Probabilità condizionale, indipendenza

$(\Omega, \mathcal{A}, P)$ sp. di probabilità.

Def: $A, B \in \mathcal{A}$ con $P(B) > 0$.

Si chiama probabilità condizionale di A rispetto a B il numero:

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

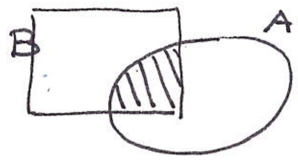
La prob. condiz. è la prob. che A si verifichi, sapendo che B si è verificato.

Se, ad es, $B \subset A$ allora $P(A \cap B) = P(B)$ e $P(A \mid B) = 1$ il che è logico perché se $B \subset A$ vuol dire che $B \Rightarrow A$, ovvero se B si è verificato, certamente (con prob. 1) si verifica A.

Risulta $P(A \cap B) = P(A \mid B)\, P(B)$, che vale anche se $P(B) = 0$.

Esempio 1

Si giocano alla roulette i numeri 3, 13, 22. I possibili risultati sono 37 (i numeri da 0 a 36).
Consideriamo la distrib. uniforme (equiprobabilità sullo spazio degli eventi). La probabilità di vincere è $\frac{3}{37}$.
Supponiamo di sapere che il gioco è truccato in modo da far uscire un n° dispari (evento B). Qual è la prob. di vincere?



Sia $A = \{3, 13, 22\}$, $B = \{1, 3, 5, 7, \dots 35\}$

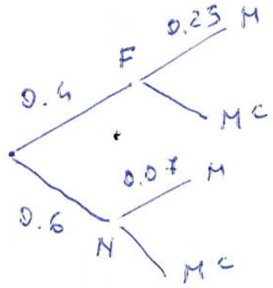
Il numero dei casi favorevoli è 2 (3, 13 dispari, mentre 22 è pari); il n° dei casi possibili è 18.

Dunque $P(A \mid B) = \frac{2}{18} = \frac{1}{9} = \frac{P(A \cap B)}{P(B)} = \frac{P(\{3, 13\})}{P(B)} =$

$= \frac{2}{37} \Big/ \frac{18}{37} = \frac{2}{18}$.

Esempio 2

Una popolazione si compone per il 40% di fumatori (F) e per il 60% di non fumatori (N). Si sa che il 25% dei fumatori e il 7% dei non fumatori sono affetti da una certa malattia respiratoria cronica (M). Qual è la prob. che un individuo scelto a caso sia affetto dalla malattia?



Si ha: $P(F) = 0.4$; $P(N) = 0.6$

$P(M|F) = 0.25$; $P(M|N) = 0.07$

Dunque: (FORMULA PROB. TOTALI)

$$P(M) = P(M \cap F) + P(M \cap N) = P(M|F) \cdot P(F) + P(M|N) \cdot P(N)$$

$$= 0.25 \cdot 0.4 + 0.07 \cdot 0.6 = \frac{25}{100} \cdot \frac{4}{10} + \frac{7}{100} \cdot \frac{6}{10} = \frac{1}{10} + \frac{42}{1000}$$

$$= 0.142.$$

FORMULA DI BAYES

Siano $A_1, \dots A_n$ eventi disgiunti ed esaustivi $(A_1 \cup A_2 \cup \dots \cup A_n) = \Omega$. allora:

$$\boxed{P(A_i|B) = \frac{P(A_i)\,P(B|A_i)}{P(B)} = \frac{P(A_i)\,P(B|A_i)}{\sum_{j=1}^{n} P(A_k)\,P(B|A_k)}}$$

Infatti:

$P(A_i \cap B) = P(A_i)P(B|A_i)$ e inoltre

$B = (A_1 \cap B) \cup (A_2 \cap B) \cup \dots (A_n \cap B)$ unione disgiunta $\Rightarrow$

$$P(B) = \sum_{j=1}^{n} P(A_j \cap B) = \sum_{j=1}^{n} P(A_j)\,P(B|A_j)$$

La formula di Bayes esprime $P(A_i|B)$ in funzione di $P(B|A_i)$

ESERCIZIO: $P(A \cap B \cap C) = P(C) \cdot P(B|C) \cdot P(A|\{B \cap C\})$

Esempio 3

Con riferimento all'esempio 2, qual è la probabilità che una persona affetta dalla malattia respiratoria sia un fumatore? Per la f. di Bayes applicata alla partizione $\Omega = F \cup N$ si ha:

$$P(F|M) = \frac{P(F)P(M|F)}{P(M)} = \frac{0.4 \cdot 0.25}{0.142} = 0.704$$

Esempio 4 (Es. 1.7)

Due giocatori di tiro al piattello sparano allo stesso bersaglio. Si sa che il primo tiratore ~~colpisce in media~~ spara in media 9 colpi durante lo stesso tempo in cui il secondo giocatore ne spara 10. La precisione dei due tiratori non è la stessa: su 10 colpi sparati dal tiratore 1, 8 colpiscono il bersaglio, dal tiratore 2 solo 7.

Durante il gioco, il bersaglio è stato colpito da un colpo, ma non si sa quale tiratore lo abbia sparato. Qual è la probabilità che il bersaglio sia stato sparato dal giocatore n° 2?

Sia A_1 = evento che un colpo è sparato dal giocatore 1

A_2 = " " 2

Possiamo supporre, tenendo conto dei dati, che

$P(A_1) = \frac{9}{10} P(A_2)$. Sia B = evento che il bersaglio sia stato colpito.

Abbiamo ancora $P(B|A_1) = 0.8$; $P(B|A_2) = 0.7$

Dalla formula di Bayes:

$$P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} = \frac{0.7 \cdot P(A_2)}{0.9 \cdot P(A_2) \cdot 0.8 + 0.7 \cdot P(A_2)}$$

$$= 0.493$$

N.B. $P(A_2)$ è incognito, ma il problema si può risolvere ugualmente.

