PROBABILITÀ

e

STATISTICA

ELENA VUK

INTRODUZIONE

La raccolta dei dati e la loro analisi sono strumenti indispensabili per comprendere la complessa realtà che ci circonda.
La statistica è "l'arte di apprendere dai dati". (S.M. Ross, 1999)
Quella parte della statistica che si occupa di illustrare e sintetizzare i dati è detta statistica descrittiva: il suo scopo è quello di ridurre il volume dei dati osservati esprimendo l'informazione rilevante in essi contenuta per mezzo di grafici ed indicatori numerici.
La parte della statistica che invece si occupa di trarre conclusioni dai dati raccolti e sintetizzati e per essi dedurre enunciati formalmente validi è la statistica inferenziale.
L'elemento fondamentale da prendere in considerazione è la casualità.

Fenomeni casuali o aleatori, cioè non completamente prevedibili a priori (che non significa che siano completamente imprevedibili) sono all'origine di variazioni.

Pertanto, per trarre conclusioni pienamente giustificate è necessario fare alcune assunzioni sulla probabilità che i dati da misurare assumano i diversi valori possibili.

L'insieme di queste ipotesi è detto modello probabilistico.

Per individuare il corretto modello astratto di probabilità che descrive i dati è indispensabile la conoscenza della teoria della probabilità.

Attraverso il calcolo delle probabilità è possibile formulare una trattazione matematica dell'incertezza.

CALCOLO DELLE PROBABILITÀ

Concetto di probabilità

Nasce nel 17° secolo con i giochi d'azzardo.

Il calcolo delle probabilità è lo studio delle proprietà statistiche che presentano gli eventi il cui verificarsi o meno (in seguito ad osservazioni o prove) non è predeterminato o comunque non prevedibile mediante le condizioni iniziali.

Tali eventi vengono detti casuali o aleatori.

Diversi sono gli approcci alla teoria della p.

Matematicamente la p. viene descritta mediante una quantità scalare in grado di caratterizzare la frequenza di ricorrenza di un dato evento al ripetersi delle prove.

TEORIA CLASSICA O A PRIORI

La probabilità viene definita a priori come il rapporto

$$p := \frac{f}{n} \in [0,1]$$

tra il numero f di casi favorevoli che un certo evento si verifichi ed il numero n dei casi possibili.

<u>Esempi</u>

- lancio di un dado (non truccato)

  6 possibili risultati ugualmente possibili e mutuamente esclusivi

  esce n° pari $\frac{3}{6}$ $\quad p = \frac{1}{2}$

  esce 5 $\quad p = \frac{1}{6}$

- estrazione di una carta da un mazzo di 52

  esce picche $\frac{13}{52}$ $\quad p = \frac{1}{4}$

  esce n° $\in [5, 10]$ $\frac{24}{52}$ $\quad p = \frac{6}{13}$

  esce asso o picche $\underbrace{(4) + (13)}_{\text{in comune asso di picche}}$ $\frac{16}{52}$ $\quad p = \frac{4}{13}$

• lancio di una moneta per 2 volte

4 possibili risultati (TT, TC, CT, CC)

ottenere 2 teste $p = \frac{1}{4}$

DIFETTO

Bisogna presupporre:

- i casi si escludono a vicenda
- i casi sono tutti ugualmente possibili (trucco!)

Inoltre tale definizione deve essere in qualche modo modificata quando il n° dei risultati possibili è infinito.

Bisogna usare il metodo della frequenza.

È possibile concepire una serie di esperimenti o prove in condizioni "abbastanza" uniformi.

È possibile postulare l'esistenza di un numero p, detto probabilità dell'evento, e approssimare p con la frequenza relativa con la quale le prove ripetute soddisfano l'evento.

TEORIA FREQUENTISTA O EMPIRICA

La probabilità di un evento è il limite cui tende la frequenza relativa di successo all'aumentare del n° delle prove.

$$p = \lim_{n \to \infty} \frac{n_A}{n}$$

dove n è il n° delle prove ed n_A è il n° di volte che si presenta un certo evento A.
(J. Bernoulli, R. Von Mises)
Al rapporto

$$\frac{\text{n° casi favorevoli}}{\text{n° casi possibili}}$$

si sostituisce

$$\frac{\text{n° esperimenti effettuati con esito favorevole}}{\text{n° complessivo esperimenti effettuati.}}$$

Il vantaggio è che non è necessario specificare che i risultati debbano essere ugualmente possibili ed incompatibili. (esperimenti con risultati discreti)
DIFETTO
- si applica solo ad esperimenti ripetibili, per i quali il limite per $n \to \infty$ abbia significato

TEORIA SOGGETTIVISTICA

La probabilità di un evento A è la misura del grado di fiducia che un individuo coerente attribuisce, secondo le sue informazione o opinioni, all'avverarsi di A.
(B. De Finetti)

Esempio
Un geologo afferma che in una certa regione vi è il 60% di probabilità che vi sia petrolio.

- interpretazione frequentista

Trovando molte regioni con caratteristiche esterne simili a quella in esame, il geologo crede che nel 60% dei casi vi sarà petrolio

- interpretazione soggettivistica

Il geologo crede che sia più verosimile che vi sia petrolio, piuttosto che non ci sia e 0.6 è la misura della fiducia nell'ipotesi che nella regione vi sia il petrolio.

In ogni caso, le regole e i metodi di calcolo nelle diverse teorie non differiscono tra loro e ciò ha permesso lo sviluppo della TEORIA ASSIOMATICA della probabilità (A. N. KOLMOGOROV, 1933)

Seguendo tale impostazione è necessario dare alcune definizioni e nozioni matematiche.

Il linguaggio è quello della TEORIA DEGLI INSIEMI.

RICHIAMI DI TEORIA DEGLI

INSIEMI

Ω : spazio, insieme universale (collezione di oggetti)

$\omega \in \Omega$: punto o elemento

$A \subset \Omega$: insieme di punti di Ω

- SOTTOINSIEME se ogni elemento di A è anche elemento di B
  $\Rightarrow A \subset B$ o $B \supset A$
- INSIEMI UGUALI A, B se $A \subset B$ e $B \subset A$
  $\Rightarrow A = B$
- INSIEME VUOTO se A non contiene punti
  $\Rightarrow A = \emptyset$
- INSIEME COMPLEMENTARE di A rispetto ad Ω :
  e' l'insieme di ~~p~~ tutti i punti di Ω che non stanno in A
  $\Rightarrow \bar{A} = A^c = \Omega - A$
- INSIEME DIFFERENZA di A e B : punti di A che non sono in B $\Rightarrow A - B$

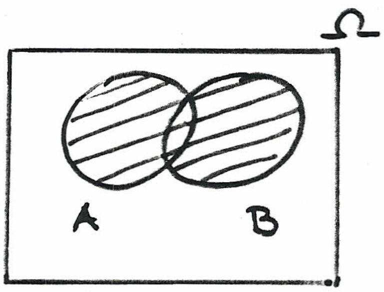
- UNIONE di A e B : insieme di punti di A o di B

$\Rightarrow A \cup B$

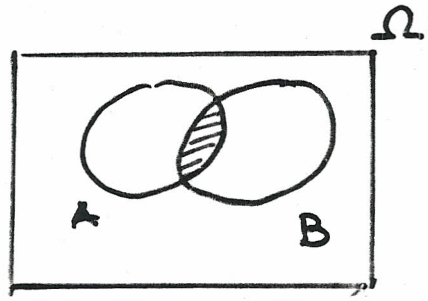
- INTERSEZIONE di A e B : insieme di punti di A e B

$\Rightarrow A \cap B$

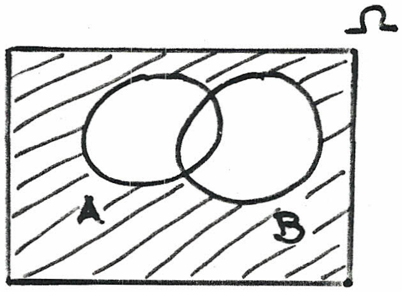
DIAGRAMMI DI VENN



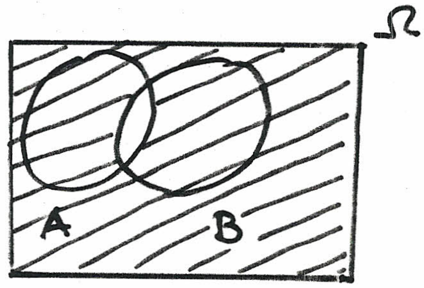
$A \cup B$



$A \cap B$



$\overline{A \cup B} = \bar{A} \cap \bar{B}$



$\overline{A \cap B} = \bar{A} \cup \bar{B}$

Leggi di De Morgan

LEGGI

- COMMUTATIVE $\quad A \cup B = B \cup A \;, \quad A \cap B = B \cap A$
- ASSOCIATIVE $\quad \begin{cases} A \cup (B \cup C) = (A \cup B) \cup C \\ A \cap (B \cap C) = (A \cap B) \cap C \end{cases}$
- DISTRIBUTIVE $\quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

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

TEOREMI

1) $\overline{(\bar{A})} = A$

2) $A \cap \Omega = A$, $A \cup \Omega = \Omega$

3) $A \cap \phi = \phi$, $A \cup \phi = A$

$A \cap A = A$, $A \cup A = A$

4) LEGGI DI DE MORGAN : $\overline{A \cup B} = \bar{A} \cap \bar{B}$

$\overline{A \cap B} = \bar{A} \cup \bar{B}$

5) $A - B = A \cap \bar{B}$

Se $\{A_\lambda\} = \{A_\lambda : \lambda \in \Lambda\}$ sottoinsiemi di Ω indicizzati da Λ.

si definisce

$\bigcup_{\lambda \in \Lambda} A_\lambda$ unione di $\{A_\lambda\}$

$\bigcap_{\lambda \in \Lambda} A_\lambda$ intersezione di $\{A_\lambda\}$

Se $\lambda = \{1, \dots, n\}$ $\Rightarrow$ $\bigcup_{i=1}^{n} A_i$, $\bigcap_{i=1}^{n} A_i$

TEOREMI DI DE MORGAN

$$\overline{\bigcup_{\lambda \in \Lambda} A_\lambda} = \bigcap_{\lambda \in \Lambda} \bar{A}_\lambda$$

$$\overline{\bigcap_{\lambda \in \Lambda} A_\lambda} = \bigcup_{\lambda \in \Lambda} \bar{A}_\lambda$$

- INSIEMI DISGIUNTI $A, B \Rightarrow A \cap B = \emptyset$

$\{A_\lambda\}_{\lambda \in \Lambda}$ sono a 2 a 2 disgiunti se $A_i \cap A_J = \emptyset \quad \forall\, i \neq J$

TEOREMI

1) se $A, B \subset \Omega$
$$\begin{cases} A = (A \cap B) \cup (A \cap \bar{B}) \\ \emptyset = (A \cap B) \cap (A \cap \bar{B}) \end{cases}$$

2) se $A \subset B$
$$\begin{cases} A \cap B = A \\ A \cup B = B \end{cases}$$

RICHIAMI DI CALCOLO COMBINATORIO

$A_1, A_2, \dots, A_k$ gruppi di oggetti
$m_1, m_2 \quad , m_k$

modi : 1 oggetto da A_1, 1 oggetto da $A_2, \dots$, 1 da A_k

$N = m_1 \cdot m_2 \cdot \dots m_k$

se $m_1 = m_2 = \dots = m_k = m \Rightarrow N = m^k$.

DISPOSIZIONI CON RIPETIZIONE
di m oggetti a gruppi di k.

$$D^{(r)}_{m,k} = m^k$$

DISPOSIZIONI SENZA RIPETIZIONE O SEMPLICI
di k oggetti scelti da m distinti con $m \geq k$.

$$D^{(s)}_{m,k} = \frac{m!}{(m-k)!}$$

conta l'ordine o differiscono per almeno un oggetto.

PERMUTAZIONI
di m oggetti distinti

$$P_m = m! = D^{(s)}_{m,m}$$

cioè disposizioni semplici con $k=m$.
differisce solo l'ordine

PERMUTAZIONI

di n oggetti non tutti distinti

n_1 di un tipo, n_2 di un secondo tipo, ..., n_k di un

k-esimo tipo con $n_1 + n_2 + \dots + n_k = n$

$$P_{n, n_1, \dots, n_k} = \frac{n!}{n_1!\, n_2! \dots n_k!}$$

COMBINAZIONI

$$C_{n,k} = \frac{D^{(s)}_{n,k}}{k!} = \frac{n!}{k!\,(n-k)!} = \binom{n}{k}$$

cioè disposizioni semplici in cui non conta l'ordine

COMBINAZIONI CON RIPETIZIONE

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

SPAZIO CAMPIONE ED EVENTI

DEF: Lo SPAZIO CAMPIONE, indicato con Ω, è la totalità di tutti i possibili risultati di un esperimento concettuale

DEF: Un EVENTO è un sottoinsieme dello spazio campione: A, B, C, ...

DEF: Lo SPAZIO (O INSIEME) DEGLI EVENTI, indicato con $\mathcal{A}$, è la famiglia di tutti gli eventi associati ad un dato esperimento.

N. B.: Un evento è sempre un sottoinsieme dello spazio campione, ma non tutti i sottoinsiemi di spazi sufficientemente grandi sono eventi, perciò la famiglia di tutti i sottoinsiemi dello spazio campione non corrisponde necessariamente allo spazio degli eventi.
Se lo s.c. Ω è finito allora lo spazio degli eventi $\mathcal{A}$ è la famiglia di tutti i sottoinsiemi di Ω, cioè $\mathcal{A}$ è l'insieme delle parti di Ω: $\mathcal{A} = \mathcal{P}(\Omega)$.

se $\dim \Omega = n$ $\qquad$ $P(\Omega) = 2^n$

#

card

PROPRIETÀ DELLO SPAZIO DEGLI EVENTI $\mathcal{A}$

i) $\Omega \in \mathcal{A}$ $\quad$ $\underline{\Omega}$ è detto EVENTO CERTO

ii) $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$

iii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

Lo spazio degli eventi viene chiamato ALGEBRA DI EVENTI.

