SPAZI DI PROB

1.1 dado + moneta $\Rightarrow \Omega = D \times M$

1.2 50 biglietti/premi 7 piacciono a Chiara, 5 a Marco, 1 solo ad entrambi

$\Omega = \{1, .., 50\}$ descrive i possibili esiti della pesca

un evento è un sottoinsieme di Ω vs esito

evento

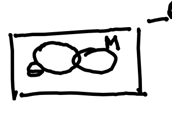
C = "il premio piacerà a C"

M

scrivere in termini di C ed M gli eventi (cioè insieme di esiti)

premio piacerà ad entrambi $C \cap M$

premio piacerà ad almeno uno dei due $C \cup M$



Ω

a nessuno dei due piacerà il premio $(C \cup M)^c$

il premio piacerà a uno solo dei due $(C \setminus (C \cap M)) \cup (M \setminus (M \cap C)) = (C \cup M) \setminus (C \cap M)$

1.3 Ω; A, B, C eventi $\Rightarrow$

tradurre in formule i seguenti eventi:

1) almeno un evento si verifica $A \cup B \cup C$

2) al più un evento si verifica $\underbrace{(A \cap B^c \cap C^c) \cup \ldots}_{} \cup \underbrace{(A \cup B \cup C)^c}_{}$ (uno o nessuno)

per dire che posso pescare anche nessuno (al più) vs esattamente

3) nessun evento si verifica $\Rightarrow$ $(A \cup B \cup C)^c$

4) tutti gli eventi si verificano $A \cap B \cap C$

5) si verifica esattamente esattamente uno

6) due eventi su tre si verificano (contemporaneamente)

$(A^c \cap B \cap C) \cup (A \cap B^c \cap C) \cup \ldots$

tradurre in termini probabilistici

7) A e C si escludono a vicenda

$P(A \cap C) = 0$

8) almeno un evento fra B e C si verifica

$P(B \cup C) = 1$

1.4 moneta lanciata due volte. Ale vince se al primo lancio esce T, Berna vince se al secondo lancio esce croce (eventi A e B)

1) Ω descrive tutti i possibili esiti dell'esp

$\{(T,T), (C,T), (T,C), (C,C)\}$ $A = \{(T,T), (T,C)\}$ $B = \{(T,C), (C,C)\}$

esprimere i seguenti eventi:

2) Ale non vince $= A^c$

Berna ——— $= B^c$

Ale e Ber vincono entrambi $A \cap B$

vince Ale ma non Berna $A \cap B^c$

almeno uno dei due vince $A \cup B$

nessuno dei due vince $(A \cup B)^c$

vince soltanto uno dei due $(A \cap B^c) \cup (A^c \cap B)$

esce testa al primo lancio ed esce croce al primo lancio

$\emptyset$

esce o testa o croce al secondo lancio

Ω

1.5

tradurre in formule gli eventi associati ad A, B, C e D

1) esattamente tre eventi su quattro si verificano

$(A \cap B \cap C \cap D^c) \cup \ldots$

2) si verifica solo C

$C \cap A^c \cap B^c \cap D^c$ ☆

3) si verifica solo C oppure si verifica solo D

☆ $\cup (D \cap C^c \ldots)$

4) almeno un evento si verifica

$A \cup B \cup C \cup D$

$P(\Omega)=1$

1.6 $\boxed{\text{prova di Bernoulli}}$: esperimento aleatorio con solo due esiti possibili (indicati con 0 e 1 e chiamati rispettivamente insuccesso e successo

vs

$\boxed{\text{delta di Dirac fissata in } x_0}$: $\delta_{x_0}(A) = \begin{cases} 1 & \text{se } x_0 \in A \\ 0 & \text{else} \end{cases}$

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

$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - \ldots + P(A \cap B \cap C)$

$(A + \cancel{x}) + (B + \cancel{x}) + (C + \cancel{x})$

• $\Omega = \{0,1\}^5$ spazio degli esiti di cinque prove di Bern.

E_n = "successo alla prova n"

scrivere in termini di $E_1, \ldots, E_5$ oppure come sottoinsiemi di Ω i seguenti eventi

$(E_n = \{(x_1, \ldots x_n, \ldots) \mid x_n = 1 \text{ e } x_i \in \{0,1\}\})$

$|E_n| = 2^4$

1) solo insuccessi

(fisso gli altri) $E_1^c \cap E_2^c \cap \ldots$ $\{(0,0,\ldots)\}$

2) solo la terza prova dà un successo

$E_1^c \cap E_2^c \cap E_3 \cap E_4^c \cap E_5^c$

3) nelle prove dispari ci sono solo successi

$E_1 \cap E_3 \cap E_5$

4) solo un successo

$(E_1 \cap E_2^c \cap E_3^c \cap \ldots) \cup (E_1^c \cap E_2 \cap E_3^c \cap \ldots) \cup \ldots$

1.7

$\Omega = \{1, \ldots, 6\}^4$ lo spazio degli esiti di quattro lanci di un dado.

$\{(x_1, x_2, x_3, x_4) \mid x_i \in \{1, 2 \ldots 6\}\}$

E_n^l = "faccia l al lancio n" $l = 1, \ldots, 6$ $n = 1, \ldots 4$

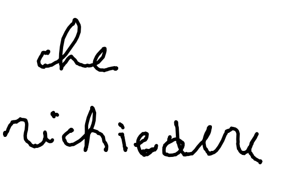
evento

determinare i seguenti eventi:

1) esce sempre 1

$E_1^1 \cap E_2^1 \cap \ldots$

$E_3^3 \setminus E_1^3 \cup \ldots$



2) solo il terzo lancio dà un 3

$E_3^3 \cap E_1^{3c} \cap E_2^{3c} \cap E_4^{3c}$

3) solo un quattro $\Rightarrow$

$(E_1^4 \cap E_2^{4c} \cap \ldots) \cup (E_1^{4c} \cap E_2^4 \cap \ldots) \cup \ldots$

1.8 una ditta riceve richieste di forniture che possono essere urgenti oppure no, e richiedere la consegna in città oppure fuori

• prob che ci sia consegna fuori città: 0.4

• urgente 0.3

• non urg in città = 0.4

• $P(C) = 0.6$ • $P(U) = 0.3$ • $P(U^c \cap C) = 0.4$

• $P(C^c) = 0.4$ • $P(U^c) = 0.7$

$\boxed{P(U \cap C^c)}$?

$P(U \cup C^c) = P(U) + P(C^c) - P(U \cap C^c)$

$\overline{U \cup C^c} = U^c \cap C \Rightarrow$

$1 - 0.4 = 0.3 + 0.4 - x$

$0.6 = 0.7 - x$

$x = \boxed{0.1}$

$\boxed{P(U^c \cup C)} = P(U^c) + P(C) - P(U^c \cap C) =$

$0.7 + 0.6 - 0.4 = \boxed{0.9}$

$P(U \cup C)$?

$P(U \cup C) = P(U) + P(C) - P(U \cap C)$

$\Omega = \{\underbrace{(U, C)}_{0.2}, \underbrace{(U, C^c)}_{0.1}, \underbrace{(U^c, C)}_{0.4}, \underbrace{(U^c, C^c)}_{0.3}\}$

$P(C^c) - 0.1$

1.9 $\boxed{(\Omega, P) \text{ spazio di prob}}$

$P(A) = 0.4$ $P(B) = 0.7$

• $P(A \cup B) = 0.4$ X

• $P(A \cup B) = 0.7$ può essere vera (V/F) $\geq$

• $P(A \cup B) = 1.1$ X

• $P(A \cap B) = 0.28$ non si può sapere (V/F)

• $P(A \cap B^c) \leq 0.3$. min (0.4, 0,3) ✓

$P(A^c \cap B) \geq 0.3$ min (0.6, ..) ✓

dado truccato $\Omega = \{1, ..4\}$

$P(1) = 2(P(2))$

$P(2) = 2\,P(3)$

$P(3) = 2P(4)$

calcolare Prob sapendo che $P(\{1\}) + \ldots + P(\{4\}) = P(\Omega) = 1$

PROB CONDIZIONATA E INDIP

"sapendo che B si è verificato, qual è la prob che..." $P(A|B) = \frac{P(A\cap B)}{P(B)}$

guardo solo quegli esiti di A che cadono anche in B

2.a
lanciamo dado => prob che esca un numero ≥ 3 sapendo che è uscito un numero pari

$A = \{3,4,5,6\}$
$B = \{2,4,6\}$ => $P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{P(\{2,4\})}{\frac{1}{2}} = \frac{2}{6}\cdot\frac{2}{1} = \frac{2}{3}$

$C = \{2,4,6\}$
$D = \{5\}$ => $P(D|C) = 0$

vale additività finita se $A_1 \cap A_2 = \emptyset$ =>

$P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$ $\cup A_3$ $+ P(A_3|B)$

$P(A|B) = 1 - P(A^c|B)$

monotonia
se $A_1 \subset A_2 \Rightarrow P(A_1|B) \leq P(A_2|B)$

Regola della catena
(segue da def)

$P(B) > 0 \Rightarrow$

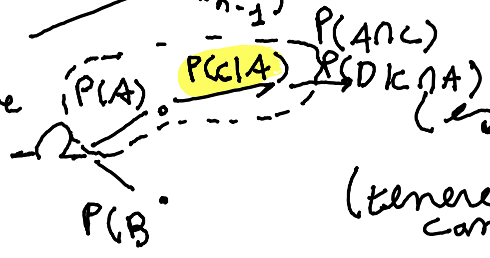
$$\boxed{P(A\cap B) = P(A|B)\cdot P(B)}$$

"sapendo che si è verificato questo, questo... e questo" segue ric da def

$$P(A_n|\underbrace{A_1\cap \dots A_{n-1}}_{>0})\cdot P(A_{n-1}|A_1 \dots)\cdot \dots P(A_2|A_1)\cdot P(A_1) =$$

$$\frac{P(A_1\cap \dots A_{n-1}\cap A_n)}{\cancel{P(A_1\cap \dots A_{n-1})}}\cdot\frac{\cancel{P(A_1\cap \dots A_{n-2}}\cap A_{n-1})}{P(\quad)}\dots\frac{\cancel{P(A_1)}}{\cancel{P(A_1)}}$$

(oss: da ogni nodo si va in una partizione di Ω)



sottoesperimenti aleatori => $\Omega = A \times B$
Ω sub est.
(estrazione senza reimmissione (tenere traccia cambiamenti) (ordine temporale)

Eventi indipendenti $A \perp\!\!\!\perp B$

$$\boxed{P(A\cap B) = P(A)P(B)}$$

(se sono disg => gran di p e.g. A = esce pari B = esce dispari)

<=> se $P(B)>0$ => $A \perp\!\!\!\perp B \Leftrightarrow P(A|B) = P(A)$

$$\boxed{\text{se } A \perp\!\!\!\perp B \Rightarrow A^c \perp\!\!\!\perp B \text{ e } A \perp\!\!\!\perp B^c}$$

