6b. Esempi di applicazione del teorema di Bayes nella diagnostica di laboratorio

Un esempio di come il teorema di Bayes in realtà possa essere applicato in modo intuitivo, è il seguente.

PROBLEMA 7
Si consideri un test destinato a rivelare la presenza nel siero di anticorpi anti-HIV. Si assuma che questo test abbia una sensibilità del 100% (il test, quindi, è positivo nel 100% dei malati). Si assuma che questo test abbia una specificità del 99.7% (il test, quindi, è negativo nel 99.7% dei soggetti sani). Si sa che la prevalenza dell'infezione da HIV è del 3 per mille (nella popolazione, su 1000 soggetti presi a caso, 3 sono infetti da virus).
Quale è il valore predittivo del test positivo?

Supponendo di effettuare il test su 1000 soggetti presi a caso i risultati saranno i seguenti: 3 soggetti presenteranno positività agli anticorpi anti-HIV (veri positivi), in quanto la prevalenza è del 3 per mille e la sensibilità del test del 100% ci assicura di individuare tutti i soggetti infetti; inoltre effettuando l'analisi su 1000 soggetti, a causa del fatto che la specificità del test è del 99.7% ci dovremo aspettare 3 positivi su 1000 soggetti sani (falsi positivi). Essendo in totale 6 i soggetti positivi, e 3 i veri positivi, avremo quindi un valore predittivo del test positivo pari a 3/6, cioè un valore predittivo del test positivo pari al 50% (p pari a 0,5).

Applicando la (6.4) avremo il calcolo espresso in dettaglio mediante il teorema di Bayes:

$$P(M+|T+) = \frac{\textit{sensibilità} \cdot \textit{prevalenza}}{\textit{sensibilità} \cdot \textit{prevalenza} + (1 - \textit{specificità}) \cdot (1 - \textit{prevalenza})} \quad \text{ovvero}$$

$$P(M+|T+) = \frac{1 \cdot 0{,}003}{(1 \cdot 0{,}003) + (0{,}003 \cdot 0{,}997)} = 0{,}501^{1} = \textit{valore predittivo del test positivo}$$

L'esempio riportato si presta a una prima importante considerazione.

Il teorema di Bayes rappresenta l'unico strumento che consente di fornire una misura quantitativa, e quindi oggettiva, del valore aggiunto fornito da un test diagnostico (per esempio un'analisi di laboratorio)[2].

Nel caso degli anticorpi anti-HIV, la differenza tra la probabilità di essere malati a posteriori (dopo avere effettuato il test, pari al 50%) e la probabilità di essere malati a priori (prima di avere effettuato il test, pari al 3 per mille) rappresenta appunto <u>il valore aggiunto</u> che il test è in grado di fornire, <u>in termini di informazione</u>, alla diagnosi clinica.

La seconda considerazione parte dal fatto che la determinazione degli anticorpi anti-HIV, un test di

[1] Il valore esatto è 0.500751126690035.
[2] Come diceva Lord Kelvin (1824-1907): "*Quando puoi misurare ciò di cui stai parlando, ed esprimerlo in numeri, puoi affermare di saperne qualcosa; se però non puoi misurarlo, se non puoi esprimerlo con numeri, la tua conoscenza sarà povera cosa e insoddisfacente: forse un inizio di conoscenza, ma non abbastanza da far progredire il tuo pensiero fino allo stadio di scienza, qualsiasi possa essere l'argomento.*"

primo livello poco costoso, consente di restringere da 1000 a 6 soli individui la rosa dei candidati ad essere sottoposti a un test di secondo livello (western-blot). Che è sì risolutivo dal punto di vista diagnostico, ma che ha un costo molto superiore a quello degli anticorpi. Eseguire la western-blot a 6 individui su mille è ragionevole, mentre eseguirla a 1000 su 1000 sarebbe uno spreco di risorse.

L'analisi bayesiana è la sola che possa fornire gli indicatori necessari per effettuare una valutazione oggettiva del rapporto costi/benefici di una strategia diagnostica, e quindi di valutarne la razionalità nei confronti di strategie diagnostiche alternative.

Dal rapporto tra costi (in termini sia economici sia di diagnosi non corrette) e benefici (sia in termini economici sia di diagnosi corrette) dei diversi percorsi diagnostici si può arrivare al consenso degli operatori su una strategia diagnostica (anche se la praticabilità in sé della strategia diagnostica continuerà ovviamente a dipendere dal contesto economico e culturale, quindi sostanzialmente dai vincoli economici e dai vincoli etici che alla strategia sono attribuiti).
La terza considerazione parte dall'osservazione che un test con sensibilità del 100% e specificità del 100% è un test che ha per definizione un valore predittivo del test positivo del 100%. Nel caso esemplificato degli anticorpi anti-HIV le caratteristiche del test erano pressoché ideali (sensibilità del 100% e specificità del 99,7%). Ma nonostante questo il valore predittivo del test positivo risultava del 50% "solamente".

In condizioni di bassa prevalenza riduzioni anche minime della specificità di un test possano comportare drastiche riduzioni del valore predittivo del test positivo.

Questo non avviene per il valore predittivo del test negativo, che risulterà uguale a 994/994 (il test è negativo in 994 sani su 994) cioè uguale al 100% (o se si preferisce uguale a 1): un test negativo consente nel caso specifico di escludere la malattia.

