sim G(p)$

DISTRIBUZIONE DI PASCAL

$\forall k \in \mathbb{N}$

$W_k \sim (p, k) \Rightarrow S_{W_k} = \{k, k+1, k+2, \dots\}$

$n \geq k$, $P(W_k = n) = P(S_{n-1} = k-1, X_n = 1) = P(S_{n-1} = k-1) \cdot P(X_n = 1)$

$$= \binom{n-1}{k-1} \cdot p^{k-1} \cdot q^{n-k} \cdot p = \binom{n-1}{k-1} \cdot p^k \cdot q^{n-k}$$

è una distribuzione di probabilità discreta con due parametri p ed n che descrive il numero di fallimenti precedenti il successo n-esimo in un processo di Bernoulli di parametro p

$$W_k = \sum_{i=1}^{k} T^{(i)} \Rightarrow E(W_k) = \sum_{i=1}^{k} (E(T^{(i)})) = \sum_{i=1}^{k} (E(T)) =$$
$$= E(T) \cdot \sum_{i=1}^{k} 1 = k/p$$

LEZIONE 13

$(\Omega = \{T, C\}^{\infty}, \mathcal{F} = \sigma[(T_n)_{n \in \mathbb{N}}], P)$ $\quad P(T) = p \in (0,1)$, $q = 1-p$

$n \in \mathbb{N}$, $X_n \sim$ Bernoulli $(p) \sim B(1, p)$

$\omega_1 = (C_1, T_2, C_3, C_4, C_5, C_6, C_7)$
$\omega_2 = (C_1, T_2, T_3, C_4, T_5, T_6, C_7)$

$X_1(\omega_1) = 0$
$X_2(\omega_1) = 1$
$X_3(\omega_1) = X_4(\omega_1) = X_5(\omega_1) = 0$

- $S_n \sim B(n,p)$ → è distribuito secondo la legge binomiale
Somma delle prime n prove

→ $S_2(\omega_1) = 1$
$S_3(\omega_1) = 1$
$S_4(\omega_1) = 2$
$S_4(\omega_2) = 3$
Conta il numero delle testa nelle prime n prove

S_n conta quante T sono uscite fino alla posizione n

- $T = W_1 \sim G(p)$

→ $T(\omega_1) = 2$
$T(\omega_2) = 3$
→ Tempo per il quale è uscito T

- $R \in \mathbb{N}$, $W_R \sim \text{Pascal}(R, p)$

→ $W_2(\omega_2) = 5$
$W_2(\omega_1) = 2$
$W_3(\omega_2) = 6$
→ è il tempo di uscita della R-esima "T"

SPETTRO → $S_{S_n} = \mathbb{N}$, $R \in \mathbb{N}$, $\mathbb{P}(S_n = R) = \binom{n}{R} p^R q^{n-R}$

$\lambda > 0$

$\cancel{p = \lambda/1}$ ← $n = 1$ | $\cancel{p_2 = \lambda/2}$ ← $n = 2$ | $\cancel{p_3 = \lambda/3}$ ← $n = 3$ | $p_4 = \lambda/4$ ← $n = 4$ — $p_5 = \lambda/5$ ← $n = 5$

Con $\lambda = 3$ scarto $n = 1, n = 2, n = 3$ finché il risultato < 1

$\lambda = 3 \Rightarrow p_4 = 0{,}75 \wedge p_5 = 0{,}6$

$n \to \infty \wedge p_n \to 0$ ← n aumenta e p decrementa ma il prodotto tra n e p è costantemente λ

$n \cdot p_n = \lambda$

$n \to \infty \Rightarrow p_n \to 0$

$$\lim_{n\to\infty} \binom{n}{k} p^k q^{n-k} = \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n!}{k!\cdot(n-k)!} \cdot \frac{\lambda^k}{n^k} \cdot \left(1-\frac{\lambda}{n}\right)^n \Big/ \left(1-\frac{\lambda}{n}\right)^k$$

$$= \frac{\lambda^k}{n^k} \lim_{n\to\infty} \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)\cdot\cancel{(n-k)!}}{n\cdot n\cdot\ldots\cdot n\cdot\cancel{(n-k)!}} \cdot \left(1-\frac{\lambda}{n}\right)^n \Big/ \left(1-\frac{\lambda}{n}\right)^k$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{\cancel{n}\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{\cancel{n}\cdot n\cdot\ldots\cdot n} \cdot \frac{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n}{\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^k \to 1}$$

$$= \frac{\lambda^k}{n!} \cdot \lim_{n\to\infty} \frac{n-1}{n} \to 1 \cdot \lim_{n\to\infty} \frac{n-2}{n} \to 1 \cdot \ldots \cdot \lim_{n\to\infty} \frac{n-k+1}{n} \to 1 \cdot \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^k}{n!} \cdot \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n$$

$$= \frac{\lambda^k}{n!} \cdot 1e^{-\lambda} \longrightarrow e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \quad \lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x$$

numero nepero $\longrightarrow$ LIMITE NOTEVOLE

Un numero Y che ha spettro $\mathbb{N}_0$ e successione di probabilità dei valori dello spettro $\longrightarrow \boxed{\frac{\lambda^n}{n!} \cdot e^{-\lambda}}$

NUMERO DI POISSON con parametro λ e si scrive: $Y \sim \text{Poisson}(\lambda) \sim \Pi(\lambda)$

[Una distribuzione di Poisson è un numero aleatorio discreto che conta il numero di occorrenze in un dato intervallo di tempo]

ESEMPIO

$Y \sim \Pi(\lambda)$ con $\lambda > 0$

$S_Y = \mathbb{N}_0$, $k \in \mathbb{N}_0$, $\mathbb{P}(Y=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

$\mathbb{P}(Y=1) = e^{-\lambda} \cdot \lambda \qquad \mathbb{P}(Y=0) = e^{-\lambda}$

$\mathbb{P}(Y=s) = e^{-\lambda} \cdot \frac{\lambda^s}{s!}$

- $k \in \{0,1,2,\dots,n\}$, $\mathbb{P}(S_n = k) = \binom{n}{k} p^k q^{n-k}$

$p_{k+1} = \binom{n}{k+1} p^{k+1} q^{n-k-1} = p/q \cdot \binom{n}{k+1} p^k q^{n-k}$

$= p/q \; p^k q^{n-k} \frac{n!}{(k+1)!\,(n-k-1)!}$

$= p/q \; p^k q^{n-k} \frac{1}{k+1} \frac{n!}{k!\,(n-k-1)!}$

$= p/q \frac{1}{k+1} p^k q^{(n-k)} (n-k) \frac{n!}{k!\,(n-k)!}$

$= p/q \frac{n-k}{k+1} p_k$

$\Rightarrow p_{k+1} = p_k \cdot p/q \cdot \frac{n-k}{k+1}$ con $k = 1, \dots, n-1$

- $Y \sim \Pi(\lambda)$, $k \in \mathbb{N}_0$, $\mathbb{P}(Y=k) = \underbrace{e^{-\lambda} \cdot \frac{\lambda^k}{k!}}_{a_k} > 0$

$a_{k+1} = e^{-\lambda} \cdot \frac{\lambda^{k+1}}{(k+1)!} = \frac{\lambda}{k+1} \cdot e^{-\lambda} \cdot \frac{\lambda^k}{k!} = \frac{\lambda}{k+1} \cdot a_k$

- $\mathbb{E}(Y) = \sum_{k=0}^{\infty} k \cdot a_k = \sum_{k=0}^{\infty} k \cdot \left(e^{-\lambda} \cdot \frac{\lambda^k}{k!}\right)$

$= e^{-\lambda} \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{(k-1)!}$

$= e^{-\lambda} \cdot \lambda \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \cdot \lambda \sum_{k'=0}^{\infty} \frac{\lambda^{k'}}{k'!} \qquad k' = k-1$

$= e^{-\lambda} \cdot \lambda \cdot e^{\lambda} = \frac{1}{\cancel{e^{\lambda}}} \cdot \lambda \cdot \cancel{e^{\lambda}} = \lambda \longrightarrow$ il parametro della legge di Poisson non è la media

ESEMPIO 5.2.1

Supponendo che il numero medio di incidenti settimanali in un particolare tratto di autostrada sia pari a 3 si vuole calcolare la probabilità che la prossima settimana vi sia almeno un incidente.
Si denota con X il numero di incidenti in quel tratto di autostrada nella settimana in esame. Poiché si può supporre che in una settimana passino un gran numero di autovetture e che ciascuna abbia una piccola probabilità di essere coinvolta in un incidente il numero di tali incidenti sarà distribuito come una variabile aleatoria di Poisson di media 3.

Y : il numero di incidenti settimanali in quel tratto

$Y \sim \Pi(\lambda = 3)$

$$P(Y \geq 1) = \sum_{k=1}^{\infty} a_k = 1 - P(X=0) = 1 - a_0 = 1 - e^{-3} \cdot \frac{3^0}{0!}$$

$$= 1 - e^{-3} \cdot \frac{1}{1} = 1 - e^{-3} \cong 0,9502$$

ESEMPIO 5.2.2

Un macchinario produce oggetti che hanno una probabilità di essere difettosi pari a 0,1.
Supponendo l'indipendenza nella qualità dei pezzi successivi, con che probabilità un campione di 10 oggetti ne conterrà al più uno difettoso? F
Il numero di pezzi difettosi è una variabile aleatoria binomiale di parametri (10, 0,1)

$$\boxed{P(S_n = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}} \qquad q = 1-p$$

S_{10} : numero degli oggetti difettosi nel campione di 10 pezzi

$S_{10} \sim B(n = 10, p = 0,1)$

$\{S_{10} = 0\} \cup \{S_{10} = 1\}$

CASO OGGETTI NON DIFETTOSI — CASO AL PIÙ UNO DIFETTOSO

$$P(F) = P_{S_{10}=0} + P_{S_{10}=1} = \binom{10}{0} \cdot 0,1^0 \cdot 0,9^{10-0} + \binom{10}{1} \cdot (0,1)^1 \cdot (0,9)^9$$

$\Rightarrow P(F) = 0,7361$

OPPURE

CASO CON VARIABILE ALEATORIA DI POISSON

$Y \sim \Pi(\lambda = n p)$

$Y \sim \Pi(\lambda = 10 \cdot 0{,}1 = 1)$

$\mathbb{P}(\{Y=0\} \cup \{Y=1\}) = q_0 + q_1 = e^{-1} \cdot \left(1 + \frac{1}{1!}\right) = 2/e = 0{,}7357$

FUNZIONE GENERATRICE DI MOMENTI

→ viene usata nella teoria della probabilità per caratterizzare in modo stretto le variabili aleatorie permettendo sia di ottenere alcuni parametri (come valore atteso e varianza) sia di confrontare due diverse var aleatorie e vedere il loro comportamento in condizioni limite

$\forall m \in \mathbb{N}, \ \mathbb{E}(X^m) = \sum_k k^m \cdot \mathbb{P}(X = k) \longrightarrow$ MOMENTI

↓

$\phi(t) = \mathbb{E}(e^{tX}) \Rightarrow \phi(0) = 1$

$\left.\frac{d\phi(t)}{dt}\right|_{t=0} = \mathbb{E}(X), \quad \left.\frac{d^2\phi(t)}{dt^2}\right|_{t=0} = \mathbb{E}(X^2)$

LEZIONE 14

$X \sim \text{Poisson}(\lambda) \ / \ X \sim \Pi(\lambda)$ → Assume solo valori interi NON negativi

$\forall x \in \mathbb{N}, \ \mathbb{P}(X = x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$

→ la distribuzione si può utilizzare per approssimazione abbastanza grande al posto della distribuzione binomiale

↓

$X_n \sim B(n, p) \ \forall n \in \mathbb{N}$

Sotto l'ipotesi che da un certo n in poi: $np = \lambda$
si può dire che:

- n abbastanza grande
- p abbastanza piccolo

$\Rightarrow \mathbb{P}(X_n = x) \simeq \frac{\lambda^x}{x!} \cdot e^{-\lambda}$

→ $n \to +\infty$, $p \to 0$

numero di prove

→ probabilità di successo

262 380
24 · 170 · 2880

ESERCIZI

Capitolo 5

10) Confrontare le probabilità esatte con l'approssimazione di Poisson nei casi seguenti.
Si intende che X è binomiale di parametri n e p

- a) $P(X=2)$ quando $n = 10$ e $p = 0{,}1$;

$\frac{\lambda^x}{x!} \cdot e^{-\lambda}$ $\lambda = n \cdot p = 1$

$\binom{n}{x} p^x (1-p)^{n-x}$

$= \binom{10}{2} (0{,}1)^2 \cdot (0{,}9)^8 = 0{,}19371$

$= \frac{\lambda^2}{2!} \cdot e^{-\lambda} = \frac{1^2}{2!} \cdot e^{-1} = \frac{1}{2} \cdot e^{-1} = \frac{1}{2} \cdot \frac{1}{e} = \frac{1}{2e} \cong 0{,}1839$

- b) $P(X=0)$ quando $n = 10$ e $p = 0{,}1$ $\Rightarrow \lambda = n \cdot p = 1$

$= \binom{10}{0} (0{,}1)^0 (0{,}9)^{10} = (0{,}9)^{10} \cong 0{,}348678$

$= \frac{\lambda^0}{0!} \cdot e^{-\lambda} = \frac{1^0}{0!} \cdot e^{-1} \cong 0{,}367879$

- c) $P(X=4)$ quando $n = 9$ e $p = 0{,}2$ $\lambda = n \cdot p = 1{,}8$

$= \binom{9}{4} (0{,}2)^4 (0{,}8)^{9-4} = \frac{9!}{4!\,(9-4)!} \cdot (0{,}2)^4 (0{,}8)^5$

$= \frac{9!}{4!\,(5)!} \cdot (0{,}2)^4 (0{,}8)^5 = 126 \cdot (0{,}2)^4 (0{,}8)^5 =$

$= 126 \cdot 0{,}0016 \cdot 0{,}32768 = 0{,}0660$

$= \frac{\lambda^4}{4!} \cdot e^{-\lambda} = \frac{1{,}8^4}{4!} \cdot e^{-1{,}8} = 0{,}4374 \cdot \frac{1}{e^{1{,}8}}$

$= 0{,}4374 \cdot 0{,}165299 = 0{,}0723$

5.2.3

Si consideri un esperimento che consiste nel contare il numero di particelle alfa emesse in un secondo da un grammo di un certo materiale radioattivo.
Si sa dalle esperienze passate che il valore medio di questa variabile aleatoria è 3,2; quale è una buona approssimazione della probabilità che nell'esperimento in esame non vengano emesse più di 2 particelle

- Si può ipotizzare che una sorgente a sia un n grande di atomi radioattivi, i quali hanno prob. di emettere particelle radioattive pari a $3{,}2/n$

si ha che $\frac{3{,}2}{n} \cdot n = 3{,}2 = \lambda$ → valore atteso della variabile aleatoria

distribuzione Poisson — probabilità di successo — numero di ripetizioni di atomi

$$P(X \leq 2) = \sum_{x=0}^{2} \binom{n}{x} \cdot \left(\frac{3{,}2}{n}\right)^{x} \cdot \left(1 - \frac{3{,}2}{n}\right)^{n-x}$$

$$P(X \leq 2) \simeq \sum_{x=0}^{2} \lambda^x / x! \cdot e^{-\lambda} = \lambda^0/0! \cdot e^{-3{,}2} + \lambda^1/1! \cdot e^{-3{,}2} + \lambda^2/2! \cdot e^{-3{,}2}$$

per i valori di x

$$= e^{-3{,}2}\left[1 + 3{,}2 + 3{,}2^2\right] \simeq 0{,}3795$$

5.2.4

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

(a) Che frazione delle giornate vedrà arrivare meno di 3 richieste?

(b) Con che probabilità in una settimana lavorativa di 5 giorni, in esattamente 3 giorni arrivano 4 richieste?

a) $X \sim \pi(5)$ → $\lambda = 5$ ← numero richieste giornaliere

$$P(X < 3) \simeq \underbrace{e^{-\lambda} \cdot \left(1 + 5 + \frac{5^2}{2}\right)}_{\text{valori di } X < 3 \quad 12{,}47} \simeq 0{,}1247$$

b) $p = P(X = 4) \simeq \lambda^x / x! \cdot e^{-\lambda} = 5^4/4! \cdot e^{-5} = \frac{625}{24} \cdot \frac{1}{e^5}$

$$= 26{,}04 \cdot \frac{1}{148{,}41} \simeq 0{,}1751$$

Y: numero di giorni lavorativi (5) in cui si verifica $X = 4$

$\Rightarrow y \sim B(5, p) \longrightarrow$ per distribuzione binomiale

$$= \binom{n}{x} p^x (1-p)^{n-x}$$

$$= \binom{5}{3} p^3 (1-p)^2$$

$$= \frac{5!}{3!(5-3)!} p^3 (1-p)^2 = \frac{120}{6 \cdot 2} p^3 (1-p)^2$$

$$= 10 \cdot 0{,}1781^3 (0{,}8219)^2 \simeq 0{,}0364$$

- $X \sim G(p)$ variabile aleatoria geometrica

$\downarrow$

Esperimento di prove indipendenti ripetute fino ad ottenere il successo. p rappresenta la probabilità di successo in una singola prova

$S_X = \mathbb{N}$

$X = 1, 2, \dots$

$\mathbb{P}(X = x) = p(1-p)^{x-1}$

$$\phi_X(t) = \mathbb{E}[e^{tX}] = \sum_{x=1}^{\infty} e^{tx} \, \mathbb{P}(X = x)$$

$$= \sum_{x=1}^{\infty} e^{tx} (1-p)^{x-1} p = p e^t \cdot \sum_{x=1}^{\infty} [e^t (1-p)]^{x-1}$$

$$= p e^t \frac{1}{1 - e^t (1-p)}$$

$\downarrow$

Valore atteso

$$\frac{d}{dt} \phi_X(t) \Big|_{t=0} = \mathbb{E}(X) \longleftarrow$$ proprietà della funzione generatrice di momenti

DIM:

$$\frac{d}{dt} \phi_X(t) \Big|_{t=0} = \frac{p e^t [1 - e^t (1-p)] - p e^t [-(1-p) e^t]}{[1 - e^t (1-p)]^2}$$

$$= \frac{\cancel{p^2} + \cancel{p} - \cancel{p^2}}{p^{\cancel{2}}} = 1/p = \mathbb{E}[X] \quad \square$$

$X \sim \pi(\lambda)$ $\quad \mathbb{E}[X] = \lambda$ $\quad$ X poisson

$$\phi_X(t) = \mathbb{E}[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x}{x!} e^{-\lambda} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} =$$

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

$$\frac{d}{d\tau} \Phi_X(\tau)\Big|_{\tau=0} = \frac{d}{d\tau}\left[e^{\lambda(e^{\tau}-1)}\right]\Big|_{\tau=0}$$

$$= e^{\lambda(e^{\tau}-1)} \cdot \lambda e^{\tau}\Big|_{\tau=0} = e^{0} \cdot \lambda \cdot e^{0} = \lambda = E(X)$$

$X \sim B(n,p)$ $E(X) = np$ X binomiale

$\Phi_X(\tau)$

ESERCIZIO

Un dado onesto viene lanciato fino all'uscita della faccia "1". Sapendo che sono stati effettuati 3 lanci senza aver visto la faccia 1, calcolare la probabilità di fare al massimo altri 3 lanci

X: variabile aleatoria che conta il numero di lanci per vedere la prima volta 1

$$X \sim G(p) \xrightarrow{} \text{probabilità che esca faccia 1}$$

$\Downarrow$

$X \sim G(1/6)$

$\Downarrow$

$$\{X>3\} \Rightarrow P(\{X \le 6\} | \{X>3\}) = P(X \le 6 | X>3)$$

$$= P(\{X \le 6\} \cap \{X>3\}) / P(X>3)$$

$$= \frac{P(X \le 6, X>3)}{P(X>3)} = P(3 < X \le 6) / P(X>3)$$

$$= P(X=4) + P(X=5) + P(X=6) / 1 - P(X=1) - P(X=2) - P(X=3)$$

ESERCIZIO

Un ragazzo lancia ripetutamente un dado e si ferma quando esce almeno una volta la faccia 1 e la faccia 2. Detto X, il numero di lanci effettuati, determinare il valore atteso

E_1 = "esce faccia 1"
E_2 = "esce faccia 2"

$X = Y + T$

$$Y = E_1 \cup E_2 \Rightarrow Y \sim G(p_y) \Rightarrow p_y = P(E_1 \cup E_2) = P(E_1) + P(E_2) = 1/6 + 1/6 = 1/3$$

$Y \sim G(1/3) \wedge T \sim G(1/6)$

$E[X] = E[Y+T] = E[Y] + E[T] = \underbrace{3+6}_{} = 9$

rispettivi valori reciproci di $Y \sim G(p)$, $T \sim G(p)$

RIEPILOGO DELLE PRINCIPALI DENSITÀ DISCRETE

NOME	DENSITÀ	LEGGE	DESCRIZIONE
BINOMIALE	$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$	$X \sim B(n,p)$	Esperimento di n prove ripetute; p rappresenta la probabilità di successo in una singola prova
BERNOULLIANA	$p(k) = p^k (1-p)^{1-k}$	$X \sim B(1,p)$	Caso particolare della densità binomiale in cui il numero di prove è pari a 1
GEOMETRICA	$p(k) = (1-p)^{k-1} p$	$X \sim G(p)$	Esperimento di prove ripetute fino ad ottenere il successo; p rappresenta la probabilità di successo in una singola prova
POISSON	$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$	$X \sim Poisson(\lambda)$ ∨ $X \sim \Pi(\lambda)$	Caso particolare binomiale in cui n viene fatto tendere $+\infty$; p, la prob. di successo di una prova è 0; λ il parametro della densità è np
IPERGEOMETRICA	$p(k) = \frac{\binom{r}{k}\binom{b}{n-k}}{\binom{r+b}{n}}$	$X \sim IG(r,b,n)$	Esperimento di n prove ripetute non indipendenti

LEZIONE 15

Numeri aleatori ottenuti:

- $B(n,p)$ con $n \in \mathbb{N}$, $p \in (0,1)$
- $G(p)$ con $p \in (0,1)$
- Pascal (k,p) con $k \in \mathbb{N}$, $p \in (0,1)$
- Poisson (λ), con $\lambda > 0$

Si introduce ora una variabile simile al binomiale

$p = \#B / \#T$

numero palline bianche

numero palline complessive

Si estrae una biglia dalla scatola

$$X_1 := \begin{cases} 1, & \text{biglia bianca} \\ 0, & \text{altrimenti} \end{cases} \Rightarrow X_1 \sim B(1,p) \sim \text{Bernoulli}(p)$$

Per ottenere una nuova variabile aleatoria diversa da X_1 (prima estrazione) chiamata VARIABILE IPERGEOMETRICA, la prima biglia va messa da parte $\rightarrow$ Indipendenza tra le prove

Sia

a: numero biglie bianche
b: numero biglie nere

$\forall n \in \mathbb{N}$, il numero delle estrazioni possibili

X: numero delle biglie bianche estratte

$\forall k \in \mathbb{N}$ $$\mathbb{P}(X=k) = \frac{\binom{a}{k} \cdot \binom{b}{n-k}}{\binom{a+b}{n}}$$

Se k bianche

$n-k$ nere uscite

esiti possibili

Densità ipergeometrica $\rightarrow$ Assegnati tre numeri interi strettamente positivi a, b, n con $n \leq a+b$, si pone $k_1 = \max(0, n-b)$, $k_2 = \min(a,n)$ e si osservi che è $k_1 \leq k_2$. Allora si definisca nel seguente modo la funzione

$$p(k) = \begin{cases} \dfrac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} & k = k_1, k_1+1, \dots, k_2 \\ 0 & \text{altrimenti} \end{cases} \longrightarrow S_X = \{\max(0, n-b), \dots, \min(n, a)\}$$

è una distr. discreta = distr. ipergeometrica

ESEMPIO

- Scatola con N batterie usate

N: numero batterie accettabili

M: numero batterie difettose

Si estrae un blocco con n batterie

X: conta il numero N di batterie accettabili

$$k \in \mathbb{N}, \quad P(X=k) = \frac{\binom{N}{k} \cdot \binom{M}{n-k}}{\binom{N+M}{n}}$$

- Si ha una cassa di componenti per assemblare un sistema

5 guasti

15 efficienti

$k \in \{\max(0, n-b), \dots, \min(n, a)\}$
$\Rightarrow k \in \{1, \dots, 6\}$

$\boxed{n=6}$ blocchi di componenti. Il sistema deve essere composto con 6 componenti.