$B = \Omega \cap B = (A^c \cup A)\cap B = (A^c\cap B)\cup(A\cap B)$ =>
$P(B) = P(A^c\cap B) + P(A\cap B) \underset{\text{ind}}{=} P(A^c\cap B) + P(B)P(A)$
=>
$P(A^c\cap B) = P(B) - P(B)P(A) = (1-P(A))P(B) = P(A^c)P(B)$

3 eventi sono ind se (ind a coppie due a due)
$P(A\cap B) = P(A)P(B)$
$P(A\cap B\cap C) = P(A)P(B)P(C)$
(ogni evento è indipendente da qualsiasi intersezione di altri eventi)

tutte coppie
tutte terne
quaterne

estrazione senza reim => $\Omega = \{(x_1, \dots x_n) \mid x_i \neq x_j \ \forall i,j\}$
$x_i \in \{1, \dots 90\}$

Formula delle prob totali

$B_1, \dots B_n$ partizioni di Ω =>

$P(A) = \sum P(A\cap B_i) = \sum P(A|B_i)P(B_i)$

se $P(B_i) > 0$ e per def reg catena

somma cammini

una volta raggiunto ramo con 0 rami, non si procede ($P(B) > 0$)

Formula di Bayes

$P(A) > 0 \wedge P(B) > 0 \Rightarrow$

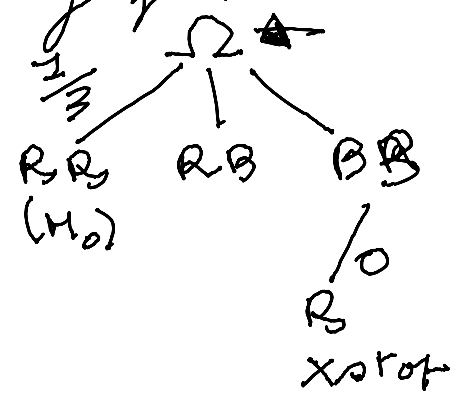
IMP

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

reg catena

e.g.

in un'urna ci sono due palline di due colori ma non si sa config =>



evento che avviene più in là nel tempo => abbiamo partizionato il sottoesperimento precedente

$$P(B_3|B_1\cap B_2) = \frac{P(B_1\cap B_2\cap B_3)}{P(B_1\cap B_2)}$$

formula prob tot

$P(B_1\cap B_2) = P(B_1\cap B_2\cap H_0) + P\dots$

$0.9‰ = 0.009$

"sapendo che nel secondo sottoesperimento aleatorio alla fine hai ottenuto quest." non so come l'ho ottenuto =>

Ω finito e esiti equiprobabili => ci riduciamo a problemi di conteggio

Calcolo combinatorio

qui albero di deriv. complementare simmetrico foglia = elem cammino

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

$n \downarrow$ $k \rightarrow$

$\binom{0}{0} = 1$

$\binom{1}{0} = 1 \quad \binom{1}{1} = 1$

$\binom{2}{0} = 1 \quad \binom{2}{1} = 2 \quad \binom{2}{2} = 1$

Stifel

$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^n b^{n-k}$

Metodo delle scelte successive

scomporre il problema in una serie di scelte successive

fissare il numero di scelte

- errori più comuni
  - non contare tutti gli elementi di A
  - contare più di una volta lo stesso elemento/stessa mano

"bisogna capire se dopo queste due scelte ho determinato in modo univoco una cinquina $(13\binom{4}{3} \cdot 12\binom{4}{2})$"

"significato per"

55	66
66	77
77	88
88	(99)
(999)	10 10 10

$|A| => |\{\ldots (22,11) \ldots\}|$

cardinalità di insieme di insieme => non conta ordine

e.g.

insieme delle doppie coppie

Ogni elemento di A può essere determinato tramite sei scelte successive

- scelta del tipo della prima coppia : 13 poss
- ——— seconda coppia : 12 ———
- ——— quinta carta : 11 ———
- ——— dei semi della prima coppia : 4 ———
- ——— sec ——— 4
- ——— del seme della quinta carta : 4 ———

"il tipo della seconda coppia ... il seme no"

ma

99 88 7 se facessi il diagramma ad albero avrei due sequenze identiche

88 99 7 => dividere per due

(cammino → insieme)

"ogni doppia coppia non viene determinata da una ed una sola sequenza ma da esattamente due sequenze distinte la ragione è che le prime due scelte sono ambigue dal momento che non esiste una prima ed una seconda coppia"

$\binom{13}{2} = \frac{13 \cdot 12}{2} = 78$

$13 \cdot 12 \cdot 11 \cdot 6 \cdot 6 \cdot 4$

$\downarrow$

$78 \cdot 11 \cdot 6 \cdot 6 \cdot 4$

$\binom{4}{2}\binom{4}{2}\binom{4}{1}$

e.g.

$\binom{13}{3} \Rightarrow \{(8,9,7), \cancel{(7,9,8)} \ldots\}$

88 99 78

(vettore)

$\boxed{DR_{n,k}}$ insieme delle disposizioni con ripetizione di k elementi di E

$|DR_{n,k}| = n^k$

ordine conta!

esprime i modi in cui possiamo disporre, in maniera ordinata ed eventualmente ripetuta un numero k di oggetti scelti da un insieme di n oggetti

estrazioni con reimmissione

$\updownarrow$

$\Omega = DR_{n,k}$

vs $|D_{n,k}| = \frac{n!}{(n-k)!}$ $n \cdot n-1 \cdots n-k+1$

$k \leq n$

estrazioni senza ripetizioni

Combinazioni

estrazioni simultanee

"pesco un sottoinsieme"

e.g. cinquina semplice

$5! = P_5$ in quanti modi differenti si possono ordinare 5 numeri

$|D_{5,5}|$

$|D_{90,5}|$

$\frac{1}{\binom{90}{5}}$

modello matematico può cambiare ma non il risul

$P(\{w\}) = \frac{1}{\binom{n}{k}}$

se ordine conta, cardinalità dell'evento cambia in n!

$|D_{n,k}| = |C_{n,k}| \cdot |P_k|$

altre perm

k per tenere conto dell'ordine

tutte poss perm

$3 \cdot 2 =$

1 2	2 3	3 1
2 1	3 2	1 3

Probabilità binomiale

non è nota

"Determiniamo la cardinalità di A_k tramite le seguenti scelte succ: che fissano

- scelta della sequenza (ordinata e con eventuali rip) delle k palline bianche estratte : $|DR_{n,k}|$ poss
- ——— n−k ——— rosse ——— : $|DR_{n,n-k}|$ poss
- scelta delle k estrazioni in cui sono uscite le palline bianche : $|C_{n,k}|$ poss

$P(A_k) = \binom{n}{k} b^k \cdot r^{n-k} / (b+r)^n$

tra le tante palline bianche... $b_1, \ldots b_i$

esperimento aleatorio con ripetizione $=> \Omega - |DR_{n,k}|$

scelta palline shift pos

amici al bar 8

$4 \cdot 3$ (scelta senza rip dei bar diversi)

e.g.

si calcoli la prob di ottenere (esattamente) 2 palline rosse estraendone 4 da un'urna che contiene 3 R 7 B.

estr con reim

$\frac{\binom{4}{2} DR_{3,2} DR_{7,2}}{DR_{10,4}}$

estr senza reim

$\frac{\binom{4}{2} D_{3,2} D_{7,2}}{D_{10,4}}$

ordine, come spazio camp, conta

estr simultanea

$\frac{\binom{3}{2} \cdot \binom{7}{2}}{\binom{10}{4}}$

Possibili sottoinsiemi due a due di palline bianche. possibili...

e.g.

estrazione lotto indovino 5/6 + jolly

fissati

dentro 6/7 shift pos

$\binom{6}{1} \cdot 6! \cdot 84$

$D_{90,7}$

"scelgo il valore del numero NON indovinato"

in che ordine escono i numeri giocati

ordine perché devo indovinare il 7°

MANO

$\Downarrow$

Ω COMBINAZIONI

e.g.

mazzo

52 carte

4 semi

13 tipi

- prob scala reale (una qualsiasi scala di 5 carte dello stesso seme)

  $4 \cdot 10$

- prob scala non stesso seme

  $4^5 \cdot 10$

- prob di avere colore (sottoinsieme di 5 carte dello stesso seme)

  $4\binom{13}{5} / \binom{52}{5}$

- prob poker : sottoinsieme di 5 carte in cui ci sono 4 carte dello stesso tipo

  $13 \cdot 4 \cdot 12$

  8888 .

  9999

metodo delle scelte successive!

- tris

  stesso tip 3, tipi diff (ORDINE NON CONTA)

  $\left(13 \cdot \binom{4}{3}\right) \cdot \binom{12}{2} \cdot 4^2$

  poiché insiem...

  possono avere seme uguale!

- coppia

  2 carte dello stesso tipo, altre tipo diverso tra loro e dalle prime due

  $\left(13 \cdot \binom{4}{2}\right) \cdot \binom{12}{3} \cdot 4^3$

e.g.

4 semi

10 tipi

mano da 10, prob di avere

asso bastoni

$\binom{39}{9} / \binom{40}{10}$

sottoinsiem

asso bastoni + asso spade

$\binom{38}{8} / \binom{40}{10}$

e.g.

si divida un mazzo di 40 carte in due mazzi da 20. Qual è la prob dell'evento

A = "il primo mazzo contiene esattamente un 7?"

$|\Omega| = \binom{40}{20}$

$|A| = \binom{39}{19}$ (se 7 per 4 semi $= 4 \cdot \binom{36}{19}$)

"per ogni sottoinsieme e per ogni carta..."

capire meglio

$\underbrace{\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{2}{3} \cdots}_{13 \text{ volte}} + \binom{13}{1} \frac{2^{12}}{3^{13}}$

che soddisfano

tutti casi possibili

$|\Omega| = \binom{3}{2}^{13}$

sotto insiemi scelte seconda partita

|si verifica 1| = |si verifica 2| = ... = 2

e.g. $\{(1,x),(1,3)\}$

generalizzazione degli eventi

Variabili aleatorie : $\Omega \to \mathbb{R}$

affermazione riguardante l'esito dell'esperimento aleatorio

tale affermazione identifica uno ed uno solo numero reale una volta noto l'esito dell'esperimento aleatorio

chiedersi "quanto vale"

ad ogni evento è possibile associare una $X = 1_A$

var che dipende da esito

V.a.c. = se assume sempre lo stesso valore numerico, qualunque sia l'esito dell'esperimento aleatorio

$$X(\omega) = a \quad \forall \omega \in \Omega \qquad P(\{\omega \in \Omega : X(\omega) = a\}) = P(X = a) = 1$$

V.a. indicatrici

$$1_A = X(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{else} \end{cases} \qquad X = \text{"vale 1 se A si verifica, 0 altrimenti"}$$

una volta noto il valore che la v.a. X ha assunto, essa va a identificare una famiglia di eventi (da lei generati)

identificare (controimmagine)

considerare i valori che può assumere v.a.