Si possono dimostrare le seguenti proposizioni:

1) $\phi \in \mathcal{A}$ $\quad$ $\underline{\phi}$ è detto EVENTO IMPOSSIBILE

2) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$

3) $A_1, \dots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

$$\bigcap_{i=1}^{n} A_i \in \mathcal{A}.$$

Esempi

- lancio del dado

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

$\text{card}\, \Omega = N(\Omega) = \#\Omega = n = 6$

$A =$ esce n° pari $= \{2, 4, 6\}$ è un evento

$\mathcal{A} = \mathcal{P}(\Omega) = (2^n)_{n=6} = 2^6 = 64$ fam. degli eventi

- lancio di 3 monete

$\Omega = \{TTT, TTC, TCT, CTT, TCC, CTC, CCT, CCC\}$

$\#\Omega = n = 8$

$\mathcal{A} = \mathcal{P}(\Omega) = (2^n)_{n=8} = 2^8 = 256$

Quando Ω è finito ma grande, le cardinalità o ampiezze sono molto laboriose da calcolare.
Si utilizzerà il CALCOLO COMBINATORIO.

- vita di una lampadina

$\Omega = \{x : x \geq 0\}$ non è finito!

A = la lampadina rimane accesa per almeno k ore, ma brucia prima di m ore

$= \{x : k \leq x \leq m\}$ è un evento $\forall k : 0 \leq k < m$.

OSS.: Non siamo interessati agli eventi, ma alla probabilità che uno di questi eventi si verifichi o meno.

ASSIOMI DELLA PROBABILITÀ

DEF: Sia Ω uno s.c. ed $\mathcal{A}$ un'algebra di eventi su Ω.

Una funzione:

$$P : \mathcal{A} \longrightarrow [0,1]$$

è una FUNZIONE DI PROBABILITÀ se valgono i seguenti assiomi:

A1) $P[A] \geq 0, \quad \forall A \in \mathcal{A}$

A2) $P[\Omega] = 1$

A3) $\forall$ successione $A_1, \dots, A_n, \dots \in \mathcal{A}$ a due a due disgiunti (cioè $A_i \cap A_j = \emptyset \quad i \neq j \quad \forall i, j$) e tali che $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ si ha:

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$$

La funzione P è una funzione d'insieme (gli elementi della funzione sono insiemi di punti anziché punti singoli).

<u>DEF</u>: La terna $(\Omega, \mathcal{A}, P)$ è detta SPAZIO DI PROBABILITÀ.

N.B.: Eventi disgiunti sono detti anche INCOMPATIBILI.

<u>OSS.</u>: Questa definizione matematica di probabilità ci dice quali funzioni di insieme possono essere chiamate funzioni di probabilità, non dice quale valore la funzione $P[\cdot]$ attribuisce ad un dato evento A. Bisognerà costruire un modello per l'esperimento casuale.

PROPRIETÀ DI $P[\circ]$

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità.

1) $P[\emptyset] = 0$

2) se $A_1, \dots, A_n \in \mathcal{A}$ sono e due a due disgiunti

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P[A_i]$$

3) $\forall A \in \mathcal{A}$ $\quad P[\bar{A}] = 1 - P[A]$ $\Leftarrow \Omega = A \cup \bar{A}$, $A \cap \bar{A} = \emptyset$

4) se $A, B \in \mathcal{A}$ $\qquad A = (A \cap B) \cup (A \cap \bar{B})$, $(A \cap B) \cap (A \cap \bar{B}) = \emptyset$

$$P[A] = P[A \cap B] + P[A \cap \bar{B}]$$

$$P[A - B] = P[A \cap \bar{B}] = P[A] - P[A \cap B]$$

5) se $A, B \in \mathcal{A}$ ed $A \subseteq B$

$$P[A] \leq P[B]$$

6) REGOLA DI ADDIZIONE

se $A, B \in \mathcal{A}$

$$\underline{P[A \cup B]} = \underline{P[A] + P[B] - P[A \cap B]}$$

7) se $A_1, \dots, A_n \in \mathcal{A}$

$$P\left[\bigcup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} P[A_i]$$

SPAZI CAMPIONARI FINITI

Per essi è possibile costruire un modello approssimato dell' esperimento in esame che ci permette di calcolare il valore della probabilità di un evento.

OSS.: Nella maggior parte dei casi si potrà assumere che i punti dello s.c. sono ugualmente probabili.

DEF.: Sia $\Omega = \{e_1, \dots, e_n, \dots\}$ uno s.c. Un sottoinsieme del tipo $A_J = \{e_J\}$, $J = 1, \dots, n, \dots$ è detto EVENTO ELEMENTARE o SEMPLICE.

Sia $\Omega = \{e_1, \dots, e_n\}$ uno s.c. FINITO costituito da n elementi distinti $e_1, \dots, e_n$. Si ha:

$$\Omega = \bigcup_{J=1}^{n} \{e_J\} \quad e \quad \{e_J\} \cap \{e_i\} = \phi \quad \forall i, J = 1, \dots, n \quad i \neq J$$

(disgiunti)

PROP.: Dati Ω finito e $\mathcal{A}$ algebra di eventi su Ω e $P: \mathcal{A} \rightarrow [0,1]$ si ha:

1) $\forall A \in \mathcal{A} \quad P[A] = \sum_{e_i \in A} P[\{e_i\}]$

2) assegnati n numeri positivi $p_1, \dots, p_n$:

$$\sum_{i=1}^{n} p_i = 1$$

ogni funzione additiva $P: \mathcal{A} \rightarrow [0,1]$ tale che $P[\{e_i\}] = p_i$, $i = 1, \dots, n$ è una FUNZIONE DI PROBABILITÀ

Se i punti campione sono <u>equiprobabili</u> allora:

1) $P[\{e_i\}] = p_i = \dfrac{1}{n} \qquad i = 1, \dots, n$

2) $\forall A \in \mathcal{A} \quad p[A] = \dfrac{\operatorname{card}(A)}{n}$

dove $n = \operatorname{card}(\Omega) = N(\Omega) = \#\Omega$.

In tal caso la funzione di probabilità è detta UNIFORME.

<u>Esempio</u>

- lancio di 2 dadi

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

$\operatorname{card}(\Omega) = 36$

$$\begin{pmatrix} 11 & 12 & \dots & & & 16 \\ 21 & & & & 25 & \vdots \\ \vdots & & & 34 & & \vdots \\ \vdots & & 43 & & & \vdots \\ \vdots & 52 & & & & \vdots \\ 61 & \dots & & & \dots & 66 \end{pmatrix}$$

$\{e\} = \{\omega\} = (i,J)$ evento elementare

$p[\{\omega\}] = \dfrac{1}{36}$

$A_7 = \text{esce } 7 = \{ (1,6); (6,1); (2,5); (5,2); (3,4); (4,3) \}$

$\operatorname{card} A_7 = \# A_7 = N(A_7) = 6$

$P[A_7] = \dfrac{\# A_7}{\# \Omega} = \dfrac{6}{36} = \dfrac{1}{6}$

Se lo s.c. Ω finito è costituito da punti non tutti equiprobabili la situazione è più complessa.

Esempio

Dado truccato

La probabilità della J-esima faccia A_J è proporzionale a J, $J = 1, \dots, 6$.

Calcolare la probabilità di ottenere:

1) una faccia pari

2) una faccia dispari

Sia $P[A_J] = \alpha J$ $\quad \alpha$ coeff. di proporzionalità

α si determina in base alla condizione di normalizzazione:

$$1 = \sum_{J=1}^{6} P[A_J] = \alpha \sum_{J=1}^{6} J = \alpha \cdot \frac{6 \cdot 7}{2} = 21\,\alpha$$

$$\Rightarrow \quad \alpha = \frac{1}{21}$$

1) $P[A_2] + P[A_4] + P[A_6] = \frac{2}{21} + \frac{4}{21} + \frac{6}{21} = \frac{4}{7} = 0,571\ldots$

2) $P[A_1] + P[A_3] + P[A_5] = \frac{1}{21} + \frac{3}{21} + \frac{5}{21} = \frac{3}{7} = 0,428\ldots$ ∎

In presenza di s.c. finiti il calcolo delle probabilità di un evento si riduce ad un problema di conteggio del n° degli elementi dell'evento.

Se $N(\Omega)$ o $N(A)$ sono grandi si utilizzerà il

CALCOLO COMBINATORIO.

PROBABILITÀ CONDIZIONATA

Accade spesso di dover calcolare la probabilità di un evento "ammesso che se ne sia verificato un altro". Questa probabilità è detta condizionata.

DEF.: Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e siano $A, B \in \mathcal{A}$ due eventi. Si definisce la probabilità condizionata di A dato B, che si indica con $P[A|B]$ la quantità scalare:

$$P[A|B] = \begin{cases} \dfrac{P[A \cap B]}{P[B]} & \text{se } P[B] > 0 \\ 0 & \text{altrimenti} \end{cases}$$

<u>Esempi</u>

- lancio di due monete

Calcolare la probabilità di:

1) avere 2 teste, data una testa sulla 1ª moneta;

2) avere 2 teste, data almeno una testa.

$\Omega = \{TT, TC, CT, CC\}$ $\quad n = 4$

esiti equiprobabili $\Rightarrow$ probabilità uniforme

Soluzione intuitiva

i) testa sulla 1ª moneta $\Rightarrow \{TT, TC\} \Rightarrow p = \frac{1}{2}$

ii) almeno una testa $\Rightarrow \{TT, TC, CT\} \Rightarrow p = \frac{1}{3}$

Scarto $A_1 = \{T \text{ sulla } 1^a \text{ moneta}\}$

$A_2 = \{T \text{ sulla } 2^a \text{ moneta}\}$

$A_1 \cup A_2 = \{\text{almeno una } T\}$

$A_1 \cap A_2 = \{\text{due } T\}$

1) $$P[A_1 \cap A_2 \mid A_1] = \frac{P[(A_1 \cap A_2) \cap A_1]}{P[A_1]} = \frac{P[A_1 \cap A_2]}{P[A_1]}$$

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

2) $$P[A_1 \cap A_2 \mid A_1 \cup A_2] = \frac{P[(A_1 \cap A_2) \cap (A_1 \cup A_2)]}{P[A_1 \cup A_2]}$$

$$= \frac{P[A_1 \cap A_2]}{P[A_1 \cup A_2]}$$

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

Tramite questo esempio possiamo osservare:

- Se Ω s.c. finito con n elementi e la probabilità è uniforme, se $P[B] > 0$ si ha:

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{\operatorname{card}(A \cap B)/n}{\operatorname{card}(B)/n} = \frac{\operatorname{card}(A \cap B)}{\operatorname{card}(B)}$$

- lancio di due dadi

Calcolare la probabilità di "fare 7" ammesso che che i dadi diano punteggi consecutivi.

$A = $ "fare 7" $= \{(1,6);(6,1);(2,5);(5,2);\underline{(3,4);(4,3)}\}$

$B = $ "punteggi consecutivi"

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

$\# A = 6 \quad , \quad \# B = 10$

$A \cap B = \{(3,4);(4,3)\}$

$\#(A \cap B) = 2$

$\Rightarrow \quad P[A \mid B] = \dfrac{P[A \cap B]}{P[B]} = \dfrac{2}{10} = \dfrac{1}{5}$

- lancio di due dadi

Calcolare la probabilità di "fare 9" ammesso che su un dado esca 3.

$A = $ "fare 9" $= \{\underline{(3,6);(6,3)};(4,5);(5,4)\}$

$B = $ "uno è 3" $= \{(1,3);(3,1);(2,3);(3,2);(3,3);$
$(3,4);(4,3);(3,5);(5,3);\underline{(3,6);(6,3)}\}$

$\# A = 4 \quad , \quad \# B = 11$

$A \cap B = \{(3,6);(6,3)\}$

$\#(A \cap B) = 2$

$\Rightarrow P[A|B] = \frac{2}{11}$.

N.B.: Non confondere la probabilità condizionata con le prove ripetute dello stesso esperimento.

<u>Lanciare un dado 2 volte.</u>

Se lancio un dado ed esce 3 e poi lo rilancio la probabilità di "ottenere 3" dopo il secondo lancio è $\frac{1}{6}$.

OSS.: Si può dimostrare che la probabilità condizionata $P[\bullet|B]$ con $P[B]>0$ è una funzione di probabilità (cioè verifica i 3 assiomi).

Si possono dimostrare le seguenti proprietà:

1) $P[\emptyset|B]=0$ <u>se $B \in \mathcal{A}$</u>

2) se $A_1, \dots, A_n \in \mathcal{A}$ a due a due disgiunti

$$P\left[\bigcup_{i=1}^{n} A_i \,|\, B\right] = \sum_{i=1}^{n} P[A_i|B]$$

3) se $A \in \mathcal{A}$

$$P[\bar{A}|B] = 1 - P[A|B]$$

4) se $A_1, A_2 \in \mathcal{A}$

$$P[A_1|B] = P[A_1 \cap A_2|B] + P[A_1 \cap \bar{A}_2|B]$$

$$P[A_1 \cup A_2|B] = P[A_1|B] + P[A_2|B] - P[A_1 \cap A_2|B]$$

5) se $A_1, A_2 \in \mathcal{A}$ e $A_1 \subseteq A_2$

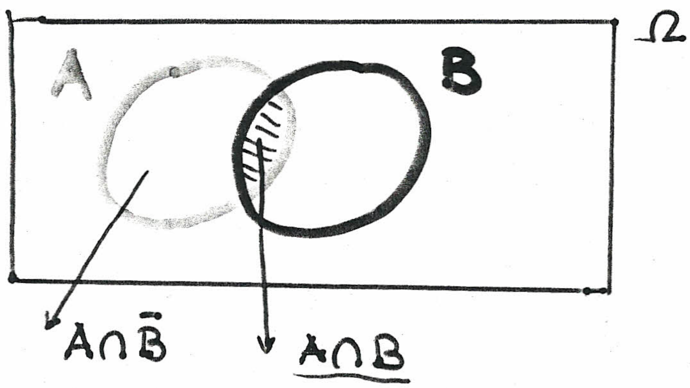
$$P[A_1 | B] \leq P[A_2 | B]$$

6) se $A_1, \dots, A_n \in \mathcal{A}$

$$P\left[\bigcup_{i=1}^{n} A_i \,\Big|\, B\right] \leq \sum_{i=1}^{n} P[A_i | B]$$

OSS. : Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità.