Il sistema funziona con almeno 4 componenti efficienti

X: numero componenti efficienti

$\rightarrow X \geq 4 \Rightarrow P(X \geq 4) = P(X=4) + P(X=5) + P(X=6)$

ADDITIVITÀ FINITA

$$= \frac{\binom{15}{4} \cdot \binom{5}{2}}{\binom{20}{6}} + \frac{\binom{15}{5} \cdot \binom{5}{1}}{\binom{20}{6}} + \frac{\binom{15}{6} \cdot \binom{5}{0}}{\binom{20}{6}}$$

$$= \frac{\binom{15}{4}\cdot\binom{5}{2} + \binom{15}{5}\cdot\binom{5}{1} + \binom{15}{6}\cdot\binom{5}{0}}{\binom{20}{6}}$$

questa è chiamata variabile IPERGEOMETRICA

$X \sim IG(a, b, n)$

$$X_k = \begin{cases} 1 & \text{se esce bianco con prob } a/a+b \\ 0 & \text{se esce nero con prob } b/a+b \end{cases}$$

$k = 1, 2, \dots, n$

$$P(X_1 = 1) = a/_{a+b} \quad \wedge \quad P(X_2 = 1) = a/_{a+b}$$

$$P(X_2 = 1) = P(\{X_2 = 1\} \cap \underbrace{[\{X_1 = 0\} \cup \{X_1 = 1\}]}_{\Omega})$$

$$= P(\{X_2 = 1\} \cap \{X_1 = 0\} \cup \{X_2 = 1\} \cap \{X_1 = 1\}) =$$

$$= P(\{X_2 = 1\} \cap \{X_1 = 0\}) + P(\{X_2 = 1\} \cap \{X_1 = 1\})$$

$$= P(X_1 = 0) \cdot P(X_2 = 1 \mid X_1 = 0) + P(X_1 = 1) \cdot P(X_2 = 1 \mid X_1 = 1)$$

$$= \frac{b}{a+b} \cdot \frac{a}{a+b-1} + \frac{a}{a+b} \cdot \frac{a-1}{a+b-1}$$

$$= \frac{ab}{(a+b)(a+b-1)} + \frac{a(a-1)}{(a+b)(a+b-1)}$$

$$= \frac{ab + a(a-1)}{(a+b)(a+b-1)} = \frac{a\cancel{(b+a-1)}}{(a+b)\cancel{(a+b-1)}} = a/_{a+b} \quad \square$$

$$\boxed{E(X) = n \cdot a/_{a+b}}$$

VALORE ATTESO = speranza matematica della distribuzione di una variabile casuale, è un indice di posizione

→ Esso rappresenta il valore previsto che si potrà ottenere in un gran numero di prove

hanno una stessa media: $X \sim B(n, p)$, $E(X) = np$
$X \sim IG(a, b, n)$, $E(X) = n \cdot a/_{a+b}$

- X è un numero aleatorio discreto con spettro S_X e valori di probabilità $(p_k)_{k \in S_X}$

$p_k = \mathbb{P}(X = k) = \mathbb{P}_X(\{k\})$ in $(\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$

$\forall B \in \mathcal{B}, \quad \mathbb{P}_X(B) = \mathbb{P}(X \in B) = \sum_{k : x_k \in B} p_k$

VAD se segue una distribuzione di probabilità discreta

- Bernoulliana
- Uniforme
- Geometrica
- Binomiale
- Pascal
- Ipergeometrica
- Poisson

ESEMPIO

$X \sim \Pi(\lambda), \quad \lambda > 0$

$S_X = \{0, 1, 2, \dots, n\}$

$k \in S_X, \quad \mathbb{P}(X = k) = e^{-\lambda} \cdot \lambda^k / k!$

$\mathbb{P}_X([0, +\infty[) = 1 - \mathbb{P}_X(\{0\})$

$(]-\infty, x])_{x \in \mathbb{R}}$ famiglia delle semirette sinistre

$B =]-\infty, x]$

$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X \leq x)$

FUNZIONE DI DISTRIBUZIONE

È la funzione che associa a ciascun valore x la probabilità del seguente evento: "che var. casuale X assume valori minori o uguali a x"

Data una variabile casuale X, la funzione che fa corrispondere ai valori di x le probabilità cumulate $P(X \leq x)$ viene detta funzione di distribuzione F_X

$F_X : \mathbb{R} \to [0, 1] \qquad F_X(x) := \mathbb{P}(X \leq x)$

Usata per calcolare la probabilità che una variabile casuale X assume valore minore o uguale a un dato specifico e quindi in questo caso bisogna calcolare le probabilità cumulate

PROPRIETÀ DA VERIFICARE

a - $0 \leq F_X(x) \leq 1$

b - $\lim_{x \to -\infty} F_X(x) = 0, \quad \lim_{x \to +\infty} F_X(x) = 1 \to$ NORMALIZZAZIONE

c - $F_X(x)$ è monotona NON decrescente : $x_1 < x_2 \Rightarrow F_X(x_1) \leq F_X(x_2)$
d - $P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$

[la funzione di distribuzione consente di stabilire la probabilità che la variabile casuale X assuma valori nell'intervallo $(x_1, x_2]$]

LEZIONE 16

$\mathcal{E} = (\Omega, F = \sigma(G), P) \longrightarrow (\mathbb{R}, B, P_X)$

$X: \Omega \xrightarrow[F\text{-misurabile}]{} \mathbb{R}$

$\omega \rightsquigarrow X(\omega)$

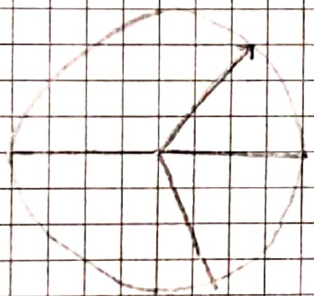
$E \in F,\ P(E)$

$B \in \mathcal{B},\ P_X(B) := P(X \in B)$

[variabili che possono assumere valori in un intervallo]

$x \in \mathbb{R},\ F_X(x) = P_X(]-\infty, x]) = P(X \leq x)$

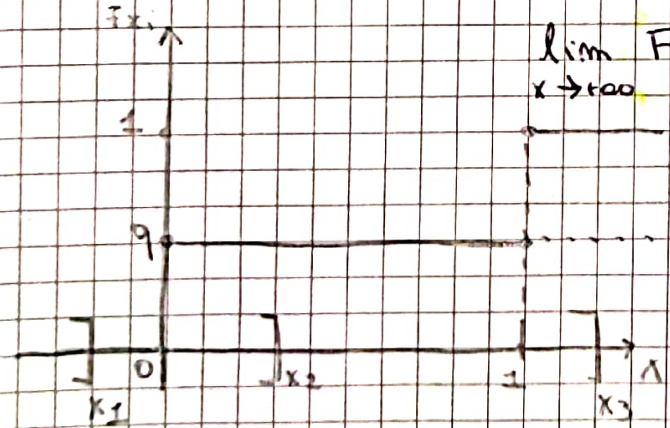
ESEMPIO



(C, B_C, λ_C) $X: C \longrightarrow \mathbb{R}$, $\omega \rightsquigarrow X(\omega) = X_C$

$X_1 = \begin{cases} 1 & , p \\ 0 & , 1-p \end{cases}$ $(p \in (0,1))$

$S_{X_1} = \{0, 1\}$



$\lim_{x \to +\infty} F_X(x) = 1$

$\lim_{x \to -\infty} F_X(x) = 0$

$F_{X_1}(x) = P(X_1) \leq x_2 = 0$

$x_1 < 0$

$0 \leq x_2 < 1$, $x_3 \geq 1$

$F_{X_1}(x_2) = P(X_1 \leq x_2) = P(X_1 = 0) = q$

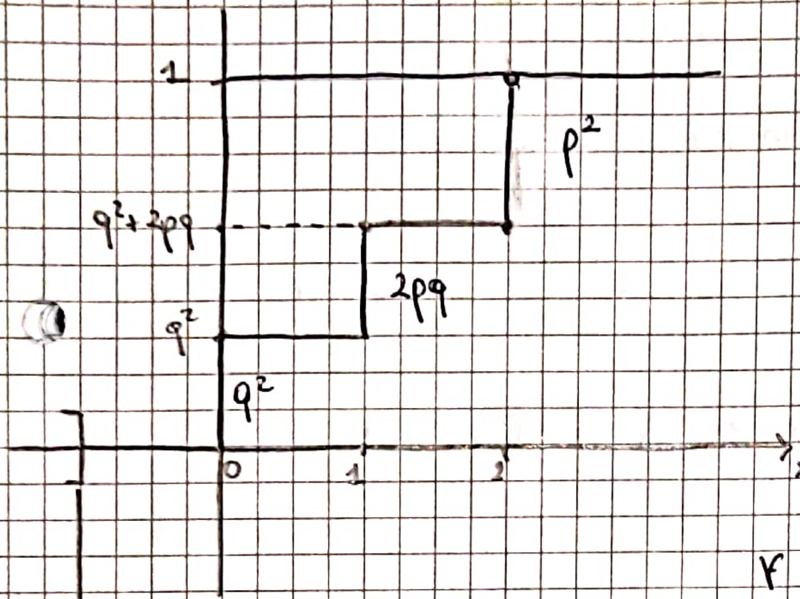
$F_{X_1}(x_3) = P(X_1 \leq x_3) = P(\{X = 0\} \cup \{X = 1\}) = P(X = 0) + P(X = 1) = 1$

$F_X(0^+) - F_X(0^-) = F_X(0) - F_X(0^-) = q - 0 > 0$

$$F_X(x) - F_X(x^-) = \begin{cases} 0, & \text{se } x \text{ non è valore spettrale} \\ p_k, & \text{se } x = x_k \in S_X \end{cases}$$

ESEMPIO

$S_2 = X_1 + X_2$ — 0 con prob q^2; 1 con prob $2pq$; 2 con prob p^2



MONOTONA CRESCENTE (NON DECRESCENTE)

qualsiasi semiretta che ha origine nei negativi è 0

esprime continuità da destra

$\forall x \in \mathbb{R}$

$$\lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x)$$

$$\lim_{\varepsilon \to 0^-} F_X(x+\varepsilon) = F_X(x)$$

$\downarrow$

NON si può dire

$F_X \longrightarrow$ NON DECRESCENTE
CONTINUA A DESTRA
da 0^+ vale 1, da 0^- vale 0

$W_1 = T \sim G(p)$

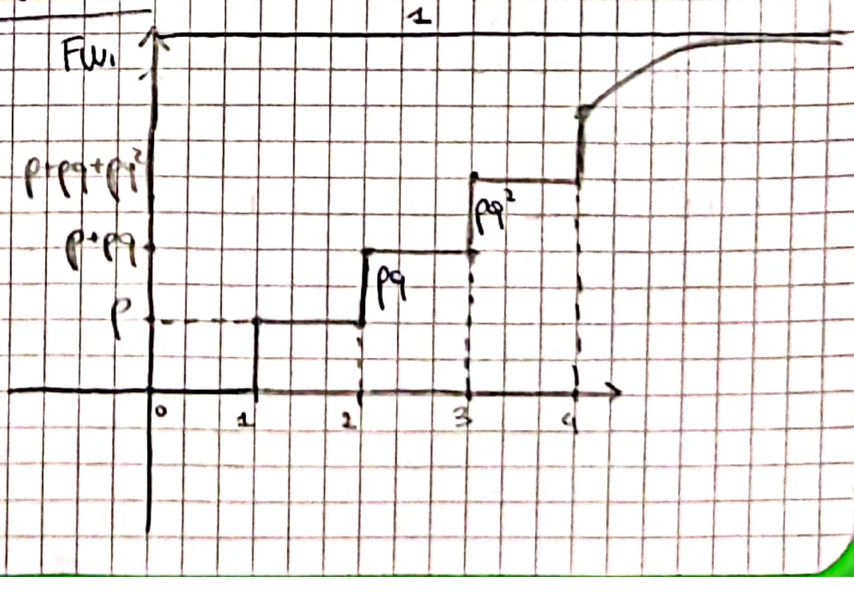
$\{1, 2, \dots\} = \mathbb{N}$

$\forall n \in \mathbb{N}$, $P(X=n) = p\,q^{n-1}$

$\downarrow$

distribuzione di probabilità di W_1

$$\boxed{\lim_{x \to +\infty} F_{W_1}(x) = 1}$$



PROPRIETÀ

Si consideri una successione di eventi $\in \mathcal{F}$:

$(E_n)_{n \in \mathbb{N}}$

- Si dice che $(E_n)_{n \in \mathbb{N}}$ è CRESCENTE $\iff$: $\forall n \in \mathbb{N}$, $E_n \subseteq E_{n+1}$
  $\to E_n \uparrow : \bigcup_{n \in \mathbb{N}} E_n$
- Si dice che $(E_n)_{n \in \mathbb{N}}$ è DECRESCENTE $\iff \forall n \in \mathbb{N}$, $E_n \supseteq E_{n+1}$
  $\to E_n \downarrow : \bigcap_{n \in \mathbb{N}} E_n$

Sia $(E_n)_{n \in \mathbb{N}}$ successione monotona

THM.

$$\lim_{n \to \infty} (\mathbb{P}(E_n)) = \mathbb{P}(\lim_{n \to \infty} (E_n)) \longrightarrow \left[\lim_{x \to x_0} f(x) = f(x_0) = f(\lim_{x \to x_0} x)\right]$$

F_X funzione distribuzione, $x \in \mathbb{R}$, $F_X(x) = \mathbb{P}(X \leq x)$

1) F_X è NON DECRESCENTE

DIM:

$$x_1 < x_2 \Rightarrow]-\infty, x_2] =]-\infty, x_1] \cup]x_1, x_2]$$

$$F_X(x_2) = \mathbb{P}(X \leq x_2) = \mathbb{P}(X \leq x_1) + \mathbb{P}(x_1 < X < x_2)$$

$$= F_X(x_1) + \mathbb{P}(x_1 < X \leq x_2)$$

$$= F_X(x_2) \geq F_X(x_1)$$

□

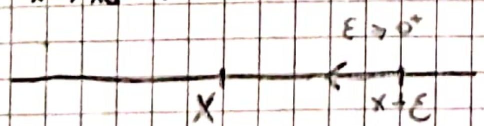
2) F_X è CONTINUA DESTRA in $\mathbb{R}$

DIM:

$$\forall x \in \mathbb{R}, \lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x)$$

Con ausilio del TEOREMA PONTE $\longrightarrow$ permette di passare dal limite per successioni al limite per funzioni

$$\lim_{x \to x_0} f(x) = \ell \longleftarrow$$



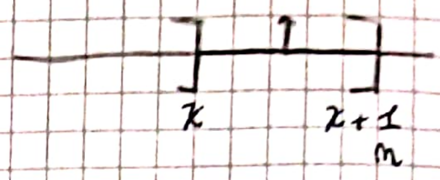
$$\boxed{\lim_{n \to \infty} f(x_n) = \ell}$$

Il Teorema ponte afferma che una funzione $f(x)$ ammette limite in un punto reale se e solo se accade che il punto in questione è di accumulazione per il dominio della funzione e che la successione delle immagini converge ad l

$\downarrow$

Comunque presa una successione definita in A e convergente in x_0, si può fare l'immagine di ogni termine della successione e la successione che ne viene fuori deve convergere ad l, per ogni successione convergente a x_0

$x < X \leq x + 1/n$



Quindi

$$\lim_{n\to\infty} F_X\left(x+\frac{1}{n}\right) = \lim_{n\to\infty} P(X \leq x + 1/n) =$$

$$= \lim_{n\to\infty} P(X \leq x) + P(x < X \leq x + 1/n)$$

$$= F_X(x) + \lim_{n\to\infty} P(x < X \leq x + 1/n)$$

$$= F_X(x) + P\left(\lim_{n\to\infty} \left\{x < X \leq x + \frac{1}{n}\right\}\right) \to 0$$

$$= F_X(x) + P(\emptyset) = F_X(x)$$

$\square$

$$\lim_{n\to\infty} F_X\left(x+\frac{1}{n}\right) = F_X(x) \Rightarrow \lim_{\varepsilon\to 0^+} F_X(x+\varepsilon) = F_X(x)$$

Fx non è decrescente + TEOREMA PONTE (per la NON DECRESCENZA)

3) $\lim_{n\to\infty} F_X(x) = 1$ $\longrightarrow$ Segue da (1) e da T. ponte

DIM

$$\lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} P(X \leq n) = P\left(\lim_{n\to\infty} \{X \leq n\}\right) = P(\Omega) = 1$$

$\square$

4) $\lim_{n\to-\infty} F_X(n) = 0$

$\square$

DIM

$$\lim_{n\to-\infty} F_X(-n) = \lim_{n\to-\infty} P(X \leq -n) = P\left(\lim_{n\to-\infty} \{X \leq -n\}\right) = P(\emptyset) = 0$$

FUNZIONE DENSIFICATA DI PROBABILITÀ

$f_X \geq 0 \longrightarrow$ ricostruire P_X

$B \in \mathcal{B}, \; P_X(B) = P(X \in B) = \int_B f_X(x)\,dx$

$\mathbb{R} \in \mathcal{B}, \; 1 = P_X(\mathbb{R}) = P(X \in \mathbb{R}) = \int_{\mathbb{R}} f_X(x)\,dx \longrightarrow$ restituisce il valore della funzione di distribuzione di X

ESEMPIO

riprendendo l'esempio della Ruota della Fortuna

$$f_{RF}(x) = \begin{cases} 1/2\pi, & 0 < x < 2\pi \\ 0, & \text{altrimenti} \end{cases}$$

$$\int_{-\infty}^{+\infty} f_{RT}(x)\,dx = \int_0^{2\pi} 1/2\pi \,dx = 1/2\pi \int_0^{2\pi} dx = 2\pi/2\pi = 1$$

LEZIONE 17

Variabili aleatorie continue

Se X è una variabile aleatoria continua, la probabilità che X assuma un certo valore x fissato è in generale zero

importante quindi calcolare la probabilità che X sia compresa fra a e b, con $a \le b$ — a, b intervallo reale

Assumono valori in un intervallo di numeri reali e vengono descritte tramite la funzione di densità di probabilità

DISCRETE

- legge di distribuzione è ricostruita mediante: $S_X, \; x \in S_X, \; P(X = x) > 0$

$B \in \mathcal{B}, \; B \cap S_X \Rightarrow P_X(B) = \sum_{x \in B \cap S_X} (P(X = x)) > 0$

→ SPETTRO

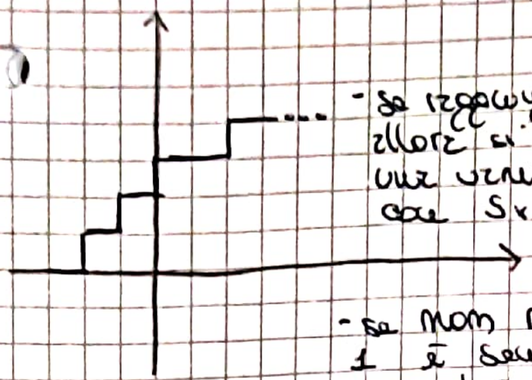
CONTINUE

- la legge P_X è ricostruita per il tramite di una funzione di densità f_X

$P_X(B) = \int_B f_X(x)\,dx$

→ SUPPORTO

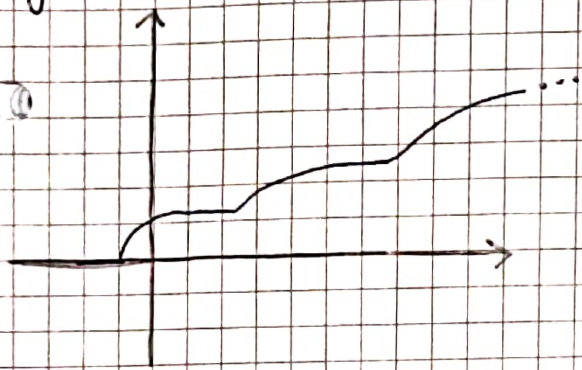
per F_X (var dis) $\longrightarrow$ MONOTONA NON DECRESCENTE