⇓

evento associato ad X

$$\{\omega \in \Omega \mid X(\omega) \in B\} \qquad B \subset \Omega$$

eventi di cui so dire con certezza se si sono verificati oppure no una volta noto il valore della X

$$\{X \in \emptyset\} = \emptyset \qquad \{X \in \mathbb{R}\} = \Omega$$

e.g. eventi generati

trovare gli eventi generati al variare di B

$$X(\omega) = a, \ \forall \omega \in \Omega \Rightarrow \{X \in B\}$$

$$P_X(B) = P(X \in B) = \begin{cases} P(\Omega) = 1 & \text{se } a \in B \\ P(\emptyset) = 0 & \text{se } a \notin B \end{cases} \qquad \text{V.a.c (delta di Dirac) } \delta_a(B)$$

$1_A \Rightarrow$

non c'è alcun elemento di Ω t.c. $X(\omega) \neq 0, 1$

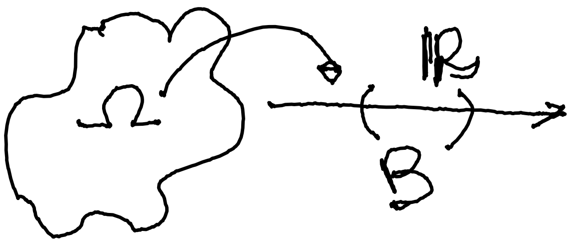
$$P_X(B) = P(X \in B) = \begin{cases} P(A) & 1 \in B, 0 \notin B \\ 0 = P(\emptyset) & 0, 1 \notin B \\ 1 - P(A) = P(A^c) & 0 \in B, 1 \notin B \\ 1 & 0, 1 \in B \end{cases}$$

V.a. di Bernoulliana

prende in input un insieme

distribuzione/legge di X

$$\boxed{P(\mathbb{R}) \to [0,1]}$$



(prob. che la v.a. prenda valori in B)

conoscere l'una è come conoscere l'altra

$$F_X \sim \boxed{X \sim P_X} = X \text{ ha distr. o legge } P_X$$

Funzione di ripartizione o CDF : $\boxed{\mathbb{R} \to [0,1]}$

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

estremo destro dell'insieme B

e.g.

V.a.i $X = 1_A$ cumulato fino a quel momento

$$F_X(x) = \begin{cases} 0 & x < 0 \\ P(A^c) & 0 \le x < 1 \\ P(A^c) + P(A) = 1 & x \ge 1 \end{cases} \qquad \text{V.a.c} \quad F_X(x) = \begin{cases} 0 & x < a \\ 1 & x \ge a \end{cases}$$

prop caratterizzanti ($\Leftarrow\Rightarrow$)

- F_X è monotona crescente (non necessariamente strettamente)
- continua a dx $\lim_{y \to x^+} F_X(y) = F_X(x) \quad \forall x \in \mathbb{R}$ (accumulato 0)
- $\lim_{x \to -\infty} F(x) = 0$
- $\lim_{x \to +\infty} F(x) = 1$

$$P(X = x) = P(X \le x) - P(X < x) = F_X(x) - F_X(x^-)$$

se continua in $x \Rightarrow = 0$

$$P(x < X \le y) = P(X \le y) - P(X \le x) = F_X(y) - F_X(x)$$

Densità discreta di X (funzione di massa di prob) $: \mathbb{R} \to [0,1]$

V.a. discrete

$p_X(x) = P(X = x)$

V.a. discreta & tabella della densità discreta

X	val.
p_X	...

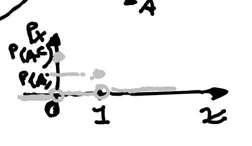
$= 1$

X v.a.d se $\exists S_X \subset \mathbb{R}$ (supporto) finito o inf numerabile t.c.

$\Rightarrow p_X(x_i) > 0 \quad \sum p_X(x_i) = 1 \quad x_i \in S_X$

$P(X \in S_X) = P(X \in \mathrm{Im}(X)) = 1$

1_A



X	0	1
p_X	$1-P(A)$	$P(A)$

Caratterizzazione delle variabili aleatorie discrete

Le seguenti affermazioni sono equivalenti tra loro

- X v.a.d
- F_X è una funzione costante a tratti e salta verso l'alto nei punti del supporto con ampiezza del salto pari a

$$F_X(x_i) - F_X(x_i -) = p_X(x_i)$$

$$\Downarrow$$

$$F_X(x) = \sum_{x_i \le x} p_X(x_i)$$

- P_X è concentrata nei punti del supporto

$$P(X \in B) = \sum_{x_i \in B} p_X(x_i) \quad \forall B \subset \mathbb{R}$$

e.g.

$$G(x) = \begin{cases} 0 & x < 0 \\ 1/2 & 0 \le x < 1 \\ 2/3 & 1 \le x < 2 \\ 11/12 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$

- $\to$ G è una funzione di ripartizione poiché
  - è monotona crescente
  - è continua a dx
  - $\lim_{x \to -\infty} G(x) = 0$
  - $\lim_{x \to +\infty} G(x) = 1$

Sia dunque X una v.a. con CDF $F_X = G$

- $\to$ X è discreta poiché G (cioè F_X) è costante a tratti e
$\Rightarrow$

X	0	1	2	3
dens disc p_X	$\frac{1}{2}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{12}$

$\{0,1,2,3\} = S_X$; $\sum = 1$

$P(X < 3) = F_X(3-) = \frac{11}{12}$

- $\to \underbrace{P_X}_{\text{distr legge}}(B) = P(X \in B) = \sum_{x_i \in B} p_X(x_i)$
- $\to P(X > \frac{1}{2}) = p_X(1) + p_X(2) + p_X(3) = \frac{1}{2} = \sum_{x_i > \frac{1}{2}} p_X(x_i)$
- $\to Y = (X-2)^2$ v.a.d. poiché
$S_Y = \{1, 0, 4\}$ $\quad p_Y(1) = p_X(1) + p_X(3)$

Y	0	1	4
p_Y	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$

Indici di sintesi di una distribuzione: μ e σ^2

$\mu_{arit} = \frac{x_1 + \ldots + x_n}{n}$ baricentro

e.g. voto dello scatto $= \frac{1}{10} + \frac{2}{10} + \cdots \frac{10}{10} = 5,5$

$\boxed{\text{Valore previsto che la v.a. assuma}}$

"expected value"

indice di Pos

media o valore atteso $\quad E[X] = \sum_i x_i \cdot p_X(x_i) = \mu_X \quad$ media pesata

e.g. $E[a] = a \cdot 1 = a \quad \forall a \in \mathbb{R} \Leftarrow$ valore atteso costante = costante

$E[1_A] = 0 \cdot P(A^c) + 1 \cdot P(A) = P(A)$

$$\boxed{E[Y] = E[h(X)] = \sum_i h(x_i) \cdot p_X(x_i)}$$

linearità del valore atteso $\quad \boxed{E[a h(X) + b g(X)] = a E[h(X)] + b E[g(X)]}$

Varianza

indice di dispersione

$$\frac{(x_1 - \mu)^2 + (x_2 - \mu)^2 + \ldots + (x_n - \mu)^2}{n}$$

$$\sigma_X^2 = \mathrm{Var}(X) = E[(X - E[X])^2] = \sum_i (x_i - E[X])^2 p_X(x_i)$$

σ = deviazione standard

$$\mathrm{Var}(X) = E[(X - E[X])^2] = E[X^2 \underbrace{- 2E[X]}_{\text{const}} X + \underbrace{E[X]^2}_{\text{const}}]$$

$= E[X^2] - 2(E[X])^2 + E[X]^2 =$ const che posso portare fuori + lin

$= E[\underbrace{X^2}_{h(x)}] - (E[X])^2 = \left(\sum_i x_i^2 p_X(x_i)\right) - (E[X])^2$

Proprietà della Varianza

- $\mathrm{Var}(X) \ge 0$
- $\mathrm{Var}(b) = 0$ $\leftarrow$ costante
- $\mathrm{Var}(\underbrace{aX + b}_{Y}) = E[(aX + b - E[aX + b])^2] =$

linearità $\to = E[(aX + b - aE[X] - b)^2] =$

$= a^2 E[(X - E[X])^2] = a^2 \mathrm{Var}(X)$

Distribuzioni notevoli discrete

- **Distribuzione uniforme discreta**

X	x_1	x_2	...	x_n
p_X	$\frac{1}{n}$	$\frac{1}{n}$		$\frac{1}{n}$

$X \sim \mathrm{Unif}(\{x_1, \ldots, x_n\})$

$E[X] = \frac{x_1 + \cdots + x_n}{n}$

$\mathrm{Var}(X) = \frac{(x_1 - E[X])^2 + \ldots + (x_n - E[X])^2}{n}$

- **Distribuzione di Bernoulli**

X	0	1
p_X	$1-p$	p

$p := P(A) \quad X \sim B(p)$

Le variabili aleatorie di Bernoulli sono tutte e sole variabili aleatorie indicatrici di un qualche evento A

funzione indicatrice del successo di (verifica) un qualche evento

$X = 1_A \quad A = \{X = 1\}$

$E[X] = p$

$\mathrm{Var}(X) = p(1-p)$

- **Distribuzione binomiale**

accade quando ho n prove di Bernoulli indipendenti e con la stessa probabilità di succ. p

A_i = "successo all'iesima prova"

e.g.

urna contenente b palline bianche ed r palline rosse

n estrazioni con reimmissione

X = "n° di palline bianche estratte"

$\Rightarrow X_i \sim B(p); \; p := \frac{b}{b+r}$

$S_X = \{0, 1, \ldots n\}$ $\leftarrow$ dimensione vettore estrazioni

$\underbrace{P(\{X = k\})}_{p_X(k)} = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, \ldots n$

si estraggono k palline bianche ed n-k palline rosse

$$\Downarrow$$

$$\boxed{X \sim B(n, p)}$$

$E[X] = n \cdot p$

$\mathrm{Var}(X) = n \cdot p(1-p)$

(e.g. chiavi)

"sapendo che al primo tentativo non ho trovato la chiave giusta ($\Rightarrow X > 1$), qual è la prob che non la trovo neanche al sec?"

$$\Rightarrow P(X > 2 \mid X > 1) = \frac{P(\{X > 2\} \cap \{X > 1\})}{P(\{X > 1\})} = \frac{P(X > 2)}{P(X > 1)}$$

V.a. continue

densità continua

PDF probability density function se proprietà caratterizzanti

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$
- $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

verifica queste prop

densità continua a differenza di quella discreta non è la prob. La prob è l'area!

e.g. $f_X(x) = \begin{cases} 50 & 0 \leq x \leq \frac{1}{50} \\ 0 & \text{else} \end{cases}$

X v.a. continua se $\exists$ una PDF f_X t.c.

$$P(a \leq X \leq b) = \int_a^b f_X(x)\,dx$$

$P(a \leq X \leq b) = P(a < X < b) + P(X = b)$

nel caso continuo i singoli punti/gli estremi non contano

Legge di X

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

$$\mathbb{P}(X \leq x) = \mathbb{P}(X < x) = \int_{-\infty}^{x} f_X(y)\,dy$$

Oss.