Siano $A, B \in \mathcal{A}$.



Si ha:

$$A = (A \cap B) \cup (A \cap \bar{B})$$

$(A \cap B)$ e $(A \cap \bar{B})$ disgiunti $\Rightarrow$

$$P[A] = P[A \cap B] + P[A \cap \bar{B}]$$

Utilizzando la def. di p.c. si ottiene:

$$P[A] = P[A|B]\,P[B] + P[A|\bar{B}]\,P[\bar{B}]$$

$$= P[A|B]\,P[B] + P[A|\bar{B}]\,(1 - P[B]) \quad \blacksquare$$

Introduciamo ora un teorema particolarmente utile negli esperimenti a più fasi.

TEOREMA DELLE PROBABILITÀ TOTALI

Sia $(\Omega, \mathcal{A}, P)$ uno spazio di probabilità e $B_1, \dots, B_n$ una collezione finita di eventi di $\mathcal{A}$ e due a due incompatibili tali che:

i) $P[B_i] > 0 \quad i = 1, \dots, n$

ii) $\Omega = \bigcup_{i=1}^{n} B_i$

Allora $\forall A \in \mathcal{A}$ si ha:

$$P[A] = \sum_{i=1}^{n} P[A \mid B_i] \, P[B_i]$$

Dimostrazione

Si noti che

$$A = \Omega \cap A = \left(\bigcup_{i=1}^{n} B_i\right) \cap A = \bigcup_{i=1}^{n} (A \cap B_i)$$

$A \cap B_i$ sono <u>incompatibili</u> poichè sottoinsiemi di B_i.

Allora:

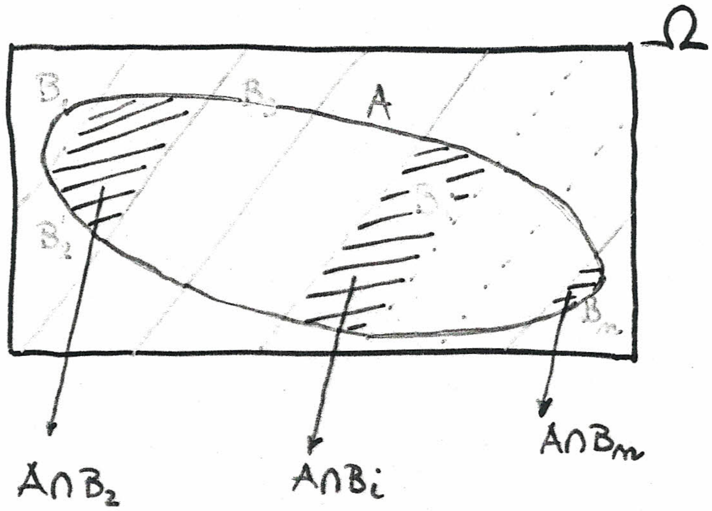
$$P[A] = P\left[\bigcup_{i=1}^{n} (A \cap B_i)\right] = \sum_{i=1}^{n} P[A \cap B_i]$$

$$= \sum_{i=1}^{n} P[A \mid B_i] \, P[B_i] \quad \blacksquare$$

dalla def. di p.c.

OSS.: Il teorema è vero se $n = \infty$.

TEOREMA PROBABILITÀ TOTALI (grafico)



$A \cap B_2 \subset B_2$

$A \cap B_3 \subset B_3$

$\vdots$

$A \cap B_i \subset B_i$

$B_1, \ldots, B_n$ disgiunti a 2 a 2

$\Rightarrow (A \cap B_1), (A \cap B_2), \ldots, (A \cap B_n)$ a 2 a 2 disgiunti

Una forma equivalente della def. di p. condizionata è la REGOLA DI MOLTIPLICAZIONE:

$$P[A \cap B] = P[B]\, P[A|B]$$

Poiché $P[A \cap B] = P[B \cap A]$, si ha anche

$$P[A \cap B] = P[B \cap A] = P[A]\, P[B|A]$$

pertanto, se $P[A]>0$, $P[B]>0$, vale la:

PRIMA FORMULA DI BAYES

$$P[A|B] = \frac{P[B|A]\, P[A]}{P[B]}$$

Data una serie di eventi che ricoprono lo spazio campione e dato un evento A, ci si può chiedere anche quale sia la probabilità condizionata di ognuno di questi eventi, dato A.

TEOREMA DI BAYES

Sia $(\Omega, \mathcal{A}, P)$ uno s. di p. e $B_1, \dots, B_n$ una collezione finita di eventi di $\mathcal{A}$ a due a due incompatibili:

i) $P[B_i] > 0 \quad i = 1, \dots, n$

ii) $\Omega = \bigcup_{i=1}^{n} B_i$

Allora $\forall A \in \mathcal{A}$ con $P[A]>0$ si ha:

$$P[B_i | A] = \frac{P[A|B_i]\, P[B_i]}{\underbrace{\sum_{J=1}^{n} P[A|B_J]\, P[B_J]}_{= P[A]}}$$

Dimostrazione

Si ottiene subito applicando la 1ª formula di Bayes e osservando che per il teorema delle probabilità totali il denominatore corrisponde a $P[A]$.

OSS.: Il teorema è vero per $n = \infty$.

Esempi

- Un giocatore frequenta un tavolo da gioco dove si alternano 2 croupier gemelli: uno è onesto (la probabilità di vincere in sua presenza è 0.5), l'altro bara e la probabilità di vincere è $p < 0.5$.
  Le presenze dei 2 gemelli sono equiprobabili.
  Un giorno il giocatore perde.
  Qual è la probabilità che il croupier sia quello disonesto?

  D = "croupier disonesto"

  $\bar{D}$ = "croupier onesto"

  V = "giocatore vince"

  $\bar{V}$ = "giocatore perde"

Per il th. delle p. totali:

$$P(\bar{V}) = P(\bar{V}|D)\,P(D) + P(\bar{V}|\bar{D})\,P(\bar{D})$$

$$= (1-p)\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} = \frac{3-2p}{4}$$

Per la 1ª formula di Bayes:

$$P(D|\bar{V}) = \frac{P(\bar{V}|D)\,P(D)}{P(\bar{V})} = \frac{(1-p)\cdot\frac{1}{2}}{\frac{3-2p}{4}} = \frac{2(1-p)}{3-2p}$$ ∎

- Un segnale può essere trasmesso da 2 canali A, B con la stessa probabilità.
  A trasmette sempre correttamente,
  B trasmette correttamente con probabilità $\frac{3}{4}$.

1) Qual è la probabilità di ricevere un segnale corretto?

2) Ricevuto un segnale corretto, qual è la probabilità che provenga da B?

A = "segnale proviene da A"

B = "segnale proviene da B"

C = "segnale corretto"

$$P(C) = P(C|A)\,P(A) + P(C|B)\,P(B) \quad \text{th. p. totali}$$

$$= 1\cdot\frac{1}{2} + \frac{3}{4}\cdot\frac{1}{2} = \frac{7}{8}$$

$$P(B|C) = \frac{P(C|B) \cdot P(B)}{P(C)} = \frac{\frac{3}{4} \cdot \frac{1}{2}}{\frac{7}{8}} = \frac{3}{7}$$ 1^a f. di Bayes

- Date 4 macchine che producono pezzi difettosi con percentuali:

$p_1 = 0{,}1\%$, $p_2 = 0{,}2\%$, $p_3 = 0{,}3\%$, $p_4 = 0{,}1\%$

1) Scelta una macchina a caso ed estratto a caso un pezzo, qual è la probabilità che sia difettoso?

2) Estratto un pezzo e visto che è difettoso, qual è la probabilità che provenga p. es. dalla 3^a macchina?

A = " pezzo difettoso"

B_i = " scelta della i-esima macchina "

<u>Ipotesi sottointese</u> : - equiprobabilità nella scelta delle macchine
- produzione dei pezzi difettosi indipendente

$$P(B_i) = \frac{1}{4}$$

$$P(A|B_i) = p_i \quad i = 1, 2, 3, 4.$$

1) $P(A) = \sum_{i=1}^{4} P(A|B_i) P(B_i)$ th. p. totali

$$= \frac{1}{4} (0{,}1 + 0{,}2 + 0{,}3 + 0{,}1)\% = \frac{1}{4} \cdot 0{,}7\% \simeq 0{,}17\%$$

2) $P(B_3|A) = \dfrac{P(A|B_3)\cdot P(B_3)}{P(A)}$ th. di Bayes

$$= \frac{\frac{1}{4}\cdot p_3}{\frac{1}{4}\cdot 0{,}7\,\%} = \frac{0{,}3}{0{,}7} \simeq 43\,\%$$

- Si hanno 5 urne numerate contenenti ognuna 10 palline.

Nella i-esima urna ci sono $\begin{cases} i \text{ palline difettose} \\ 10-i \text{ palline non difettose} \end{cases}$

Scelta un'urna a caso ed estratta una pallina,

1) qual è la probabilità che la pallina sia difettosa?

2) estratta una pallina e visto che è difettosa, qual è la probabilità che provenga (p. es.) dall'urna n° 5?

$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$	$\boxed{5}$
1-9	2-8	3-7	4-6	5-5

Intuitivamente ci sono 15 palline difettose su 50.

$\Rightarrow p = \frac{3}{10}$ (dom. 1)

A = "pallina difettosa"

B_i = "i-esima urna"

$P[B_i] = \frac{1}{5}$

$$P[A|B_i] = \frac{i}{10}$$

$$P[A] = \sum_{i=1}^{5} P[A|B_i]\,P[B_i] \qquad \text{th. p. totali}$$

$$= \sum_{i=1}^{5} \frac{i}{10} \cdot \frac{1}{5} = \frac{1}{50} \sum_{i=1}^{5} i = \frac{1}{50} \cdot \frac{5 \cdot 6}{2} = \frac{3}{10}$$

$$\left(\sum_{J=1}^{n} J = \frac{n(n+1)}{2} \right)$$

2) $$P[B_5|A] = \frac{P[A|B_5] \cdot P[B_5]}{P[A]} \qquad \text{th. di Bayes}$$

$$P[A] = \sum_{i=1}^{5} P[A|B_i] \cdot P[B_i]$$

$$= \frac{\frac{5}{10} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{1}{3}$$

In generale

$$P[B_k|A] = \frac{\frac{k}{10} \cdot \frac{1}{5}}{\frac{3}{10}} = \frac{k}{15} \qquad k = 1, \dots, 5$$

OSS. : In termini non condizionati tutti i B_i sono equiprobabili ($1/5$), condizionati al verificarsi dell'evento A non lo sono più ($\frac{i}{15}$).

REGOLA DEL PRODOTTO

Sia $(\Omega, \mathcal{A}, \mathbb{P})$ uno s. di p. e $B_1, \dots, B_n$ una collezione finita di eventi di $\mathcal{A}$: $P[B_1 \cap \dots \cap B_{n-1}] > 0$, allora

$$P[B_1 \cap \dots \cap B_n] = P[B_1] \cdot P[B_2 | B_1] \cdot P[B_3 | B_1 \cap B_2] \dots$$
$$\dots P[B_n | B_1 \cap \dots \cap B_{n-1}]$$

(no dim.)

__Esempio__

Data un'urna contenente 10 palline (3 nere, 7 bianche)
Si estrae una pallina, si registra il suo colore, si rimette nell'urna aggiungendone altre due dello stesso colore.

Qual è la probabilità di estrarre una pallina nera in ognuno dei primi 3 tentativi?

B_i = estrazione pallina nera all'i-esimo tentativo

$$P[B_1 \cap B_2 \cap B_3] = P[B_1] \cdot P[B_2 | B_1] \cdot P[B_3 | B_1 \cap B_2]$$

$$= \frac{3}{10} \cdot \frac{5}{12} \cdot \frac{7}{14} = \frac{1}{16}$$

INDIPENDENZA

In generale $P[A|B]$ è diversa da $P[A]$.
Perciò, sapere che l'evento B si sia verificato, modifica di solito la probabilità che si sia verificato A.

<u>DEF.</u>: Sia $(\Omega, \mathcal{A}, P)$ uno s. di p. Due eventi $A, B \in \mathcal{A}$ si dicono INDIPENDENTI se:

$$P[A \cap B] = P[A] \cdot P[B]$$

- Se $A, B \in \mathcal{A}$ sono indipendenti in $(\Omega, \mathcal{A}, P)$, sono indipendenti $\underline{A \text{ e } \bar{B}}$, $\underline{\bar{A} \text{ e } B}$, $\underline{\bar{A} \text{ e } \bar{B}}$.

  Verifichiamo l'indipendenza di A e $\bar{B}$.
  Poiché $A = (A \cap B) \cup (A \cap \bar{B})$
  e $(A \cap B) \cap (A \cap \bar{B}) = \phi$ (disgiunti)

$\Rightarrow P[A] = P[A \cap B] + P[A \cap \bar{B}]$

allora

$$P[A \cap \bar{B}] = P[A] - P[A \cap B] \overset{\text{INDIP.}}{=} P[A] - P[A] \cdot P[B]$$

$$= P[A] (1 - P[B]) = P[A] \cdot P[\bar{B}]$$

(per esercizio gli altri 2 casi).

OSS.: Sia l'indipendenza che la probabilità condizionata vengono usate per calcolare $P[A \cap B]$.

- $P[A \cap B] = P[A] \cdot P[B]$ se A, B indip.
- $P[A \cap B] = P[A|B] \, P[B]$ se A, B non indip.

dove la 2ª è valida in generale purché $P[A|B]$ sia definito.