- se raggiunge 1, allora si tratta di una variabile discreta con S_X finito

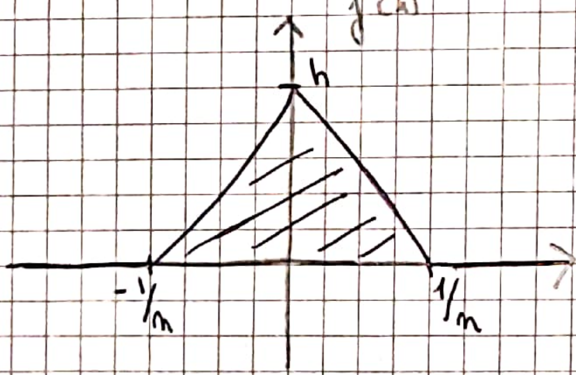
- se non raggiunge mai 1, 1 è sempre discreta ma il supporto ha la cardinalità numerabile

per $f_X(x)$ (var cont)



- se arriva a 1 è un intervallo limitato

Si immaginiamo, ora di avere questa situazione



$b = 2/n$

$\frac{h \frac{2}{n}}{2} = 1 = \frac{h \cdot b}{2} = 1$

$\Downarrow$

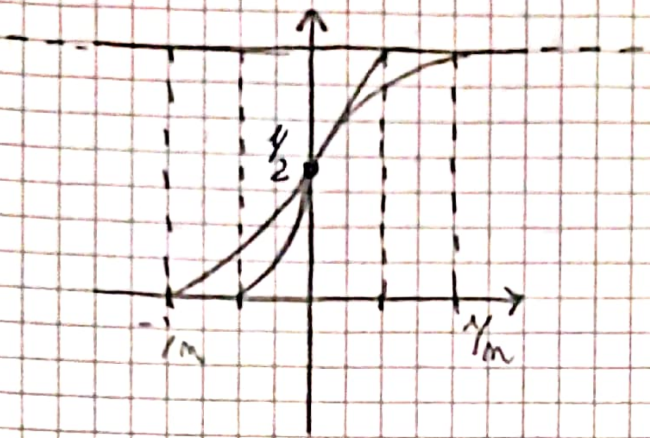
$h = 2/b = \frac{2}{\frac{2}{n}} = 2 \cdot \frac{n}{2} = n \Rightarrow \underline{h = n}$

- CONTINUE

$B \in \mathcal{B}$, $P_X(B) = \int_B f_X$

$B =]-\infty, x]$, $P_X(]-\infty, x]) = P(X \le x) = F_X(x) \leftarrow$ è una primitiva della funzione di densità di probabilità

$\hookrightarrow = \int_{-\infty}^{x} f_X(y)\, dy$



VARIABILI CONTINUE

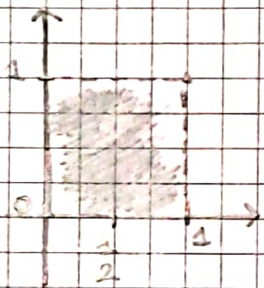
- 3 variabili continue fondamentali:

  a) var con legge uniforme
  b) var con legge esponenziale
  c) GAUSSIANE/NORMALI

LEGGE UNIFORME → distribuzione di probabilità uniforme attribuisce la stessa probabilità a tutti i punti appartenenti a un dato intervallo → densità di probabilità PIATTA

$X \sim U(0,1)$

$$f_x(x) = \begin{cases} 1, & \text{se } x \in (0,1) \\ 0, & \text{altrimenti} \end{cases}$$



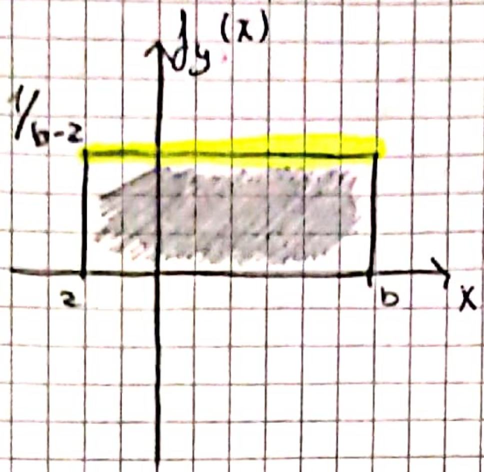
$$P_X(]0,\varepsilon[) = \int_0^\varepsilon 1\, dx = 1 \int_0^\varepsilon dx = \varepsilon$$

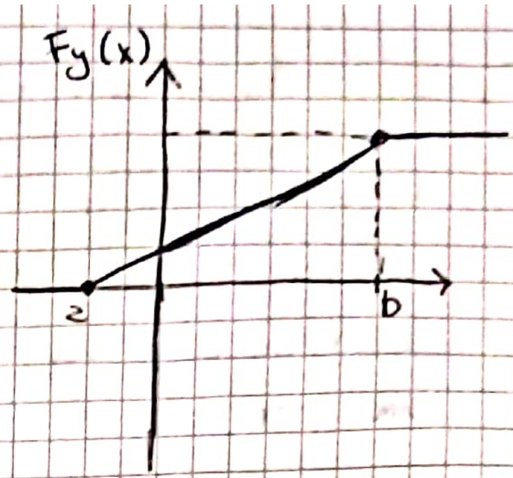
- $X \leq 0$, $F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \int_{-\infty}^{x} 0\,dy = 0$
- $0 < x < 1$, $F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \cancel{\int_{-\infty}^{0} f_X(y)\,dy}^{\,0} + \int_0^x f_X(y)\,dy =$

  $= \int_0^x f_X(y)\,dy = \int_0^x 1\,dy = x$
- $X > 1$, $F_X(x) = 1$

LEGGE UNIFORME GENERICA

$Y \sim U(a,b)$ con $a<b$ $\Rightarrow f_Y(x) = \begin{cases} 1/b-a, & \text{se } x \in (a,b) \\ 0, & \text{altrimenti} \end{cases}$





VALORE ATTESO

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \cdot f_X(x)\,dx = \int_0^1 x \cdot 1 \cdot dx = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 =$$

$$= \underbrace{\frac{1}{2}(1-0)}_{\text{primitive}} = 1/2 \qquad \frac{1}{2}x^2\Big|_0^1 \quad \frac{1}{2}(1^2) - \frac{1}{2}(0^2) = \frac{1}{2}(1-0) = \frac{1}{2}$$

$$\mathbb{E}(y) = \mathbb{E}[a+(b-a)X] = \mathbb{E}(a) + \mathbb{E}[(b-a)X] = a + \mathbb{E}[(b-a)X] =$$

$$= a+(b-a)\,\mathbb{E}(X) = a+(b-a)\frac{1}{2} = \frac{a+b}{2}$$

$t \in \mathbb{R}$, e^{tx} allora $\phi_X(t) = \mathbb{E}(e^{tx}) = \int_0^1 e^{tx} \cdot f_X(x)\,dx$

con $t \neq 0$ $= \int_0^1 e^{tx} \cdot 1\,dx = \frac{e^{tx}}{t}\Big|_0^1 = 1/t\,(e^t - 1)$

in generale $\longrightarrow$ $\phi_X(0) = \mathbb{E}(e^{0x}) = \mathbb{E}(1) = 1$

$$\lim_{t\to 0} \frac{e^t - 1}{t} = \lim_{t \to 0} \frac{e^t}{1} = 1$$

per DE L'HOPITAL

$$\phi'(t) = \frac{e^t \cdot t - (e^t - 1)}{t^2} = \frac{te^t - e^t + 1}{t^2}$$

$$\phi'(0) = \lim_{t\to 0} \frac{te^t + 1 - e^t}{t^2} = \frac{0}{0} \Rightarrow \lim_{t \to 0} \frac{\cancel{e^t} + te^t - \cancel{e^t}}{2\cancel{t}} =$$

$$= \lim_{t\to 0} \frac{e^t}{2} = \frac{1}{2}$$

- $\mathbb{E}(X^2) = \int_0^1 x^2 \, dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}(1-0) = 1/3 \rightarrow X \sim U(0,1)$

- $\mathbb{E}(Y^2) = \mathbb{E}\{[a+(b-a)X]^2\} = \mathbb{E}[a^2+(b-a)^2X^2+2a(b-a)X] =$

$= a^2 + (b-a)^2\,\mathbb{E}(X^2) + 2a(b-a)\,\mathbb{E}(X)$

$= a^2 + \frac{(b-a)^2}{3} + \frac{2a(b-a)}{2} =$ $\quad \frac{1}{3} \cdot \frac{0}{3} : \frac{1}{3}$

$= \frac{3a^2+b^2+a^2-2ab+3ab-3a^2}{3} = \frac{a^2+b^2+ab}{3}$

- $\mathbb{D}^2(X) := \mathbb{E}[X - \mathbb{E}(X)]^2$

VARIANZA → è la media aritmetica dei quadrati delle differenze tra ogni valore X_i della distribuzione e un valore medio preso come riferimento

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

$= \mathbb{E}(X^2) - 2\mathbb{E}(X)\cdot\mathbb{E}(X) + \mathbb{E}^2(X)$

$= \mathbb{E}(X^2) - \mathbb{E}^2(X) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$

LEZIONE 18

RIEPILOGO

$(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{\mathcal{F}\text{-mis}} (\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$

X: numero aleatorio

- $\mathbb{P}$ misura gli elementi di $\mathcal{F}$
- $\mathbb{P}_X$ misura gli elementi di $\mathcal{B}$

ha come argomento un insieme $B \in \mathcal{B}$, $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$

funzione distribuzione per variabili aleatorie DISCRETE ← da cui:

$\forall x \in \mathbb{R}, \ ]-\infty, x], \ \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x) := F_X(x)$

Tramite $F_X(x)$ è possibile ricostruire la misura di probabilità $\mathbb{P}_X$

NUMERI: DISCRETI | CONTINUI

DISCRETI

* $\exists\ S_X = \{x_1, x_2, \dots\}$
SPETTRO

$p_1 = \mathbb{P}(X = x_1) > 0$
$p_2 = \mathbb{P}(X = x_2) > 0$
$\vdots$
$p_n = \mathbb{P}(X = n) > 0$
$= 1$

* $\mathbb{E}(X) = \sum_{x_k \in S_X} x_k \cdot \mathbb{P}(X = x_k)$

CONTINUI

$\exists\ f_X \geq 0$ su $\mathbb{R}$ tale che:

$F_X(x) = \int_{-\infty}^{x} f(y)\, dy$;

$= f_X(x) = \frac{d}{dx} F_X(x)$

$\mathbb{P}(X \in B) = \int_B f_X(x)\, dx$

$\mathbb{E}(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx =$

$= \int_{S_X} x \cdot f_X(x)\, dx$

Vale per entrambi, invece: $\mu_X = \mathbb{E}(X)$

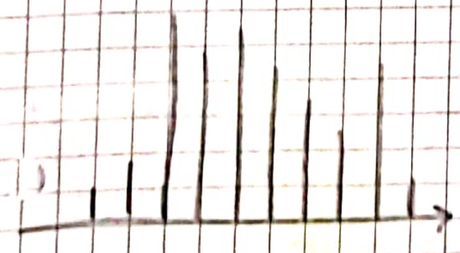
VARIANZA
$$\left[\begin{array}{l} = \mathbb{E}(aX+b) = a \cdot \mathbb{E}(X) + b \\ = D^2(X) = \mathbb{E}[((X-\mu_X)^2)] < +\infty \\ = \mathbb{E}(X^2) - \mathbb{E}^2(X) \end{array}\right.$$

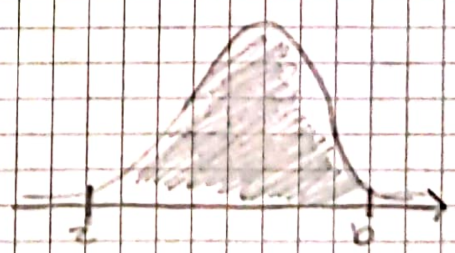
In sostanza, la differenza che intercorre tra questi due tipi di variabili/numeri aleatori è che nei numeri discreti ci sono punti dell'asse x che sono maggiori di 0, invece nei continui tutti i numeri sono maggiori di 0

Una variabile quantitativa si dice DISCRETA quando può assumere un insieme finito o NUMERABILE di valori, mentre si dice CONTINUA quando può assumere tutti i valori compresi in un intervallo reale

VARIABILI ALEATORIE

DISCRETE | CONTINUE





VARIABILE ALEATORIA CONTINUA ESPONENZIALE

* $X \sim Esp(\alpha)$, $\alpha > 0$ ← X distribuita in maniera esponenziale

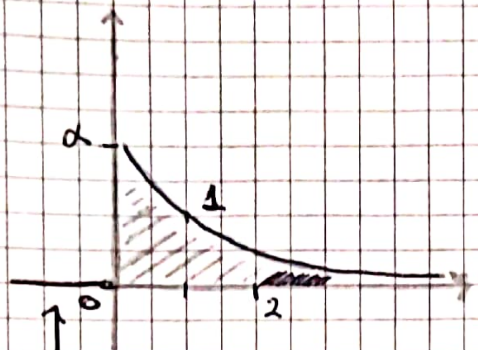
$$f_X(x) = \begin{cases} \alpha \cdot e^{-\alpha x}, & x \geq 0 \\ 0, & \text{altrimenti} \end{cases}$$

→ $f(x;\alpha)$ = funzione di densità col α parametro di scala

* descrive la "durata di vita" di un fenomeno che non invecchia ⇒ ossia la distribuzione esponenziale è PRIVA DI MEMORIA

↓

può rappresentare il tempo di attesa prima che si verifichi un certo evento casuale



La densità è nulla ai fini probabilistici

- $P(X>2) = \int_2^{+\infty} \alpha e^{-\alpha x}\, dx = -e^{-\alpha x}\Big|_2^{+\infty} = e^{-\alpha x}\Big|_{+\infty}^{2} =$

$= e^{-2\alpha} - \underbrace{\lim_{x\to\infty} e^{-\alpha x}}_{0} = e^{-2\alpha}$

- $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(y)\, dy \overset{\boxed{x<0}}{=} 0$ CASO X NEGATIVO

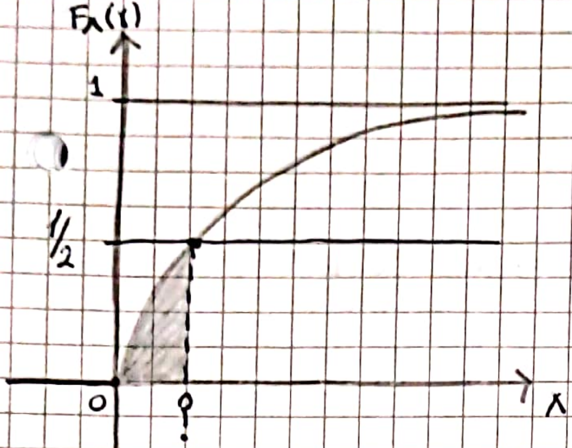
$\overset{\boxed{x>0}}{=} \int_0^x \alpha e^{-\alpha y}\, dy = e^{-\alpha y}\Big|_x^0 = 1 - e^{-\alpha x}$ CASO X POSITIVO

$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\alpha x}, & x > 0 \\ 0, & \text{altrimenti} \end{cases}}$$

- $E(X) = \int_{\mathbb{R}} x \cdot f_X(x)\, dx = \int_0^{+\infty} x \cdot \alpha e^{-\alpha x}\, dx$

per integrazione per parti $= \underbrace{-e^{-\alpha x} \cdot x\Big|_0^{+\infty}}_{0} + \int_0^{+\infty} e^{-\alpha x}\, dx = 1/\alpha \cdot \underbrace{\int_0^{+\infty} \alpha e^{-\alpha x}\, dx}_{1}$

$= 1/\alpha$



$\mathbb{E}(X) = 1/\alpha$]

preso un punto sull'asse verticale, questo punto intercetterà la curva esponenziale UNA sola volta $\longrightarrow$ $F_X(?) = \frac{1}{2}$

$$= 1 - e^{-\alpha \cdot ?} = \frac{1}{2} \iff e^{-\alpha \cdot ?} = \frac{1}{2} - 1$$

$$= -e^{-\alpha \cdot ?} = -\frac{1}{2} \iff e^{-\alpha ?} = \frac{1}{2} \iff$$

$$= \ln(e^{-\alpha ?}) = -\ln(2) \iff -\alpha \cdot ? = -\ln(2) \iff$$

$$= \alpha \cdot ? = \ln(2) \iff ? = \frac{1}{\alpha} \cdot \ln(2)$$

Questo valore è chiamato MEDIANA]

[N.B. $\longrightarrow$ Nel caso uniforme MEDIA e MEDIANA coincidono]

- $$\phi_X(\tau) := \mathbb{E}(e^{\tau x}) = \int_0^{+\infty} e^{\tau x} \cdot f_X(x)\,dx$$

$$= \int_0^{+\infty} e^{\tau x} \cdot \alpha e^{-\alpha x}\,dx = \int_0^{+\infty} \alpha e^{-x(\alpha-\tau)}\,dx =$$

$$= \frac{\alpha}{\alpha-\tau} \underbrace{\int_0^{+\infty} (\alpha-\tau) e^{-(\alpha-\tau)x}\,dx}_{1} \Rightarrow \boxed{\phi_X(\tau) = \frac{\alpha}{\alpha-\tau}}$$

moltiplico e divido per $\alpha-\tau$

$\longrightarrow \tau < \alpha \iff \alpha-\tau > 0$

PROPRIETÀ ESPONENZIALE

"ASSENZA DI USURA" $\longrightarrow$ una variabile casuale esp. non ricorda il passato ma si comporta come se fosse nuova

- $\mathbb{P}(X > s+t \mid X > t) = \mathbb{P}(X > s)$

$$= \frac{\mathbb{P}(X>s+t,\ X>t)}{\mathbb{P}(X>t)} = \frac{\mathbb{P}(X>t+s)}{\mathbb{P}(X>t)} = \frac{1-F_X(s+t)}{1-F_X(t)}$$

$$= \frac{e^{-\alpha(s+t)}}{e^{-\alpha t}} = \frac{e^{-\alpha s}\,\cancel{e^{-\alpha t}}}{\cancel{e^{-\alpha t}}} = e^{-\alpha s} = 1 - F_X(s) = \mathbb{P}(X>s)$$

LEZIONE 19

RIEPILOGO

- Legge di distribuzione di un numero aleatorio

DISCRETO

- Bernoulli $B(p)$, $p \in (0,1)$
- Binomiale $B(n,p)$, $n \in \mathbb{N}$, $p \in (0,1)$
- Geometrica $G(p)$, $p \in (0,1)$
- Poisson $\Pi(\lambda)$, $\lambda > 0$
- Pascal $W(k,p)$, $k \in \mathbb{N}$, $p \in (0,1)$

CONTINUO

- Uniforme discreto $U(x_1, \dots, x_n)$, $x_1, \dots, x_n \in \mathbb{R}$
- Uniforme $U(a,b)$, $a < b \in \mathbb{R}$
- Esponenziale $Esp(\alpha)$, $\alpha > 0$, $[0,+\infty[$ supporto

- FUNZIONE DI GAUSS

↓

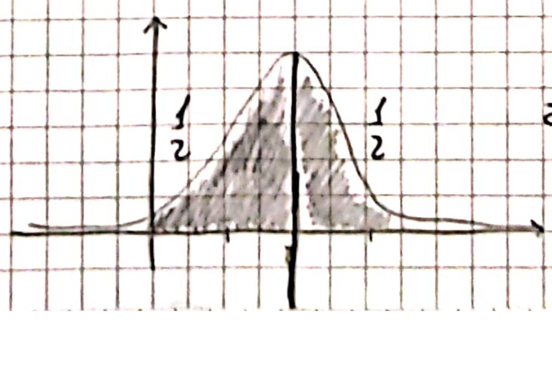
è la funzione d'onda dello stato fondamentale dell'oscillatore atomico quantistico

- funzione di densità di probabilità GAUSSIANA

$\forall x \in \mathbb{R}$, $f_X(x, a, b^2) = \frac{1}{\sqrt{2\pi b^2}} e^{-(x-a)^2/2b^2} > 0$

$a \in \mathbb{R}$, $b^2 > 0$

$]-\infty, +\infty[$ supporto funzione gaussiana



area $= \frac{1}{2} + \frac{1}{2} = 1$

- $\mathbb{E}(X) = \int_{-\infty}^{+\infty} x\, f_X(x, a, b)\, dx]$ con $a=0$, $b^2=1 \Rightarrow b=1$

un numero gaussiano con $a=0$ e $b^2=1$ si chiama STANDARD

$a \to \mu$, $b^2 \to \sigma^2$

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

- si prende $\tau \in \mathbb{R}$, $\phi_Z(\tau) = \mathbb{E}(e^{\tau Z}) = \int_{-\infty}^{+\infty} e^{\tau x} f_Z(x)\, dx =$

$$= \int_{-\infty}^{+\infty} e^{\tau x} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}\, dx$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x^2 - 2\tau x)}\, dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x^2 - 2\tau x + \tau^2 - \tau^2)}\, dx$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}[(x^2 - 2\tau x + \tau^2) - \tau]}\, dx$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x-\tau)^2} \cdot e^{\frac{1}{2}\tau^2}\, dx =$$

$$= e^{\frac{1}{2}\tau^2} \cdot \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x-\tau)^2}\, dx}_{} =$$

$\to$ funzione densità gaussiana di $N(\tau, 1)$ e vale 1

$$= e^{\frac{1}{2}\tau^2} \cdot \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x-\tau)^2}\, dx = e^{\frac{1}{2}\tau^2}$$

$\to$ con $\tau = 0$ si ha $\phi_Z(0) = \mathbb{E}(e^{0Z}) = \mathbb{E}(1) = 1 = \mathbb{E}(Z^0)$

$$\mathbb{E}(Z) = d\phi_Z(\tau)/dx \Big|_{\tau=0} = \phi_Z(\tau) \cdot \tau \Big|_{\tau=0} = \tau \Big|_{\tau=0} = 0$$

$$\mathbb{E}(Z^2) = d^2\phi_Z(\tau)/d\tau^2 \Big|_{\tau=0} = \phi_Z(\tau)' \cdot \tau + \phi_Z(\tau) \cdot 1 \Big|_{\tau=0} =$$

$$= \phi_Z(0) = 1$$

$\Rightarrow D^2(Z) = \mathbb{E}(Z^2) - \mathbb{E}^2(Z) = 1 - 0 = 1$

risultato finale $\longrightarrow$ $Z \sim N(0,1)$

$\mathbb{E}(Z) \leftarrow$ $\quad \to D^2(Z)$

OPERAZIONE DI STANDARDIZZAZIONE

- $X \sim N(a, b^2)$ → scarto dalla media: $X - \mathbb{E}(X) / \sqrt{\mathbb{D}^2(X)}$