Applicando la (6.6) avremo il calcolo espresso in dettaglio mediante il teorema di Bayes:

$$P(M\text{-}|T\text{-}) = \frac{specificità \cdot (1 - prevalenza)}{specificità \cdot (1 - prevalenza) + (1 - sensibilità) \cdot prevalenza} \quad \text{ovvero}$$

$$P(M\text{-}|T\text{-}) = \frac{0{,}997 \cdot 0{,}997}{(0{,}997 \cdot 0{,}997) + (0 \cdot 0{,}003)} = 1{,}00 = \textit{valore predittivo del test negativo}$$

In condizioni di bassa prevalenza associate a bassa specificità di un test, aumenta il valore predittivo del test negativo.

Vediamo ora due altri esempi di applicazione del teorema di Bayes.

PROBLEMA 8
La probabilità che una persona con più di cinquant’anni senza sintomi abbia un cancro colorettale è dello 0,3%. Se una persona ha un cancro colorettale, c’è una probabilità del 50% che abbia il test del sangue occulto nelle feci positivo; se non ha un cancro colorettale, c’è una probabilità del 3% che abbia comunque il test del sangue occulto nelle feci positivo. Immaginate una persona sopra i cinquant’anni, asintomatica, sottoposta a screening e con il test del sangue occulto nelle feci positivo.
Quale è la probabilità che abbia veramente un cancro colorettale?

Analizziamo innanzitutto le tre affermazioni:
- "*la probabilità che una persona con più di cinquant'anni senza sintomi abbia un cancro del colorettale è dello 0,3%*", significa che la prevalenza della malattia è del 3 per mille.
- "*se una persona ha un cancro colorettale, c'è una probabilità del 50% che abbia il test del sangue occulto nelle feci positivo*", significa che il test ha un sensibilità del 50%;
- "*se non ha un cancro colorettale, c'è una probabilità del 3% che abbia comunque il test del sangue occulto nelle feci positivo*", significa che il test ha una specificità del 97%;

Per calcolare il valore predittivo del test positivo applichiamo la (6.4)

$$P(M+|T+) = \frac{0{,}50 \cdot 0{,}003}{(0{,}50 \cdot 0{,}003) + (0{,}03 \cdot 0{,}997)} = 0{,}048 = \textit{valore predittivo del test positivo}$$

e per calcolare il valore predittivo del test negativo applichiamo la (6.6)

$$P(M-|T-) = \frac{0{,}97 \cdot 0{,}997}{(0{,}97 \cdot 0{,}997) + (0{,}5 \cdot 0{,}003)} = 0{,}998 = \textit{valore predittivo del test negativo}$$

In altre parole:
- un risultato positivo del test comporta una probabilità di essere malato del 4,8%, quindi su 1000 persone con il test positivo saranno 48 i veri positivi e 952 i falsi positivi;
- un risultato negativo del test comporta una probabilità di non essere malato del 99,8%, quindi su 1000 persone con il test negativo saranno 998 i veri negativi e 2 i falsi negativi.

PROBLEMA 9
Un marcatore tumorale ha le seguenti caratteristiche: (i) è positivo in 95 su 100 pazienti con il cancro; (ii) è negativo in 95 su 100 pazienti senza in cancro; (iii) in media, 5 persone su una popolazione di 1000 hanno un cancro non ancora diagnosticato del tipo che il marcatore tumorale in questione rileva.
Se il test è prescritto a un paziente selezionato casualmente in questa popolazione e l'esito è positivo, quale è la probabilità che il paziente abbia realmente il cancro?

Analizziamo innanzitutto le tre affermazioni:
- "*(i) è positivo in 95 su 100 pazienti con il cancro*", significa che il test ha un sensibilità del 95%;
- "*(ii) è negativo in 95 su 100 pazienti senza in cancro*", significa che il test ha una specificità del 95%;
- "*(iii) in media, 5 persone su una popolazione di 1000 hanno un cancro non ancora diagnosticato del tipo che il marcatore tumorale in questione rileva*", significa che la prevalenza della malattia è del 5 per mille.

Per calcolare il valore predittivo del test positivo applichiamo la (6.4)

$$P(M+|T+) = \frac{0{,}95 \cdot 0{,}005}{(0{,}95 \cdot 0{,}005) + (0{,}05 \cdot 0{,}995)} = 0{,}087 = \textit{valore predittivo del test positivo}$$

e per calcolare il valore predittivo del test negativo applichiamo la (6.6)

$$P(M\text{-}|T\text{-}) = \frac{0{,}95 \cdot 0{,}995}{(0{,}95 \cdot 0{,}995) + (0{,}05 \cdot 0{,}005)} = 0{,}9997 = \textit{valore predittivo del test negativo}$$

In altre parole:
- un risultato positivo del test comporta una probabilità di essere malato del 8,7%, quindi su 1000 persone con il test positivo saranno 87 i veri positivi e 914 i falsi positivi;
- un risultato negativo del test comporta una probabilità di non essere malato del 99,97%, quindi su 1000 persone con il test negativo saranno 1000 i veri negativi e 0 i falsi negativi.

L'analisi del problema 8 e del problema 9 conferma il fatto che il valore predittivo del test positivo è fortemente penalizzato da condizioni di bassa prevalenza della malattia e/o diminuzione della specificità. Questi argomenti sono discussi in maggior dettaglio nella parte relativa a teorema e strategie diagnostiche.