TEOREMA Sia $(\Omega, \mathcal{A}, P)$ uno s. di p. e $A, B \in \mathcal{A}$.

Le seguenti condizioni sono EQUIVALENTI :

1. $P[A \cap B] = P[A]\, P[B]$.
2. se $P[B] > 0 \Rightarrow P[A | B] = P[A]$
3. se $P[A] > 0 \Rightarrow P[B | A] = P[B]$

Dimostrazione

$1 \Rightarrow 2$ $\quad P[A|B] = \dfrac{P[A \cap B]}{P[B]} \overset{1)}{=} \dfrac{P[A] \cdot P[B]}{P[B]} = P[A]$

$2 \Rightarrow 3$ $\quad P[B|A] \underset{1^a \text{ F. Bayes}}{=} \dfrac{P[A|B]\, P[B]}{P[A]} \overset{2)}{=} \dfrac{P[A]\, P[B]}{P[A]} = P[B]$

$3 \Rightarrow 1$ $\quad P[A \cap B] = P[B|A]\, P[A] \underset{3)}{=} P[A]\, P[B]$

Esempio

- lancio di 2 dadi

Se A = " fare dispari "

B = " fare 1 sul 1° dado "

C = " fare 7 "

Dire se sono indipendenti A e B, A e C, B e C.

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

$\# \Omega = 36$

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

$\# A = 18$

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

$\# B = 6$

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

$\# C = 6$

$$P[A] = \frac{\# A}{\# \Omega} = \frac{18}{36} = \frac{1}{2}$$

$$P[B] = \frac{\# B}{\# \Omega} = \frac{6}{36} = \frac{1}{6}$$

$$P[C] = \frac{\# C}{\# \Omega} = \frac{6}{36} = \frac{1}{6}$$

$A \cap B = \{(1,2); (1,4); (1,6)\}$

$\# A \cap B = 3$

$A \cap C = \{(1,6); (6,1); (2,5); (5,2); (3,4); (4,3)\}$

$\# A \cap C = 6$

$B \cap C = \{(1,6)\}$

$\# B \cap C = 1$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{\# A \cap B}{\# B} = \frac{3}{6} = \frac{1}{2} = P[A] \quad \text{si}$$

$$P[A|C] = \frac{P[A \cap C]}{P[C]} = \frac{\# A \cap C}{\# C} = \frac{6}{6} = 1 \neq P[A] \quad \text{no}$$

$$P[B|C] = \frac{P[B \cap C]}{P[C]} = \frac{\# B \cap C}{\# C} = \frac{1}{6} = P[B] \quad \text{si}$$

OSS.: Sia $(\Omega, \mathcal{A}, P)$ uno s. di p. e dati A, B, C tre eventi in $\mathcal{A}$, se A è indipendente da B e da C

$\overset{?}{\Rightarrow}$ A è indipendente da $B \cap C$? No!

Esempio

- lancio 2 dadi

A = "fare 7" = $\{(1,6); (6,1); (2,5); (5,2); (3,4); (4,3)\}$

B = "esce 4 sul 1° dado" = $\{(4,1); (4,2); (4,3); (4,4); (4,5); (4,6)\}$

C = "esce 3 sul 2° dado" = $\{(1,3); (2,3); (3,3); (4,3); (5,3); (6,3)\}$

$\# A = \# B = \# C = 6$

$P[A] = P[B] = P[C] = \frac{6}{36} = \frac{1}{6}$

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} = P[A]$$

$$P[A \mid C] = \frac{P[A \cap C]}{P[C]} = \frac{\frac{1}{36}}{\frac{1}{6}} = \frac{1}{6} = P[A]$$

$$P[A \mid B \cap C] = \frac{P[A \cap B \cap C]}{P[B \cap C]} = \frac{\frac{1}{36}}{\frac{1}{36}} = 1 \neq P[A].$$

DEF. Sia $(\Omega, \mathcal{A}, P)$ uno s. di p. TRE eventi $A, B, C \in \mathcal{A}$ si dicono INDIPENDENTI se

$$P[A \cap B \cap C] = P[A]\, P[B]\, P[C]$$
$$P[A \cap B] = P[A]\, P[B]$$
$$P[A \cap C] = P[A]\, P[C]$$
$$P[B \cap C] = P[B]\, P[C]$$

OSS. Se 3 eventi A, B, C sono indipendenti, allora ognuno di essi è indipendente da qualunque evento si possa costruire con gli altri due.

Esempio

Dimostrare che A e $B \cup C$ sono indipendenti.

$$\begin{aligned} P[A \cap (B \cup C)] &= P[(A \cap B) \cup (A \cap C)] \\ &= P[A \cap B] + P[A \cap C] - P[A \cap B \cap C] \\ &= P[A]\,P[B] + P[A]\,P[C] - P[A]\,P[B \cap C] \\ &= P[A] \left\{ P[B] + P[C] - P[B \cap C] \right\} \\ &= P[A]\, P[B \cup C] \end{aligned}$$

DEF.: Siano $(\Omega, \mathcal{A}, P)$ uno s. di p. e $A_1, \dots, A_n$ n eventi in $\mathcal{A}$.

$A_1, \dots, A_n$ sono INDIPENDENTI se

$$P[A_i \cap A_j] = P[A_i]\, P[A_j] \quad \text{per } i \neq j$$

$$P[A_i \cap A_j \cap A_k] = P[A_i]\, P[A_j]\, P[A_k] \quad \text{per } i \neq j,\ j \neq k,\ k \neq i$$

$$\vdots$$

$$P\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} P[A_i]$$

OSS.: Eventi DISGIUNTI (o INCOMPATIBILI o MUTUAMENTE ESCLUSIVI) **NON** sono eventi INDIPENDENTI.

$\forall A, B \in \mathcal{A}$ disgiunti si ha $A \cap B = \emptyset$

$\Rightarrow P[A \cap B] = 0.$

Se A e B fossero indipendenti

$$P[A \cap B] = P[A]\,P[B]$$

$\Rightarrow \quad P[A]\,P[B] = 0 \quad \Rightarrow \quad P[A]$ o $P[B] = 0$

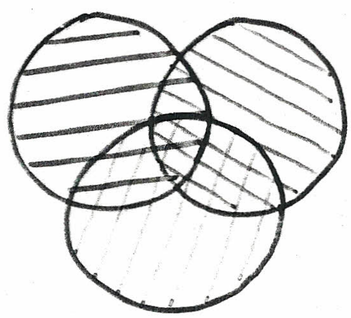
$\Rightarrow$ A o B è l'evento impossibile.

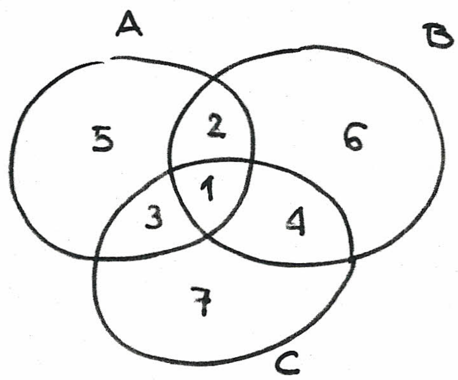
Esercizi di ricapitolazione

Dimostrare la formula:

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

$$-p(A\cap B) - p(A\cap C) - p(B\cap C)$$





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

$$-p((A\cup B)\cap C)$$

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

ma

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

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

Se gli eventi sono indipendenti

- $P(A\cup B\cup C) = p(A)\,p(\bar{B})\,p(\bar{C}) + p(B)\,p(\bar{C}) + p(C)$

perché

$$P(A\cap B) = p(A)\cdot p(B)$$

$$p(A\cap C) = p(A)\cdot p(C)$$

$$p(B\cap C) = p(B)\,p(C)$$

$$p(A\cap C\cap B) = p(A)\cdot p(B)\cdot p(C)$$

e raccogliendo si ottiene:

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

$$- P(A)P(B) - P(A)\cdot P(C) - P(B)P(C)$$

$$= P(C) + P(B)\left[1 - P(C)\right] +$$

$$P(A)\left[1 + P(B)P(C) - P(B) - P(C)\right]$$

$$= P(C) + P(B)P(\bar{C}) + P(A)\left[(1 - P(C)) -\right.$$

$$\left.- P(B)\left[1 - P(C)\right]\right]$$

$$= P(C) + P(B)P(\bar{C}) + P(A)\left[P(\bar{C}) - P(B)P(\bar{C})\right]$$

$$= P(C) + P(B)P(\bar{C}) + P(A)P(\bar{C})\left[1 - P(B)\right]$$

$$= P(A)P(\bar{B})P(\bar{C}) + P(B)P(\bar{C}) + P(C).$$

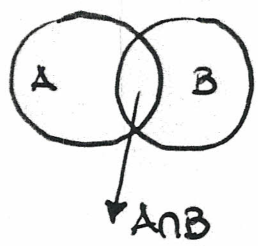
Esempio

In una popolazione : 28% fuma sigarette (= A)

7% " sigari (= B)

5% " entrambe (= $A \cap B$)

Calcolare la % di chi NON fuma né sigari, né sigarette.



$P(A \cup B) = P(A) + P(B) - P(A \cap B)$ fuma sigarette O sigari

$= 0.28 + 0.07 - 0.05 = 0.30$

$P(\overline{A \cup B}) = 1 - P(A \cup B) = 0.7 \Rightarrow 70\%$

Esercizi

1) I pezzi prodotti da una certa macchina possono avere 2 tipi di difetti D_1, D_2.

Si sa:

$P[D_1] = 0.1$ presenza del primo difetto

$P[\bar{D}_2] = 0.8$ assenza del secondo difetto

$P[D_1 \cap D_2] = 0.01$ presenza di entrambi i difetti

Calcolare la probabilità che il pezzo scelto non abbia alcun difetto.

Devo cioè calcolare $P[\bar{D}_1 \cap \bar{D}_2]$.

$P[\bar{D}_1] = 0.9$; $P[D_2] = 0.2$

Per la regola di addizione

$$P[D_1 \cup D_2] = P[D_1] + P[D_2] - P[D_1 \cap D_2]$$

$$= 0.1 + 0.2 - 0.01 = 0.29$$

Legge di De Morgan

$$P[\bar{D}_1 \cap \bar{D}_2] = P[\overline{D_1 \cup D_2}] = 1 - P[D_1 \cup D_2]$$

$$= 1 - 0.29$$

$$= 0.71$$

Calcolare la probabilità di ottenere ALMENO ($\neq$ esattamente!) 2 numeri uguali

$\Omega = \{1, \dots, 6\}$ per 1° lancio

casi possibili $= 6^3 = 216$

casi favorevoli $= ?$

1) $i\,i\,j$ $\quad j = 1, \dots, 6 \quad i = 1, \dots, 6 \quad 6 \times 6 = 36$

2) $i\,j\,i$ $\quad$ " $\quad$ " $\quad$ "

3) $j\,i\,i$ $\quad$ " $\quad$ " $\quad$ "

$\Rightarrow$ totale $= 36 + 36 + 36 = 108$

$\Rightarrow p = \frac{108}{216} = \frac{1}{2}$ non va bene!

Il caso $i\,i\,j$ con $j = i$ viene contato in 2) e in 3)

perciò totale $= 36 + 30 + 30 = 96$

$\Rightarrow p = \frac{96}{216} = \frac{4}{9}$ ok!

I problemi del Cavalier de Méré
(giocatore d'azzardo)
sottoposti a B. Pascal (1623-1662)

1. Trovare il più piccolo numero intero n tale che lanciando n volte 1 DADO, la probabilità di avere almeno un 6 sia maggiore di $1/2$;

2. Trovare il più piccolo numero intero m tale che lanciando m volte 2 DADI, la probabilità di avere almeno un (6, 6) sia maggiore di $1/2$.

1. In un singolo lancio

A = "fare 6"

$$P[A] = \frac{1}{6} \quad , \quad P[\bar{A}] = \frac{5}{6}$$

B = almeno un 6 in n lanci

$\bar{B}$ = nessun 6 in n lanci

$$P[\bar{B}] = \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdots \frac{5}{6}}_{n \text{ volte}} = \left(\frac{5}{6}\right)^n$$

$$P[B] = 1 - P[\bar{B}] = 1 - \left(\frac{5}{6}\right)^n$$

Trovare n : $P[B] > \frac{1}{2}$.

$$1 - \left(\frac{5}{6}\right)^n > \frac{1}{2} \quad \Rightarrow \left(\frac{5}{6}\right)^n < \frac{1}{2}$$

$m = 1 \quad \left(\frac{5}{6}\right)^1 = 0,8\bar{3}$

$m = 2 \quad \left(\frac{5}{6}\right)^2 = 0,694$

$m = 3 \quad \left(\frac{5}{6}\right)^3 = 0,578$

$m = 4 \quad \left(\frac{5}{6}\right)^4 \cong 0,482$

$\Rightarrow P[B]\Big|_{m=4} \cong 0,518$

2. In un lancio di 2 dadi

$A =$ fare $(6,6)$

$P[A] = \frac{1}{36} \quad , \quad P[\bar{A}] = \frac{35}{36}$

$B =$ almeno un $(6,6)$ in m lanci

$\bar{B} =$ nessun $(6,6)$ in m lanci

$P[\bar{B}] = \left(\frac{35}{36}\right)^m$

$P[B] = 1 - \left(\frac{35}{36}\right)^m$

Trovare m : $P[B] > \frac{1}{2} \quad \Rightarrow \quad \left(\frac{35}{36}\right)^m < \frac{1}{2}$

$\vdots$

$m = 24 \quad \left(\frac{35}{36}\right)^{24} \simeq 0.51 \quad \Rightarrow \quad P[B] \cong 0.49$

$m = 25 \quad \left(\frac{35}{36}\right)^{25} \simeq 0.49 \quad \Rightarrow \quad P[B] \cong 0.51$

Esempio

Data un'urna con 11 palline: 5 nere e 6 bianche.

Si estraggono 2 palline.