PROPRIETÀ

a) $\mathbb{E}\left[\frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{D}^2(X)}}\right] = 1/\sqrt{\mathbb{D}^2(X)} \cdot [\mathbb{E}(X) - \mathbb{E}(X)] = 0$ MEDIA NULLA

b) $\mathbb{D}^2\left[\frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{D}^2(X)}}\right] = 1^*$ CALCOLARE LA VARIANZA VARIANZA UNITARIA

La standardizzazione riconduce una variabile aleatoria distribuita con μ e σ^2 a una v.a. con distribuzione standard dove $\mu = 0$ e $\sigma^2 = 1$

- X ammette finita la varianza:

$y = X + c$, con $c \in \mathbb{R}$

$\mathbb{D}^2(y) = \mathbb{E}\{[y - \mathbb{E}(y)]^2\} = \mathbb{E}\{[X + c - \mathbb{E}(X+c)]^2\}$

$= \mathbb{E}[(X + \cancel{c} - \mathbb{E}(X) - \cancel{c})]^2 = \mathbb{E}[(X - \mathbb{E}(X))]^2$

$= \mathbb{D}^2(X)$

Sia ora $d \in \mathbb{R}$, $T = X/d$, allora

$\mathbb{D}^2(T) = \mathbb{E}\{[T - \mathbb{E}(T)]^2\} = \mathbb{E}\{[X/d - \mathbb{E}(X)/d]^2\}$

$= \mathbb{E}\{[X - \mathbb{E}(X)/d]^2\} = \mathbb{E}\{[X + \mathbb{E}(X)]^2/d\}$

$= 1/d^2 \cdot \mathbb{E}\{[X - \mathbb{E}(X)]^2\} = \frac{\mathbb{D}^2(X)}{d^2}$

ESEMPI

- $\mathbb{D}^2\left(\frac{X-2}{3}\right) = \frac{1}{9}\mathbb{D}^2(X-2) = \frac{1}{9}\mathbb{D}^2(X)$
- $\mathbb{D}^2(4X) = 16\,\mathbb{D}^2(X)$

→ In generale è $\mathbb{D}^2(cX + d) = \mathbb{D}^2(cX) = c^2 \cdot \mathbb{D}(X)$

Ritornando a b: $\mathbb{D}^2\left[\frac{X - \mathbb{E}(X)}{\sqrt{\mathbb{D}^2(X)}}\right] = \frac{1}{\mathbb{D}^2(X)} \cdot \mathbb{D}^2[X - \mathbb{E}(X)] =$

$= \frac{\mathbb{D}^2(X)}{\mathbb{D}^2(X)} = 1^*$

Ritornando al caso Gaussiano:

$Z \sim N(0,1) \longrightarrow \left[Y = \dfrac{X - \mathbb{E}(X)}{D^2(X)} \right] \sim N(0,1)$

$\dfrac{X-\mu}{\sigma}$

- Si prende ora $Z \sim N(a,b)$ con $b>0 \Rightarrow a+bZ$

$$\begin{cases} \mathbb{E}(a+bZ) = a + b\,\cancel{\mathbb{E}(Z)}^{0} = a \\ D^2(a+bZ) = D^2(bZ) = b^2 \cdot \cancel{D^2(Z)}^{1} = b^2 \end{cases}$$

LEZIONE 20

ESEMPIO 5.5.1

Sia X una variabile aleatoria normale $\longrightarrow X \sim N(3, 16)$ (μ, σ^2)

Si trovino (a): $\mathbb{P}(X < 11)$
(b): $\mathbb{P}(X > -1)$
(c): $\mathbb{P}(2 < X < 7)$

$Z \sim N(0,1)$

- $x \in \mathbb{R}$, $f_X(x) = \dfrac{1}{\sqrt{2\pi 16}} \cdot e^{-(x-3)^2/2\cdot 16}$

- $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy = \mathbb{P}(X \leq x)$

OP. STANDARDIZZAZIONE

$Y = \dfrac{X-3}{\sqrt{16}} = \dfrac{X-3}{4} \sim Z$

a) $\mathbb{P}(X < 11) = \mathbb{P}(X - 3 < 8) = \mathbb{P}(X-3 < 11-3)$

$= \mathbb{P}\left(\dfrac{X-3}{4} < \dfrac{11-3}{4}\right) = \mathbb{P}(Y < 2) \sim Z$

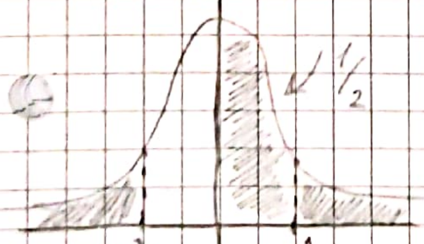
$\mathbb{P}(Y<2) = 0,9772 \longleftarrow$ Tabella distribuzione normale standard

b) $\mathbb{P}(X > -1) = 1 - \mathbb{P}(X \leq -1) = 1 - \mathbb{P}(Y \leq -1)$

$\dfrac{-1-3}{4} = -1$

$= 1 - [0,5 - \mathbb{P}(0 < Y < 1)] = 0,5 + \mathbb{P}(0 < Y < 1)$

$\simeq 0,84$



c) $P(2 < X < 7) = P(-1/4 < y < 1) = F_z(1) - F_z(-1/4)$

$= P(-1/4 < y \le 0) + P(0 < y < 1)$

$0{,}5987 + 0{,}34 \simeq 0{,}9387$

$\frac{2-3}{4} = -\frac{1}{4}$ $\qquad$ $\frac{7-3}{4} = 1$

punti standardizzati

ESERCIZIO 5.5.3

La potenza W dissipata da una resistenza è proporzionale al quadrato della differenza di potenziale V ai suoi capi.

$W = rV^2$ con r costante $\longrightarrow$ $V \sim N(6,1)$

Sia $r = 3$ e si suppone che V sia normale di media 6 e deviazione standard 1

(a) $E(W)$

(b) $P(W > 120)$

(b) $P(W > 120) = P(3V^2 > 120) = P(V^2 > 40)$

$= P(V > \sqrt{40}) = P(V - 6 > \sqrt{40})$

0.5 $\quad$ $= P(y \le \sqrt{40} - 6) = 1 - P(y \le \sqrt{40} - 6) =$

$y = \frac{V-6}{1^2}$ $\quad$ $= P(y \le 0{,}3246) \simeq 0{,}3727$

$= V - 6$

(a) $E(W) = 3 \cdot E(V^2) = \underline{E(V^2) + D^2(V)} = 3 \cdot (36+1) = 3 \cdot 37 = 111$ w

$E(V) = 6, \; D^2(V) = 1$

$D^2(V) = E(V^2) - E^2(V)$

$Z \sim N(0,1)$ $\qquad$ $F_Z(x) = \int_{-\infty}^{x} \frac{1}{2\pi} \cdot e^{-y^2/2} \, dy$

$y = z^2$

$$F_y(y) = \begin{cases} 0, & y \le 0 \\ F_y(\sqrt{y}) - \frac{1}{2}, & y > 0 \end{cases}$$

$y > 0 \longrightarrow \mathbb{P}(Y \le y) = \mathbb{P}(Z^2 \le y) = \mathbb{P}(-\sqrt{y} \le Z \le \sqrt{y})$

$= 2\mathbb{P}(0 \le Z \le \sqrt{y}) = F_Z(\sqrt{y}) - \frac{1}{2}$

$F_Y(y) = F_Z(\sqrt{y}) - F_Z(-\sqrt{y})$

$f_Y(y) = f_Z(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} + f_Z(-\sqrt{y}) \cdot \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{y}} f_Z(\sqrt{y}) =$

$= \frac{1}{\sqrt{y}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{-y/2}$

OSSERVAZIONE 4.2.1

Quando si è a conoscenza della funzione di massa di probabilità di una variabile aleatoria discreta, oppure la funzione di densità di probabilità di una continua, oppure ancora quando si conosca la funzione di ripartizione/distribuzione di una variabile aleatoria qualsiasi, si hanno abbastanza informazioni da poter calcolare la probabilità di ogni evento che dipende solo da tale variabile aleatoria.
In questo caso si può dire che si conosce la distribuzione o legge della variabile aleatoria considerata $\square$

ESEMPIO 5.1.1

Una azienda produce dischetti per PC che sono difettosi con probabilità 0.01
Questi dischetti sono poi venduti in confezioni da 10 pezzi, con la garanzia di rimborso in caso vi sia più di un pezzo difettoso
Che percentuale delle confezioni viene ritornata?
Se si comprano tre confezioni, qual è la probabilità di ritornarne esattamente una?

X: numero di pezzi difettosi in una scatola da 10

$X \sim B(n,p) \Rightarrow \begin{matrix} n = 10 \\ p = 0{,}01 \end{matrix} \Rightarrow X \sim B(10, 0.01)$

$\mathbb{P}(X > 1) = 1 - \mathbb{P}(X=0) - \mathbb{P}(X=1)$

$= 1 - \binom{10}{0} \cdot 0.01^0 \cdot 0.99^{10} - \binom{10}{1} \cdot 0.01^1 \cdot 0.99^9$

ESEMPIO 5.4.1

Sia X una variabile aleatoria uniforme nell'intervallo $[0, 10]$
Si trovino le probabilità che

(a) : $2 < X < 9$
(b) : $1 < X < 4$
(c) : $X < 5$
(d) : $X > 6$

$X \sim U(0, 10)$

a) - $\mathbb{P}(2 < X < 9) = \mathbb{P}(X > 2) \cup \mathbb{P}(X < 9)$

$$= \frac{1}{10-0} \int_2^9 dx = \frac{9-2}{10-0} = \frac{7}{10}$$

b) - $\mathbb{P}(1 < X < 4) = \mathbb{P}(X > 1) \cup \mathbb{P}(X < 4)$

$$= \frac{1}{10-0} \int_1^4 dx = \frac{4-1}{10-0} = \frac{3}{10}$$

c) - $\mathbb{P}(X < 5)$

$$= \frac{1}{10-0} \int_0^5 dx = \frac{5-0}{10-0} = \frac{5}{10} = \frac{1}{2}$$

d) - $\mathbb{P}(X > 6)$

$$= \frac{1}{10-0} \int_6^{10} dx = \frac{10-6}{10-0} = \frac{4}{10} = \frac{2}{5}$$

ESEMPIO 5.4.2

Ad una certa fermata passa un autobus ogni 15 minuti a cominciare dalle 7. Se un passeggero arriva alla fermata in un momento casuale con distribuzione uniforme tra le 7 e le 7.30
Si calcoli con che probabilità dovrà aspettare il prossimo autobus per (a) meno di 15 minuti, (b) almeno 12 minuti

(a) - Sia X l'istante in cui questa persona arriva alla fermata

$X \sim U(0, 30)$

Siccome il passeggero deve aspettare meno di 5 minuti solo se arriva tra le 7.10 e le 7.15 oppure tra le 7.25 e le 7.30

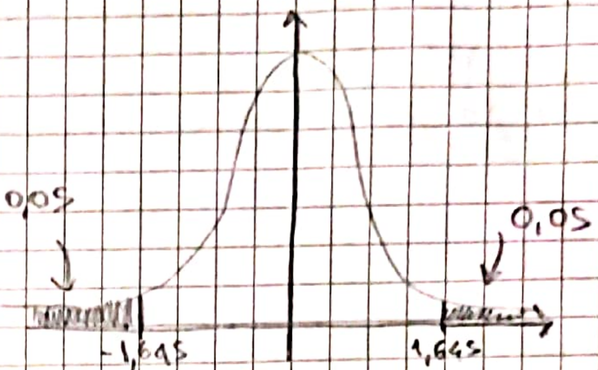
$$\mathbb{P}(10 < X < 15) + \mathbb{P}(25 < X < 30) = \frac{15-10}{30} + \frac{30-25}{30} = \frac{5}{30} + \frac{5}{30} = \frac{1}{3}$$

LEZIONE 21

QUANTILE SUPERIORE

$\forall \alpha \in (0,1)$, z_α è il quantile superiore della normale standard se accade che: $\mathbb{P}(Z \geq z_\alpha) = \alpha$

- In particolare:
  - $\alpha = 0{,}05$ e dalle Tavole di distribuzione gaussiana $z_{0,05} \cong 1{,}645$
  - $\alpha = 0{,}025 \Rightarrow z_{0,025} \cong 1{,}96$
  - $\alpha = 0{,}01$



ESEMPIO 4.3.1

Da un gruppo di 12 batterie, di cui 3 nuove, 4 usate e 5 difettose, ne vengono scelte tre a caso.
Siano X e Y rispettivamente il numero di batterie nuove e usate tra quelle scelte.
La funzione di massa di probabilità congiunta, $p(i,j)$ è data dai valori seguenti

$$p(0,0) = \binom{5}{3} \Big/ \binom{12}{3} = \frac{10}{220}$$

$S_X = \{0,1,2,3\}$
$S_Y = \{0,1,2,3\}$

$|\Omega| = 16$

$$p(0,2) = \mathbb{P}(X=2, Y=0) = \binom{4}{2} \cdot \binom{5}{2} \Big/ \binom{12}{3} = \frac{30}{220}$$

⋮

X: "numero di batterie nuove"
Y: "numero batterie usate"

$$p(0,0) = \mathbb{P}(X=0, Y=0) = \frac{\binom{5}{3} \longrightarrow \text{\# batt difettose}}{\binom{12}{3} \longrightarrow \text{\# tot batterie},\ \longrightarrow \text{\# batt scelte}}$$

$$p(1,0) = \mathbb{P}(X=1, Y=0) = \binom{3}{1} \binom{5}{2} \Big/ \binom{12}{3} = \frac{30}{220}$$

$$p(2,0) = \mathbb{P}(X=2, Y=0) = \frac{\binom{3}{2} \cdot \binom{5}{1}}{\binom{12}{3}} = \frac{15}{220}$$

Queste probabilità possono essere convenientemente presentate in forma tabellare

$X \backslash Y$	0	1	2	3
0	$p_{0,0}$	$p_{0,1}$	$p_{0,2}$	$p_{0,3}$
1	$p_{1,0}$	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$
2	$p_{2,0}$	$p_{2,1}$	$p_{2,2}$	$p_{2,3}$
3	$p_{3,0}$	$p_{3,1}$	$p_{3,2}$	$p_{3,3}$

$S_{X,Y} = \{(0,0), (0,1), (0,2), (0,3),$
$(1,0), (1,1), (1,2)$
$(2,0), (2,1)$
$(3,0)\}$

$X \backslash Y$	0	1	2	3	
0	10/220	40/220	30/220	4/220	→ 84
1	30/220	60/220	18/220	0	→ 108
2	15/220	12/220	0	0	→ 27
3	1/220	0	0	0	→ 1
	↓ 56	↓ 112	↓ 48	↓ 4	= 220

la somma di tutte le probabilità deve valere 1

= 220

$\mathbb{P}(X=0) = \mathbb{P}(\{X=0\} \cap \Omega) =$

$= \mathbb{P}(\{X=0\} \cap \{Y=0\} \cup \{Y=1\} \cup \{Y=2\} \cup \{Y=3\}) =$

$= \mathbb{P}(\{X=0\} \cap \{Y=0\} \cup \{X=0\} \cap \{Y=1\} \cup \{X=0\} \cap \{Y=2\} \cup$

$\cup \{X=0\} \cap \{Y=3\}) = \mathbb{P}(X=0, Y=0) + \mathbb{P}(X=0, Y=1) +$
$+ \mathbb{P}(X=0, Y=2) + \mathbb{P}(X=0, Y=3)$

↓

distribuzioni marginali dei due numeri aleatori in cui $X=0$

ESEMPIO 4.3.2

All'interno di una certa popolazione, il 15% delle coppie non ha figli, il 20% ne ha uno, il 35% ne ha due e il 30% ne ha tre

Inoltre ogni bambino, indipendentemente da tutti gli altri può essere maschio o femmina con pari probabilità.
Se si seleziona una famiglia a caso e si denotano con X e Y il numero di femmine e di maschi presenti tra i figli in tale famiglia si ottiene la funzione di massa di probabilità mostrata in 1.2

Y: "Numero figli maschi"

X: "Numero figli femmine"

$P(X=0, Y=0) = P(\text{nessun figlio}) = 0{,}15$
$P(X=1, Y=0) = P(1 \text{ figlio, femmina}) = P(1 \text{ figlio}) P(1 \text{ femmina} \mid 1 \text{ figlio})$
$= 0{,}20 \times 0{,}5 = 0{,}1$
$P(X=2, Y=0) = P(2 \text{ figli, femmine}) = 0{,}35 \times 0{,}5^2 = 0{,}0875$
$P(X=3, Y=0) = P(3 \text{ figli, femmine}) = 0{,}30 \times 0{,}5^3 = 0{,}0375$

TABELLA COMPLETA 1.2

	0	1	2	3	
0	0,15	0,10	0,09	0,03	→ 0,3750
1	0,10	0,17	0,11	0	→ 0,3875
2	0,09	0,11	0	0	→ 0,2
3	0,03	0	0	0	→ 0,03
	↓ 0,37	↓ 0,38	↓ 0,2	↓ 0,03	

FUNZIONE DI DENSITÀ DI PROBABILITÀ CONGIUNTA

due variabili aleatorie sono congiuntamente continue se esiste una funzione non negativa $f(x,y)$ definita per tutti gli x e y, avente la proprietà tale che: preso $B \in \mathcal{B}^2$, $P([X,Y]) \in B =$

$$= \iint_B f_{X,Y}(x,y)\,dx\,dy$$

$\Rightarrow \quad X, Y \quad \Omega \longrightarrow \mathbb{R}^2$
$\qquad\qquad\quad \omega \rightsquigarrow (X(\omega), Y(\omega))$

Borel

$\xi : (X, Y) : (\Omega, F, P) \xrightarrow{F\text{-mis}} (\mathbb{R}^2, \mathcal{B}^2, P_{X,Y})$

$B \in \mathcal{B}, \; P_{X,Y}(B) = P[(X,Y) \in B]$

$B =]-\infty, x] \times]-\infty, y] \longrightarrow F_{X,Y}(x,y) = P(X \le x, Y \le y)$
$= \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x,y)\,dx\,dy$

$$\Rightarrow F_{X,Y}(x,y) = \mathbb{P}(X \le x, Y \le y)$$
$$= \mathbb{P}(X \in B, Y \in B)$$
$$= \int_B \int_B f(x,y)\,dx\,dy$$
$$= \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)\,dx\,dy \qquad f(x,y) = \frac{\partial^2 F(x,y)}{\partial a\,\partial b}$$

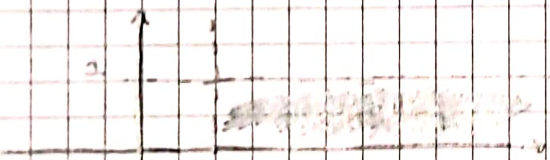
in tutti i punti in cui le derivate parziali sono definite

- $\forall x \in \mathbb{R},\ f_X(x)\,\cancel{dx} = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,\cancel{dx}\,dy$ → **DENSITÀ MARGINALI**
- $\forall y \in \mathbb{R},\ f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\,dx$

ESEMPIO 4.3.3

Siano X e Y due variabili aleatorie congiuntamente continue con densità di prob congiunta data da

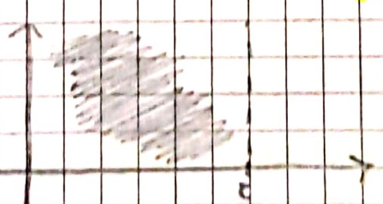
$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & x>0,\ y>0 \\ 0 & \text{altrimenti} \end{cases}$$



- $\mathbb{P}(X>1, Y<1) = ?$

$$= \int_0^1 dy \int_1^{\infty} 2e^{-x}e^{-2y}\,dx = 2\int_0^1 e^{-2y}\,dy \int_1^{+\infty} e^{-x}\,dx$$
$$= 2\int_0^1 e^{-2y}\,dy \cdot \left[e^{-x}\Big|_{+\infty}^{1}\right] = 2e^{-1}\int_0^1 e^{-2y}\,dy$$
$$= \cancel{2}e^{-1}\left[\frac{1}{\cancel{2}}e^{-2y}\Big|_1^0\right] = e^{-1}(e^{-2}-1)$$

- $\mathbb{P}(X<2) = ?$



La variabile aleatoria Y può assumere qualsiasi valore, quindi y si integra su tutto $\mathbb{R}$

$$= \int_0^{2} e^{-x} \left(\int_0^{+\infty} 2e^{-x} e^{-2y} \, dx \, dy \right) =$$

$$= 2 \int_0^{2} e^{-x} \int_0^{+\infty} e^{-2y} \, dy = 1 - e^{-2}$$

Ritornando al caso discreto

	0	1	
0	1/4	1/4	→ 1/2
1	1/4	1/4	→ 1/2
	↓ 1/2	↓ 1/2	

$P(X=0, Y=0) = 1/4$
$P(X=0) = 1/2$
$P(Y=0) = 1/2$ → eventi INDIPENDENTI

$P(X=0) \cdot P(Y=0) = \frac{1}{4}$

LEZIONE 22

- Riepilogo

Distribuzione / legge CONGIUNTA

CASO: DISCRETO

$(x,y) \in S_{(x,y)}, \; p_{(x,y)}(x,y) > 0$

$- \sum_{(x,y) \in S_{(x,y)}} p(x,y) = 1$

$- (x,y) \in S_{(x,y)}$

$p_{(x,y)}(x,y) = p_X(x) \cdot p_Y(y)$

$\Updownarrow$

X e Y sono INDIPENDENTI

CONTINUO

$(x,y) \in \mathbb{R}, \; f_{(x,y)}(x,y) \geq 0$

$\forall B \in \mathcal{B}, \; \int_B f_{(x,y)}(x,y) \, dx \, dy$

$\forall x, y \in \mathbb{R}, \; F_{(x,y)}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t) \, ds \, dt$

$\int_{\mathbb{R}^2} f_{X,Y}(x,y) \, dx \, dy = 1$

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

$\Updownarrow$

$F_{(x,y)}(x,y) = F_X(x) \cdot F_Y(y)$

ESEMPIO 4.3.5

$$f(x,y) = \begin{cases} 2e^{-x} e^{-2y}, & x>0, y>0 \\ 0, & \text{altrimenti} \end{cases} \Rightarrow F(x,y) = e^{-x} \cdot 2e^{-2y}$$