Se X è una v.a.c, sua densità f_X non è unica

e.g. $f_X(x) = \begin{cases} 50 & 0 \leq x \leq \frac{1}{50} \\ 0 & \text{else} \end{cases} \Rightarrow \int_B f_X(x)\,dx = \int_B h(x)\,dx$

$h(x) = \begin{cases} 50 & 0 \leq x \leq \frac{1}{50} \\ 12 & x = 72 \\ 0 & \text{else} \end{cases}$

$S_X = \{x \in \mathbb{R} \mid f_X(x) > 0\}$ (prop caratterizzanti)

$P_X(x) = P(X = x) = 0$

$\int x = \frac{x^2}{2} \Rightarrow D(\frac{x^2}{2}) = x$

CDF $\rightarrow F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy \Rightarrow$ F_X continua non solo a destra

$\Downarrow$

$\mathbb{P}(X \leq x)$

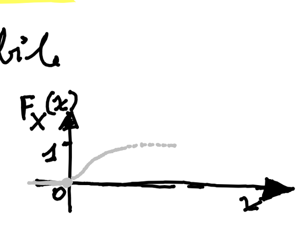
$F_X(x) = F_X(x^-) = \mathbb{P}(X \leq x) = \mathbb{P}(X < x)$

Dalla funzione di ripartizione alla densità continua

$F'_X(x) = f_X(x) \quad \forall x$ in cui F_X è derivabile

e.g. $F_X(x) = \begin{cases} 0 & x \leq 0 \\ (1-e^{-x})^2 & x \geq 0 \end{cases}$

v.a.c



densità $\Rightarrow f_X(x) = F'_X(x)$

deriviamo nei punti $\Rightarrow$ in cui si può derivare

una delle densità possibili di questa v.a.

l'imp è che sia definita per tutti i punti

nei punti di raccordo non ci interessa se la derivata è continua o meno

$f_X(x) = \begin{cases} 0 & x \leq 0 \\ 2(1-e^{-x}) \cdot e^{-x} & x > 0 \end{cases}$

è possibile modificare il valore di f_X in un numero finito o numerabile di punti

e.g. $f_X(0) = 0$

$\mathbb{P}(X > 1) = \int_1^{+\infty} f_X(x)\,dx = F_X(+\infty) - F_X(1) = 1 - F_X(1) = 1 - (1-e^{-1})^2$

$\mathbb{P}(1 < X < 2) = \int_1^2 f_X(x)\,dx = F_X(2) - F_X(1) = (1-e^{-2})^2 - (1-e^{-1})^2$

$D\left(\frac{3}{16}(1-e^{-(x-1)}) + \frac{13}{16}\right) = \frac{3}{16} \cdot D(\underbrace{1-e^{-x-1}}_{f(x)}) = \frac{3}{16} \cdot -1 \cdot -1 \cdot e^{-x-1}$

e.g.

Funzioni di variabili aleatorie continue

$Y = h(X)$

$F_Y(y) = \mathbb{P}(Y \leq y) = \underbrace{\mathbb{P}(h(X) \leq y)}$

obiettivo: esprimerla in termini della funzione di ripartizione di X

e.g.

$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 1 - e^{-x} & x \geq 0 \end{cases}$

v.a.c.

incognita CAPIRLO PRIMA POSSIBILE

$S_X = [0, +\infty)$

$S_Y = [1, +\infty)$ "$e^0, \ldots$"

$Y = e^X \Rightarrow F_Y(y) = \mathbb{P}(e^X \leq y)$

sua densità?

$F(y) = 0 \quad y \leq 1$

- se $y \leq 0 \Rightarrow$ disuguaglianza mai verificata

$\{e^X \leq y\} = \emptyset \Rightarrow F_Y(y) = 0 \quad \forall y \leq 0$

- $y > 0 \Rightarrow$

$e^X \leq y \iff X \leq \log Y$

altrimenti log non ha senso

$F_Y(y) = \mathbb{P}(X \leq \log y) = F_X(\log y) =$

$= \begin{cases} 0 & \log y \leq 0 \\ 1 - e^{-\log y} & \log y \geq 0 \end{cases} = \begin{cases} 0 & y \leq 1 \\ 1 - e^{\log \frac{1}{y}} & y \geq 1 \end{cases} =$

$= \begin{cases} 0 & y \leq 1 \\ 1 - \frac{1}{y} & y \geq 1 \end{cases}$

$f_Y = \begin{cases} 0 & y < 1 \\ +y^{-2} & y > 1 \end{cases}$

F_X è CDF di una v.a.c. se

- monotona crescente
- continua (tutti si devono raccordare / ci deve essere continuità purché sia una funz integral
- $\lim_{x \to -\infty} F_X(x) = 0$
- $\lim_{x \to +\infty} F_X(x) = 1$

e.g. $F_X(x) = \begin{cases} 0 & x \leq 0 \\ ax^2 & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases} \Rightarrow a \cdot 1 = 1$

$S_X = [0,1]$

$\Downarrow$

$Y = X^4$ è v.a.c? densità

$S_Y = [0,1]$

$\Rightarrow F_Y(y) = \begin{cases} 0 & 0 \leq y \\ ? & 0 < y < 1 \\ 1 & y \geq 1 \end{cases}$

$F_Y(y) = \mathbb{P}(X^4 \leq y) =$ $y \geq 0$ altrimenti non ha senso

$= \mathbb{P}(-\sqrt[4]{y} \leq X \leq \sqrt[4]{y}) =$

se $\sqrt[3]{y}$ allora non vincolo

$= F_X(\sqrt[4]{y}) - F_X(\underbrace{-\sqrt[4]{y}}_{<0}) =$

$= F_X(\sqrt[4]{y}) = (\sqrt[4]{y})^2 \quad 0$

$\Rightarrow$

$0 < y < 1 \Rightarrow 0 < \sqrt[4]{y} < 1$

$\begin{cases} \sqrt{y} & 0 < y < 1 \\ 1 & \end{cases}$

$\Downarrow$

$f_Y(y) = \begin{cases} 0 & 0 \leq y \\ \frac{1}{2\sqrt{y}} & 0 < y \leq 1 \\ & y > 1 \end{cases}$

Media o valore atteso v.a.c.

$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx \qquad \mathbb{E}[Y] = \mathbb{E}[h(X)] = \int_{-\infty}^{+\infty} h(x) f_X(x)\,dx$

Varianza

Vale linearità del valore atteso

$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f_X(x)\,dx$

Varianza $= \sigma^2$

deviazione standard $= \sigma$

$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \int_{-\infty}^{+\infty} x^2 f_X(x)\,dx - \mathbb{E}[X]^2$

Proprietà della varianza

$\text{Var}(X) \geq 0$

$\text{Var}(aX+b) = a^2\,\text{Var}(X)$

Distribuzioni continue notevoli

- Uniforme

$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases} \qquad S_X = [a,b]$

$P_X\left(\biguplus_n A_n\right) = \sum_n \int_{A_n} f_X(y)\,dy$

e.g. $B = [-1, 0] \cup (2, 3] \cup [5, +\infty)$

$P_X(B) = \mathbb{P}(\underbrace{X \in [-1,0]}_{-1 \leq X \leq 0}) + \mathbb{P}(2 < X \leq 3) + \mathbb{P}(X \geq 5) =$

$= \int_{-1}^{0} f_X(y)\,dy + \int_2^3 f_X(y)\,dy + \int_5^{+\infty} f_X(y)\,dy$

e.g.

X = tempo di vita di un componente elettronico

$f_X(x) = \begin{cases} 0 & x \leq 0 \\ e^{-x} & x > 0 \end{cases}$

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

$= \underbrace{\int_{-\infty}^{0} f_X(y)\,dy}_{0 \Rightarrow F_X(x) = 0 \quad x \leq 0} + \int_0^x f_X(y)\,dy =$

$= [-e^{-y}]_0^x = -e^{-x} + 1 \quad x > 0$

e.g.

$f_R(x) = \begin{cases} cxe^{-x^2} & x > 0 \\ 0 & \text{else} \end{cases} \Rightarrow \int_{-\infty}^{+\infty} cxe^{-x^2} = 1$

v.a.c.

$= \int_0^{+\infty} cxe^{-x^2} = 1 =$

$= -\frac{c}{2} \int_0^{+\infty} -2x e^{-x^2} = 1$

$c\left[e^{-x^2}\right]_0^{+\infty} = -2$

$c[0 - 1] = -2$

$\Rightarrow c = 2$

$\Downarrow$

$f_R(x) = \begin{cases} 2xe^{-x^2} & x > 0 \\ 0 & \text{else} \end{cases}$

$F_R(0) = 0$

$F_R(x) = \int_0^x 2ye^{-y^2}\,dy =$

$= \left[-e^{-y^2}\right]_0^x =$

$= -e^{-x^2} + 1$

$F_R(x) = \begin{cases} 0 & x \leq 0 \\ -e^{-x^2} + 1 & x > 0 \end{cases}$

Metodo per parti: $\int f(x) \cdot g(x) = F(x) g(x) - \int F(x) g'(x)$

$P(R > 2) = 1 - F_R(2) = \cancel{1} + e^{-4} - \cancel{1} = e^{-4}$

$X = \begin{cases} 1 & \text{se } R < 2 \\ 0 & \text{else} \end{cases} \Rightarrow S_X = \{0, 1\}$

X	0	1
P_X	$1-e^{-4}$	e^{-4}

$-\int_0^{+\infty} x \cdot 2xe^{-x^2} = \underbrace{\left[e^{-x^2} \cdot x\right]_0^{+\infty}}_{0} - \underbrace{\int_0^{+\infty} e^{-x^2}}_{\frac{\sqrt{\pi}}{2}} = \frac{\sqrt{\pi}}{2}$

$X \sim B(\underbrace{1 - F_R(2)}_{\text{è una prob}})$

$\int_0^1 a(1-x^3) + \int_{-1}^{0} \frac{1}{4} = 1 \Rightarrow$

e.g.

$f_X(x) = \begin{cases} 0 & |x| \geq 1 \\ 1/4 & -1 < x < 0 \\ a(1-x^3) & 0 \leq x < 1 \end{cases} \Rightarrow \int_0^1 a(1-x^3) = \frac{3}{4}$

v.a.c.