Qual è la probabilità che siano 1 bianca e 1 nera.

La 1ª estrazione avviene su 11.

$p(N) = \frac{5}{11} \qquad p(B) = \frac{6}{11}$

Alla 2ª estrazione le palline sono 10. (5 bianche - 5 nere
oppure
4 nere - 6 bianche).

$p(B \mid N) = \frac{6}{10}$

$p(N \mid N) = \frac{4}{10}$

$p(B \mid B) = \frac{5}{10}$

$p(N \mid B) = \frac{5}{10}$

La p che siano 1B e 1N = (1B - 1N) o (1N - 1B).

$$p = \frac{5}{11} \cdot \frac{6}{10} + \frac{6}{11} \cdot \frac{5}{10}$$

$$= \frac{6}{11}.$$

Lo spazio comprende $10 \cdot 11 = 110$ elementi tutti equiprobabili.

$6 \times 5 = 30$ casi in cui la 1ª è bianca e la 2ª è nera.

$5 \times 6 = 30$ " " " " " nera " " " " bianca.

$$\frac{30 + 30}{110} = \frac{6}{11}.$$

Esempio

In una scatola a sono:

5 transistor rotti ($= R$)

10 " difettosi ($= D$ funzionano per un po' e poi si rompono)

25 " buoni ($= B$)

Scelgo a caso un transistor.

Qual è la probabilità che sia buono se inizialmente funziona.

In totale sono 40 $\begin{cases} 5\ R \\ 35\ (D+B) \end{cases}$

Spazio degli esiti ridotto: 35 $\begin{cases} 10\ D \\ 25\ B \end{cases}$

$$p = \frac{25}{35} = 5/7$$

$$P(B|\bar{R}) = \frac{P(B \cap \bar{R})}{P(\bar{R})} = \frac{P(B)}{P(\bar{R})} = \frac{P(B)}{1-P(R)} = \frac{\frac{25}{40}}{1-\frac{5}{40}} = \frac{25}{35} = \frac{5}{7}$$

Esercizio

Uno studente svolge un test (tempo a disposizione 1 ora)

La probabilità di terminare il test in meno di x ore

è pari a $\frac{x}{2}$, $0 \le x \le 1$.

Dopo tre quarti d'ora lo studente non ha finito il test.

Qual è la probabilità che lo studente utilizzi l'ora intera?

$A_x = \{$ lo studente conclude il test in meno di x ore, $0 \le x \le 1 \}$

$B = \{$ lo studente usa l'ora intera $\}$

$$P(B) = P(\bar{A}_1) = 1 - P(A_1) = 1 - \left(\frac{x}{2}\right)_{x=1} = \frac{1}{2}$$

$\bar{A}_{\frac{3}{4}} = \{$ al tempo $3/4$ d'ora lo studente non ha terminato il test $\}$

$$P(B \mid \bar{A}_{3/4}) = \frac{P(B \cap \bar{A}_{3/4})}{P(\bar{A}_{3/4})} = \frac{P(B)}{1 - P(A_{3/4})} = \frac{\frac{1}{2}}{1 - \left(\frac{x}{2}\right)_{3/4 = x}}$$

$$= \frac{\frac{1}{2}}{1 - \frac{3}{8}} = \frac{4}{5} = 0.8$$

Esempio

In una prova a risposte multiple m lo studente conosce la risposta con probabilità p, indovina la risposta con probabilità $1-p$.

Se prova ad indovinare, risponde correttamente con probabilità $\frac{1}{m}$.

Qual è la probabilità condizionata che sapesse la risposta ad una domanda a cui ha risposto correttamente?

$A = \{$sceglie la risposta giusta$\}$

$B = \{$sa la risposta$\}$

- $P(B|A) = \dfrac{P(A|B)\,P(B)}{P(A)} = \dfrac{1 \cdot p}{P(A)}$

ma

$$P(A) = \underset{\substack{\| \\ 1}}{P(A|B)}\ \underset{\substack{\| \\ p}}{P(B)} + \underset{\substack{\| \\ \frac{1}{m}}}{P(A|\bar{B})} \cdot \underset{(1-p)}{P(\bar{B})}$$

$$= p + \frac{(1-p)}{m}$$

$$\Rightarrow P(B|A) = \frac{mp}{mp+1-p}$$

se $p = \frac{1}{2}$, $1-p = \frac{1}{2}$, $m = 5$

$\uparrow$ numero delle alternative

$$P(B|A) = \frac{5}{6}.$$

Esempio

Un' assicurazione pensa che la popolazione possa essere suddivisa tra persone inclini e non inclini a provocare incidenti.

Dati: • I (inclini) ha un incidente in un anno con $p_1 = 0.4$

• NI (non inclini) " " " " " " $p_2 = 0.2$

• 30% della popolazione è incline.

Qual è la probabilità che un nuovo assicurato abbia un incidente in un anno?

$A = \{\text{avere un incidente in un anno}\}$

$B = \{\text{essere incline}\}$

$$P(A) = P(A|B)\,P(B) + P(A|\bar{B})\,P(\bar{B})$$

$P(B) = 0.3$; $P(\bar{B}) = 0.7$

$P(A|B) = p_1 = 0.4$; $P(A|\bar{B}) = p_2 = 0.2$

$$\Rightarrow P(A) = 0.4 \times 0.3 + 0.2 \times 0.7 = 0.26$$

Se ciò avviene, qual è la probabilità che la persona appartenga al gruppo I?

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)\,P(B)}{P(A)} = \frac{0.4 \times 0.3}{0.26} = \frac{6}{13} \approx 0.46$$

Esercizio

Un mazzo di 52 carte è suddiviso casualmente in 4 mazzi da 13 carte ciascuno.
Calcolare la probabilità che ci sia un asso in ogni mazzetto.

$A_i = \{$ l'i-esimo mazzetto ha <u>esattamente</u> un asso $\}$

$$P = P(A_1 \cap A_2 \cap A_3 \cap A_4)$$

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1)\, P(A_2 | A_1)\, P(A_3 | A_1 \cap A_2) \cdot$$

per la regola del prodotto $\cdot P(A_4 | A_1 \cap A_2 \cap A_3)$

dove

$P(A_1) = 1$ se A_1 asso di cuori in un mazzo

$P(A_2 | A_1) = \frac{39}{51}$ se A_2 " " " e picche in 2 diversi m.

$P(A_3 | A_1 \cap A_2) = \frac{26}{50}$ se A_3 " " " " ", quadri in 3 diversi

$P(A_4 | A_1 \cap A_2 \cap A_3) = \frac{13}{49}$ se A_4 " " " " " ", fiori in 4 div

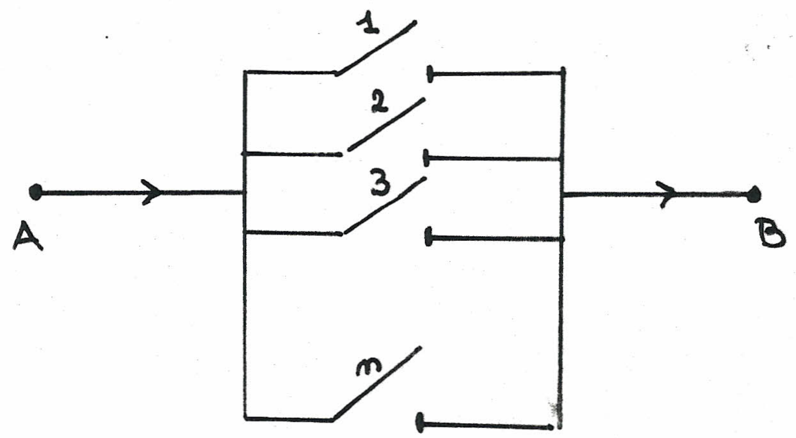
$$\Rightarrow P = 1 \cdot \frac{39}{51} \cdot \frac{26}{50} \cdot \frac{13}{49} \approx 0.105$$

cioè c'è il 10,5% di possibilità che in ogni mazzetto vi sia un asso.

3) Un sistema costituito da n componenti si dice in PARALLELO se funziona fintanto che ALMENO UNO dei componenti funziona.

Sia dato un sistema in parallelo per il quale il j-esimo componente funziona con probabilità P_j, INDIPENDENTEMENTE da tutti gli altri con $j = 1, \dots, n$

Qual è la probabilità che il sistema funzioni?



$A_j = \{$ il componente j-esimo funziona $\}$ $\quad j = 1, \dots, n$

$$
\begin{aligned}
P[S \text{ funziona}] &= 1 - P[S \text{ non funziona}] \\
&= 1 - P[\text{nessun componente funziona}] \\
&= 1 - P[\bar{A}_1 \cap \bar{A}_2 \cap \dots \cap \bar{A}_n] \\
&= 1 - P[\bar{A}_1] \cdot P[\bar{A}_2] \dots P[\bar{A}_n] \\
&= 1 - (1 - p_1) \cdot (1 - p_2) \dots (1 - p_n) \\
&= 1 - \prod_{i=1}^{n} (1 - p_i)
\end{aligned}
$$

VARIABILI ALEATORIE

Nei problemi di calcolo delle probabilità si è spesso condotti a considerare delle quantità che sono funzioni del risultato di un fenomeno casuale. La nozione di variabile casuale viene usata per descrivere gli eventi e misure i risultati di una prova, i quali possono assumere differenti valori anche quand le prove vengono eseguite sotto le stesse condizioni.

Esempi di v.a. o casuali sono: la durata della vita umana, il modulo della velocità di una molecola di un gas, il valore numerico degli esiti di un lancio di un dado.

Nei primi due casi la v.a. è continua, nel terzo è discreta.

<u>DEF.</u> : Dato $(\Omega, \mathcal{A}, P)$ uno s. di p., si dice VARIABILE ALEATORIA o CASUALE un'applicazione

$$X : \Omega \to \mathbb{R}$$

tale che

$$\forall x \in \mathbb{R} \quad A_x = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}$$

cioè A_x è un evento.

Una variabile casuale è perciò una funzione di ω tale che si possa calcolare

$$P[\{\omega ; X(\omega) \leq x\}],$$

cioè tale che abbia senso calcolare la probabilità che X prenda valori più piccoli di x.

Per semplicità $(X \leq x) \doteq \{\omega \in \Omega : X(\omega) \leq x\}$

X fa corrispondere un n° reale ad ogni esito dell'esperimento

Esempio : Lancio di 2 dadi

$X((i,j)) = i + j$ dove

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

X assegna ad ogni lancio la somma dei punteggi dei 2 dadi, è v.a. che assume i valori $2, 3, \dots, 12$.

DEF. : Dato $(\Omega, \mathcal{A}, P)$ s.p. e X v.a., chiamiamo FUNZIONE DI RIPARTIZIONE di X la funzione :

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

$$F(x) = P[X \leq x] = P[\{\omega \in \Omega : X(\omega) \leq x\}]$$

Esempio : Lancio 2 dadi.

IPOTESI : probabilità uniforme (non truccati)

se $x < 2$ $\quad (X \leq x) = \emptyset \quad P[X \leq x] = 0 \qquad F(x) = 0$

se $2 \leq x < 3$ $\quad (X \leq x) = \{(1,1)\} \qquad F(x) = \frac{1}{36}$

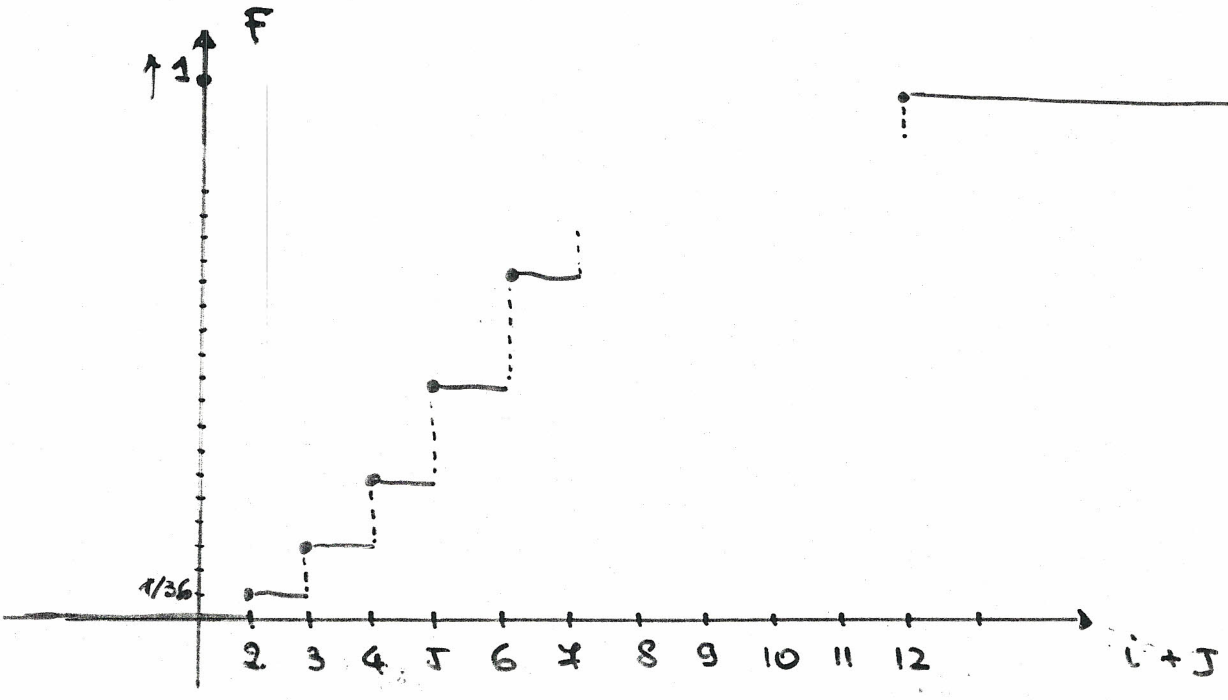
se $3 \leq x < 4$ $\quad (X \leq x) = \{(1,1); (1,2); (2,1)\} \qquad F(x) = \frac{3}{36}$

se $4 \leq x < 5$ $\quad (X \leq x) = \{(1,1); (1,2); (2,1); (1,3); (3,1); (2,2)\} \qquad F(x) = \frac{6}{36}$

⋮

se $11 \leq x < 12$ $\quad (X \leq x) = \Omega \setminus \{(6,6)\} \qquad F(x) = \frac{35}{36}$

se $x \geq 12$ $\quad (X \leq x) = \Omega \qquad F(x) = 1$



PROPRIETÀ DI F

i) $\lim_{x \to -\infty} F(x) = 0 \quad , \quad \lim_{x \to +\infty} F(x) = 1$

ii) $F(\cdot)$ è f. monotona non decrescente, cioè

se $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2) \qquad \forall x_1, x_2 \in \mathbb{R}$

iii) $P[X > x] = 1 - P[X \leq x] = 1 - F(x)$

iv) $F(\cdot)$ è f. continua a destra, cioè:

$$\lim_{h \to 0^+} F(x+h) = F(x) \quad , \quad \forall x \in \mathbb{R}$$

v) $P[x_1 < X \leq x_2] = F(x_2) - F(x_1) \quad , \quad \forall x_1, x_2 \in \mathbb{R}, \ x_1 < x_2$

La f. di r. descrive la distribuzione dei valori della v. a. X.