$\Downarrow$

$X \sim Esp(1) \wedge Y \sim Esp(2)$

4.3.5 DISTRIBUZIONE CONDIZIONALE

partendo dall'esempio 4.3.2

X: numero di femmine
Y: numero di maschi

X \ Y	0	1	2	3	
0	0,15	0,1	0,08	0,03	→ 0,37
1	0,1	0,17	0,11	0	→ 0,38
2	0,09	0,11	0	0	→ 0,2000
3	0,03	0	0	0	→ 0,03
	↓ 0,37	↓ 0,38	↓ 0,2	↓ 0,037	

la famiglia selezionata ha 1 figlia

$$\mathbb{P}(Y=0 \mid X=1) = \frac{\mathbb{P}(Y=0, X=1)}{\mathbb{P}(X=1)} = 0{,}1000 / 0{,}3875 = 0{,}2581$$

$$\mathbb{P}(Y=1 \mid X=1) = \frac{0{,}1750}{0{,}3875} = 0{,}4516$$

$$\mathbb{P}(Y=2 \mid X=1) = \frac{0{,}1125}{0{,}3875} = 0{,}29$$

$$\mathbb{P}(Y=3 \mid X=1) = \frac{0}{0{,}3875} = 0$$

→ in questo caso si parla di distribuzione condizionale

Si ricorda che presi due eventi E, F con $\mathbb{P}(F)>0$, la probabilità di E condizionata a F è data dall'espressione

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

Siano X e Y due variabili aleatorie discrete con funzione di massa di probabilità congiunta $p(\cdot,\cdot)$. Si dice funzione di massa di probabilità condizionata di X dato Y e si indica con $p_{X|Y}(\cdot,\cdot)$ la funzione così definita:

$$p_{X|Y}(x|y) = \mathbb{P}(X=x \mid Y=y) = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(Y=y)} = p(x,y) / p_Y(y)$$

$p(0,0) = 0{,}4 \qquad p(1,0) = 0{,}1$

$p(0,1) = 0{,}2 \qquad p(1,1) = 0{,}3$

$\longrightarrow$ data la presenza di una fialz, vi sono 23 possibilità su 31 che vi sia rimasto un maschio

DISTRIBUZIONE / LEGGE CONDIZIONATA

$f_{X,Y}(x,y)$

$B \in \mathcal{B} \in \mathcal{B}^2 \qquad \mathbb{P}(X,Y \in B) = \int_B f_{X,Y}(x,y)\,dx\,dy$

$x \in \mathbb{R} \quad (x, y_0) \qquad \ldots \qquad f_X(x) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dy$

$y \in \mathbb{R}, \ (x_0, y) \qquad f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dx$

$f_{X|Y}(x|y) = f_{X,Y}(x,y) / f_Y(y) \quad ; \quad f_{Y|X}(y|x) = f_{X,Y}(x,y) / f_X(x)$

ESEMPIO 4.3.P

È data la seguente densità congiunta di X e Y

$$f(x,y) = \begin{cases} \frac{12}{5} x (2-x-y) & , \ 0<x<1,\ 0<y<1 \\ 0 & , \ \text{altrimenti} \end{cases}$$

Si calcoli la densità condizionata di X rispetto a Y = y per $0<y<1$

$y_0 \in (0,1)$

$$f_{X|Y}(x, y_0) = f_{X,Y}(x,y_0) / f_Y(y_0) = \frac{\frac{12}{5} x (2-x-y_0)}{\boxed{\int_{\mathbb{R}} f_{X,Y}(x,y_0)\,dx}}$$

$$\int_0^1 2{,}4 \cdot x(2-x-y)\,dx$$

$$= \frac{\frac{12}{5} x (2-x-y_0)}{\frac{8}{5} - \frac{6}{5} y}$$

$$= 2{,}4 \left[\int_0^1 2x\,dx - \int_0^1 x^2 - y \int_0^1 x\,dx \right. \longrightarrow$$

$$= 2{,}4\left[x^2\Big|_0^1 - \frac{x^3}{3}\Big|_0^1 - \frac{x^2}{2}\Big|_0^1\right] =$$

$$= 2{,}4\left(1 - \frac{1}{3} - \frac{y}{2}\right) = 2{,}4\cdot\left(\frac{2}{3} - \frac{y}{2}\right) = \frac{12}{5}\left(\frac{2}{3} - \frac{y}{2}\right) =$$

$$= \frac{8}{5} - \frac{6}{5}y$$

VALORE ATTESO

- discreto:

$$\mathbb{E}(X) = \sum_{x \in S_X} x_i \cdot p_X(x_i) = \sum_{x_i \in S_X} x_i \, \mathbb{P}(X = x_i) \qquad 4.4.1$$

Si tratta della media pesata dei possibili valori di X

- il valore atteso di X è definito solo se la serie 4.4.1 converge in valore assoluto, ovvero deve valere

$$\sum_{x \in S_X} |x_i| \, \mathbb{P}(X = x_i) < +\infty$$

Se la serie diverge allora si dice che X non ha valore atteso

- continuo: **DEF. 4.4.2**

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \, f_X(x) \, dx \iff \mathbb{E}(|X|) < +\infty \cdot \int_{\mathbb{R}} |x| \, f_X(x) \, dx$$

ESEMPIO

$X \sim U(0, 1.5)$

$$f_X(x) = \begin{cases} \frac{1}{1.5}, & 0 < x < 1.5 \\ 0, & \text{altrimenti} \end{cases}$$

$$\mathbb{E}(X) = \int_0^{1,5} x \cdot \frac{1}{1,5} \, dx = \frac{1}{1,5} \cdot \frac{1}{2} x^2 \Big|_0^{1,5} = \frac{2,25}{3} = 0,75$$

$\Rightarrow$ $\mathbb{E}(X)$ ha la stessa unità di misura della variabile aleatoria X

Corollario 4.5.2

Per ogni coppia di costanti reali a e b, sia $g(X) = aX + b$

allora, X ammette la media => g(X) ammette la media

- $\mathbb{E}(aX+b) = \int_{\mathbb{R}} (ax+b)\, f_X(x)\, dx$

$= \int_{\mathbb{R}} [ax \cdot f_X(x) + b\, f_X(x)]\, dx$

$= \int_{\mathbb{R}} ax \cdot f_X(x)\, dx + \int_{\mathbb{R}} b\, f_X(x)\, dx =$

$= a \int_{\mathbb{R}} x \cdot f_X(x)\, dx + b \cdot \int_{\mathbb{R}} f_X(x)\, dx =$

$= a\mathbb{E}(X) + b\,\mathbb{E}(1) = a\mathbb{E}(X) + b$

$$\mathbb{E}[aX+b] = a\mathbb{E}[X] + b \qquad \square$$

- Siano X e Y due numeri aleatori congiuntamente distribuiti e $f_{(X,Y)}(x,y)$ funzione di densità di prob congiunta

- Sia g una qualunque funzione di due variabili, allora esiste $\mathbb{E}[g(X,Y)]$

$$\mathbb{E}[g(X,Y)] = \begin{cases} \sum_x \sum_y g(x,y)\,p(x,y) & \text{CASO DISCRETO} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)\, f(x,y)\, dx\, dy & \text{CASO CONTINUO} \end{cases}$$

$\mathbb{E}(X+Y) = \int_{\mathbb{R}} (x+y) \cdot f_{X,Y}(x,y)\, dx\, dy =$

$-\infty \qquad\qquad +\infty$

$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x+y)\, f_{X,Y}(x,y)\, dx\, dy =$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x \cdot f_{X,Y}(x,y)\, dx\, dy + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{X,Y}(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{+\infty} x \left[\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dy \right] dx + \int_{-\infty}^{+\infty} y \left[\int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx \right] dy =$$

$$= \underbrace{\int_{-\infty}^{+\infty} x\, f_X(x)\, dx + \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy}_{\text{funzione di densità marginale una volta per } x \text{ e una volta per } y} = \mathbb{E}(X) + \mathbb{E}(Y)$$

funzione di densità marginale una volta per x e una volta per y

ESEMPIO 4.5.5.

Un'impresa edile ha recentemente sottoposto i suoi preventivi per tre gare, per degli appalti che le darebbero profitti per 10000, 20000, 40000 mila dollari.
Se le probabilità di vittoria su questi appalti sono 0,2, 0,8 e 0,3. Qual è il profitto totale medio che farà l'azienda?

X_1, X_2, X_3 i profitti percepiti per i tre lavori:

X_1	X_2	X_3
10000	20000	40000
0,2	0,8	0,3
0,8	0,2	0,7

$$\mathbb{E}(X_1 + X_2 + X_3) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3)$$

$$= (10000 \cdot 0{,}2) + (20000 \cdot 0{,}8) + (40000 \cdot 0{,}3)$$

$$= 2000 + 16000 + 12000 = 30000$$

profitto atteso del proprietario

LEZIONE 23

Media di somma di numeri aleatori

$$\mathbb{E}\left(\sum_{i=1}^{N} (X_i) \right) = \sum_{i=1}^{N} \left(\mathbb{E}(X_i) \right)$$

ESEMPIO 4.5.6

Una segretaria ha finito di scrivere una pila di N lettere, e ha appena compilato le buste con gli indirizzi, quando

Tutto il materiale le cade per terra e si mischia. Se si inseriscono le lettere nelle buste in maniera del tutto casuale, qual è il numero medio di lettere che capitano nelle buste corrette?
Sia X il numero di lettere che finiscono nelle buste giuste

X: numero di lettere nelle buste giuste

$\mathbb{E}(X)$?

$$i = 1, 2, \dots, n \qquad B(1, p_i) \sim X_i = \begin{cases} 1, & \text{se nella busta "}i\text{" c'è la lettera corrispondente} \\ 0, & \text{altrimenti} \end{cases}$$

$$X = \sum_{i=1}^{N} X_i \ ; \qquad \mathbb{E}(X) = \sum_{i=1}^{N} \mathbb{E}(X_i) = \sum_{i=1}^{N} p_i^{*} = \sum_{i=1}^{N} \mathbb{P}(X_i = 1)$$

$$* \ p_i = \mathbb{P}(X_i = 1) = \frac{1}{N}$$

$$\longrightarrow \frac{1}{N} \sum_{i=1}^{N} 1 = \frac{\cancel{N}}{\cancel{N}} = 1 \longrightarrow$$ indipendentemente dal numero di lettere presenti; in media vi sarà una sola lettera nelle buste giuste

ESEMPIO 4.5.7

In un prodotto commerciale vengono inseriti dei buoni sconto a regalo. Vi sono 20 tipi diversi di buoni, e in ogni confezione se ne trova uno qualsiasi con pari probabilità.
Se si aprono 10 confezioni, qual è il valore atteso del numero di tipi diversi di buoni sconto che si trovano?

X: "numero di tipi diversi di buoni che si trovano nelle 10 confezioni"

$\forall i = 1, 2, \dots, 20$ allora $X = X_1 + \dots + X_{20}$

$$X_i := \begin{cases} 1, & \text{se il tipo } i\text{-esimo di buoni è presente nelle 10 confezioni} \\ 0, & \text{altrimenti} \end{cases}$$

$$X = \sum_{i=1}^{20} X_i \Rightarrow \mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{20} X_i\right) = \sum_{i=1}^{20} \left(\mathbb{P}(X_i = 1)\right)$$

$$\left[\mathbb{P}(X_i = 1) = 20 \cdot \left[1 - \left(\frac{19}{20}\right)^{10}\right] \right]$$

$$= \sum_{i=1}^{20} \mathbb{P}(X_i = 1) = 20 \cdot \left[1 - \left(\frac{19}{20}\right)^{10}\right] \cong 8,025$$

OSSERVAZIONE 4.51

Vi è una interessante proprietà della media che emerge quando si vuole predire con il minore errore possibile il valore che verrà assunto da una variabile aleatoria.
Si suppone di voler predire il valore di X. Se si sceglie un numero reale c e si dice che X sarà uguale a c, il quadrato dell'errore che si commette è $(X-c)^2$

La Media è il MIGLIOR predittore di X rispetto all'errore quadratico medio

Si mostra di seguito che la media dell'errore al quadrato è MINIMIZZATA se per c si sceglie il valore della media di X.

- $\mu = \mathbb{E}[X]$

$$\mathbb{E}[(X-c)^2] = \mathbb{E}[(X-\mu+\mu-c)^2]$$
$$= \mathbb{E}[(X-\mu)^2 + 2(X-\mu)(\mu-c) + (\mu-c)^2]$$
$$= \mathbb{E}[(X-\mu)^2] + 2(\mu-c)\,\mathbb{E}[X-\mu] + (\mu-c)^2$$
$$= \mathbb{E}[(X-\mu)^2] + (\mu-c)^2 \longrightarrow \mathbb{E}(X-\mu) = \mathbb{E}[X] - \mu = 0$$
$$\geq \mathbb{E}[(X-\mu)^2] \longrightarrow = \mathbb{D}^2(X)$$
$$\longrightarrow c = \mathbb{E}(X) \quad \mathbb{D}^2(X)$$

Perciò la migliore previsione di X in termini di minimizzazione dell'errore quadratico medio, è il suo valore atteso.

$\square$

SCARTO ASSOLUTO MEDIO

- $\underset{c\in\mathbb{R}}{\operatorname{argmin}} (\mathbb{E}(|X-c|^2)) = \mathbb{E}(X) \longrightarrow$ MEDIA
- $\underset{c\in\mathbb{R}}{\operatorname{argmin}} (\mathbb{E}(|X-c|)) = M_e \longrightarrow$ MEDIANA
- $\underset{c\in\mathbb{R}}{\operatorname{argmin}} (\mathbb{E}(|X-c|^0)) = M_o \longrightarrow$ MODA

- Media = centro ordine 2
- Mediana = centro ordine 1
- Moda = centro ordine 0

VARIANZA

- Data una variabile aleatoria X, di cui sia nota la distribuzione, sarebbe molto utile se si potessero riassumere le caratteristiche fondamentali della sua distribuzione con quantità sintetiche come la media $\mathbb{E}[X]$.
  Tuttavia $\mathbb{E}[X]$ è il "baricentro" dei valori possibili di X e non coglie la variabilità di questi valori.

- Siano W, Y e Z così:

$$W := 0 \text{ con probabilità } 1 \longrightarrow \mathbb{E}(W)=0, \quad \mathbb{D}^2(W)=0$$

$$Y := \begin{cases} -1, & \text{con prob } 1/2 \\ 1, & \text{con prob } 1/2 \end{cases} \longrightarrow \mathbb{E}(Y)=0, \quad \mathbb{D}^2(Y)=1$$

$$Z := \begin{cases} -100, & \text{con prob } 1/2 \\ 100, & \text{con prob } 1/2 \end{cases} \longrightarrow \mathbb{E}(Z)=0, \quad \mathbb{D}^2(Z)=10000$$

Hanno media nulla, ma vi è molta più variabilità in Y che non in W, e ancora di più in Z

$$\left[\mathbb{D}^2[X] = \mathbb{E}\{[X-\mathbb{E}(X)]^2\}\right]$$

~~ESEMPIO~~ 4.6.1

Sia X una variabile aleatoria con media μ. La varianza di X che si denota con $Var(X)$ è la quantità:

$$Var(X) = \mathbb{E}[(X-\mu)^2] \quad \text{con } \mu = \mathbb{E}(X)$$

ESEMPIO 4.6.1

- Si calcoli la varianza del punteggio di un dado non truccato.
  Sia X il punteggio realizzato dal dado

$$X = Z \sim Ud(1,2,3,4,5,6) \longrightarrow Ud = \mathbb{P} = 1/n$$

$$\Rightarrow \mathbb{P}(X=i) = \frac{1}{6}$$

$$\mathbb{E}(Z) = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{1}{6}\sum_{i=1}^{6} i = \frac{1}{\cancel{6}} \cdot \frac{\cancel{6}\cdot 7}{2} = \frac{7}{2}$$

$$\mathbb{E}(Z^2) = \sum_{i=1}^{6} i^2 \cdot \frac{1}{6} = \frac{1}{6}\sum_{i=1}^{6} i^2 = (1+4+9+16+25+36)\cdot\frac{1}{6} = \frac{91}{6}$$

$$\Rightarrow Var(X) = \mathbb{E}[(X-\mu)^2] = \mathbb{E}(X)^2 - \mathbb{E}(X^2) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{91}{6} - \frac{49}{4}$$

$$= 35/12 \longrightarrow \left[7/2 - \sqrt{35/12},\ 7/2 + \sqrt{35/12}\right] \simeq (2,5)$$

- Una utile identità che riguarda la varianza è la seguente

$\hookrightarrow \forall a,b \in \mathbb{R} : Var(aX+b) = a^2 Var(X)$

DIM

$\mu = \mathbb{E}(X)$ e $\mathbb{E}[aX+b] = a\mathbb{E}[X]+b = a\mu+b$

$$\Rightarrow Var(aX+b) = \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2]$$
$$= \mathbb{E}[(aX+b-a\mu-b)^2]$$
$$= \mathbb{E}[a^2(X-\mu)^2]$$
$$= a^2 Var(X)$$

$\square$

- la quantità $\sqrt{Var(X)}$ è detta DEVIAZIONE STANDARD della variabile aleatoria X

LA COVARIANZA E LA VARIANZA DELLA SOMMA DI VARIABILI ALEATORIE

Siano assegnate due variabili aleatorie X e Y di media μ_x e μ_y rispettivamente.
La loro covarianza $Cov(X,Y)$ è la quantità:

4.7.1 $Cov(X,Y) := \mathbb{E}[(X-\mu_x)(Y-\mu_y)]$ con $\mu_x = \mathbb{E}(X)$ e $\mu_y = \mathbb{E}(Y)$

$$Cov(X,Y) = \mathbb{E}[XY - \mu_x Y - \mu_y X + \mu_x\mu_y]$$
$$= \mathbb{E}(XY) - \mu_y\mathbb{E}[X] - \mu_x\mathbb{E}[Y] + \mu_x\mu_y$$
$$= \mathbb{E}(XY) - \mu_x\mu_y \cancel{- \mu_x\mu_y + \mu_x\mu_y}$$
$$= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Dalla formula 4.7.1 si deducono alcune semplici proprietà, quali:

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

\+ $Cov(X,X) = Var(X) \rightarrow$ la covarianza generalizza il concetto di varianza

FORMULA

Si suppone che: $X = \sum_{i=1}^{n} \alpha_i X_i$, $Y = \sum_{j=1}^{m} \beta_j Y_j$

$$Cov(X,Y) = Cov\left(\sum_{i=1}^{n} \alpha_i X_i, \sum_{j=1}^{m} \beta_j Y_j\right)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \cdot Cov(X_i, Y_j)$$

LEZIONE 24

RIEPILOGO

X, Y variabili aleatorie congiuntamente distribuite

$\Rightarrow \exists f_{X,Y}(x,y) \vee \exists P_{(X,Y)}(x,y) = P(X=x, Y=y)$

- $Cov(X,Y) := \mathbb{E}[(X-\mu_X)(Y-\mu_Y)]$

BILINEARITÀ $\longrightarrow$ $X = \sum_{i=1}^{n} \alpha_i X_i, \quad Y = \sum_{i=1}^{m} \alpha_i Y_i$

$$Cov(X,Y) = Cov\left(\sum_{i=1}^{n} \alpha_i X_i, \sum_{i=1}^{m} \alpha_i Y_i\right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \, Cov(X_i, Y_j)$$

Due numeri aleatori U e V sono NON CORRELATI se e solo se:

$$Cov(U,V) = 0 \iff \mathbb{E}(U \cdot V) = \mathbb{E}(U)\,\mathbb{E}(V)$$

THM 4.7.3

Se X e Y sono variabili aleatorie indipendenti, allora

$\mathbb{E}[XY] = \mathbb{E}(X)\mathbb{E}(Y)$ $\longrightarrow$ ciò vale a dire che due variabili aleatorie indipendenti sono NON CORRELATI

$\Downarrow$

$Cov(X,Y) = 0$

L'indipendenza implica la NON correlazione

DIM (DISCRETO)

Si prova che $\mathbb{E}[XY] = \mathbb{E}(X)\mathbb{E}(Y)$

$$\mathbb{E}[XY] = \sum_i \sum_j x_i y_j \, P(X = x_i, Y = y_j) =$$

$= $ per l'indipendenza $= \sum_i \sum_j x_i y_j \, \mathbb{P}(X = x_i)\, \mathbb{P}(Y = y_j)$

$= \sum_i x_i \, \mathbb{P}(X = x_i) \sum_j y_j \, \mathbb{P}(Y = y_j)$

$= \mathbb{E}[X]\, \mathbb{E}[Y]$ $\square$

I casi in cui una o entrambe le variabili aleatorie siano continue si provano in maniera analoga

DIM (CONTINUE)

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

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy \, f_X(x)\, f_Y(y)\, dx\, dy =$$

$$= \int_{-\infty}^{+\infty} x f_X(x)\, dx \cdot \int_{-\infty}^{+\infty} y f_Y(y)\, dy = \mathbb{E}[X]\, \mathbb{E}[Y]$$ $\square$

ESEMPIO 4.7.1

Si determini la varianza del numero di teste di 10 lanci indipendenti di una moneta non truccata

$S = Z_{1,1} + Z_{1,2} + \ldots + Z_{1,10}$ $\quad \forall i = 1, \ldots, 10$

$$Var(S) = Var\left(\sum_{i=1}^{10} Z_{1,i}\right) = \sum_{i=1}^{10} Var(Z_{1,i})$$

$$= \sum_{i=1}^{10} Var(Z_1) = Var(Z_1) \cdot \sum_{i=1}^{10} 1 = 10 \cdot Var(Z_1)$$

$$= 10 \cdot \frac{35}{12} = \frac{175}{6}$$

ESEMPIO 4.7.2

Si determini la varianza del numero di teste nel lancio di 10 monete

$S_{10} = X_{1,1} + X_{1,2} + \ldots + X_{1,10}$

$$Var(S_{10}) = Var\left(\sum_{i=1}^{10} X_{1,i}\right) = \sum_{i=1}^{10} Var(X_{1,i})$$

$$= \sum_{i=1}^{10} Var(X_1) = 10 \cdot p(1-p) = 10 \cdot \frac{1}{4} = \frac{5}{2}$$

FUNZIONE GENERATRICE DEI MOMENTI

- La funzione generatrice dei momenti di una variabile aleatoria X, è definita per tutti i t reali per i quali il valore atteso di e^{tx} ha senso dall'espressione

$$\phi(t) := \mathbb{E}[e^{tx}] = \begin{cases} \sum_x e^{tx} p(x) & X \text{ discreta} \\ \int_{-\infty}^{+\infty} e^{tx} f(x)\,dx & X \text{ continua} \end{cases}$$

Tutti i momenti di cui è dotata X possono essere ottenuti derivando più volte nell'origine la funzione $\phi(t)$

$$\phi'(t) = \frac{d}{dt}\mathbb{E}[e^{tx}] = \mathbb{E}\left[\frac{d}{dt} e^{tx}\right] = \mathbb{E}[Xe^{tx}]$$

da cui $\phi'(0) = \mathbb{E}[X]$

- Un'altra importante proprietà di ϕ è che la funzione generatrice dei momenti della somma di variabili aleatorie indipendenti è il prodotto delle f. generatrici delle singole variabili aleatorie

PROPOSIZIONE 4.9.1

Se X e Y sono var aleatorie indipendenti con ϕ_X e ϕ_Y allora $\phi_{X+Y} = \phi_X(t) \cdot \phi_Y(t)$

DIM

$$\phi_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}(e^{tX} \cdot e^{tY}) \leftarrow$$

X, Y indipendenti $\Rightarrow$ e^{tX}, e^{tY} indipendenti

$$= \mathbb{E}(e^{tX}) \cdot \mathbb{E}(e^{tY}) = \phi_X(t) \cdot \phi_Y(t)$$

$\square$

LA LEGGE DEBOLE DEI GRANDI NUMERI

PROPOSIZIONE 4.9.1 (Disuguaglianza di Markov)

Se X è una variabile aleatoria che non è mai negativa, allora $\forall a > 0$

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

DIM: $\longrightarrow$

Si dà la dimostrazione nel caso che X sia continua con densità f

$$-\ \mathbb{E}(X) = \int_0^{\infty} x f_X(x)\,dx = \underbrace{\int_0^{a} x f_X(x)\,dx + \int_a^{+\infty} x f_X(x)\,dx}_{\geq 0}$$

$$\Rightarrow \mathbb{E}(X) \geq \int_a^{\infty} x f_X(x)\,dx \geq \int_a^{\infty} a f_X(x)\,dx = a\int_a^{+\infty} f_X(x)\,dx$$

$$\Rightarrow \mathbb{E}(X) \geq a\,\mathbb{P}(X \geq a) = \frac{\mathbb{E}(X)}{a} \geq \frac{a\,\mathbb{P}(X\geq a)}{a} =$$

$$= \frac{\mathbb{E}(X)}{a} \geq \mathbb{P}(X \geq a) \qquad \square$$

PROPOSIZIONE 4.9.2 (Disuguaglianza di Chebyshev)

Se X è una variabile aleatoria con media μ e varianza σ^2 allora per ogni $r>0$,

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

DIM

$$\mathbb{P}(|X-\mu_X| \geq r) = \mathbb{P}[(X-\mu_X)^2 \geq r^2] \rightarrow \text{eventi equiprobabili}$$

$$\Rightarrow \leq \frac{\mathbb{E}[(X-\mu)^2]}{r^2} = \frac{\sigma^2}{r^2} \qquad \square$$

L'importanza delle disuguaglianze di Markov e Chebyshev, sta nel fatto che permettono di limitare la probabilità di eventi rari che riguardano variabili aleatorie di cui si conosca solo la media, oppure la media e la varianza. Naturalmente tali probabilità possono essere calcolate esattamente e non vi è necessità di ricorrere all'utilizzo di maggiorazioni.