$\Rightarrow S_X = (-1, 1)$

$a\left(\int_0^1 1 - \int_0^1 x^3\right) = \frac{3}{4}$

$\left(1 - \frac{1}{4}\right) = \frac{3a}{4} \Rightarrow \boxed{a = 1}$

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

$x \leq 0 \quad \int_{-1}^{x} f_X(y)\,dy = \left[\frac{1}{4}y\right]_{-1}^{x} = \frac{1}{4}x + \frac{1}{4}$

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

$= \frac{1}{4} + [y]_0^x - \left[\frac{y^4}{4}\right]_0^x =$

$= \frac{1}{4} + x - \frac{x^4}{4}$

$F_X(x) = \begin{cases} 0 & x \leq -1 \\ \frac{1}{4}x + \frac{1}{4} & -1 \leq x \leq 0 \\ \frac{1}{4} + x - \frac{x^4}{4} & 0 \leq x \leq 1 \\ 1 & x \geq 1 \end{cases}$

$P(X \leq \frac{1}{2}) = \frac{1}{4} + \frac{1}{2} - \frac{1}{64}$

$P(-\frac{1}{4} \leq X \leq \frac{1}{4}) = F_X(\frac{1}{4}) - F_X(-\frac{1}{4}) = \left(\frac{1}{4} + \cancel{\frac{1}{4}} - \frac{1}{4^5}\right) - \left(-\frac{1}{4^2} + \cancel{\frac{1}{4}}\right) =$

$P(X = 0) = 0$

$P(X \leq 3) = 1$

$= \frac{1}{4} - \frac{1}{4^5} + \frac{1}{4^2} = \frac{4^4 - 1 + 4^3}{4^5} = \frac{256 + 64 - 1 = 319}{1024}$

$E[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \int_{-1}^{0} \frac{x}{4}\,dx + \int_0^1 x(1-x^3)$

$E[X^2] = \int_{-\infty}^{+\infty} x^2 f_X(x)\,dx = \int_{-1}^{0} \frac{x^2}{4}\,dx + \int_0^1 x^2(1-x^3)$

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

e.g.

$F_X(x) = \begin{cases} 0 & x < 0 \\ (1-e^{-x})^2 & x \geq 0 \end{cases} \Rightarrow$

v.a.c

$P(X > 1) = 1 - F_X(1) = 1 - (1 - \frac{1}{e})^2$

$f_X(x) = \begin{cases} 0 & x < 0 \\ 2(1-e^{-x})e^{-x} = 2(e^{-x} - e^{-2x}) & x \geq 0 \end{cases}$

e.g.

$F_X(x) = \begin{cases} 0 & x < 1 \\ 1 - e^{1-x} & x \geq 1 \end{cases} \Rightarrow S_X = (1, +\infty)$

$f_X(x) = \begin{cases} 0 & x \leq 1 \\ e^{1-x} & x > 1 \end{cases}$

$E[X] = \int_{-\infty}^{+\infty} x \cdot e^{1-x}\,dx = \int_1^{+\infty} x e^{1-x}\,dx = -\left(\left[e^{1-x} \cdot x\right]_1^{+\infty} - \int_1^{+\infty} e^{1-x}\right) =$

$-(-1 + [e^{1-x}]_1^{\infty}) = 1 - [e^{1-x}]_1^{+\infty} = 2$

$Y = e^X \Rightarrow S_Y = (e, +\infty) \Rightarrow F_Y(y) = 0 \quad y < e$

$F_Y(y) = \mathbb{P}(e^X \leq y) = \mathbb{P}(X \leq \ln y) \Rightarrow$

$F_Y(y) = \begin{cases} 0 & y < e \\ \underbrace{1 - e^{1-\ln y}}_{1 - \frac{e}{y}} & y \geq e \end{cases}$

$\Downarrow$

$f_Y(y) = \begin{cases} 0 & y \leq e \\ -e \cdot -1 \cdot y^{-2} & \text{else} \end{cases}$

e.g.

$f_X(x) = \begin{cases} \frac{2x}{\lambda} e^{-\frac{x^2}{\lambda}} & \text{per } x > 0 \\ 0 & \text{per } x \leq 0 \end{cases}$

v.a.c.

$\Rightarrow \int_0^{+\infty} \frac{2x}{\lambda} e^{-\frac{x^2}{\lambda}} = 1 = \left[e^{-\frac{x^2}{\lambda}}\right]_0^{+\infty} = -1 \Rightarrow$

$e^{-\frac{\infty}{\lambda}} - \cancel{1} = -\cancel{1}$

$e^{-\frac{\infty}{\lambda}} = 0 \Rightarrow$ $\boxed{\text{se } \lambda > 0 \Rightarrow e^{-\infty} = 0}$ se $\lambda < 0 \Rightarrow$ IMP

e.g.

$f_X(x) = \begin{cases} \lambda e^{-(x-\lambda)} & x > \lambda \\ 0 & x \leq \lambda \end{cases} \Rightarrow$

$\int_\lambda^{+\infty} \lambda e^{-(x-\lambda)} = 1 \Rightarrow$

$-\lambda \int_\lambda^{+\infty} -e^{-(x-\lambda)} = 1 \Rightarrow$

$-\lambda \left[e^{-(x-\lambda)}\right]_\lambda^{+\infty} = 1 \Rightarrow -\lambda\left[\cancel{e^{-\infty+\lambda}} - 1\right] = 1$

$\boxed{\lambda = 1}$

$f_X(x) = \begin{cases} e^{-x+1} & x > 1 \\ 0 & x \leq 1 \end{cases} \qquad S_X = (1, +\infty)$

$F_X(x) = \begin{cases} 0 & x \leq 1 \\ 1 - e^{-x+1} & x \geq 1 \end{cases}$

$P(X > 3) = 1 - F_X(3) = \cancel{1} - (\cancel{1} - e^{-3+1}) = e^{-2}$

$\int_1^x e^{-y+1}\,dy = -\left[e^{-y+1}\right]_1^x = 1 - e^{-x+1}$

$P(X^3 > 27) = P(X > 3) = e^{-2}$

$E[3X^2 - 1] = \int_{-\infty}^{+\infty} (3x^2 - 1) f_X(x)\,dx = \int_1^{+\infty} (3x^2 - 1)(e^{-x+1})$

e.g.

$f_X(x) = \begin{cases} \frac{3}{2}x^2 & -k < x < k \quad (-1 < x < 1) \\ 0 & \text{else} \end{cases}$

$\frac{3}{2} \int_{-k}^{k} x^2 = 1 \Rightarrow \left[\frac{x^3}{3}\right]_{-k}^{k} = \frac{2}{3} \qquad \frac{k^3}{3} + \frac{k^3}{3} = \frac{2}{3} \Rightarrow \frac{2k^3}{3} = \frac{2}{3}$

$k = 1$

$F_X(x) = \begin{cases} 0 & x \leq -1 \\ \frac{x^3}{2} + \frac{1}{2} & \\ 1 & x \geq 1 \end{cases}$

$\frac{3}{2} \int_{-1}^{x} y^2\,dy = \left[\frac{y^3}{2}\right]_{-1}^{x} = \frac{x^3}{2} + \frac{1}{2}$

$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x \cdot \frac{3}{2}x^2 = \frac{3}{2} \int_{-1}^{1} x^3 = \frac{3}{2}\left[\frac{x^4}{4}\right]_{-1}^{1} = 0$

$S_{X^2} = S_X$

$E[X^2] = \int_{-1}^{1} x^2 \cdot \frac{3}{2}x^2 = \frac{3}{2}\left[\frac{x^5}{5}\right]_{-1}^{1} = \frac{3}{2}\left(\frac{1}{5} + \frac{1}{5}\right) = \frac{3}{5}$

e.g.

$f_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{1000} e^{-\frac{1}{1000}x} & x \geq 0 \end{cases}$

X = "tempo di vita di una lampadina" in ore

$P(X > 31 \cdot 24) = 1 - F_X(31 \cdot 24) = \int_{31 \cdot 24}^{+\infty} \frac{1}{1000} \cdot e^{-\frac{1}{1000}x}\,dx$

$Y = 5X \Rightarrow$

$F_Y(y) = P(5X \leq y) = \mathbb{P}(X \leq y/5) = F_X(y/5)$

$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\frac{1}{1000}x} & x \geq 0 \end{cases} \Rightarrow F_Y(y) = \begin{cases} 0 & y < 0 \\ 1 - e^{-\frac{y}{5000}} & y \geq 0 \end{cases}$

$-\left(\int_0^x -\frac{1}{1000} e^{-\frac{1}{1000}y}\,dy\right) = -\left[e^{-\frac{1}{1000}y}\right]_0^x$

$\Downarrow$

$f_Y(y) = \begin{cases} 0 & y < 0 \\ e^{-\frac{y}{5000}} \cdot \frac{1}{5000} & \end{cases}$

Vettori aleatori

due o più v.a. che riguardano lo stesso esperimento

$(X_1, X_2, \dots, X_n): \Omega \to \mathbb{R}^n$ vettore le cui componenti sono v.a.

$P_{(X,Y)}: \mathcal{P}(\mathbb{R}^2) \to [0,1]$

$P_{(X,Y)}(B) = \mathbb{P}((X,Y) \in B) \quad \forall B \subset \mathbb{R}^2$

legge / distribuzione

$P_{(X,Y)}$ legge congiunta di X e Y (contiene + info)

P_X (legge) marginale di X

P_Y — Y

esistono densità discrete congiunte tra loro diverse ma con le stesse marginali

$B = B_1 \times B_2 \Rightarrow$

$\mathbb{P}((X,Y) \in B) = \mathbb{P}(\{X \in B_1\} \cap \{Y \in B_2\}) =$ notazione

$\mathbb{P}(X \in B_1, Y \in B_2)$

funzione di ripartizione congiunta $F_{(X,Y)}(x,y) = P_{(X,Y)}((-\infty, x] \times (-\infty, y]) \quad \forall (x,y) \in \mathbb{R}^2$

Indipendenza di Variabili a.

estendere la nozione di indipendenza già definita per gli eventi:

$P_{(X,Y)}(B_1 \times B_2) = P_X(B_1) \cdot P_Y(B_2) \iff X \perp\!\!\!\perp Y$

se ind $\Rightarrow$ mi individua un'unica congiunta

se sono ind la distr congiunta si fattorizza nel prodotto delle marginali

n v.a. si dicono indipendenti se gli eventi da esse generati

$\{X_1 \in B_1\}, \dots \{X_n \in B_n\}$ sono ind

$\boxed{\text{se } X \perp\!\!\!\perp Y \Rightarrow f(X) \perp\!\!\!\perp g(Y)}$

$\mathbb{P}(\underbrace{f(X) \in B_1}_{X \in f^{-1}(B_1)}, g(Y) \in B_2) = \mathbb{P}(f(X) \in B_1)\mathbb{P}(g(Y) \in B_2)$

e.g.

$S_X = \{2,3\} \Rightarrow$ $B_1 = \{4\}$

$S_{f(X)} = \{4, 9\}$ $\{f(X) \in B_1\} = \{2\}$

$\{X \in f^{-1}(\{4\})\} = \{X \in \{2\}\} = \{2\}$

(ind per ogni B_1, B_2 fix...)