Spesso si preferisce introdurre un'altra funzione detta funzione di densità.

In questo caso bisogna distinguere tra CASO DISCRETO cioè quando i valori della v.a. X sono in numero finito o numerabile e CASO CONTINUO cioè quando i valori della v.a. X appartengono ad $\mathbb{R}$ o un suo intervallo.

<u>DEF.</u> : Dato $(\Omega, \mathcal{A}, P)$ s.p. e X v.a. discreta.

Definiamo FUNZIONE DI DENSITÀ DISCRETA la funzione

$$f : \mathbb{R} \to [0,1]$$

$$f(x) = P[\,X(\omega) = x\,] = P[\,\{\omega \in \Omega : X(\omega) = x\}\,]$$

x è detto PUNTO MASSA, $\in X(\Omega)$.

<u>Esempio</u> : Lancio di 2 dadi dove $X = i + J$.
(equiprobabilità)

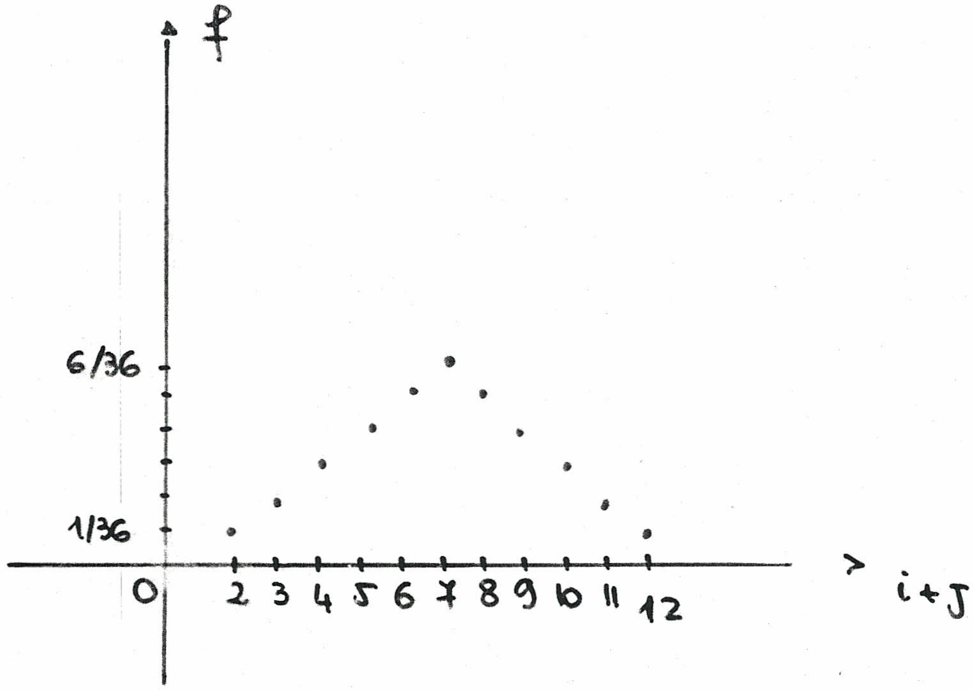
se $x = 2$ $\quad P[X=2] = P[\{(1,1)\}] = \frac{1}{36}$

se $x = 3$ $\quad P[X=3] = P[\{(1,2);(2,1)\}] = \frac{2}{36}$

$\vdots$

se $x = 11$ $\quad P[X=11] = P[\{(5,6);(6,5)\}] = \frac{2}{36}$

se $x = 12$ $\quad P[X=12] = P[\{6,6\}] = \frac{1}{36}$



PROPRIETÀ DI f DISCRETA

i) $f(x_n) \geq 0 \quad , \forall x_n \in X(\Omega)$

ii) $\displaystyle\sum_{x_n \in X(\Omega)}^{!} f(x_n) = 1$

TEOREMA

Dato $(\Omega, \mathcal{A}, P)$ s.p. Assegnato la f. di r. di X v.a. discreta, è possibile calcolare la f. di densità f e viceversa

Infatti

nota $f(\cdot) \Rightarrow F(x) = P(X \leq x)$

$$= \sum_{(J:\, x_J \leq x)}^{!} f(x_J)$$

$\exists$ perchè F è non decrescente

nota $F(\cdot) \Rightarrow f(x_J) = F(x_J) - \lim_{h \to 0^+} F(x_J - h)$

(discontinua)

$= P[X = x_J]$ ampiezza del salto

Esercizio

Data X v.a. discreta che può assumere i valori 1, 2, 3, 4.

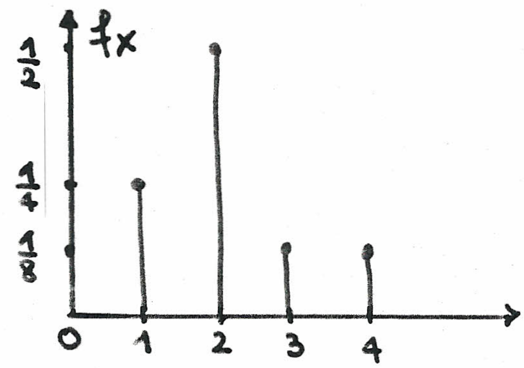
Sapendo che $P(1) = \frac{1}{4}$, $P(2) = \frac{1}{2}$, $P(3) = \frac{1}{8}$ quanto vale $P(4)$?

Disegnare il grafico di f_X.

Calcolare la F_X e disegnarne il grafico.

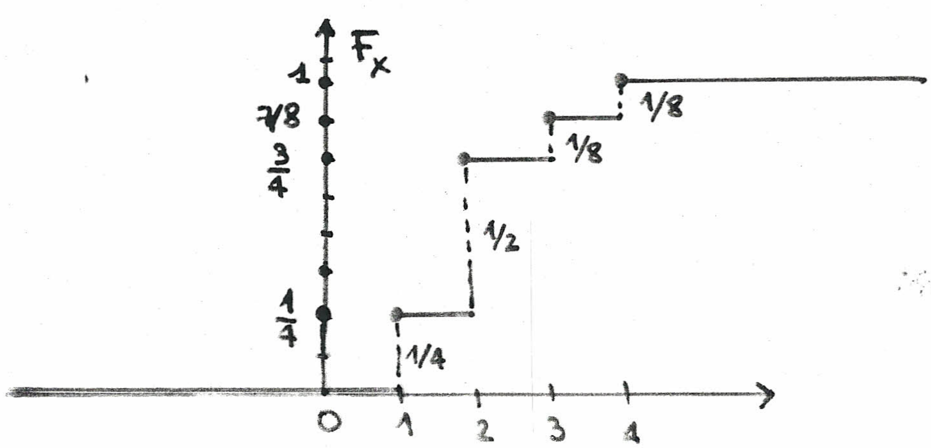
Poichè $\sum_{i=1}^{m} f(x_i) = 1$ e $P(X=x) = f(x)$

$$P(4) = 1 - \frac{1}{4} - \frac{1}{2} - \frac{1}{8} = \frac{1}{8}$$



$$F(x) = \sum_{i:\, x_i \le \bar{x}} f(x_i)$$

$$F(x) = \begin{cases} 0 & -\infty < x < 1 \\ \frac{1}{4} & 1 \le x < 2 \\ 3/4 & 2 \le x < 3 \\ 7/8 & 3 \le x < 4 \\ 1 & 4 \le x < \infty \end{cases}$$

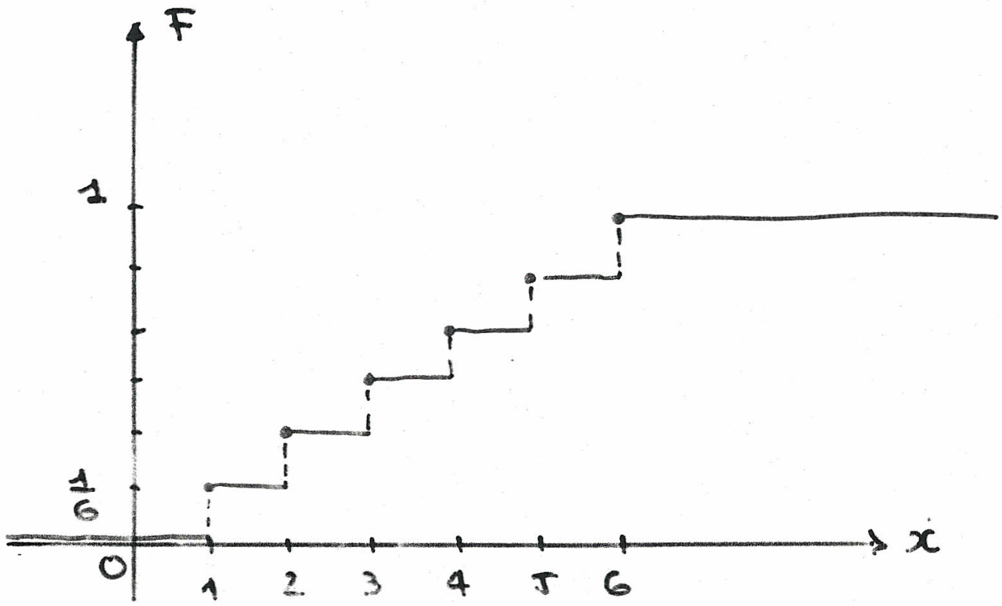


<u>Esempio</u> : Lancio del dado

X = n° faccia in alto

$$f(x) = \frac{1}{6} \qquad x = x_k \qquad k = 1, 2, 3, 4, 5, 6$$

$$F(x) = P[X \le x] = \begin{cases} 0 & x < 1 \\ \frac{1}{6} & 1 \le x < 2 \\ \frac{2}{6} & 2 \le x < 3 \\ \frac{3}{6} & 3 \le x < 4 \\ \frac{4}{6} & 4 \le x < 5 \\ \frac{5}{6} & 5 \le x < 6 \\ 1 & x \ge 6 \end{cases}$$



F è costante tra un punto massa e l'altro ed ha un salto nei punti massa se $x = 2.5$ $\quad F(2.5) = \sum_{(j \,:\, x_j \leq 2.5)}' f(x_j)$

$$= f(1) + f(2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$$

se $x = 3$ $\quad f(3) = F(3) - \lim_{h \to 0^+} F(3-h)$

$$= \frac{3}{6} - \frac{2}{6} = \frac{1}{6}$$

DEF. : Dato $(\Omega, \mathcal{A}, P)$ s.p. e X v.a.

X è detta v.a. continua se esiste

$$f : \mathbb{R} \to [0, +\infty)$$

tale che

$$\forall x \in \mathbb{R} \qquad F(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)\, dt$$

La funzone f è detta FUNZIONE DI DENSITÀ DI PROBABILITÀ di X e la f. di r. $F(\cdot)$ è detta assolutamente continua (può essere scritta come l'integrale della sua derivata!)

PROPRIETÀ DI f CONTINUA

i) $f(x) \geq 0 \quad \forall x \in \mathbb{R}$ in cui è definita $F'(x)$

ii) $\displaystyle\int_{-\infty}^{+\infty} f(t)\,dt = 1$

iii) $\forall a, b \in \mathbb{R}\ ,\ a < b$

$$P[a \leq X(\omega) \leq b] = \int_a^b f(t)\,dt = F(b) - F(a)$$

La definizione di probabilità nel caso continuo presuppone l'esistenza di un'opportuna funzione $f(x)$, il cui integrale sull'intervallo (a,b) fornisce la probabilità che la variabile casuale continua X assuma valori $\in (a,b)$

$\Rightarrow \forall x \in \mathbb{R} \quad P(X=x)=0$ se X v.a. continua.

1. Nel continuo l'espressione "evento di probabilità nulla" non è sinonimo di "evento impossibile", come accade nel discreto.

   Nel continuo è significativo solo calcolare la probabilità che X assuma valori in un dato intervallo: questa è una prima sostanziale differenza tra v. discrete e continue.

2. Se X v.a. continua

$$P[X \leq a] = P[X < a]$$

$$P[X \geq a] = P[X > a]$$

$$P[a<X<b] = P[a \leq X < b] = P[a < X \leq b] = P[a \leq X \leq b]$$

3. La densità $f(x)$ non rappresenta la probabilità $P[X=x]$. (X v.a. continua).

   Infatti $P(X=x]$ è sempre nulla $\forall x \in \mathbb{R}$, mentre $f(x)$ non è dappertutto nulla.

   La funzione $f(x)$ non è una probabilità, è solo il suo integrale su un intervallo che ha il significato di probabilità.

Nel caso discreto invece, la distribuzione di probabilità $f(x_j)$ è per definizione la probabilità $P[X = x_j]$.