LEZIONE 25

ESEMPIO 4.9.1

Il numero di pezzi prodotti da una fabbrica durante una settimana è una variabile aleatoria di media 50.
(a) Cosa si può dire sulla probabilità che la media di produzione ne superi i 75 pezzi?
(b) Se la varianza pari a 25, cosa si può dire sulla probabilità che la produzione sia compresa tra 40 e 60 pezzi?

$X =$ variabile aleatoria che indica il numero di pezzi prodotti in una settimana

- $\mathbb{E}(X) = 50$

(a) $\mathbb{P}(X \geq 75) \longrightarrow$ Markov

$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a} \Rightarrow \mathbb{P}(X \geq 75) \leq \frac{50}{75} = \frac{2}{3} \approx 0{,}667$

(b) $Var(X) = 25$

$\mathbb{E}(X) = 50$ $\longrightarrow \mathbb{P}(40 \leq X \leq 60) = ?$

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

$\Rightarrow \mathbb{P}(|X - 50| \geq 10) \leq \frac{25}{10^2} = \frac{1}{4}$

$\Downarrow$

$\mathbb{P}(40 < X < 60) = \mathbb{P}(|X - 50| \leq 10) \geq 1 - \frac{1}{4} = \frac{3}{4} = 0{,}75$

LDGN

La legge debole dei grandi numeri è un risultato di convergenza

$(X_n)_{n \in \mathbb{N}}$

$X_n \longrightarrow X$

- numeri aleatori indipendenti e identica distribuzione

$\lim_{n \to \infty} f_n(x) = f(x)$

- $\mathbb{E}(X_1^2) < +\infty$ ($\Rightarrow Var(X_1) < +\infty$) $= Var(X_2) = Var(X_3) = \sigma^2$, $\exists$ finita $\mathbb{E}(X_1) = \mathbb{E}(X_2) = \mathbb{E}(X_3) = \mu$

$\forall n \in \mathbb{N}$, $S_n = \sum_{i=1}^{n} X_i$, $\frac{S_n}{n}$

Th: $\frac{S_n}{n} \xrightarrow[\text{prob}]{\text{convergenza}} \mu$

DIM

$\forall n \in \mathbb{N}$, $\mathbb{E}(S_n) = \mathbb{E}\left(\sum_{i=1}^{n} X\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n\mu$

$\mathbb{E}(S_n/n) = \frac{1}{n} \cdot \mathbb{E}(S_n) = \frac{\cancel{n}\mu}{\cancel{n}} = \mu$

- $Var(S_n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} [Var(X_i)] = n\sigma^2$

- $Var\left(\frac{S_n}{n}\right) = \frac{1}{n^2} Var(S_n) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$

Grazie alla disuguaglianza di Chebyshev, la legge debole dei grandi numeri è provata tramite un enunciato che afferma che la media aritmetica di n copie indipendenti di una variabile aleatoria tende al valore atteso di quest'ultima per n che tende all'infinito

→ Tale convergenza si precisa dicendo che scelto un ε piccolo, la media aritmetica si discosta dal valore atteso per più di ε con probabilità che tende a zero, quando n tende all'infinito

$$\varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq Var\left(\frac{S_n}{n}\right) \Big/ \varepsilon^2 = \frac{\sigma^2}{n} \Big/ \varepsilon^2 = \frac{\sigma^2}{n\varepsilon^2}$$

$$\lim_{n \to \infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \lim_{n \to \infty} \sigma^2/\varepsilon^2/n = 0$$

$$\downarrow$$

$$\lim_{n \to \infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq 0$$

$$\downarrow$$

$$\lim_{n \to \infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0 \iff \frac{S_n}{n} \longrightarrow \mu$$

TEOREMA DI BERNOULLI

$\forall n \in \mathbb{N}, \quad X_n \sim B(1, p) \ ; \quad X_n : \begin{pmatrix} 1 & 0 \\ p & q \end{pmatrix}$

- $\mathbb{E}(X_n) = p \ , \quad Var(X_n) = pq$

$(X_n)_{n \in \mathbb{N}} \longrightarrow$ successione di variabili aleatorie i.i.d con media e varianza comune

$$\frac{S_n}{n} \xrightarrow{prob} p$$

$$\downarrow$$

la frequenza relativa tende alla probabilità dell'evento

LEZIONE 26

- STATISTICA DESCRITTIVA

↓

Si intende un insieme di tecniche e strumenti finalizzati ad assolvere uno dei principali compiti assegnati alla statistica

↓

Descrivere, rappresentare e sintetizzare in maniera opportuna un insieme di dati relativamente ad una o più caratteristiche di una popolazione di interesse

- COLLETTIVO STATISTICO → Si intende la totalità dei casi ovvero dei membri sui quali è possibile rilevare uno o più caratteri che rivestono particolare importanza per il fenomeno che si sta studiando

Popolazione ⟶ Spazio campione (insieme unità statistiche)

Carattere ⟶ Numero aleatorio

ESEMPIO 1

Durante il semestre viene proposto agli studenti dei corsi di laurea offerti in Ateneo il questionario sulla valutazione della struttura didattica nella sua complessità

⟶ popolazione = studenti

Per lo studio in questione ci si serve di un questionario avente un certo numero di domande

↓

In questo caso ciascuna domanda del questionario corrisponde ad un CARATTERE

Per la maggior parte delle domande lo studente può rispondere tra 4 possibili risposte tra loro alternative:

a) Decisamente no
b) Più no che sì
c) Più sì che no
d) Decisamente sì

CARATTERI QUALITATIVI ORDINALI

Esempio 2

In un allevamento di bufale da latte si vuole mettere in relazione la produzione giornaliera di latte con la grandezza delle mammelle e con lo stato di salute delle stesse

→ popolazione = bufale dell'allevamento ciascuna delle quali è un membro

→ caratteri da rilevare = 1) produzione latte
2) misura mammelle
3) stato di salute mammelle

↓

Le diverse espressioni con le quali si manifesta un carattere si chiamano MODALITA'

- Si può pensare ad una popolazione come ad un insieme. La cardinalità della popolazione è detta TAGLIA.
- Un carattere viene designato con una lettera latina maiuscola, come variabili aleatorie
- I valori rilevati vengono designati con la stessa lettera ma in minuscolo

→ Se, allora, y rappresenta il carattere sotto studio non continuo, la sequenza

$$\underline{y} = (y_1, y_2, \dots, y_N)$$

rappresenta l'intera rilevazione dei dati

bisogna "SINTETIZZARE" i dati ottenuti in maniera quantitativa: 1) MEDIA ANALITICA 2) I CENTRI

CARATTERI QUANTITATIVI

MEDIA ANALITICA → INDICI DI POSIZIONE
→ OP. DI SINTETIZZARE

1) $\underline{y} = (y_1, \dots, y_N)$ → bisogna sostituire a y_i, con $i = 1, \dots, N$, lo stesso numero reale y

2) Bisogna specificare un criterio C (funzione di circostanza) rispetto al quale si vuole ottenere la valutazione di tendenza centrale.

• La valutazione della funzione di circostanza C su y deve coincidere con la valutazione di C su $y^* = (y, \dots, y)$

$$\boxed{C(y_1, \dots, y_n) = C(y, \dots, y)}$$

a) MEDIA ARITMETICA

Si sceglie come funzione di circostanza C la somma dei dati:

$$C(y_1, y_2, \dots, y_n) = y_1 + y_2 + \dots + y_N$$

$$C(y_1, y_2, \dots, y_n) = C(y, y, \dots, y)$$

$$\iff y_1 + y_2 + \dots + y_N = y + y + \dots + y \quad \} N \text{ volte}$$

$$\iff y_1 + y_2 + \dots + y_N = N \cdot y$$

In definitiva, la soluzione dell'equazione di circostanza è la media aritmetica dei dati

$$\boxed{\bar{y} = y_1 + y_2 + \dots + y_N / N = \frac{1}{N} \sum_{i=1}^{N} y_i}$$

b) MEDIA ARMONICA

Si sceglie come funzione di circostanza C la somma dei reciproci dei dati

$$C(y_1, y_2, \dots, y_N) = \frac{1}{y_1} + \frac{1}{y_2} + \dots + \frac{1}{y_N}$$

$$C(y_1, y_2, \dots, y_N) = C(y, y, \dots, y)$$

$$\iff \frac{1}{y_1} + \frac{1}{y_2} + \dots + \frac{1}{y_N} = \frac{1}{y} + \frac{1}{y} + \dots \frac{1}{y} \quad \} N \text{ volte}$$

$$\iff \left(\frac{1}{y_1} + \frac{1}{y_2} + \dots + \frac{1}{y_N}\right) = N/y$$

In definitiva, la soluzione dell'equazione di circostanza è la media armonica dei dati

$$\boxed{M_a = \frac{N}{\frac{1}{y_1} + \dots + \frac{1}{y_N}} = \left(\frac{\frac{1}{y_1} + \dots + \frac{1}{y_N}}{N}\right)^{-1}} = \left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{y_i}\right)^{-1}$$

e) MEDIA GEOMETRICA

Si sceglie come funzione di circostanza e il prodotto dei dati:

$$c(y_1, y_2, \dots, y_N) = y_1 \cdot y_2 \cdot \dots \cdot y_N$$

$$c(y_1, y_2, \dots, y_N) = c(y, y, \dots, y)$$

$\iff y_1 \cdot y_2 \cdot \dots \cdot y_n = y \cdot y \cdot \dots \cdot y$ { N-volte

In definitiva la soluzione dell'equazione di circostanza è la media geometrica dei dati:

$$\boxed{M_g = \sqrt[N]{y_1 \cdot y_2 \cdots y_N} = \sqrt[N]{\prod_{i=1}^{N} y_i}}$$

I CENTRI

Sia $\underline{y} = (y_1, y_2, \dots, y_N)$ una rilevazione dati di un carattere quantitativo Y, e sia $\forall x \in \mathbb{R},\ d(x, \underline{y}) \geq 0$ → $f^{(n)}(x) = d(x, \underline{y})$

una funzione che si ritiene adatta a rappresentare la **DISTANZA** di un generico valore reale x da tutti gli elementi della rilevazione dati $\underline{y}$

Si definisce centro di una rilevazione dati $\underline{y}$, e lo si indica con $\xi(\underline{y})$, il punto minimo assoluto della funzione $d(x, \underline{y})$

$$\xi(\underline{y}) = \underset{x \in \mathbb{R}}{\operatorname{argmin}}\ d(x, \underline{y})$$

- sono considerate le seguenti funzioni di distanza → $f(x) = d(x, \underline{y}) = \begin{cases} \sum_{i=0}^{N} |x - y_i|^0, & r = 0 \\ \left(\sum_{i=0}^{N} |x - y_i|^r\right)^{\frac{1}{r}}, & r > 0 \end{cases}$

- La funzione $d_r(x, \underline{y})$ è detta distanza di ordine r e il suo punto MINIMO ASSOLUTO è detto centro di ordine r

$\longrightarrow \xi_r := \arg\min_{x \in \mathbb{R}} f_r(x)$ (CENTRO DI ORDINE R)

- CASO $r=2$ $\longleftarrow$ COINCIDE CON LA MEDIA ARITMETICA $\frac{1}{N}\sum_{i=1}^{N} y_i$

$$\xi_2 = \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$$

DIM

$$f_2(x) = \sqrt{\sum_{i=1}^{N}(x-y_i)^2}$$

$$\arg\min_{x\in\mathbb{R}} f_2(x) = \arg\min_{x\in\mathbb{R}} \sum (x-y_i)^2$$

$$\forall x \in \mathbb{R},\ f(x) = \sum_{i=1}^{N}(x-y_i)^2$$

$\downarrow$

ricerca punti stazionari

$$f'(x) = \frac{2}{N}\sum_{i=1}^{N}(x-y_i) \Rightarrow f'(x)=0$$

$$\Leftrightarrow \sum_{i=1}^{N}(x-y_i)=0 \Leftrightarrow x = \frac{1}{N}\sum_{i=1}^{N} y_i = \bar{y}$$

$\downarrow$

ricerca dei punti di minimo e massimo di $f(x)$

$$f''(x) = \frac{2}{N}\sum_{i=1}^{N} 1 = 2 \Rightarrow f''(\bar{y}) = 2 > 0,$$

ciò implica che $\bar{y}$ è un punto di minimo relativo per $f(x)$

$\downarrow$

Ricerca minimo assoluto di $f(x)$

$\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = +\infty$, la derivata prima di $f(x)$ è definita in $\mathbb{R}$ e $f(x)$ ammette un unico punto di minimo relativo

$\downarrow$

In definitiva $\longrightarrow$ $\xi_2(\underline{y}) = \arg\min_{x\in\mathbb{R}} d_2(x,\underline{y}) = \arg\min_{x\in\mathbb{R}} f(x) = \bar{y}$

$\square$

LEZIONE 27

CENTRI

$\underline{y} = (y_1, y_2, \dots, y_N)$

ξ è quel numero reale che sta alla minima distanza da tutti i dati

$\hookrightarrow \forall x \in \mathbb{R},\ f_r(x) = \begin{cases} \sum_{i=1}^{N} |x - y_i|^0 \\ \sum_{i=1}^{N} |x - y_i|^{1/r}, & r > 0 \end{cases}$

$\Rightarrow \xi_r := \underset{x \in \mathbb{R}}{\arg\min}\ f_r(x) \longrightarrow$ è il centro di ordine r

$\Rightarrow$ il centro di ordine r, con $r = 1$, ovvero $\xi_1(\underline{y})$ coincide con la **MEDIANA**

$$\xi_1(\underline{y}) = \underset{x \in \mathbb{R}}{\arg\min}\ d_1(x, \underline{y}) = \arg\min \frac{1}{N} \sum_{i=1}^{N} |x - y_i| = M_1$$

$\Rightarrow$ il centro di ordine r, con $r = 2$, ovvero $\xi_2(\underline{y})$ coincide con la **MEDIA ARITMETICA**

$$\boxed{x \in \mathbb{R},\quad f(x) = \sqrt[2]{\frac{1}{N} \sum_{i=1}^{N} (x - y_i)^2}}$$

$\Rightarrow$ il centro di ordine r con $r = 0$, ovvero $\xi_0(\underline{y})$, coincide con la **moda** della rilevazione dati

$\hookrightarrow$ Sia x_1 uno dei dati e sia m_1 il numero delle volte che esso si trova in $\underline{y}$

$\vdots$

Sia x_k uno dei dati e sia m_k il numero delle volte che esso si trova in $\underline{y}$

$\hookrightarrow$ con $x_1 \neq x_2 \neq \dots \neq x_k$

$$f_0(x) = \begin{cases} N, & x \notin \{x_1, x_2, \dots, x_k\} \\ N - m_1, & x = x_1 \\ \vdots \\ N - m_k, & x = x_k \end{cases} = \sum_{i=1}^{N} |x - y_i|^0$$

$$\boxed{\xi = \underset{x \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^{N} |x - y_i|^0 = \min\{N - n_1, \dots, N - n_k\} = N - e}$$

con $e = \max(n_1, n_2, \dots, n_k)$ □

STATISTICA DESCRITTIVA $\longrightarrow$ RAPPRESENTARE I DATI

$y = (y_1, y_2, \dots, y_N)$

$\hookrightarrow$ realizzazione dati

MODALITÀ	F. ASSOLUTA	F. RELATIVE	F. ACCUMULATIVA RELATIVA	INTENSITÀ
x_1	n_1	$f_1 = n_1/N$	$F_1 = f_1$	$x_1 f_1$
x_2	n_2	$f_2 = n_2/N$	$F_2 = F_1 + f_2 = f_1 + f_2$	$x_2 f_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_k	n_k	$f_k = n_k/N$	$F_k = F_{k-1} + f_k = 1$	$x_k f_k = \bar{y}$

distribuzione di frequenze

$$\bar{y} = \sum_{i=1}^{N} y_i / N = \frac{x_1 n_1 + \dots + x_k n_k}{N} =$$

$$= x_1 \frac{n_1}{N} + \dots + x_k \frac{n_k}{N} = x_1 f_1 + \dots + x_k f_k = \bar{y}$$

ESEMPIO

$N = 16$ TAGLIA = 16

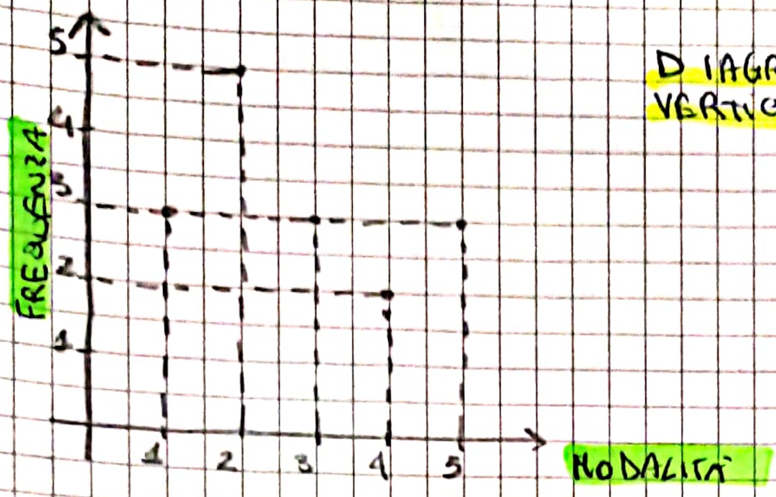
2, 1, 2, 5, 3, 2, 4, 4, 1, 2, 5, 3, 3, 5, 1, 2

x_i	n_i	f_i	F_i	$x_i f_i$
1 (Q_0)	3	3/16	3/16	3/16
2 (Moda)	5	5/16	8/16 (Q_2, Me, Q_1)	10/16
3	3	3/16	11/16	9/16
4	2	2/16	13/16 (Q_3)	8/16
5 (Q_4)	3	3/16	16/16 = 1	15/16

Grafico della distribuzione di frequenze

- M_o(MODA) = 2 poiché 2 ha la frequenza assoluta più alta
- M_e(MEDIANA) = 2 poiché, vedendo nella colonna di F_i è il primo valore che supera 0,5 o è uguale

DIAGRAMMA A BASE VERTICALE



- CLASSIFICAZIONE DEI CARATTERI

CARATTERI
- QUALITATIVO
  - NOMINALI
  - ORDINALI
- QUANTITATIVO
  - DISCRETO
  - CONTINUO

- Le distribuzioni di frequenza possono essere rappresentate in forma grafica con scelta eseguita opportunamente rispetto al tipo di carattere

→ Per i caratteri qualitativi e quantitativi discreti, oltre al diagramma CIRCOLARE, la rappresentazione grafica più usata è quella del diagramma a barre verticali

→ Per i soli caratteri quantitativi continui è opportuno utilizzare la rappresentazione grafica a istogramma

La differenza qualitativa tra un istogramma e un diagramma a barre verticali è che negli istogrammi le barre verticali devono essere contigue.