Vettore aleatorio discreto

se sue componenti sono discrete

$S_{((X,Y))} \subseteq S_X \times S_Y$

$\{(x,y) \mid P_{(X,Y)}(x,y) > 0\}$

$P_{(X,Y)}: \mathbb{R}^2 \to [0,1], \ 0 \le \cdot \le 1$

$P_{(X,Y)}(x,y) = \mathbb{P}(X = x, Y = y) = \mathbb{P}((X,Y) = (x,y))$

densità discreta congiunta di X e Y

P_X e P_Y densità discrete marginali

$P_{(X,Y)}(x,y) = \mathbb{P}(X = x \mid Y = y)\mathbb{P}(Y = y)$ calcolare $P_{(X,Y)}(x,y)$ tramite la regola della catena

Prop

- $P_{(X,Y)}(x,y) = 0 \quad \forall (x,y) \notin S_X \times S_Y$
- $\sum_{x_i \in S_X} \sum_{y_j \in S_Y} P_{(X,Y)}(x,y) = 1 = \sum_{(x_i, y_j) \in S_X \times S_Y} P_{(X,Y)}(x_i, y_j)$
- $0 \le P_{(X,Y)}(x_i, y_j) \le 1$
- $\mathbb{P}((X,Y) \in B) = \sum_{(x_i, y_j) \in B} P_{(X,Y)}(x_i, y_j) \quad \forall B \subset \mathbb{R}^2$

Condizioni da verificare

Relazione tra

Densità discreta congiunta e densità discrete marginali

(come ottenere marginali dalla congiunta) (grazie a formula delle Prob totali)

$P_X(x_i) = \sum_j P_{(X,Y)}(x_i, y_j) \quad \forall x_i \in S_X$ + Tabella della densità discreta congiunta

x_i | y_j | legge del vettore (X,Y) | P_X | 1

$X \perp\!\!\!\perp Y \iff P_{(X,Y)}(x_i, y_j) = P_X(x_i) P_Y(y_j) \quad \forall (x_i, y_j) \in S_X \times S_Y$

e.g. $X \perp\!\!\!\perp Y$

X \ Y	-1	5	10	P_X
0	0.12	0.12	0.16	0.4
5	0.18	0.18	0.24	0.6
P_Y	0.3	0.3	0.4	1

$\mathbb{P}(X < Y) =$

$\mathbb{P}(|XY| \ge 5) =$

$\mathbb{P}(X + Y > 5) =$

$U = |XY| \Rightarrow S_U = \{0, 5, 25, 50\}$

$V = X + Y \Rightarrow S_V = \{-1, 5, 10, 4, 15\}$

(X,Y)	(U,V)
(0,-1)	(0,-1)
(0,5)	(0,5)
(0,10)	(0,10)
(5,-1)	(5,4)
(5,5)	(25,10)
(5,10)	(50,15)

$P_{(U,V)}(0,-1) = P_{(X,Y)}(0,-1)$

(altre possibili coppie = 0)

Valore atteso e Varianza di una funzione di (X,Y)

$\mathbb{E}[h(X,Y)] = \sum_{i,j} h(x_i, y_j) P_{(X,Y)}(x_i, y_j)$ doppia sommatoria (somma pesata di tutti elem in tabella)

$\mathrm{Var}(h(X,Y)) = \mathbb{E}[(h(X,Y) - \mathbb{E}[h(X,Y)])^2] =$

$= \sum_{i,j} (h(x_i, y_j) - \mathbb{E}[h(X,Y)])^2 P_{(X,Y)}(x_i, y_j)$

• Corol.

$\mathrm{Var}(h(X,Y)) = \mathbb{E}[(h(X,Y))^2] - \mathbb{E}[h(X,Y)]^2$

• corollari da formula valore atteso

- $\boxed{\mathbb{E}[aX + bY]} = \sum_{i,j} (a x_i + b y_j) P_{(X,Y)}(x_i, y_j) =$

v.a.d.

$= \sum_{i,j} a x_i \underbrace{P_{(X,Y)}(x_i, y_j)}_{P_X(x_i)} + \dots$

$= \boxed{a\mathbb{E}[X] + b\mathbb{E}[Y]}$

- $\boxed{X \perp\!\!\!\perp Y \Rightarrow \mathbb{E}[XY] =}$

$= \sum_{i,j} x_i y_j P_{(X,Y)}(x_i, y_j) = \sum_{i,j} x_i y_j P_X(x_i) \cdot P_Y(y_j) =$ indip

$= (\sum_i x_i P_X(x_i)) \cdot (\sum_j y_j P_Y(y_j)) = \boxed{\mathbb{E}[X]\mathbb{E}[Y]}$

Indici di sintesi della distribuzione di un vett. a.d

• vettore delle medie: $(\mathbb{E}[X], \mathbb{E}[Y])$ indica centro distribuzione max densità

• matrice delle covarianze

$\mathrm{Cov}(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] =$ indica come e quanto variano assieme

$= \sum_{i,j} \underbrace{(x_i - \mathbb{E}[X])(y_j - \mathbb{E}[Y])}_{h(x_i,y_j)} P_{(X,Y)}(x_i, y_j)$

se $\mathrm{Cov}(X,Y) = 0 \Rightarrow X$ e Y si dicono scorrelate

$\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$

$\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

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

$\sum x_i y_j \cdot P_{(X,Y)}(x_i, y_j)$

se X e Y scorrelate $\Rightarrow \mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

$P(Z = 1 \cap X = 1) =$

"... $P(\overset{Z}{XY} = 1 \mid X = 1) P(X = 1) =$

$= P(Y = 1 \mid X = 1) P(X = 1) =$

$= \dfrac{P(Y = 1) \cancel{P(X = 1)}}{\cancel{P(X = 1)}} \cdot P(X = 1)$

e.g.

X_1 \ X_2	-1	0	1
-1	$\frac{p^2}{2}$	$\frac{2p(1-p)}{5}$	$\frac{(1-p)^2}{2}$
0	…		
1			

$\Rightarrow$ per far sì che sia densità discreta (congiunta) $\Rightarrow$

$0 \le P_{X,Y}(x_i, y_j) \le 1$

$\sum P_{X,Y}(x_i, y_j) = 1$

$\Downarrow$

$\begin{cases} 0 \le \frac{p^2}{2} \le 1 \\ \vdots \\ p^2 + 2p(1-p) + (1-p)^2 = 1 \end{cases}$

$\Downarrow$

$0 \le p \le 1$

$M = X_1 \cdot X_2 \Rightarrow$ 0, 1, -1

$\mathbb{E}[M] = \sum x_i y_j P_{(X,Y)}(x_i, y_j)$

$\mathrm{Var}(M) = \sum (x_i y_j)^2 P_{(X,Y)}(x_i, y_j) - (\mathbb{E}[M])^2$

M	-1	0	1
P_M	$(1-p)^2$	$2p(1-p)$	p^2

e.g. 3 nuovi, 2 usati ma funzionanti, 2 difettosi

$S_X = \{0, \dots, 3\}$ num componenti nuovi pescati

$S_Y = \{0, \dots, 2\}$ num comp usati ma funzionanti pescati

$|\Omega| = \binom{7}{3}$

$P_X(x_i) = \dfrac{\binom{4}{3 - x_i}}{|\Omega|}$

$\binom{3}{2} \dfrac{3 \cdot 2 \cdot 4}{7 \cdot 6 \cdot 5}$ 4 3

3 - 0	3
3 - 1	2
3 - 2	1
3 - 3	0

$\binom{3}{0} \dfrac{4 \cdot 3 \cdot 2}{7 \cdot 6 \cdot 5} + \binom{3}{1} 3 \dfrac{3 \cdot 4 \cdot 3}{7 \cdot 6 \cdot 5} + \binom{3}{2} 3 \dfrac{3 \cdot 2 \cdot 4}{7 \cdot 6 \cdot 5} + \binom{3}{3} \dfrac{3 \cdot 2}{7 \cdot 6 \cdot 5}$

$\dfrac{4}{35} + \dfrac{12}{35} + \dfrac{1}{35} + \dfrac{3 \cdot 3 \cdot 2}{7 \cdot 6 \cdot 5}$ $\dfrac{18}{35}$

$P_{(X,Y)}(x_i,y_j) = \begin{cases} \binom{3}{x_i}\binom{2}{y_j}\binom{2}{3-(x_i+y_j)} \Big/ \binom{7}{3} & \text{se } x_i+y_j = 1,2,3 \\ 0 & \end{cases}$

imp che sia 0!

Z = numero di componenti pescati funzionanti

$S_Z = \{0, 1, \ldots 3\}$ IMP!

$P_Z(z_i) = \dfrac{\binom{5}{z_i}\binom{2}{3-z_i}}{\binom{7}{3}}$

3 nuovi
2 usati ma OK
2 KO

OPPURE

$3-1=2$
$3-2=1$
$3-3=0$

$Z = X+Y$

$S_Z = \{1,2,3\}$

$E[X+Y] \Rightarrow lin \Rightarrow E[X]+E[Y]$

Prob che l'apparecchio funzioni $= P(Z=3)$

e.g

D = lancio di un dado a tre facce $= \{1,2,3\}$ risultato

T = numero di teste ottenute $= \{0,1,2,3\}$

D	1	2	3
P_D	1/3	1/3	1/3

la somma di una riga porta a 1...

$\binom{2}{0}\frac{1}{4} = \binom{2}{2}\frac{1}{4}$

$\binom{2}{1}\frac{1}{2}\frac{1}{2} = \frac{2}{4}$

$\binom{3}{t_i} \cdot \frac{1}{2}^{t_i} \frac{1}{2}^{3-t_i}$

$\binom{3}{0} \cdot \cancel{1} \cdot \left(\frac{1}{2}\right)^3$

$\begin{matrix}3\\1\end{matrix} \quad \frac{1}{2} \cdot \frac{1}{2}^2$

$\left(\frac{1}{8}\right) + 3 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} + \frac{1}{8} = \frac{8}{8}$

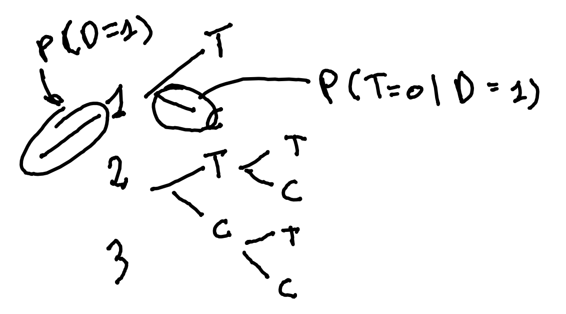
$\begin{matrix}3\\2\end{matrix} \quad \begin{matrix}3\\3\end{matrix}$

$P_{d_i t_j} = \binom{d_i}{t_j} \frac{1}{2}^{t_j} \cdot \frac{1}{2}^{d_i-t_j}$ se $t_j \le d_i$

0 else

USARE REGOLA CATENA

$P(T=t_i | D=d_j) \underbrace{P(D=d_j)}_{1/3}$



e.g un dado a quattro facce viene lanciato due volte

X = prodotto dei valori apparsi nei due lanci $\Rightarrow S_X = \{1\ 2\ 3\ 4, 6, 8, 9, 12, 16\}$

Y = valore massimo che appare nei due lanci $\{1, 2, 3, 4\}$

$\boxed{\text{USARE REGOLA CATENA}}$ (ricordarsi di moltiplicare/normalizzare per l'evento precedente!)