Pertanto distribuzioni discrete e densità continue sono oggetti matematici di tipo diverso, non confrontabili tra loro; lo strumento che consente di confrontare v.a. discrete e v.a. continue sono invece le rispettive funzioni di ripartizione.

CASO DISCRETO

$$P[X<a] = P[X \leq a] - P[X=a] = F(a) - f(a)$$

CASO CONTINUO

$$P[X<a] = P[-\infty < X < a] = P[-\infty < x \leq a] = F(a) - \underbrace{F(-\infty)}_{=0} = F(a)$$

Esempio

X v.a. continua : rappresenta la durata di una conversazione telefonica.

Si assume

$$F(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0, \quad \lambda > 0 \end{cases}$$

allora

$$f(x) = F'(x) = \lambda e^{-\lambda x}$$

Se la durata è misurata in minuti

$$P[5 < X < 10] = \int_5^{10} \lambda e^{-\lambda x} dx$$

primitiva $-e^{-\lambda x}$

$$= \left[- e^{-\lambda x} \right]_{x=5}^{x=10}$$

$$= e^{-5\lambda} - e^{-10\lambda}$$

ma anche

$$P[5 < X < 10] = F(10) - F(5)$$

$$= (1 - e^{-10\lambda}) - (1 - e^{-5\lambda})$$

$$= e^{-5\lambda} - e^{-10\lambda}$$

Introduciamo ora degli indicatori numerici che possono misurare le caratteristiche di una v.a. X.

<u>DEF.</u>: Data X v.a., chiamiamo VALORE ATTESO o MEDIA di X il numero μ_x o $E[X]$ definito come:

a) $E[X] = \sum_i x_i f(x_i)$ se X discreta

b) $E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx$ se X continua

In a) $E[X]$ è definito come serie, ammesso che tale serie sia assolutamente convergente; in b) è definito come integrale se esso esiste.

<u>Osservazione</u>:

La <u>media</u> di X è una misura che indica dove sono "centrati" i valori di X ed è quindi detta centro di gravità o baricentro di una densità.

<u>Esempi</u>

1) Lancio di 2 dadi: $X = i + j$

$$E[X] = 2\cdot\frac{1}{36} + 3\cdot\frac{2}{36} + \dots + 11\cdot\frac{2}{36} + 12\cdot\frac{1}{36} = 7$$

2) durata conversazioni telefoniche:

X v.a. continua con

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0,\ \lambda > 0 \\ 0 & x < 0. \end{cases}$$

$$E[X] = \int_{-\infty}^{+\infty} \lambda x\, e^{-\lambda x} dx = -\int_0^{\infty} -\lambda x\, e^{-\lambda x} dx =$$

$$= \cancel{\left| -x\, e^{-\lambda x} \right|_0^{\infty}} + \int_0^{\infty} e^{-\lambda x} dx = \left(-\frac{1}{\lambda} e^{-\lambda x}\right)\Big|_0^{\infty} = \frac{1}{\lambda}.$$

OSSERVAZIONE

Se X v.a. discreta e finita, $X(\Omega)$ ha n elementi $x_1, \dots, x_n$ e la probabilità è uniforme

$$E[X] = \sum_{i=1}^{n} x_i \overset{f(x_i)}{\frac{1}{n}} = \frac{x_1 + \dots + x_n}{n}$$

è detto VALORE MEDIO di X.

OSSERVAZIONI

- $E[X] = \sum_i x_i f(x_i)$ dove la serie è assolutamente convergente

rappresenta la media pesata di tutti i possibili valori che X v.a. discreta può assumere ognuno pesato con la probabilità che X lo assuma $\left(f(x_i) = P(X = x_i) \right)$

ANALOGIA CON IL BARICENTRO

Data X v.a. discreta che può assumere i valori $-1, 0, 1, 2$.

$x_1 = -1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$

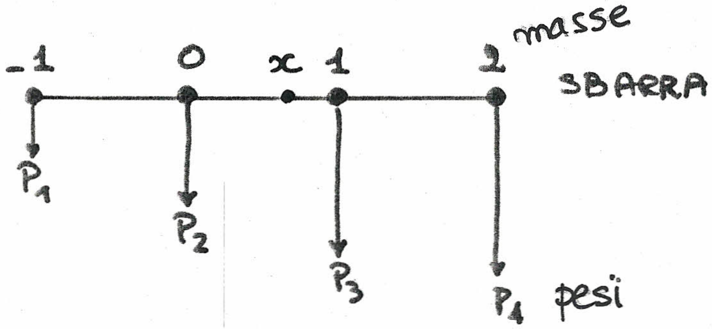
con probabilità

$P(-1) = 0.10$, $P(0) = 0.25$, $P(1) = 0.30$, $P(2) = 0.35$

(N.B. $\sum_i P(X = x_i) = 1$)

$$E[X] = \sum_{i=1}^{4} x_i f(x_i) = -1 \cdot 0{,}1 + 0 \cdot 0{,}25 + 1 \cdot 0{,}3 + 2 \cdot 0{,}35$$

$$= -0{,}1 + 0{,}3 + 0{,}7 = 0.9$$



Il baricentro o centro di gravità è il punto in cui la sbarra rimane in equilibrio.

$\Rightarrow$ $E[X]$ è il punto rispetto al quale è nulla la somma dei momenti dei pesi delle singole masse.

$$\sum_{i=1}^{4} (x_i - E[X])\, p(x_i) = 0$$

$$\Rightarrow (1+x)\cdot 0{,}1 + x\cdot 0{,}25 - (1-x)\cdot 0{,}3 - (2-x)\cdot 0{,}35 = 0$$

$$0{,}1 + x(0{,}1 + 0{,}25 + 0{,}3 + 0{,}35) - 0{,}3 - 0{,}7 = 0$$

$$x - 0{,}9 = 0 \Rightarrow x = 0{,}9.$$

N.B. Ricorda che $\forall a \in \mathbb{R}$ una funzione $f(x)$ è simmetrica rispetto ad a se

$$f(x) = f(2a - x)$$

Sia X v.a.
Se $\exists\, E[X]$ e $f_X(x)$ è simmetrica rispetto ad a

$$\Rightarrow E[X] = a$$

<u>C. CONTINUO</u>

$$E[X] = \int_{-\infty}^{a} x f(x)\, dx + \int_{a}^{+\infty} x f(x)\, dx$$

$y = 2a - x \qquad dy = -dx$

$$= \int_{-\infty}^{a} x f(x)\, dx + \int_{a}^{-\infty} (2a-y) f(2a-y) (-dy)$$

$$= \int_{-\infty}^{a} x f(x)\, dx + \int_{-\infty}^{a} (2a-y) \underbrace{f(2a-y)}_{\overset{\text{SIMM.}}{=} f(y)}\, dy$$

$$= \cancel{\int_{-\infty}^{a} x f(x)\, dx} + 2a \int_{-\infty}^{a} f(y)\, dy - \cancel{\int_{-\infty}^{a} y f(y)\, dy}$$

ma se f è simmetrica rispetto ad a

$$\int_{-\infty}^{a} = \int_{a}^{\infty} = 1/2$$

$$= 2a \cdot \frac{1}{2} = a$$

Sebbene $E[X]$ fornisca la media pesata di tutti i possibili valori di X, non dà informazioni riguardo alla dispersione di questi valori.

Ad esempio se W, Y, Z sono definite.

$W := 0$ con $P = 1$

$$Y := \begin{cases} -1 & \text{con } P = 1/2 \\ 1 & \text{"} \quad \text{"} \end{cases}$$

$$Z := \begin{cases} -100 & \text{con } P = 1/2 \\ +100 & \text{"} \quad \text{"} \end{cases}$$

Tutte hanno media nulla $E[W] = E[Y] = E[Z] = 0$, ma vi è molta più variabilità in Y che non in W (che è addirittura costante) e ancor di più in Z.

Poichè i valori di X sono distribuiti attorno alla sua media $\mu_X = E[X]$, un modo per misurare la loro variabilità potrebbe essere quantificare la loro distanza da μ_X, cioè calcolando p. es. $E[\,|X-\mu_X|\,]$.

Il modulo è scomodo.

Perciò si preferisce $E[(X-\mu_X)^2]$.

Vedremo che $var[X] := E[(X-\mu_X)^2]$ = varianza sarà l'indicatore numerico atto allo scopo.

Perciò

~~Per~~ Per dare ~~[illegible]~~ una misura della dispersione della densità di X in un intorno del suo valore atteso utilizziamo un altro indicatore numerico:

DEF. : Sia X v.a e μ_X la sua media $(\mu_X = E[X])$

Chiamiamo VARIANZA di X il numero reale positivo

$$var[X] = \sigma_X^2 = \sum_i (x_i - \mu_X)^2 f(x_i) \qquad \text{se } X \text{ discreta}$$

$$var[X] = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_x)^2 f(x)\,dx \qquad \text{se } X \text{ continua}$$

DEF. Sia X v.a. con varianza var[X].

Il numero reale positivo

$$\sigma_X = \sqrt{var[X]}$$

è detto DEVIAZIONE STANDARD o SCARTO QUADRATICO MEDIO di X.

Esempi

1) Lancio di 2 dadi con equiprobabilità. $X = i + j$

$$var[X] = (2-7)^2 \cdot \frac{1}{36} + (3-7)^2 \cdot \frac{2}{36} + \dots + (11-7)^2 \cdot \frac{2}{36} +$$

$$+ (12-7)^2 \cdot \frac{1}{36} = \frac{35}{6}$$

2) durata delle telefonate. $f(x) = \begin{cases} 0 & x < 0 \\ \lambda e^{-\lambda x} & x \geq 0, \lambda > 0 \end{cases}$

$$var[X] = \int_{-\infty}^{+\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} dx \qquad E[x] = \frac{1}{\lambda}$$

$$= \int_0^{\infty} \left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} dx$$

$$= \int_0^{\infty} \left(x^2 - \frac{2}{\lambda} x + \frac{1}{\lambda^2}\right) \lambda e^{-\lambda x} dx =$$

$$= -\int_0^{\infty} -\lambda e^{-\lambda x} x^2 dx - 2\int_0^{\infty} x e^{-\lambda x} dx + \frac{1}{\lambda}\left(\frac{-1}{\lambda}\right)\int_0^{\infty} -\lambda e^{-\lambda x} dx$$

$$= -\cancel{\left[e^{-\lambda x} x^2 \right]_0^{\infty}} + \int_0^{\infty} e^{-\lambda x} \cancel{2x}\, dx - 2\int_0^{\infty} x\, \cancel{e^{-\lambda x}} dx +$$

$$+ \frac{1}{\lambda}\left[-\frac{1}{\lambda} e^{-\lambda x}\right]_0^{\infty} = -\frac{1}{\lambda^2}(-1) = \underline{\underline{\frac{1}{\lambda^2}}}.$$

<u>OSSER.</u> : Come la media rappresenta il centro di gravità o baricentro, così la varianza si può interpretare come il momento d'inerzia della densità rispetto ad un asse verticale passante per il baricentro.

DEF.: Data una funzione $g : \mathbb{R} \to \mathbb{R}$.

Se X v.a. $\Rightarrow Y = g(X)$ è v.a.

Chiamiamo VALORE ATTESO di $g(\cdot)$ della v.a. X.

$$E[g(X)] = \sum_i g(x_i)\, f(x_i) \qquad X \text{ discreta}$$

$$E[g(X)] = \int_{-\infty}^{+\infty} g(x)\, f(x)\, dx \qquad X \text{ continua}$$

OSSERVAZIONE

- Se $g(X) = X$

$\Rightarrow E[g(X)] = E[X] = \mu_X$

- Se $g(X) = (X - \mu_X)^2$

$\Rightarrow E[g(X)] = E[(X-\mu_X)^2] = var[X] = \sigma_X^2$

Esempi

1) Sia X v.a. che assume i valori 0, 1, 2 con probabilità

$P[0] = 0.2$, $P[1] = 0.5$ $P[2] = 0.3$

Si consideri $Y = g(X) = X^2$

$\Rightarrow$ Y assume i valori 0, 1, 4.

$E[Y] = 0 \cdot 0.2 + 1 \cdot 0.5 + 4 \cdot 0.3 = 1.7.$

2) Il tempo per trovare un guasto in un impianto elettrico (in ore) è una v.a. X continua con

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{altrove} \end{cases}$$

L'interruzione di x ore provoca un danno economico $Y = g(x) = x^3$.

$$E[Y] = \int_{\cancel{-\infty}\,0}^{\cancel{+\infty}\,1} x^3 \cdot 1 \, dx = \frac{1}{4} x^4 \Big]_0^1 = 0.25.$$

PROPRIETÀ

i) $E[c] = c$ con c costante

$$E[c] = \int_{-\infty}^{+\infty} c\, f(x)\, dx = c \underbrace{\int_{-\infty}^{+\infty} f(x)\, dx}_{=1} = c$$

ii) $E[c\, g(X)] = c\, E[g(X)]$ con c costante.

iii) $E[c_1 g_1(X) + c_2 g_2(X)] = c_1 E[g_1(X)] + c_2 E[g_2(X)]$

con c_1, c_2 costanti

iv) se $\forall x \in \mathbb{R}$ $g_1(X) \leq g_2(X)$

$\Rightarrow E[g_1(X)] \leq E[g_2(X)]$

v) DISUGUAGLIANZA DI MARKOV

Data $g: \mathbb{R} \to \mathbb{R}$: $g(x) \geq 0 \ \forall x \in \mathbb{R} : \exists E[g(x)]$

allora

$$P[g(x) \geq r] \leq \frac{1}{r} E[g(x)] \quad , \forall r > 0$$

Dim: X v.a. continua con f. di densità $f(x)$

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x)\,dx = \int_{\{x : g(x) \geq r\}} g(x) f(x)\,dx + \overbrace{\int_{\{x : g(x) < r\}} g(x) f(x)\,dx}^{\geq 0}$$

$$\geq \int_{\{x : g(x) \geq r\}} g(x) f(x)\,dx \geq \int_{\{x : g(x) \geq r\}} r\, f(x)\,dx \geq r\, P[g(x) \geq r]$$

vi) DISUGUAGLIANZA DI CHEBYSHEV

Dato X v.a. con valore atteso μ_x e varianza σ_x^2 si ha

$$P[|X - \mu_x| \geq \eta\, \sigma_x] \leq \frac{1}{\eta^2} \quad , \forall \eta > 0$$

Dim.:

Posto $g(x) = (x - \mu_x)^2$, $r = \eta^2 \sigma_x^2$

$$P[(X-\mu_x)^2 \geq \eta^2 \sigma_x^2] \leq \frac{1}{\eta^2 \sigma_x^2} \underbrace{E[(X-\mu_x)^2]}_{= \sigma_x^2} = \frac{1}{\eta^2}$$

ma anche

$$P[(X-\mu_x)^2 \geq \eta^2 \sigma_x^2] = P[|X - \mu_x| \geq \eta\, \sigma_x]$$

Tale disuguaglianza pone un vincolo alla probabilità degli eventi.