→ Inoltre negli istogrammi la frequenza delle classi di modalità rappresenta l'area del rettangolo e non l'altezza come nel diagramma a barre

Per qualsiasi $j \in \{1, 2, \dots, k\}$, se b_j rappresenta l'ampiezza della j-esima classe di modalità, l'altezza h_j del relativo rettangolo è ottenuta mediante la formula inversa per l'area:

$$j = (1, 2, \dots, k), \quad h_j = n_j / b_j$$

- SINTETIZZARE

- INDICI DI POSIZIONE → $\underline{y} = (y_1, y_2, \dots, y_N)$

con un unico valore di TENDENZA CENTRALE, ossia un valore che fornisca un'indicazione di massima sulla localizzazione di y. Ciò è utile per istituire un confronto del fenomeno studiato con altri fenomeni dello stesso tipo

per i caratteri qualitativi è possibile fare ricorso ai due indici moda e mediana

→ per i caratteri quantitativi ci sono almeno altri due approcci teorici in grado di fare ottenere indici di tendenza centrale:

a) MEDIE ANALITICHE
b) CENTRI

- INDICI DI DISPERSIONE

→ servono per descrivere sinteticamente la misura con la quale una rilevazione dati di un carattere quantitativo è distante da una sua tendenza centrale

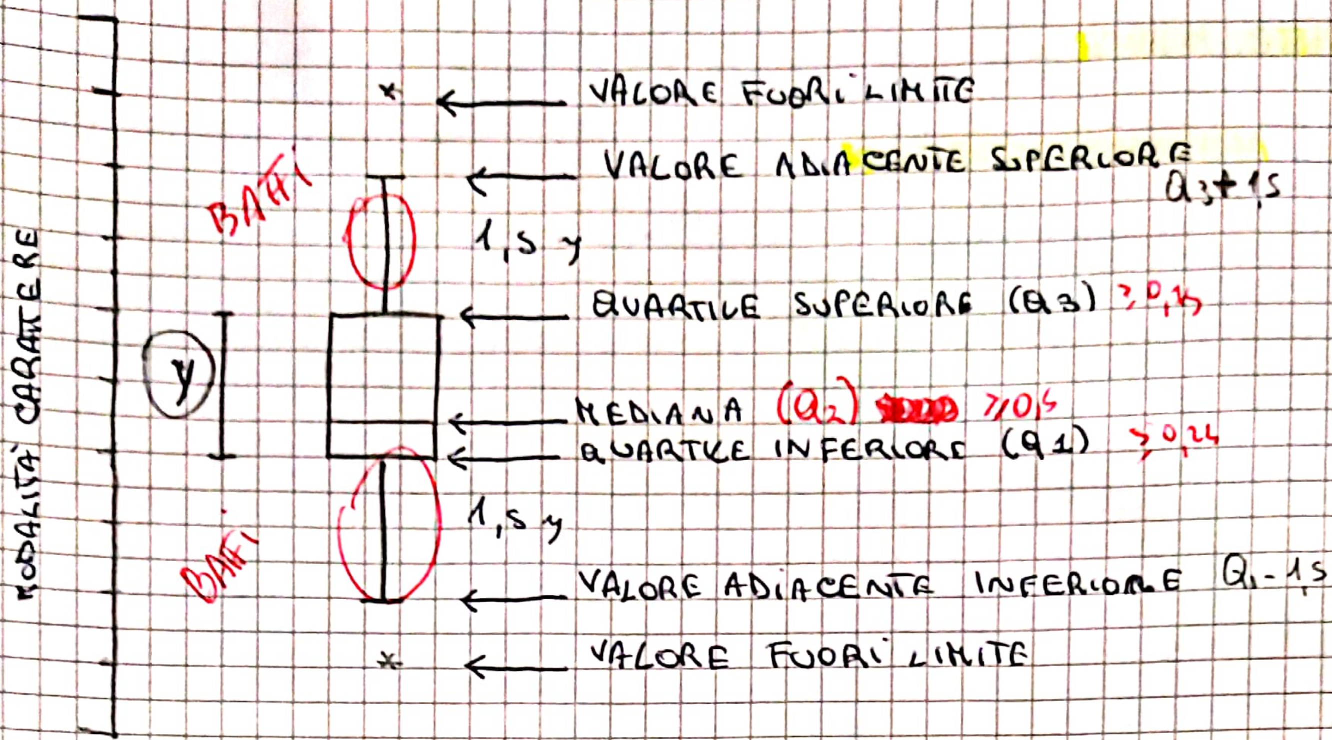
La dispersione esprime la bontà o l'inadeguatezza di un indice di tendenza centrale quale descrittore di una distribuzione di frequenza

→ C. QUALITATIVI = indice di ricchezza ⇒ conteggio del numero delle modalità presenti nella rilevazione dati

LEZIONE 28

DIAGRAMMA SCATOLA CON BAFFI

Un metodo grafico per rappresentare una distribuzione di frequenze che mette in risalto anche la dispersione intorno alla mediana è il grafico della scatola con baffi o anche BOX PLOT



La linea interna alla scatola rappresenta la mediana della distribuzione. Le linee esterne della scatola rappresentano il primo e il terzo quartile

↓

La distanza (y) è una misura di dispersione della distribuzione

→ il 50% dei dati si trovano compresi tra questi due valori

↓

le distanze tra ciascun quartile e la mediana forniscono informazioni relativamente alla forma della distribuzione

↓

Le linee che si allungano da Q_1 e Q_3 sono i valori ADIACENTI. Il VAI (valore adiacente inferiore) è il valore più piccolo tra i dati che risulta maggiore o uguale a $Q_1 - 1,5\,y$ e il VAS (valore adiacente superiore) è il valore più grande tra i dati che risulta minore o uguale a $Q_3 + 1,5\,y$

↓

Invece, i valori fuori limite costituiscono una anomalia rispetto alla maggior parte dei valori osservati e pertanto è necessario identificarli per poterne analizzare le caratteristiche

INDICI DI DISPERSIONE

→ serve per descrivere sinteticamente la misura con la quale una rilevazione dati di un carattere quantitativo è distante da una sua tendenza centrale

→ esprime la bontà o la inadeguatezza di un valore di tendenza centrale

- Sia $y = (y_1, y_2, \dots, y_m)$ una rilevazione dati di un carattere quantitativo

Sono INDICI DI DISPERSIONE:

a) differenza tra il dato più grande da quello più piccolo

$$T = y_{(N)} - y_{(1)} = Q_4 - Q_0$$

b) la differenza tra il primo e il terzo quartile

$$\gamma = Q_3 - Q_1$$

c) la media del valore assoluto delle differenze dei dati dalla loro mediana Q_2

$$S_{Q_2} = \frac{1}{N} \sum_{i=1}^{N} |y_i - Q_2|$$

↓

scarto medio assoluto associato alla mediana Q_2

d) la media del valore assoluto delle differenze dei dati dalla loro media aritmetica $\bar{y}$

$$S_{\bar{y}} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \bar{y}|$$

f) la radice quadrata della varianza (deviazione std)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2}$$

e) la media del quadrato delle differenze dei dati dalla loro media (varianza)

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2$$

LA DISTRIBUZIONE DELLE STATISTICHE CAMPIONARIE

DEFINIZIONE 6.6.1

Un insieme $X_1, X_2, \dots, X_n$ di variabili aleatorie indipendenti, tutte con la stessa distribuzione F, si dice campione o CAMPIONE ALEATORIO della distribuzione F

$$X = (X_1, X_2, \dots, X_n) \longleftarrow \text{CAMPIONE ALEATORIO}$$
i.i.d.

$$\Downarrow$$

$$F_X(x_1, x_2, \dots, x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n)$$

- Il termine statistica indica una variabile aleatoria che è semplicemente una funzione dei dati di un campione

DEFINIZIONE

Si dice parametrica del campione casuale semplice un numero aleatorio che ha la distribuzione uguale a quella comune a $X_1, X_2, \dots, X_n$

$$F_X(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} F_X(x_i) \longrightarrow T = g(X) = \sum_{i=1}^{N} X_i$$

$$\Rightarrow \prod_i = g_i(X) = X_i$$

DEFINIZIONE

Si dice STATISTICA ogni funzione misurabile / calcolabile di un campione casuale semplice

$$X \sim B(p) \longrightarrow \sum_{i=1}^{n} X_i = p$$

Distribuzione di probabilità di statistiche

statistica = indica una variabile aleatoria che è semplicemente una funzione dei dati di un campione

I due principali esempi di statistiche sono la MEDIA CAMPIONARIA e la VARIANZA CAMPIONARIA

LA MEDIA CAMPIONARIA

- Si consideri una popolazione di elementi a ciascuno dei quali è associata una grandezza numerica.
Sia $X_1, X_2, \dots, X_n$ un campione di dati estratto da questa popolazione. Si suppone che i valori numerici associati a ciascun campione, siano variabili aleatorie i.i.d.
Si denota con μ e σ^2 rispettivamente la media e la varianza della popolazione

$$\text{M.C.} = \boxed{\bar{X} := \frac{X_1 + X_2 + \dots + X_n}{n}} \qquad (6.2.1)$$

Si noti che $\bar{X}$ è una funzione delle variabili aleatorie $X_1, X_2, \dots, X_n$

Essa è una statistica e a sua volta è una variabile aleatoria

Quanto valgono il valore atteso della media campionaria e la sua varianza?

- $\mu = \mathbb{E}(X)$ esiste finita

$$\Rightarrow \mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] =$$

$$= \frac{\mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]}{n} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{\cancel{n}\mu}{\cancel{n}} = \mu \qquad \square$$

- $Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$

$$= \frac{Var(X_1) + Var(X_2) + \dots + Var(X_n)}{n} = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} Var(\bar{X}) = \frac{n\,Var(X)}{n^2} = \frac{\cancel{n}\sigma^2}{n^2} = \frac{\sigma^2}{n} \qquad \square$$

La media campionaria ha quindi lo stesso valore atteso della distribuzione da stimare, mentre la sua varianza risulta ridotta di un fattore n

→ Si deduce che $\bar{X}$ è centrata attorno a μ, mentre la sua variabilità si riduce sempre di più con l'aumentare di n

IL TEOREMA DEL LIMITE CENTRALE

Siano $X_1, X_2, \dots, X_n$ delle variabili aleatorie indipendenti e identicamente distribuite tutte con media μ e varianza σ^2.
Allora se n è grande, la somma $X_1 + X_2 + \dots + X_n$ è approssimativamente normale con media $n\mu$ e varianza $n\sigma^2$.

Si può anche normalizzare la somma in modo da ottenere una distribuzione approssimativamente normale standard

$$T = \frac{X_1 + X_2 + \dots + X_n - n\mu}{\sigma\sqrt{n}} = \boxed{\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \sim N(0,1)}$$

≈ approssimativamente distribuita come

LEZIONE 24

ESERCIZIO 6.31

Una compagnia di assicurazioni ha 25000 polizze auto attive. Il risarcimento annuo per ogni singolo assicurato è una variabile aleatoria con media 320 e deviazione standard 540.
Quanto vale la probabilità che in un anno la richiesta di indennizzi superi gli 8,3 milioni?

X: richieste annuale complessive di indennizzi
Y: richieste annuale per singolo assicurato

$E(Y) = 320$
$Var(Y)$: 540

X ha distribuzione normale con media $320 \times 25000 = 8 \times 10^6$ e deviazione standard $540\sqrt{25000} \approx 8{,}54 \times 10^4$

$Z \sim N(0,1)$

$$P(X > 8{,}3 \cdot 10^6) = P\left(\frac{X - 8 \times 10^6}{8{,}54 \times 10^4} > \frac{8{,}3 \times 10^6 - 8 \times 10^6}{8{,}54 \times 10^4}\right)$$

$$\approx P\left(Z > \frac{0{,}3 \times 10^6}{8{,}54 \times 10^4}\right) \approx P(Z > 3{,}51) \approx 0$$

DISTRIBUZIONE MEDIA CAMPIONARIA

Sia $X_1, X_2, \dots, X_n$ un campione proveniente da una popolazione di media μ e varianza σ^2.
Il TLC permette di approssimare la distribuzione della media campionaria

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

Siccome il prodotto di una variabile aleatoria normale per una costante è ancora normale, allora quando n è grande, $\bar{X}$ è gaussiana

v.c. con valore atteso μ e dev. standard $\sigma/\sqrt{n}$ allora

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

ESEMPIO 6.3.9

Una popolazione formata da operai maschi, presenta dei pesi corporei di media 167 e dev. std 27

(a) da un campione di 36 elementi, quanto vale la probabilità che la media campionaria stia tra 163 e 171?

X = pesi operaio

$\mu = \mathbb{E}(X) = 167$, $\sigma = 27$

$n = 36$ e $X = (X_1, \dots, X_{36})$

$\mathbb{P}(163 < \bar{X} < 171)$

$$= \mathbb{P}\left(\frac{163-167}{27/6} < \frac{\bar{X}-167}{4.5} < \frac{171-167}{27/6}\right)$$

$$= \mathbb{P}\left(-0,8889 < \frac{\bar{X}-167}{4,5} < 0,8889\right)$$

$$= 2\mathbb{P}\left(0 < \frac{\bar{X}-167}{4,5} < 0,8889\right)$$

$$= 2\mathbb{P}\left(\frac{\bar{X}-167}{4,5} < 0,8889\right) = \cancel{2 \cdot 0,7389}$$

$$= 2 \cdot 0,3133 \approx 0,63$$

(b) E se si selezionano 144 operai?

$n = 144$ $\qquad \frac{\bar{X} - 167}{2,25} = Z \sim N(0,1)$

$P(163 < \bar{X} < 171) =$

$= P\left(\frac{-4}{27/12} < \frac{\bar{X} - 167}{27/12} < \frac{4}{27/12}\right)$

$= P\left(-1,7778 < \frac{\bar{X} - 167}{2,25} < 1,7778\right)$

$= 2P\left(0 < \frac{\bar{X} - 167}{2,25} < 1,7778\right)$

$= 2 \cdot 0,4641 \simeq 0,93$

QUANDO UN CAMPIONE È ABBASTANZA NUMEROSO?

Il teorema del limite centrale lascia aperta la questione di quanto grande debba essere la numerosità del campione n, affinché l'approssimazione normale sia valida. Ad esempio, se la distribuzione della popolazione è normale, allora $\bar{X}$ sarà a sua volta normale.

FUNZIONE GENERATRICE DEI MOMENTI:

L'indipendenza di numeri aleatori:

$|t| < t_0$, $\phi_X(t) := E(e^{tX})$

- X_1, X_2 indipendenti $\longrightarrow$ $X_1 + X_2$ sono dotati di f.g.m.

$|t| < \min(t_{01}, t_{02})$, $\phi_{X_1+X_2}(t) = [E\, e^{t(X_1+X_2)}]$

$= E[e^{tX_1} \cdot e^{tX_2}] = E(e^{tX_1}) \cdot E(e^{tX_2})$

$= \phi_{X_1}(t_1) \cdot \phi_{X_2}(t_2)$

In generale se $X_1, X_2, \dots, X_n$ indipendenti e esiste la loro f.g.m. nell'intorno dell'origine

$|t| < \min\{t_0, \dots, t_n\}$, $\phi_{\Sigma X_i}(t) = \prod_{i=1}^{n} \phi_{X_i}(t)$

- Se $X_1 \sim X_2 \sim \dots \sim X_n \Rightarrow \phi_{\Sigma X_i}(t) = \prod_{i=1}^{n} \phi_{X_i}(t) = \phi_X^n(t)$

- $\underline{X} = (X_1, X_2, \dots, X_n)$ di $X \sim N(\mu, \sigma^2)$

$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \Rightarrow t \in \mathbb{R}$, $\phi_{\bar{X}}(t) = \phi_{\Sigma X_i}(t)$:

$= \phi_{\Sigma X_i}(t/n) = \phi_X^n(t/n) \longrightarrow$ f.g.m. della media campionaria $\bar{X}$

LA VARIANZA CAMPIONARIA

- Sia $X_1, X_2, \dots, X_n$ un campione aleatorio, proveniente da una distribuzione di media μ e varianza σ^2
Sia $\bar{X}$ la sua media campionaria

Si introduce una seconda statistica

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

VARIANZA CAMPIONARIA

la sua radice quadrata $S = \sqrt{S^2}$ prende il nome di deviazione standard campionaria

$E[S^2] \longrightarrow \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$ con

$\bar{x} = \sum_{i=1}^{n} x_i / n \Rightarrow S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2 \right)$

$(n-1)S^2 = \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$

LEZIONE 30

STIMA PARAMETRICA

- Si considera un campione aleatorio $X_1, X_2, \dots, X_n$ estratto da una distribuzione F_θ che dipende da un vettore di parametri incogniti θ

- Si dice STIMATORE per una funzione del parametro della densità, $\psi(\theta)$, una statistica utilizzata per la stima di $\psi(\theta)$

- $\bar{X}$ è uno stimatore per la media μ di una densità
$\bar{X}$ $\longrightarrow$ $\mu \in \mathbb{R}$, $E(\bar{X}) = \mu$, o $G(\mu)$ regione parametrica

$\bar{X}$ stimatore CORRETTO per μ

- DA WIKI

uno stimatore è una funzione che associa ad ogni possibile campione un valore del parametro da stimare

- X ha varianza σ^2 non nota $\quad \bar{X}^{(2)} = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

$\sigma^2 > 0$ (R. parametrica)

$\forall\, i = 1, \dots, n$

$* \ \mathbb{E}(X_i^2) = \mathbb{E}(X^2) = \sigma^2 + \mu^2$

$* \ \mathbb{E}(\bar{X}^{(2)}) = Var(\bar{X}) + \mathbb{E}^2(\bar{X}) = \frac{\sigma^2}{n} + \mu^2$

$* \ \mathbb{E}\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \mathbb{E}\left[n\left(\bar{X}^{(2)} - \bar{X}^2\right)\right] =$

$= \mathbb{E}\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right] = \mathbb{E}\left(\sum_{i=1}^{n} X_i^2\right) - n\sum_{i=1}^{n}(X^2)$

$= \sum_{i=1}^{n}\mathbb{E}(X_i^2) - \left(n\left(\frac{\sigma^2}{n} + \mu^2\right)\right) = n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2} =$

$= \sigma^2(n-1)$

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

$\sigma^2 > 0$ (R. parametrica), $\mathbb{E}(S_c^2) = \sigma^2 \cdot \frac{n-1}{n}$

$\lim_{n\to\infty} \mathbb{E}(S_c^2) = \lim_{n\to\infty} \sigma^2 \cdot \frac{n-1}{n} \iff \mathbb{E}(S_c^2) = \sigma^2$

- Uno stimatore $T = g(X)$ si dice corretto per la stima di $\psi(\theta)$ se e solo se $\theta \in H$, $\mathbb{E}(T) = \mathbb{E}_T(\theta) = \psi(\theta)$

$X = (X_1, X_2, \dots, X_n) \longrightarrow$ da X gaussiana di media μ incognita e di varianza σ^2

$T = \frac{X_1 + X_n}{2}$

$\forall \mu \in \mathbb{R}$, $\mathbb{E}_T(\mu) = \frac{1}{2}[\mathbb{E}(X_1) + \mathbb{E}(X_n)] = \frac{1}{2}(\mu + \mu) = \mu$

$\forall \mu \in \mathbb{R}$, $\mathbb{E}\bar{X}(\mu) = \mu$

Sono entrambi stimatori corretti

$Var(\bar{X}) = \frac{\sigma^2}{n}$ $\quad Var(T) = \frac{1}{4}[Var(X_1) + Var(X_n)]$

$= \frac{1}{4} 2\sigma^2 = \frac{\sigma^2}{2}$

Trovare un criterio per confrontare due stimatori.

RISCHIO QUADRATICO MEDIO $\longrightarrow$ indice di dispersione statistico

$\theta \in \Theta$, $\boxed{R_T(\theta) = \mathbb{E}\{[T - \varphi(\theta)]^2\}}$

$[T - \varphi(\theta)]^2$ = ERRORE QUADRATICO MEDIO

è la somma della varianza dello stimatore con il quadrato della sua distorsione

$\mathbb{E}(T) = \mu_T$

$$[T - \varphi(\theta)]^2 = [(T - \mu_T)^2 + (\mu_T - \varphi(\theta))]^2$$

$$= (T - \mu_T)^2 + [\mu_T - \varphi(\theta)]^2 + 2(\mu_T - \varphi(\theta)) \cdot (T - \mu_T) =$$

$$= \mathbb{E}[(T - \mu_T)^2] + \mathbb{E}[\mu_T - \varphi(\theta)]^2 + \underbrace{2(\mu_T - \varphi(\theta)) \cdot \mathbb{E}[(T - \mu_T)]}_{=0} =$$

$$= Var(T) + [\mu_T - \varphi(\theta)]^2 = Var(T) + d^2(\theta)$$

$\Rightarrow \mu_T - \varphi(\theta) = d(\theta) \Rightarrow$ DISTORSIONE

DEFINIZIONE 1

Siano T e S due stimatori per $\varphi(\theta)$. Si dice che S è preferibile a T e si scrive $S \preceq T \iff$ in tutta la regione parametrica, il rischio quadratico medio di S è NON maggiore del rischio quadratico medio di T

$\longrightarrow$ $S \preceq T \iff \theta \in \Theta$, $\boxed{R_S(\theta) \leq R_T(\theta)}$

DEFINIZIONE 2

Siano T e S due stimatori per $\varphi(\theta)$. Si dice che S è STRETTAMENTE preferibile a T e si scrive $S \prec T \iff S \preceq T$ ed esiste un $\theta_0 \in \Theta$ in corrispondenza del quale il rischio di S è minore strettamente del rischio di T:

$\longrightarrow$ $S \prec T \iff S \preceq T \;\land\; \boxed{\exists \theta_0 \in \Theta : R_S(\theta_0) < R_T(\theta_0)}$

DEFINIZIONE 3

Si dice che uno stimatore T è ammissibile ai fini della stima di $\varphi(\theta)$ se non esiste un altro stimatore strettamente preferibile a T

LEZIONE 31

$\mathbb{E}_T(\theta)$

$\mathbb{E}_T : \Theta \longrightarrow \mathbb{R}$

$R_T(\theta) = Var_T(\theta) - [E_T(\theta) - \psi(\theta)]^2$

$\theta \in \Theta \qquad := \mathbb{E}\{[T - \psi(\theta)]^2\}$

valore di un parametro che massimizza la densità congiunta / della distr. di prob.