$P(Y=y_j | X=x_i)$

$P(Y=4 | X=16) = 1$

1 1	1	1
1 2	2	2
1 3	3	3
1 4	4	4
2 1	2	2
2 2	2	4
2 3	3	6
2 4	4	8
3 1	3	3
3 2	3	6
3 3	3	9
3 4	4	12
4 1	4	4
4 2	4	8
4 3	4	12
4 4	4	16

$16 \Rightarrow 4$
$12 \Rightarrow 4$
$9 \Rightarrow 3$
$8 \Rightarrow 4$
$6 \Rightarrow 3$
$4 = 4$ or 2
3 3
2 2
1 1

X 8 12 16

Y
1
2
3
4 $\quad \frac{2}{16} \quad \frac{2}{16} \quad \frac{1}{16} \Big| \frac{7}{16}$

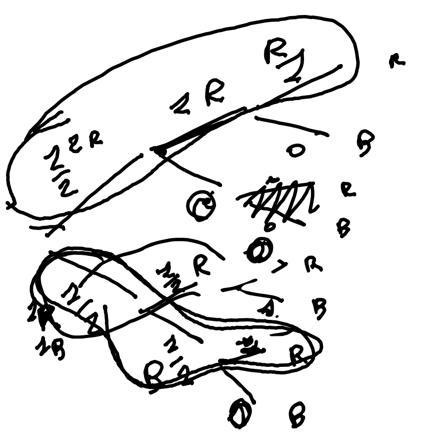
$\frac{1}{16}$

MAX

Y	1	2	3	4
P_Y	1/16	3/16	5/16	7/16

PRODOTTO

X	1	2	3	4	6	8	9	12	16
P_X	1/16	2/16	2/16	3/16	2/16	2/16	1/16	2/16	1/16



$P(R_2 | R_1 \cap (A \cup A^c))$

$= \dfrac{P(R_1 \cap R_2)}{P(R_1 \cap A) + \ldots}$

$\dfrac{1/2}{\frac{1}{4} + \frac{1}{4}} \qquad \dfrac{1/2}{3/4} \qquad \frac{1}{2} \cdot \cancel{\frac{4}{3}}$

$\underbrace{P(Y=4 | X=4)}\ \underbrace{P(X=4)}_{3/16}$

"Come se diventasse nuovo spazio campionario"

$\underbrace{4,1}_{1/3}\ \underbrace{2,2}_{1/3},\ \underbrace{1\ 4}_{1/3}$

$\frac{2}{3} \cdot \frac{3}{16}$

~~

e.g

X = risultato del lancio del dado

Y = "n° della pallina estratta dall'urna con X palline

$S_X = \{1, \ldots 4\}$ $\qquad S_Y = \{1, \ldots 4\}$

$P_{(X,Y)}(x_i, y_j) = \begin{cases} \frac{1}{x_i} \cdot \frac{1}{4} & \\ 0 & \text{se } y_j > x_i \end{cases}$

Y \ X	1	2	3	4
1	1/4	1/8	1/12	1/16
2	0	1/8	1/12	1/16
3	0	0	1/12	1/16
4	0	0	0	1/16
	1/4	1/4	1/4	1/4

2

~ ~~

e.g X = risultato primo lancio $\qquad S_X = \{1, 2, 3\}$

Y = ——— secondo $\qquad S_Y$ = same

X Y	U	V
1 1	1	0
1 2	2	1
1 3	3	2
2 1	2	1
2 2	4	0
2 3	6	1
3 1	3	2
3 2	6	1
3 3	9	0

2/9

1/9 $\qquad$ 1/3 $\qquad$ 2/3 $\qquad$ 1/3

analizzare e sommare tutte le possibili coppie $\quad$ 1/3 $\quad$ 1/3 $\quad$ 1/3

Se $(X,Y) = (1,1) \Rightarrow (U,V) = 1,0 \Rightarrow$
Se $(X,Y) = (1,2) \Rightarrow (U,V) = (2,1) \Rightarrow P(2,1) = \frac{1}{9}$
se $(X,Y) = (2,1) \Rightarrow (U,V) = (2,1) \Rightarrow P(2,1) = \frac{2}{9}$

$S_U = \{1, 2, 3, 4, 6, 9\}$

$S_V = \{0, 1, 2\}$

$P(U=u_i | V=v_j) P(V$

V	0	1	2
P_V	1/3	4/9	2/9

Catene di Markov

una sequenza di variabili aleatorie è detta

indipendenti e identicamente distribuite (i.i.d)

se
- le variabili hanno tutte la stessa distribuzione di prob.
- sono indipendenti

dipendenza a catena: la variabile aleatoria X_{n+1} sarà influenzata direttamente solo da quella che la precede, ovvero X_n e tramite essa sarà influenzata da tutte le v.a. precedenti. Una volta nota X_n, la conoscenza supplementare dei valori $X_1, \dots, X_{n-1}$ non darà alcuna ulteriore informazione riguardo il valore di X_{n+1}

Catena di Markov: successione di variabili aleatorie $(X_n)_n$ (ovvero un processo stocastico a tempo discreto) che verifica le seguenti prop.
- le v.a. $X_1, \dots, X_n, \dots$ sono discrete e il loro supporto è contenuto nello stesso insieme S discreto

$S_{X_1} \subset S, \dots, S_{X_n} \subset S, \dots$ (in questo caso una X_i potrebbe non assumere tutti i valori del supporto)

S si chiama lo spazio degli stati della catena di Markov

- (Prop di Markov: dip a "catena")

$$\underbrace{\mathbb{P}(X_{n+1}=j}_{\text{futuro}} \mid \underbrace{X_1=i_1, \dots, X_{n-1}=i_{n-1}}_{\text{passato}}, \underbrace{X_n=i}_{\text{presente}}) = \mathbb{P}(X_{n+1}=j \mid X_n=i)$$

"il valore futuro è j"

$\mathbb{P}(F \mid P_a \cap P_r) = P(F \mid P_{res})$

$\forall_{i,j,n}$ $\pi_{ij}(n) = \mathbb{P}(X_{n+1}=j \mid X_n=i)$

probabilità di transizione all'istante n dallo stato i allo stato j

regole meccanismo con cui la catena di Markov va avanti nel tempo. Se io so che all'istante n sono allo stato i, questo mi dice qual è la prob di saltare nello stato j

se uniforme => tutte uguali

Catene di Markov omogenee e a stati finiti (default da ora in poi)

$(X_n)_n$ è omogenea (nel tempo) se le prob di transizione non dipende da n ma solo dai valori assunti => (da i a j) π_{ij} => matrice di transizione $\Pi_{N \times N}$
- S insieme finito $\{1, \dots N\}$

- $0 \le \pi_{ij} \le 1$ // perché sono delle prob
- $\sum_{j=1}^{N} \pi_{ij} = 1$ $\forall$ riga iesima

non può fare < 1 perché significherebbe che si andrebbe anche in un altro stato che non ho elencato

& "densità discreta di X_{n+1} sapendo che $X_n = i$"

Legge condizionata della v.a. $X_{n+1} \mid (X_n = i)$

$X_{n+1} \mid X_n=i$	1	2	...	N
$P_{X_{n+1} \mid X_n=i}$	π_{i1}	π_{i2}		π_{iN}

Rappresentazione grafica di una catena di Markov
- ogni stato $i \in S$ corrisponde ad un nodo del grafo
- $\pi_{ij} > 0$ => arco orientato da nodo i al nodo j con il valore π_{ij} riportato sull'arco

somma archi uscenti da un nodo = 1

Probabilità di transizione in più passi

$$\boxed{\mathbb{P}(X_{n+1}=j \mid X_n=i) = \pi_{ij}}$$

$$P(X_{n+m}=j \mid X_n=i) = \pi_{ij}^{(m)}$$

m passi

per $m=0$,

prob che all'istante presente sia in j sapendo che sono in i

$\pi_{ij}^{(0)} = P(X_n=j \mid X_n=i) = \begin{cases} 1 & \text{se } i=j \\ 0 & \text{else} \end{cases}$

$=> \boxed{\Pi^{(0)} = I_N}$ casi base

per $m=1$,

$\pi_{ij}^{(1)} = \pi_{ij}$ $\boxed{\Pi^{(1)} = \Pi}$

$$\boxed{\underbrace{\Pi \cdot \ldots \cdot \Pi}_{m \text{ volte}} = \Pi^{(m)}}$$

prob 1 passo

itero meccanismo

e.g. $m=2$

$\pi_{ij}^{(2)} = \sum_{k=1}^{N} \pi_{ik}\pi_{kj}$ somma delle probabilità di tutti i possibili cammini

calcolo diretto di $\pi_{ij}^{(m)}$ a partire dal grafo
- a partire dal grafo, si trovano tutti i cammini che portano da i a j in m passi
- la prob di ogni cammino è il prodotto delle prob lungo gli archi del cammino stesso
- $\pi_{ij}^{(m)}$ è la somma delle prob dei cammini che portano da i a j in m passi

scelta del bar in comune scelta del bar diverso

$\frac{\binom{3}{2} 4 \cdot 3}{4^3}$

Classi comunicanti

Fissati due stati $i, j \in S$ (non necessariamente $i \ne j$)

$\boxed{j \text{ è accessibile da } i}$ (in termini di grafo orientato significa che il nodo i è connesso con il nodo j)

se $\exists\, m \ge 0$ t.c. $\pi_{ij}^{(m)} > 0$ $\boxed{i \rightsquigarrow j}$

$i \rightsquigarrow i$ $\forall i$ $(\pi_{ii}^{(0)} = 1 > 0)$

$i \ne j => i \rightsquigarrow j \iff \exists\, m \ge 1$ ed $\exists$ un cammino $i_1 \to i_2 \to \dots i \to m+1$ in m passi t.c. $i_1 = i$, $i_{m+1} = j$ e $\pi_{i_1 i_2}\pi_{i_2 i_3} \cdots \pi_{i_m i_{m+1}} > 0$

i e j si dicono $\boxed{\text{comunicanti}}$ se

$i \rightsquigarrow j$ e $j \rightsquigarrow i$ (i è fortemente connesso con j)

$\Downarrow$

$i \leftrightsquigarrow j$

Si chiama $\boxed{\text{classe comunicante}}$ un sottoinsieme di S costituito da tutti gli stati tra loro comunicanti

rel di equiv (componente fortemente connessa)

classe di equiv => partizione dell'insieme dei nodi (ogni nodo appartiene ad una ed una sola classe comunicante)

(ogni stato $i \in S$ è comunicante con se stesso)

$i \ne j => i \leftrightsquigarrow j \iff \exists$ un cammino chiuso (stato di partenza coincide con stato di arrivo) che passa per i e j

$(X_n)_n$ è $\boxed{\text{irriducibile}}$ se $\boxed{\exists \text{ un'unica classe comunicante}}$, che è quindi data dall'$\boxed{\text{insieme } S \text{ stesso}}$

Legge di X_n $\vec{P}_{X_n} = [P_{X_n}(1) \; . \; . \; . \; P_{X_n}(N)]$

$$\boxed{\vec{P}_{X_n} = \vec{P}_{X_1}\Pi^{n-1}}$$

$n-1$ passi

in particolare

$$\boxed{P_{X_n}(j)} = \mathbb{P}(X_n = j) = \boxed{\sum_{i=1}^{N} P_{X_1}(i)\, \pi_{ij}^{(n-1)}}$$

$\vec{P}_{X_{n+m}} = \vec{P}_{X_m}\Pi^n$

e.g. distribuzione iniziale

"La catena di Markov parte dallo stato 2" $X_1 = 2$

X_1	1	2	0 ..	N
P_{X_1}	0	1	0	0

lancio una moneta

T => 2, C => 3 label etichette

1	2	3	4 .. N
0	1/2	1/2	0 ... 0

Distribuzione invariante

$\vec{P}_X = \vec{\pi}$
- (ogni sua componente è compresa tra 0 e 1 (poiché sono prob))
- somma delle sue componenti è uguale a 1

$\vec{\pi}$ è una distribuzione invariante o stazionaria o di equilibrio per Π

se $\boxed{\vec{\pi} = \vec{\pi}\,\Pi}$

è sufficiente richiedere questa cond per n passi: se dopo 1 passo appena rimane lo stesso => lo rimane anche per i passi successivi

e.g.