OSS: $P(A_1|B) + P(A_2|B) = 1$ ..

INDIPENDENZA STOCASTICA

Def: $A, B \in \mathcal{A}$ sono indipendenti se

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

Si dice, inoltre, che $A_1, \dots A_n \in \mathcal{A}$ sono a due a due indip. se

$$P(A_i \cap A_j) = P(A_i) \cdot P(A_j) \quad \forall\, i \neq j.$$

$A_1, \dots A_n$ sono una famiglia di eventi indip. se $\forall\, k \leq n$, per ogni scelta di indici $i_1, \dots i_k$:

$$P(A_{i_1} \cap A_{i_2} \cap \dots A_{i_k}) = P(A_{i_1})\, P(A_{i_2}) \cdots P(A_{i_k})$$

Es:

$\Omega = \{1, 2, 3, 4\}$ con la distrib. uniforme:

$P(\{1\}) = P(\{2\}) = P(\{3\}) = P(\{4\}) = \frac{1}{4}$.

$A_1 = \{1, 4\}$ $\quad A_2 = \{2, 4\}$ $\quad A_3 = \{3, 4\}$

Risulta: A_1, A_2, A_3 sono indip a due a due, ma:

$$P(A_1 \cap A_2 \cap A_3) = P(\{4\}) = \frac{1}{4} \neq \frac{1}{8} = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

Se A e B sono indipendenti, dalla def. segue:

$$P(A|B) = P(A) \qquad (P(B) > 0)$$

D'altra parte: $P(B|A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{P(A) \cdot P(B)}{P(A)} = P(B)$

Dunque A, B indip $\Rightarrow \begin{cases} P(A|B) = P(A) \\ \quad e \\ P(B|A) = P(B) \end{cases}$

Ovvero: sapere se B (A) si è verificato o meno non aggiunge alcuna informazione sulla previsione che si verifichi A (B).

A, B indip $\Rightarrow A^c, B^c$ indip; A, B^c indip; A^c, B indip.

Esempio (Lancio di una moneta)

Il lancio di una moneta dà testa con prob. p, $0 \le p \le 1$ e croce con prob. $1-p$. La moneta viene lanciata n volte. Qual è la prob. di ottenere testa nei primi k lanci (e croce negli altri $n-k$)?

Possiamo supporre che i lanci siano indipendenti l'uno dell'altro. Sia $\Omega = \{\omega = (\omega_1, \dots \omega_n),\ \omega_i = 0, \text{ oppure } 1\}$ (1 se viene testa, 0 croce). Se $p = \frac{1}{2}$, allora

$$P(\{\omega\}) = \frac{1}{2^n} = \frac{1}{\#\Omega}.$$

Se $p \neq \frac{1}{2}$, poniamo per $i = 1, \dots n$: $A_i = \{\omega \mid \omega_i = 1\}$ (evento che il lancio i-esimo dia ~~croce~~ testa); dunque $P(A_i) = p$.

Se A = evento di cui cerchiamo la prob, risulta

$$A = A_1 \cap \dots \cap A_k \cap A_{k+1}^c \cap \dots \cap A_n^c$$

[redacted]

Allora, per l'indip. degli eventi:

$$P(A) = P(A_1) \cdot \dots \cdot P(A_k) \cdot P(A_{k+1}^c) \cdots P(A_n^c) =$$

$$= \underbrace{p \cdot p \cdot \dots \cdot p}_{k \text{ volte}} \cdot \underbrace{(1-p) \cdot \dots \cdot (1-p)}_{n-k \text{ volte}}$$

$$= p^k (1-p)^{n-k}$$

[redacted]

+ Esempio 2 estraz. senza rimpiazzo
B_1, B_2 dipendenti

ESERCIZIO 1

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

Infatti:

$$(1) = \frac{P(A\cap B\cap C)}{P(C)}$$

$$(2) = \frac{P(\cancel{A\cap C})}{P(C)}\cdot\frac{P(B\cap A\cap C)}{P(\cancel{A\cap C})} = \frac{P(B\cap A\cap C)}{P(C)}$$

ESERCIZIO 2

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

Infatti:

$$\overset{(2)}{\cancel{P(A\cap B\cap C)}} = \frac{P(A\cap B\cap C)}{P(\cancel{B\cap C})}\cdot\frac{P(\cancel{B\cap C})}{\cancel{P(C)}}\cdot\cancel{P(C)} = (1)$$

Esempio

Un certo Test del sangue è efficace al 99% nell'individuare una certa malattia. Si possono verificare però dei "falsi positivi" con probabilità p (ovvero una persona sana che si sottopone al test, con prob. p risulta erroneamente malata).

Se l'incidenza della malattia sulla popolazione è dello 0.5%, qual è la prob. che un soggetto sia malato, condizionata al fatto che il Test sia positivo?

$M = \{\text{soggetto è malato}\}$, $E = \{\text{il Test è positivo}\}$

$$P(M|E) \overset{\text{Bayes}}{=} \frac{P(E|M)\,P(M)}{P(E|M)\,P(M) + P(E|M^c)\,P(M^c)}$$

$$\frac{0.99 \cdot 0.005}{0.99 \cdot 0.005 + p \cdot 0.995} \doteq$$

Se $p = 0.001 \Rightarrow P(M|E) = 0.8326$

$p = 0.0001 \Rightarrow \quad = 0.9802$

$p = 0.005 \Rightarrow \quad = 0.4987$

$p = 0.01 \Rightarrow \quad = 0.3322$