FORMA EQUIVALENTE

$$P[\mu_x - \eta\sigma_x < X < \mu_x + \eta\sigma_x] = P[|X-\mu_x| < \eta\,\sigma_x] \geq 1 - \frac{1}{\eta^2}$$

Più la varianza è piccola, più è piccola la probabilità che X prenda valori lontani dalla sua media.

Esempio

Sia X v.a. corrispondente al n° di pezzi prodotti da una fabbrica in una settimana con $\mu_x = 50$.

1) calcolare $P[X \geq 75]$

2) Sapendo che $\sigma_x^2 = 25$, calcolare $P[40 < X < 60]$.

1. Applico Markov.

$$P[X \geq 75] \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3} \qquad (g(x) = x)$$

2. $$P[40 < X < 60] = P[-10 < X - 50 < 10]$$

$$= P[|X-50| < 10]$$

$$= 1 - P[|X-50| \geq 10]$$

ma

$$P[|X-50| \geq 10] = P[(X-50)^2 \geq 100] \leq \frac{E[(X-50)^2]}{100} =$$

$$= \frac{\sigma_x^2}{100} = \frac{25}{100} = \frac{1}{4}$$

$$\Rightarrow P[40 < X < 60] \geq \frac{3}{4}. \quad (\text{Chebyshev})$$

N.B.: la funzione di densità $f(x)$ di X non è nota.

PROPRIETÀ DI VAR[X]

i) $\underline{var[X] = E[(X - E[X])^2]} = E[(X-\mu_X)^2]$

segue dalla definizione di varianza e di $E[g(X)]$.

ii) $\underline{var[X] = E[X^2] - (E[X])^2} = E[X^2] - \mu_X^2$

$$E[(X - E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2]$$
$$= E[X^2] - 2E[X]\cdot E[X] + (E[X])^2$$
$$= E[X^2] - (E[X])^2$$

iii) $\underline{var[aX] = a^2 var[X]}$, $\forall a \in \mathbb{R}$

$$var[aX] = E[(aX - E[aX])^2] \doteq$$

poichè $E[aX] = aE[X]$, $\forall a \in \mathbb{R}$

$$\doteq E[(aX - aE[X])^2] = E[a^2(X - E[X])^2]$$
$$= a^2 E[(X - E[X])^2] = a^2 var[X].$$

iv) $var[a+X] = var[X]$, $\forall a \in \mathbb{R}$

$$var[a+X] = E[((a+X) - E[a+X])^2]$$
$$= E[(\cancel{a} + X - \underbrace{E[a]}_{=\cancel{a}} - E[X])^2] = E[(X - E[X])^2]$$
$$= var[X].$$

N.B. per la iii)

$$var[X+X] = 4\, var[X].$$

Oltre al valore atteso e alla varianza, altre quantità possono misurare le caratteristiche di una v.a.: sono i valori attesi di potenze della v.a. detti momenti.

<u>DEF.</u>: Sia X v.a con media $\mu_X = E[X]$.
Chiamiamo (se esiste) MOMENTO DI ORDINE n ($n \in \mathbb{N}$) di X la quantità:

$$\mu'_n = E[X^n]$$

N.B. $\mu'_1 = E[X] = \mu_X$.

<u>DEF.</u>: Sia X v.a. con media μ_X.
Chiamiamo MOMENTO CENTRALE DI ORDINE n di X la quantità:

$$\mu_n = E[(X-\mu_X)^n]$$

N.B. $\mu_1 = E[(X-\mu_X)] = E[X] - \mu_X = 0$

$$\mu_2 = E[(X-\mu_X)^2] = \text{var}[X]$$

<u>OSSERVAZIONE</u>: Tutti i momenti centrali dispari di X sono 0 se la funzione di densità di X è simmetrica rispetto a μ_X.
La quantità $\check{X} = X - \mu_X$ è detta v.a. centrata.

Data X v.a. e f_X simmetrica rispetto ad un valore reale $\boxed{a = \mu_x}$.

$\Rightarrow$ se $\exists\ \mu_3$ (momento centrale di ordine 3)

$$\boxed{\mu_3 = 0}$$

C. CONTINUO

$$\mu_3 = \int_{-\infty}^{+\infty} (x-\mu_x)^3 f(x)\,dx = \int_{-\infty}^{a} (x-\mu_x)^3 f(x)\,dx +$$

$$\int_{a}^{+\infty} (x-\mu_x)^3 f(x)\,dx = \int_{-\infty}^{a} (x-\mu_x)^3 f(x)\,dx +$$

$$\int_{-\infty}^{a} (2a-y-\mu_x)^3 f(y)\,dy =$$

$x = 2a-y$ e $f(y) = f(2a-y)$, $a = \mu_x$

$$= \int_{-\infty}^{a} (x-\mu_x)^3 f(x)\,dx + \underbrace{\int_{-\infty}^{a} (2\mu_x - y - \mu_x)^3 f(y)\,dy}_{= -\int_{-\infty}^{a} (y-\mu)^3 f(y)\,dy} \equiv 0$$

$\mu_x = E[x] = a$

In generale pertanto

$$\mu_{2n+1} = 0, \quad n \in \mathbb{N}$$

<u>OSS</u>: Il momento centrale μ_3 è un indice di asimmetria della curva $f(x)$; il momento centrale μ_4 è un indice di appiattimento della curva $f(x)$ attorno al suo punto di massimo.

In statistica i primi due momenti sono di grande importanza.

Non sapendo in generale quale sia la distribuzione con cui si lavora, è necessario almeno conoscere la posizione e avere una qualche idea sulla sua dispersione.

Perciò è utile trovare una funzione capace di darci una rappresentazione di tutti i momenti.

<u>DEF.</u>: Data X v.a., definiamo $\forall\, t \in I = [-h, h], h>$

FUNZIONE GENERATRICE DEI MOMENTI di X

la quantità

$$m(t) = E[e^{tX}]$$

se tale valore è finito.

N.B.: E' il valore atteso di e^{tX}.

$m(t) = E[e^{tx}] = \sum_x e^{tx} f(x)$ se X è discreta

$m(t) = E[e^{tx}] = \int_{-\infty}^{+\infty} e^{tx} f(x)\, dx$ se X è continua

Se esiste $m(t)$ essa è derivabile con continuità in un intorno dell'origine.

Se deriviamo $m(t)$ rispetto a t :

$$\frac{d}{dt} m(t) = \int_{-\infty}^{+\infty} x\, e^{tx} f(x)\, dx$$

⋮ k volte

$$\frac{d^k}{dt^k} m(t) = \int_{-\infty}^{+\infty} x^k e^{tx} f(x)\, dx$$

$$\underbrace{\lim_{t\to 0} \frac{d^k}{dt^k} m(t)}_{= m^{(k)}(0)} = \int_{-\infty}^{+\infty} x^k f(x)\, dx = E[X^k] = \mu'_k$$

<u>Esempio</u> : Durata telefonate X. v.a.

$$f(x) = \begin{cases} 0 & x<0 \\ \lambda e^{-\lambda x} & x \geq 0, \quad \lambda > 0 \end{cases}$$

$$m(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \int_0^{\infty} \lambda e^{tx} e^{-\lambda x} dx$$

$$= \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx = \lambda \left[-\frac{1}{(\lambda - t)} e^{-(\lambda-t)x} \right]_0^{\infty}$$

$$= \frac{\lambda}{\lambda - t}, \quad |t| < \lambda.$$

$$m'(t) = \frac{d}{dt} m(t) = \frac{\lambda}{(\lambda - t)^2}$$

$$\underline{m'(0)} = \frac{1}{\lambda} = E[X] = \mu_1'$$

$$m''(t) = \frac{d^2}{dt^2} m(t) = \frac{2\lambda}{(\lambda - t)^3}$$

$$\underline{m''(0)} = \frac{2}{\lambda^2} = E[X^2] = \mu_2'$$

Άρα

$$var[X] = \overset{\mu_2'}{E[X^2]} - \overset{(\mu_1')^2}{(E[X])^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

$$var[X] = m''(0) - [m'(0)]^2$$

ESEMPIO

X v.a. continua :

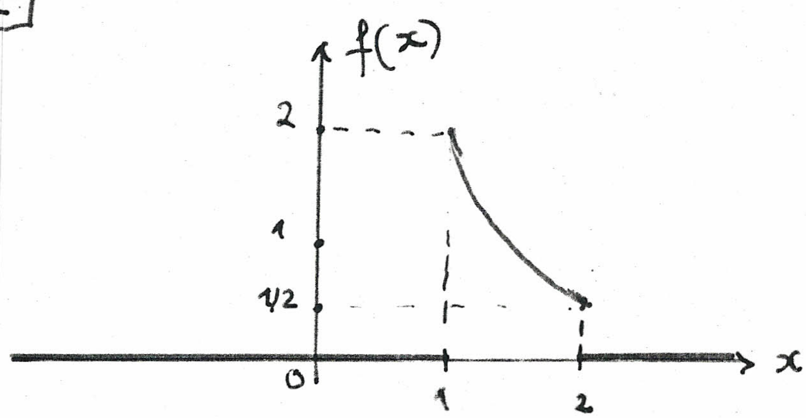
$$f(x) = \begin{cases} \frac{c}{x^2} & 1 \leq x \leq 2 \\ 0 & \text{altrove} \end{cases}$$

1. determinare $c \in \mathbb{R}$: f sia una densità e disegnare il grafico di $f(x)$;
2. determinare $F(x)$ e grafico;
3. calcolare $P(1.5 < X < 2)$
4. calcolare μ_x, σ_x^2.

1) $\forall x \in \mathbb{R}$ $f(x) \geq 0 \Rightarrow \underline{c \geq 0}$

$$1 = \int_{-\infty}^{+\infty} f(x)\,dx = \int_1^2 \frac{c}{x^2}\,dx = c\left[-\frac{1}{x}\right]_1^2 = c\left(-\frac{1}{2}+1\right) = \frac{c}{2}$$

$\Rightarrow \boxed{c = 2}$



2) $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

per $x < 1$ $F(x) = 0$

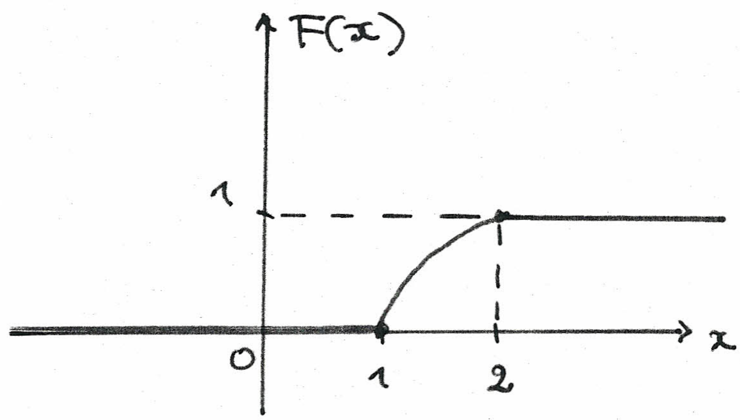
per $1 \leq x \leq 2$ $F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{1} f(t)\,dt}_{0} + \int_1^x f(t)\,dt$

$$= \int_1^x \frac{2}{t^2}\,dt = 2\left[-\frac{1}{t}\right]_1^x = 2\left(-\frac{1}{x}+1\right) = 2 - \frac{2}{x}$$

per $x \geq 2$ $\quad F(x) = \int_{-\infty}^{x} f(t)\,dt = \underbrace{\int_{-\infty}^{1} f(t)\,dt}_{=0} + \underbrace{\int_1^2 f(t)\,dt}_{=1} + \underbrace{\int_2^x f(t)\,dt}_{=0} = 1$

Riassumendo

$$F(x) = \begin{cases} 0 & x < 1 \\ 2 - \frac{2}{x} & 1 \leq x \leq 2 \\ 1 & x \geq 2 \end{cases}$$



3) $P(1.5 < X < 2) = \int_{\frac{3}{2}}^{2} \frac{2}{x^2}\,dx = 2\left(-\frac{1}{x}\right)_{\frac{3}{2}}^{2} = 2\left(-\frac{1}{2}+\frac{2}{3}\right) = \frac{1}{3}$

<u>oppure</u> $\quad = F(2) - F\left(\frac{3}{3}\right) = 1 - \frac{2}{3} = \frac{1}{3}$

4) $\mu_x = \int_1^2 x \cdot \frac{2}{x^2}\,dx = 2\int_1^2 \frac{1}{x}\,dx = 2\ln x \Big|_1^2 = 2\ln 2 = \ln 4$

$\sigma_x^2 = E[X^2] - \mu_x^2$ $\quad$ FORMULA UTILE •

$$E[X^2] = \int_1^2 x^2 \frac{2}{x^2}\, dx = \int_1^2 2\, dx = 2x\Big|_1^2 = 2$$

$$\sigma_X^2 = 2 - (\ln 4)^2$$

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x) f(x)\, dx$$ FORMULA GENERALE

oppure

$$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 f(x)\, dx$$

$$= \int_1^2 (x - \ln 4)^2 \cdot \frac{2}{x^2}\, dx$$