



TEOREMA DI BAYES

nella diagnostica
di laboratorio

Il *teorema di Bayes* lega la misura di probabilità condizionata di un evento, detta “a posteriori”, alla misura di probabilità dello stesso evento, detta “a priori”.

Dati due eventi A e B

$P(A/B)$ viene detta probabilità “a posteriori” poichè permette di calcolare la probabilità di A, sapendo che si è verificato o si verificherà B.

$$P(B/A) \neq P(A/B)$$

probabilità di B condizionata da A \neq probabilità di B condizionata da A

$P(A)$ e $P(B)$ sono probabilità a priori

$A \cap B = B \cap A$ $P(A \cap B) = P(B \cap A)$ ma:

$$P(A \cap B) = P(A) P(B|A)$$

$$P(B \cap A) = P(B) P(A|B)$$

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

Il *teorema di Bayes* è detto anche teorema della probabilità delle cause, e si può applicare in moltissimi campi giuridico, economico e sociale.

In quello medico è importante valutare la probabilità che la causa di un dato reale e oggettivo (l'esito del test) sia proprio la presenza della malattia

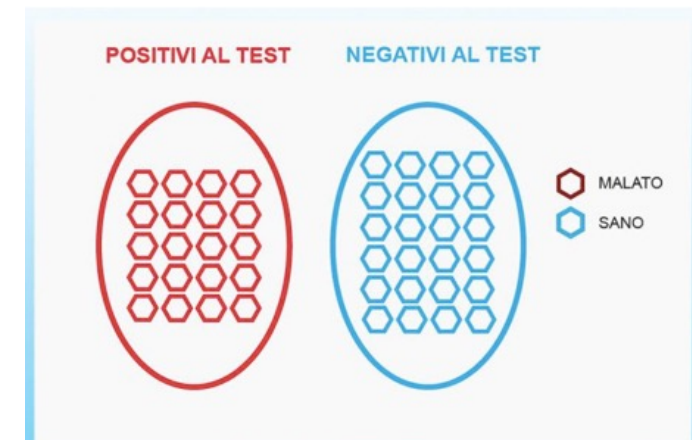
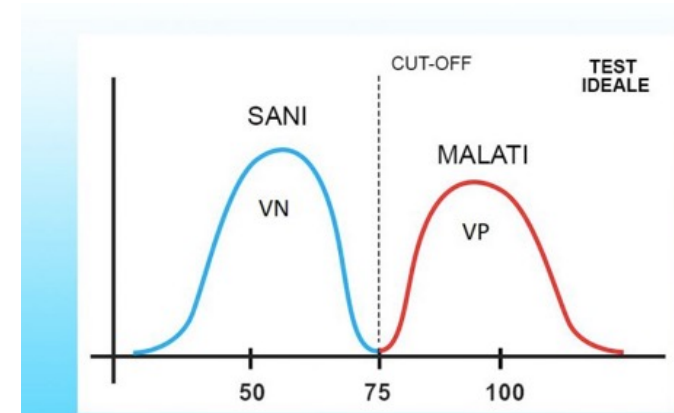
La diagnostica medica è di tipo sperimentale pertanto i test diagnostici reali hanno un margine di errore non trascurabile, risultando positivi per alcuni soggetti sani (falsi positivi) e negativi per alcuni soggetti malati (falsi negativi), come si vede nella figure successive.

Situazione ideale:

- ***pazienti malati sono veri positivi VP***
- ***pazienti sani sono veri negativi VN.***

Non abbiamo né falsi positivi FP (pazienti in realtà sani ma che risultano positivi al test), né falsi negativi FN (pazienti malati che risultano negativi al test).