STIMATORI DI MASSIMA VEROSIMIGLIANZA

Una qualunque statistica il cui scopo sia quello di dare una stima di un parametro θ si dice stimatore di θ. Gli stimatori sono quindi variabili aleatorie. Il valore deterministico assunto da uno stimatore è detto stima

CLASSE particolare di stimatori ⟶ STIMATORI DI MASSIMA VERISIMIGLIANZA

$X = (X_1, X_2, \dots, X_n)$ e $f(x_1, x_2, \dots, x_n | \theta)$

funzione di massa congiunta di X

questi valori denotano le possibili realizzazioni di f

$f_X(x, \theta) = f_{X_1}(x_1, \theta) \cdot f_{X_2}(x_2, \theta) \cdot \dots \cdot f_{X_n}(x_n, \theta)$
$= f_X(x_1, \theta) \cdot f_X(x_2, \theta) \cdot \dots \cdot f_X(x_n, \theta)$] θ incognito

LIKELIHOOD

$$\boxed{L(\theta; x) = \prod_{i=1}^{n} f_X(x_i; \theta)} = f_X(x_1, x_2, \dots, x_n; \theta)$$

funzione di massima verosimiglianza

la stima di massima verosimiglianza θ è definita come il valore di θ che rende massima $f_X(x_1, \dots, x_n; \theta)$ quando i valori osservati sono $x_1, \dots, x_n$

restituisce $\boxed{\hat{\theta}_{MV} := \underset{\theta \in \Theta}{\operatorname{argmax}} \; L(\theta, x)}$

Nel calcolare il valore di θ che massimizza f, conviene spesso usare il fatto che le due funzioni $f(x_1, x_2, \dots, x_n, \theta)$ e $\log[f(x_1, \dots, x_n; \theta)]$ assumono il massimo in corrispondenza di entrambe dello stesso valore di θ

(log-likelihood)

$X \sim Esp(\alpha) \qquad \alpha > 0$

$$f_X(x;\alpha) = \alpha \cdot e^{-\alpha x}, \qquad x > 0$$

$s = \sum_{i=1}^{n} x_i$

$$L(\alpha, x) = \prod_{i=1}^{n} f_X(x_i, \alpha) = \prod_{i=1}^{n} \alpha e^{-\alpha x_i} = \alpha^n \cdot e^{-\alpha s}$$

$$\frac{d L(\alpha, x)}{d(\alpha)} = n\alpha^{n-1} \cdot e^{-\alpha s} - s\alpha^n \cdot e^{-\alpha s} =$$

$$= e^{-\alpha s} \cdot \alpha^{n-1} (n - \alpha s)$$

$$\frac{d L(\alpha, x)}{d\alpha} = 0 \iff n - \alpha s = 0 \iff \boxed{\alpha = \frac{n}{s}}$$

punto stazionario

$$\frac{d^2 L(\alpha, x_i)}{d\alpha^2} = -s e^{-\alpha s} \cdot \alpha^{n-1} (n - \alpha s) + e^{-\alpha s} (n-1) \cdot \alpha^{n-2} (n - \alpha s) + - s e^{-\alpha s} \cdot \alpha^{n-1} =$$

$$= e^{-\alpha s} \cdot \alpha^{n-2} [-s\alpha (n - \alpha s) + (n-1)(n - \alpha s) - s\alpha]$$

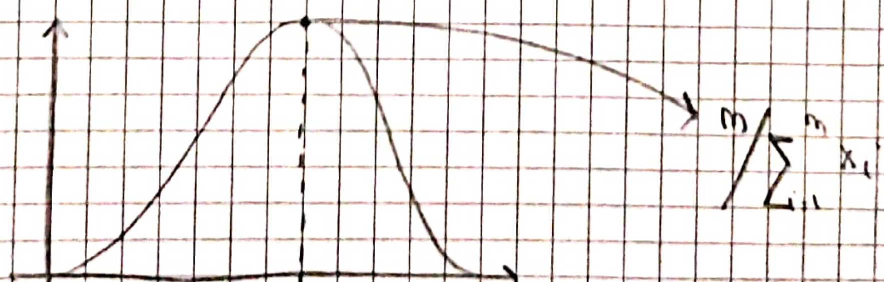
Se $\alpha = \frac{n}{s} \Rightarrow e^{-\alpha s} \cdot \alpha^{n-2} [-s\alpha \cancel{(n - \alpha s)} + (n-1)\cancel{(n - \alpha s)} - s\alpha]$

$$\Rightarrow \frac{d^2 L(\alpha, x_i)}{d s^2} = \boxed{-n \left(\frac{n}{s}\right)^{n-2} e^{-n} < 0}$$

Quindi questo valore non è solo il massimo assoluto ma anche il massimo relativo

$$\lim_{\alpha \to 0^+} L(\alpha, x_i) = 0, \quad \lim_{\alpha \to \infty} L(\alpha; x) = 0 \Big] \text{ MASSIMO ASSOLUTO}$$

$$\hat{\alpha}_{MV} = n/s = n \Big/ \sum_{i=1}^{n} x_i \Rightarrow \hat{\alpha}_{MV} = n \Big/ \sum_{i=1}^{n} x_i = (\bar{x})^{-1}$$



- $X \sim B(1, p) \qquad p \in (0,1)$

$$L(p, x) = \prod_{i=1}^{n} \mathbb{P}(X = x_i) = \prod_{i=1}^{n} p^{x_i} \cdot q^{1-x_i}$$

$x = 0, 1 \qquad \mathbb{P}(X=0) = q \qquad \mathbb{P}(X=x) = p^x \cdot q^{1-x}$

$\mathbb{P}(X=1) = p$

$$L(p,x) = \sum_{i=1}^{n} \ln\left(p^{x_i} \cdot q^{1-x_i}\right) = \sum_{i=1}^{n} \ln(p^{x_i}) + \sum_{i=1}^{n} \ln(q^{1-x_i})$$

$$= s \cdot \ln(p) + (n-s)\ln(q) = s \cdot \ln(p) + (n-s) \cdot \ln(1-p)$$

$$= \frac{d\, L(p,x)}{d(p)} = \frac{s}{p} + (n-s)\frac{1}{1-p}$$

$$= \frac{s}{p} + \frac{n-s}{1-p}$$

$$\frac{d\, L(p,x)}{d(p)} = 0 \iff \frac{s}{p} = \frac{n-s}{1-p} \iff s(1-p) = (n-s)p \iff$$

$$\iff s - \cancel{sp} = np - \cancel{sp} = s = np \iff \boxed{p = \frac{s}{n}} \longrightarrow$$ CANDIDATO A ESSERE MAX ASSOLUTO

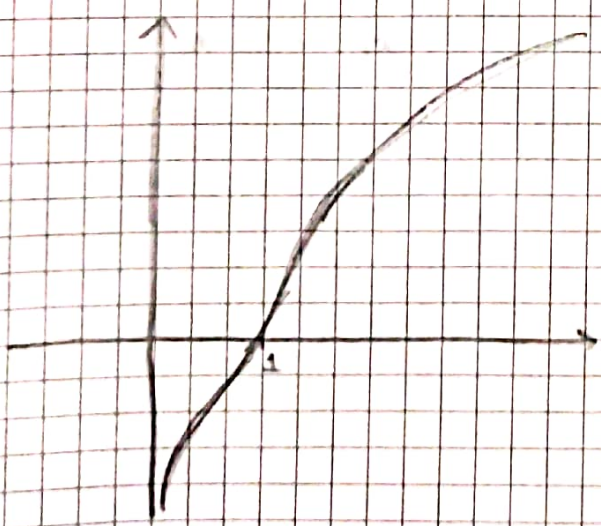
$$p = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\frac{d^2 L(p,x)}{d^2 p} = -\frac{s}{p^2} - \frac{(n-s)}{(1-p)^2} < 0$$

$$= L(p,x) = n \ln(p) + (n-s)\ln(1-p)$$

$$\lim_{p \to 0^+} L(p, x_i) = -\infty \ ; \quad \lim_{p \to 1^-} L(p, x) = -\infty$$

$$\hat{p}_{MV} = \frac{1}{n}\sum_{i=1}^{n} x_i$$



LEZIONE 32

- $x \in S_x$, $\hat{\theta}_{MV} = \underset{\theta \in \Theta}{\arg\max}\ L(\theta, \underline{x})$

$= \underset{\theta \in \Theta}{\arg\max}\ \ell(\theta, x)$

$x \in S_x$, $\theta \in \Theta$, $L(\theta, x) = \prod_{i=1}^{n} f_X(x_i; \theta)$

$X \sim N(\mu, \sigma^2)$

$-\prod_{i=1}^{n} f_X(x_i; \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\sigma^2}\right)^{\frac{1}{2}} \cdot e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$

① $= \left(\frac{1}{\sqrt{2\pi}}\right)^{2} \left(\frac{1}{\sigma^2}\right)^{n/2} \cdot e^{-1/2\sigma^2 \sum_{i=1}^{n} x_i} := L$

$-\ell(\mu, \sigma^2, x) = -n \ln \sqrt{2\pi} - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$

- μ incognito, σ^2 noto

$\ell(\mu, x) = n \ln(\sqrt{2\pi}) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$

$-\frac{d\, \ell(\mu, x)}{d\mu} = \frac{2}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)$

$\Downarrow$

② eq. punti stazionari $\longrightarrow \sum_{i=1}^{n} (x_i - \mu) = 0 \iff$

$\iff \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu \iff n\mu = \sum_{i=1}^{n} X_i \iff \boxed{\mu = \overline{X}}$

è l'unico massimo assoluto

$-\frac{d^2 \ell(\mu, x)}{d^2\mu} = -\frac{n}{\sigma^2} < 0$ sempre negativa

$\hat{\mu}_{MV} = \overline{X} \wedge \hat{M}_{VM} = \overline{X}$ stimatore della massima verosimiglianza

e

- μ noto, σ^2 incognita $\qquad \beta = \sigma^2$

$$\ell(\beta, x) = n \cdot \ln(\sqrt{2\pi}) - \frac{n}{2} \ln \beta - \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta}$$

$$-\frac{d\,\ell(\beta, x)}{d\beta} = -\frac{n}{2} \cdot \frac{1}{\beta} + \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta^2}$$

$\downarrow$

eq. punti stazionari $\rightarrow \frac{n}{\cancel{2}} \cdot \frac{1}{\cancel{\beta}} = \frac{1}{\cancel{2}} \sum_{i=1}^{n} (x_i - \mu)^2 \frac{1}{\beta^{\cancel{2}}} \iff$

$$\iff n = \sum_{i=1}^{n} (x_i - \mu)^2 \frac{1}{\beta} \iff \boxed{\beta = \frac{1}{n} \sum_{i=0}^{n} (x_i - \mu)^2}$$

$$-\frac{d^2 \ell(\beta, x)}{d^2\beta} = \frac{n}{2} \cdot \frac{1}{\beta^2} - \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu) \frac{2}{\beta^3}$$

$$\frac{d^2 \ell(s_d^2, \bar{x})}{d^2\beta} = \frac{(s_d^2)^2}{2} \left[n - 2 \cancel{\sum_{i=1}^{n} (x_i - \mu)^2} \cdot n / \cancel{\sum_{i=1}^{n} (x_i - \mu)^2} \right]$$

$$= -\frac{(s_d^2)^2}{2} \cdot n < 0$$

$\longrightarrow \hat{\sigma}^2_{MV} = s^2_d \Rightarrow \hat{S}^2_{MV} = S^2_d$

- μ incognita, σ^2 incognita $\qquad \beta = \sigma^2$

$$\ell(\mu, \beta, x) = n \ln(\sqrt{2\pi}) - \frac{n}{2} \ln(\beta) - \frac{1}{2\beta} \sum_{i=1}^{n} (x_i - \mu)^2$$

derivate parziali

$$-\frac{d\,\ell(\mu, \beta, x)}{d\mu} = \frac{1}{\beta} \sum_{i=1}^{n} (x_i - \mu)$$

$$-\frac{d\,\ell(\mu, \beta, x)}{d\beta} = -\frac{n}{2} \cdot \frac{1}{\beta} + \frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta^2}$$

$$\begin{cases} \frac{1}{\beta} \sum_{i=1}^{n} (x_i, y) = 0 \\ +\frac{1}{2\beta} \left[-n + \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta} \right] = 0 \end{cases}$$

SISTEMA PER PUNTI STAZIONARI

$\iff$

$$\Leftrightarrow \begin{cases} \sum_{i=1}^{n} (x_i - \mu) = 0 \\ -n + \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta} = 0 \end{cases} \Leftrightarrow \begin{cases} \mu = \bar{x} \\ n = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot \frac{1}{\beta} \end{cases}$$

$$\Leftrightarrow \begin{cases} \mu = \bar{x} \\ \beta = s^2 d \end{cases}$$

punto stazionario: $(\bar{x}, s^2 d)$
dove $s^2 d = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

$$\Downarrow$$

$$\begin{cases} \hat{\mu}_{MV} = \bar{X} \\ \sigma^2 = s^2 d \end{cases} \Rightarrow \begin{cases} \hat{\mu}_{MV} = \bar{X} \\ \hat{S}^2_{MV} = S^2 d \end{cases}$$

$$\begin{bmatrix} A = x_1! \cdot \ldots \cdot x_n! \\ s = \sum x_i \end{bmatrix}$$

$X \sim \Pi(\lambda)$, $\lambda > 0$

$$\prod_{i=1}^{n} f_X(x_i, \lambda) = \prod_{i=1}^{n} e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!} =: L((\lambda, \underline{x}))$$

$$\ell(\lambda, x) = -\ln(A) - n\lambda + s \ln(\lambda)$$

Obiettivo: $\underset{\lambda > 0}{\text{argmax}} \left[-\ln(A) - n\lambda + s\ln(\lambda) \right]$

$$\frac{d\ell(\lambda, x)}{d\lambda} = -n + s/\lambda$$

$$\frac{d\ell(\lambda, x)}{d\lambda} = 0 \Leftrightarrow n = \frac{s}{\lambda} \Leftrightarrow \lambda = \frac{s}{n} = \bar{x}$$

$$\frac{d^2 \ell(\lambda, x)}{d^2\lambda} = -\frac{s}{\lambda^2} < 0$$

$$\hat{\lambda}_{MV} = \bar{x}, \quad \hat{\lambda}_{MV} = \bar{X}$$

METODO DEI MOMENTI

- è un metodo di ricerca degli stimatori
↓
in base al metodo dei momenti: uno stimatore deve soddisfare una condizione che caratterizza uno o più suoi momenti empirici

- θ ha dimensione $k = (\theta_1, \theta_2, \dots, \theta_k)$ — k parametri incogniti

Si sceglie $p \in \mathbb{N}$, $p \geq k$

Poi, $J_1, J_2, \dots, J_p \in \mathbb{N}$

sistema di p equazioni nei k parametri incogniti per ottenere

$\Rightarrow \mu'_{J_1} = \mathbb{E}(X^{J_1}) = f_{J_1}(\theta_1, \theta_2, \dots, \theta_k)$

$\vdots$

$\Rightarrow \mu'_{J_p} = \mathbb{E}(X^{J_p}) = f_{J_p}(\theta_1, \theta_2, \dots, \theta_k)$

funzione di momenti

$$\begin{cases} \theta_1 = g_1(\mu'_{J_1}, \dots, \mu'_{J_k}) \\ \theta_2 = g_2(\mu'_{J_1}, \dots, \mu'_{J_k}) \\ \vdots \\ \theta_k = g_k(\mu'_{J_1}, \dots, \mu'_{J_k}) \end{cases}$$

$\downarrow$

$\hat{\theta}_{1,MV} = g_1(\bar{X}^{J_1}, \dots, \bar{X}^{J_k})$

$\vdots$

$\hat{\theta}_{k,MV} = g_k(\bar{X}^{J_1}, \dots, \bar{X}^{J_k})$

- $X \sim U(0,b)$ b incognita $\Theta = b > 0$

$\mu'_1 = \mathbb{E}(X) = \frac{b}{2}$

$b = 2\mu'_1 \Rightarrow \hat{B}_{MM} = 2\bar{X} \wedge \hat{\theta}_{MV} = X_{(n)}$

LEZIONE 33

Metodi per la costruzione di stimatori

→ a - Metodo della massima verosimiglianza

→ b - Metodo dei momenti

$\underline{X} = (X_1, X_2, \dots, X_n)$, F_X , $\theta \in \Theta$

a

$L(\theta, x)$

- $\hat{\theta}_{MV} : \arg\max_{\theta \in \Theta} L(\theta, x) = \arg\max \ell$

↓

STIMA

- $\hat{\Theta}_{MV} := \arg\max L(\Theta, X)$

↓

STIMATORE

b

$$\begin{cases} \mu'_{j_1} = f_{j_1}(\theta_1, \theta_2, \dots, \theta_k) \\ \vdots \\ \mu'_{j_k} = f_{j_k}(\theta_1, \theta_2, \dots, \theta_k) \end{cases}$$

$\Updownarrow$

$$\begin{cases} \theta_1 = g_1(\mu'_{j_1}, \dots, \mu'_{j_k}) \\ \vdots \\ \theta_k = g_k(\mu'_{j_1}, \dots, \mu'_{j_k}) \end{cases}$$

$\Updownarrow$

$$\begin{cases} \hat{\theta}_{1,MV} = g_1(\bar{X}^{j_1}, \dots, \bar{X}^{j_k}) \\ \vdots \\ \hat{\theta}_{k,MV} = g_k(\bar{X}^{j_1}, \dots, \bar{X}^{j_k}) \end{cases}$$

- Uno stimatore è una statistica

$\forall n \in \mathbb{N}$, $T_n = g(X_n)$

$X = (X_1, X_2, \dots, X_n)$ c.c.s.

e.g. $\bar{X} = \sum_{i=1}^{n} X_i$

$\to (T_n)_{n \in \mathbb{N}}$ successione di stimatori

↓

è uno stimatore consistente per la stima di θ se e solo se $T_n \xrightarrow{\text{converge in prob.}} \theta$

THM

Condizione sufficiente per avere $(T_n)_{n \in \mathbb{N}}$ consistente è

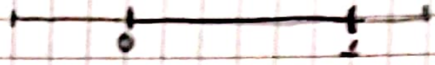
e 1) T_n asintoticamente corretto

2) $\forall n \in \mathbb{N}$, $\exists\, Var(T_n) < +\infty$ che è infinitesima al crescere di n

TMM

Se $(T_n)_{n\in\mathbb{N}}$ è consistente per θ e g è una funzione continua, allora $[g(T_n)]_{n\in\mathbb{N}}$ è consistente per $g(\theta)$

- $X \sim G(p)$, $p \in (0,1)$



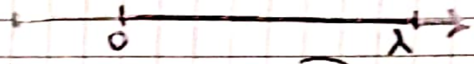
$j_1 = 1$

$\mu'_1 = 1/p \Rightarrow \exists g_1(p)$

$\downarrow$

$p = 1/\mu'_1 \Rightarrow \hat{p}_{MM} = \frac{1}{\bar{X}}$ è consistente per p

- $X \sim P(\lambda)$, $\lambda > 0$



$|t| < t_0$, $\phi_X(t) \Rightarrow$ Tutti i momenti campionari sono consistenti.

$j_1 = 1$

$\mu'_1 = \lambda \wedge \lambda = \mu'_1$, $\hat{\lambda}_{MM} = \bar{X}$

$\downarrow \qquad\qquad \downarrow$

$g_x \qquad\qquad g$

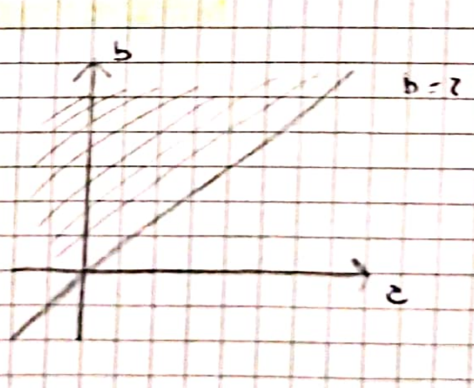
$E(X) = Var(X) = \lambda \longrightarrow Var(X) = \mu'_2 - (\mu'_1)^2 \Rightarrow \mu'_2 = Var(X) + (\mu'_1)^2$

$$\begin{cases} \mu'_1 = \lambda \\ \mu'_2 = \lambda + \lambda^2 \end{cases} \Rightarrow \lambda^2 + \lambda + \mu'_2 = 0$$

$$\lambda_{1,2} = \frac{-1 \pm \sqrt{1+4\mu'_2}}{2}$$

$$2\ \hat{\lambda}_{MM} = \frac{-1 + \sqrt{1+4\bar{X}^2}}{2}$$

- $X \sim U(a,b)$



$j_1 = 1$, $j_2 = 2$

$$\mu'_1 = E(X) = \int_a^b \frac{x}{b-a}\,dx$$

$$= \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\frac{x^2}{2}\Big|_a^b$$

$$= \frac{b^2-a^2}{2(b-a)} = \frac{b+a}{2}$$

$$\mu'_2 = \mathbb{E}(X^2) = \int_a^b \frac{x^2}{b-a}\,dx = \frac{1}{b-a}\int_a^b x^2\,dx = \frac{1}{b-a}\,\frac{x^3}{3}\Big|_a^b .$$

$$= \frac{a^2+ab+b^2}{3}$$

$$\begin{cases} \mu'_1 = \dfrac{a+b}{2} \\ \mu'_2 = \dfrac{a^2+ab+b^2}{3} \end{cases} \Leftrightarrow \begin{cases} a+b = 2\mu'_1 \\ a^2+ab+b^2 = 3\mu'_2 \end{cases} \Leftrightarrow \begin{cases} a+b = 2\mu'_1 \\ (a^2+2ab+b^2)-ab = 3\mu'_2 \end{cases}$$

$$\Leftrightarrow \begin{cases} a+b = 2\mu'_1 \\ 4(\mu'_1)^2 - ab = 3\mu'_2 \end{cases} \Leftrightarrow \begin{cases} a+b = 2\mu'_1 \\ ab = 4(\mu'_1)^2 - 3\mu'_2 \end{cases}$$

$a+b = S \quad , \quad ab = P$

$z^2 - Sz + P = 0 \qquad z_{1,2} = \dfrac{S \pm \sqrt{S^2-4P}}{2}$

$z_1 + z_2 = S/1 \,, \quad z_1 z_2 = P \qquad a = \dfrac{S - \sqrt{S^2-4P}}{2} \,, \quad b = \dfrac{S \pm \sqrt{S^2-4P}}{2}$

$$a = \frac{2\mu'_1 - \sqrt{4(\mu'_1)^2 - 16(\mu'_1)^2 + 12\mu'_2}}{2} = \frac{2\mu'_1 - \sqrt{12[\mu'_2 - (\mu'_1)^2]}}{2}$$

$$\begin{cases} a = \mu'_1 - \sqrt{3[\mu'_2 - (\mu'_1)^2]} \\ b = \mu'_1 + \sqrt{3[\mu'_2 - (\mu'_1)^2]} \end{cases} \Rightarrow \begin{cases} \hat{A}_{MM} = \bar{X} - \sqrt{3[\bar{X}^{(2)} - (\bar{X})^2]} \\ \hat{B}_{MM} = \bar{X} + \sqrt{3[\bar{X}^{(2)} - (\bar{X})^2]} \end{cases}$$