	T	C
T	1/2	1/2
C	1/2	1/2

=> $\vec{\pi} \cdot \Pi = \vec{\pi}$

Poi se prendiamo in particolare le righe => la formula

$\vec{P}_{X_1} = \vec{P}_{X_1}\Pi$ => X_n ha distribuzione $\vec{P}_{X_1}$

$$\boxed{\vec{P}_{X_n}} = \vec{P}_{X_1}\Pi^{n-1} = \underbrace{\vec{P}_{X_1}\Pi}_{\vec{P}_{X_1}}\Pi^{n-2} = \ldots = \boxed{\vec{P}_{X_1}}$$

e.g.

$$\Pi = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$\vec{P}_{X_1} = [\tfrac{1}{2} \; \tfrac{1}{4} \; 0 \; 0 \; 0 \; \tfrac{1}{4}]$

$\pi_{12}^{(2)} := \mathbb{P}(X_{m+2}=2 \mid X_m=1)$ $\forall m \in \mathbb{N}$

$\pi_{12}^{(2)} = \sum_k \sum_h \sum_l \pi_{1k}\pi_{k2}\pi_{1h}\pi_{hl}\pi_{l2}$

$\vec{P}_{X_3} = \vec{P}_{X_1}\Pi^{3-1}$

$$\underbrace{P(X_3=2)}_{} = \sum_{i=1}^{N} P_{X_1}(i) \cdot \pi_{i2}^{3-1} = \frac{1}{2}\pi_{12}^{(2)} + \frac{1}{4}\pi_{22}^{(2)} + \frac{1}{4}\pi_{62}^{(2)}$$

è il secondo elemento del vettore riga

$\pi_{12}^{(2)} = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ 1 cammino $1 \to 1 \to 2$

$\pi_{22}^{(2)} = 0$ (no cammini di len due da 2 a 2)

$\pi_{62}^{(2)} = 0$

si calcolano graficamente!

$\boxed{\text{distribuzione iniziale = dove ci troviamo}}$ =>

"supponiamo di partire dalla pagina numero 1" oppure parto con una pallina nell'urna

X_1	1	2	3 ...	N
P_{X_1}	1	0	0	0

Prob di passare da uno stato all'altro

e.g.

	1	2	3	4	5	6
1	1/6	5/6				
2	1/6	1/6	4/6			
3	1/6	0	2/6	3/6		
4	1/6			3/6	2/6	
5	1/6				4/6	1/6
6	1/6					5/6

prob che esca un numero + grande di due

Se esce 1 il sistema si sposta in tale stato. Se esce un numero maggiore dello stato presente, il sistema si sposta nello stato successivo. altri casi => rimane nello stato presente.

Sapendo che all'istante n il sistema si trova nello stato 3, qual è la prob che il sistema si trovi nello stato 4 all'istante $n+2$?

$=> \pi_{3,4}^{(2)} = P(X_{n+2}=4 \mid X_n=3)$

(meccanismi non cambiano nel tempo)

distribuzione invariante del sistema =>

$$\begin{cases} \vec{\pi} = \vec{\pi}\,\Pi \\ \sum \pi_i = 1 \end{cases}$$

$\Downarrow$

vettore riga per i-esima colonna

$$\begin{cases} \pi_1 = \pi_1 \cdot \frac{1}{6} + \pi_2 \cdot \frac{1}{6} + \pi_3 \cdot \frac{1}{6} + \ldots \pi_6 \cdot \frac{1}{6} \\ \pi_2 = \pi_1 \cdot \frac{5}{6} + \pi_2 \cdot \frac{1}{6} \\ \pi_3 = \frac{4}{6} \cdot \pi_2 + \frac{2}{6} \cdot \pi_3 \\ \vdots \\ 1 = \pi_1 + \pi_2 + \ldots \pi_6 \end{cases}$$

$\frac{1}{6} \cdot \sum \pi_i$ (con $\sum \pi_i = 1$)

$\vec{\pi} = (\frac{1}{6}, \dots, \frac{1}{6})$

e.g.

6 palline, 2 B 4 R

3? 3?

X_n = numero di palline rosse contenute nell'urna A dopo N estrazioni

$S = \{1, 2, 3\}$

	1	2	3
1	1/3	2/3	0
2	2/9	5/9	2/9
3		2/3	1/3

R B B | R R R | R R B | B R R

Sapendo che nell'urna A ci sono due palline rosse, qual è la prob che dopo tre estrazioni ce ne siano tre?

$\pi_{23}^{(3)} = P(X_{n+3}=3 \mid X_n=2)$

$1 => \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$

$2 => \frac{2}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{5}{9}$

$3 => \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$

estrazione di due palline con stesso colore non cambia conf

riga 3 è riga 1 dove urna A e B si scambiano

densità discreta di X_0

$$P(X_0 = x) = \frac{\binom{4}{x}\binom{2}{3-x}}{\binom{6}{3}}$$ branch di possibili sottoinsiemi

$x \in [1, 2, 3]$ =>

X_0	1	2	3
P_{X_0}	1/5	3/5	1/5

$\vec{P}_{X_2} = \vec{P}_{X_1}\Pi$

$\vec{P}_{X_1} = \vec{P}_{X_0} \cdot \Pi$ =>

$\vec{P}_{X_2} = \vec{P}_{X_0} \cdot \Pi^2$ itero due volte

$P(X_1 = 2 \mid X_2 = 1)$ =>

Bayes => per ristabilire ordine

$$P(X_1=2 \mid X_2=1) = \frac{\overbrace{P(X_2=1 \mid X_1=2)}^{\pi_{21} \text{ per def}} \cdot P(X_1=2)}{P(X_2=1)}$$

nota

$S = \{1, \dots, 4\}$

$P(X_0 = 1) = 1$ lancio un dado a 4 facce =>

pedina 1 => 2 se ris dispari, 4 ris pari

$\vec{P}_{X_2} = \vec{P}_{X_0} \cdot \Pi^2$

ma sappiamo che $\vec{P}_{X_0}$ ha legge costante => la sfruttiamo =>

$$\mathbb{P}(X_2 = j) = \sum_{i}^{4} P_{X_0}(i)\, \pi_{ij}^{(2)} = \pi_{1j}^{(2)} \Rightarrow \text{graf.}$$

singola componente della dist

1 solo per $i = 1$

$\pi_{11}^{(2)} = \ldots$

$\pi_{12}^{(2)} = \ldots$

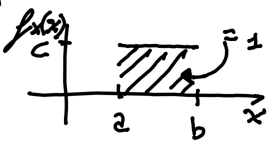
$\vdots$

V.a.c distribuzioni notevoli

• distribuzione uniforme continua

$S_X = [a,b]$

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases}$$



$c = \frac{1}{b-a}$

$X \sim Unif(a,b)$

$$F_X = \begin{cases} 0 & x \le a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x \ge b \end{cases}$$

$$E[X] = \int_a^b \frac{x}{b-a} = \frac{b+a}{2}$$

• distribuzione normale (o gaussiana) di media μ e varianza σ^2

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

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

X ha distribuzione normale standard se

$$X \sim \mathcal{N}(0,1) \sim Z$$

$$\varphi = f_Z \qquad \Phi = F_Z$$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

Standardizzazione

$$X \in \mathcal{N}(\mu, \sigma^2)$$

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

$$\Phi(x) = F_Z(x) = \mathbb{P}(Z \le x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le x\right) = F_X(\sigma x + \mu)$$

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \mathbb{P}\left(Z \le \frac{x-\mu}{\sigma}\right) = F_Z\left(\frac{x-\mu}{\sigma}\right) = \Phi(\dots)$$

Proprietà di Φ

$$\Phi(0) = \frac{1}{2}$$

$$\Phi(-x) = 1 - \Phi(x) \quad \forall x \ge 0$$

e.g. $X \sim \mathcal{N}(9.99, (0.012)^2)$

• trovare percentuale di provette che fanno traboccare (nella forma $1 - \Phi(x)$)

$$P(\{X > 10\}) = 1 - F_X(10)$$

oppure standardizzare subito $P(X > 10) = P\left(\frac{X-\mu}{\sigma} > \frac{10-\dots}{\dots}\right) = P\left(Z > \frac{10-\dots}{\dots}\right)$

$$F_X(10) = \mathbb{P}(X \le 10) = \mathbb{P}\left(\frac{X-9.99}{0.012} \le \frac{10-\dots}{\dots}\right) = \Phi\left(\frac{10-\dots}{\dots}\right)$$

• determinare l t.c. la percentuale di provette che contengono una qty di liquido inferiore a l sia 10%

$$F_X(l) = \frac{1}{10}. \qquad \Phi^{-1}(0.1) \simeq -1.282$$

$$\Phi\left(\frac{l-\dots}{\dots}\right) = \frac{1}{10} \Rightarrow \frac{l-\dots}{\dots} = \Phi^{-1}$$

e.g. $X \sim \mathcal{N}(\mu, \sigma^2)$

• $\mu = 4$, $\sigma^2 = 0.09 \Rightarrow P(3.95 < X < 4.05) =$

$$= P\left(\frac{3.95-\dots}{\sqrt{\dots}} < Z < \frac{4.05-\dots}{\sqrt{\dots}}\right) =$$

$$\Phi\left(\frac{0.5}{\sqrt{0.09}}\right) - \left(-\frac{0.5}{\sqrt{0.09}}\right)$$

• $\mu = 4$, $\sigma^2 = 0.09 \Rightarrow P(X < 4) =$

$$= P\left(Z < \frac{4-4}{\dots} = 0\right)$$

$$= \Phi(0) = \frac{1}{2}$$

• 5% sotto il minimo tollerato

12% superiore al max tollerato

$$F_X(3.95) = \frac{5}{100}$$

$$1 - F_X(4.05) = \frac{12}{100}$$

$$\Rightarrow \begin{cases} \Phi\left(\frac{3.95-\mu}{\sigma}\right) = \frac{5}{100} \\ \Phi\left(\frac{4.05-\mu}{\sigma}\right) = \frac{88}{100} \end{cases}$$