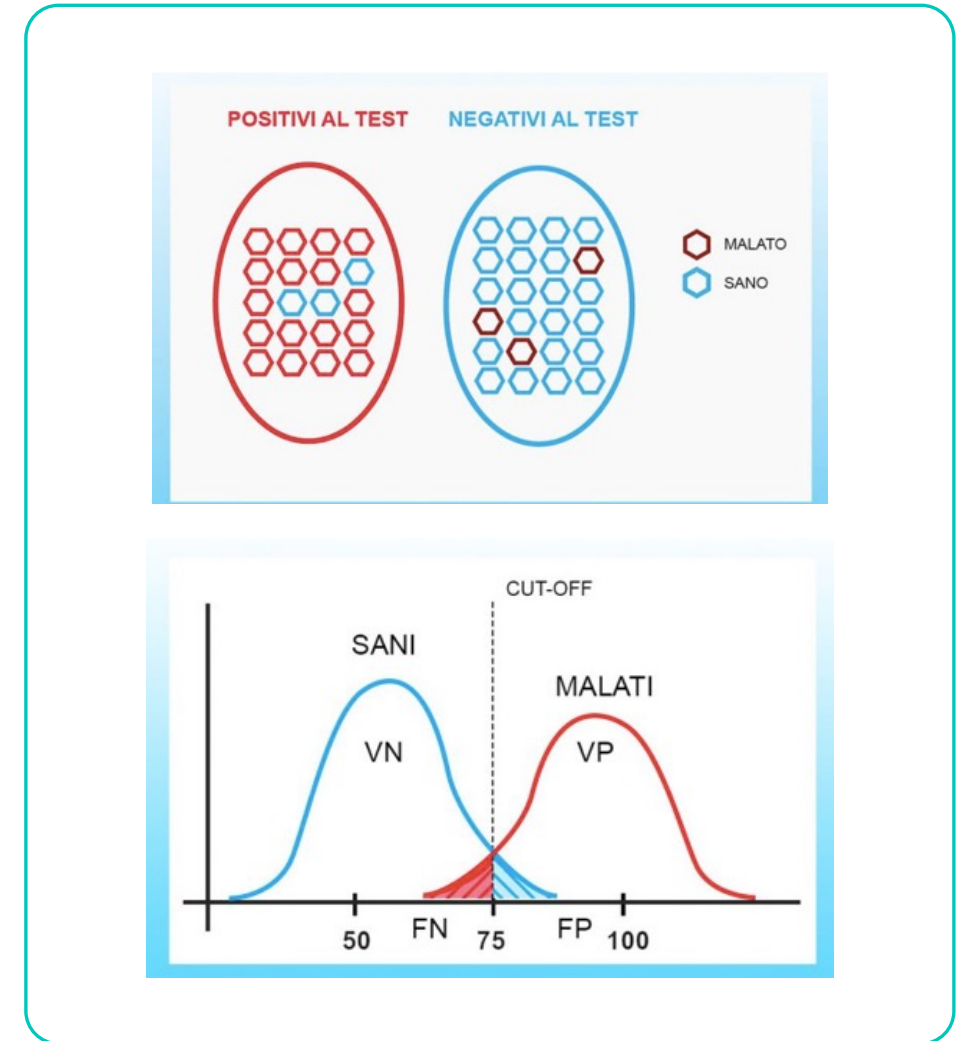
Dopo aver scelto un opportuno valore di cut-off, supponiamo di sottoporre 44 persone a un test, di cui 20 sono malate e 24 sane. Il test ideale le classifica tutte correttamente, come rappresentato in figura



Rappresentazione più verosimile della realtà è la seguente: -

- pazienti sani in cui il test dà un valore maggiore del cut-off scelto, e vengono classificati falsamente come positivi **FP**,
- pazienti malati in cui il test dà un valore maggiore del cut-off, classificati dunque falsamente come negativi **FN**.

Notiamo che scegliendo un valore di cut-off maggiore, quindi spostando a destra la linea tratteggiata, si avranno meno falsi positivi. Viceversa, scegliendo un cut-off minore, spostando la linea a sinistra, si otterranno meno falsi negativi.



Diamo ora tre definizioni molto importanti per misurare l'affidabilità di un test diagnostico:

ACCURATEZZA = $(VP+ VN)/TOT$ PAZIENTI misura la percentuale di diagnosi corretta, ovvero quanti pazienti vengono classificati correttamente, ma non dice nulla su come il test sbaglia (FP e FN).

SENSIBILITÀ = VP/TOT MALATI è la proporzione di malati che vengono classificati in maniera corretta. Se il test è molto sensibile riuscirà a individuare anche basse concentrazioni di virus, dando meno risultati *falsi negativi*. In altri termini, la sensibilità è la probabilità che un individuo malato risulti positivo al test, $P(+ | M)$

Un test è sensibile al 100% quando tutti i malati risultano positivi, quindi nessun malato sfuggerà al test.

SPECIFICITÀ = VN/TOT SANI è la proporzione di sani classificati correttamente. Un test altamente specifico riesce a non confondere il vero obiettivo, il virus, con altre forme simili, e darà quindi meno esiti *falsi positivi*. In altri termini, la specificità è la probabilità che un individuo sano risulti negativo al test, $P(- | S)^2$.

Un test è specifico al 100% quando tutti i sani risultano negativi, quindi nessun paziente sano avrà una diagnosi errata.



The infographic is set within a rounded rectangular frame with a teal border. On the right side, there is a cartoon illustration of a man with glasses and a dark blue shirt, holding a black pointer stick. In the top right corner, there is a circular logo with the word 'MED' and a white arrow pointing to the right. The background is light blue.

SENSIBILITÀ

Probabilità che il test sia positivo se un individuo è malato. Capacità di classificare correttamente i soggetti malati.

ALTA SENSIBILITÀ: pochi falsi negativi.

SPECIFICITÀ

Probabilità che il test sia negativo se un individuo è sano. Capacità di classificare correttamente i soggetti sani.

ALTA SPECIFICITÀ: pochi falsi positivi.

La dott.ssa Giulia sospetta che il proprio paziente Mario abbia una certa malattia. Per validare la sua ipotesi lo sottopone a un test diagnostico efficace al 99%, cioè in grado di rilevare la presenza (o l'assenza) della malattia nel 99% dei casi.

Sapendo che solo l'1% degli abitanti del paese di Mario hanno quella determinata malattia, qual è la probabilità che, se il test dovesse dare esito positivo, Mario sia effettivamente malato?



In altri termini, quanto è affidabile il risultato del test?

Cerchiamo quindi di formalizzare il problema, vedendo una semplice applicazione del teorema di Bayes

In termini probabilistici: noi conosciamo $P(+|M)$, la sensibilità, e vogliamo trovare $P(M|+)$, il valore predittivo del test. Con $P(M)$ indichiamo la prevalenza della malattia, ovvero quanto è diffusa la malattia nella popolazione, con $P(S)$ la probabilità di essere sani, con $P(+)$ la probabilità che il test risulti positivo, e applichiamo il teorema di Bayes:

$$P(M|+) = \frac{P(+|M)P(M)}{P(+)} = \frac{P(+|M)P(M)}{P(+|M)P(M) + P(+|S)P(S)} = \frac{0.99 * 0.01}{0.99 * 0.01 + 0.01 * 0.99} = 0.5 = 50\%$$

50% valore predittivo positivo del test è un risultato piuttosto scadente nonostante il test abbia sensibilità e specificità elevate. La probabilità che un individuo positivo sia malato non è soddisfacente, perché questo valore dipende in larga misura anche da quanto è diffusa la malattia nella popolazione. Per rispondere alla domanda iniziale si potrebbe procedere in due modi: o si utilizza un test con performance migliori, o si ripete il test più volte

Perché fare due tamponi per capire se si è guariti dal Coronavirus?

Fino a non molto tempo fa, per essere sicuri che un paziente affetto da COVID-19 fosse guarito, si effettuavano 2 tamponi a distanza di 24 ore l'uno dall'altro. Se l'esito di entrambi era negativo allora si poteva affermare con una certa sicurezza che il paziente fosse guarito

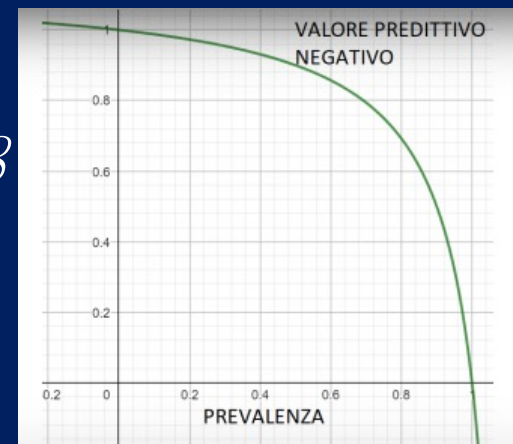
$$P(S|-) = \frac{0.95 * 0.75}{0.95 * 0.75 + 0.05 * 0.25} = 0.98 = 98\%$$

Supponiamo che il test diagnostico utilizzato abbia specificità e sensibilità del 95% e calcoliamo, $P(S|-)$, il *valore predittivo negativo* del test, assumendo una prevalenza del 25%. Quindi il paziente ha probabilità del 98% di essere guarito.

Ripetiamo ora il test, applicando di nuovo il teorema di Bayes, aggiornando opportunamente la probabilità che il paziente sia sano. La probabilità è salita ben al 99,8%! Inoltre, come possiamo intuire, il valore predittivo negativo in funzione della prevalenza è decrescente, quindi meno è diffusa la malattia più il valore predittivo si avvicina al 100%.

$$P(S|-) = \frac{P(-|S)P(S|-)}{P(-|S)P(S|-)+P(-|M)P(1-S|-)} = \frac{0,95*0,98}{0,95*0,98+0,05*0,02} = 0,998 = 99,8\%$$

La probabilità è salita a 99,8%



$P(-|M)$ e $P(-|S)$ sono, rispettivamente, le probabilità che il test risulti positivo sapendo che il paziente è malato (sensibilità), e che il test risulti negativo sapendo che paziente è sano (specificità). Si tratta di probabilità condizionate perché sappiamo che l'evento a destra della barretta verticale si è verificato

ECOGRAFIA

Si considerino i seguenti eventi relativi alla nascita di un bambino:

M il nascituro è maschio

F il nascituro è femmina

EM l'ecografia prevede "maschio"

EF l'ecografia prevede "femmina"



Si consideri dapprima la probabilità che due eventi si verifichino entrambi (*probabilità congiunta*):

$$P(M,EM)$$

Valgono le seguenti espressioni:

$$P(M,EM) = P(M | EM) P(EM)$$

$$P(M,EM) = P(EM | M) P(M)$$

Teorema di Bayes

Quindi:

$$P(M | EM) P(EM) = P(EM | M) P(M)$$

ossia

$$P(M | EM) = \frac{P(EM | M) P(M)}{P(EM)}$$

ESEMPIO

Supponiamo

$$P(M) = 0.5$$

$$P(F) = 0.5$$

$$P(EM | M) = 0.9$$

$$P(EM | F) = 0.05$$

- e di conseguenza

$$P(EF | M) = 0.1$$

$$P(EF | F) = 0.95$$

$$P(EM) = P(EM | M) P(M) + P(EM | F) P(F) = 0.9 \cdot 0.5 + 0.05 \cdot 0.5 = 0.475$$

$$P(EF) = 1 - P(EM) = 0.525$$

Possiamo ora applicare la formula di Bayes

$$P(M | EM) = \frac{P(EM | M) P(M)}{P(EM)} = \frac{0,9 \cdot 0,5}{0,475} = 0,947$$

$$P(F | EF) = \frac{P(EF | F) P(F)}{P(EF)} = \frac{0,9 * 0,5}{0,525} = 0,904$$

GRAZIE



I. T. "G. DELEDDA" Lecce

I T DELEDDA LECCE

CAPIROLA BENEDETTA

CESARIO BEATRICE

DAINI GIOELE IACOPO

LORENZO MATTEO

PETRACHI GIULIA

RESCIO GIORGIA