8a. Teorema di Bayes, rapporto di verosimiglianza (likelihood ratio, LR) e odds

Se partiamo dalla forma classica (5.13) del teorema di Bayes

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

che possiamo anche riscrivere come

$$P(A|B) = \frac{P(B|A)}{P(B)} \cdot P(A) \qquad (8.1)$$

e se consideriamo il rapporto

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

globalmente come un *fattore* di moltiplicazione, la (8.1) si può semplificare riscrivendola come

$$P(A|B) = fattore \cdot P(A) \qquad (8.2)$$

che può essere letta come

$$P(\text{della causa}|\text{dato l'effetto}) = \ fattore \cdot P(\text{della causa}) \qquad (8.3)$$

ovvero come

$$P(\text{a posteriori}) = \ fattore \cdot P(\text{a priori}) \qquad (8.4)$$

Questa ultima espressione è particolarmente interessante in relazione all'uso, diffuso nel mondo anglosassone, di utilizzare in luogo della probabilità P gli *odds*, che sono la probabilità espressa dal bookmaker come rapporto tra la vincita e la posta giocata.

La relazione che intercorre tra *odds* e *probabilità* è

$$odds = P / (1 - P) \qquad (8.5)$$

e inversamente

$$probabilità = odds / (1 + odds) \qquad (8.6)$$

PROBLEMA 12
Trasformare in odds la probabilità dell'80% cioè P = 0,8[1]

[1] *odds* = 0,8 / (1 – 0,8) = 0,8 / (0,2) = 4 cioè = *4:1*

PROBLEMA 13
Trasformare in probabilità gli odds 4:1[2]

La (8.1) può essere riformulata in termini di odds come

$$O(A|B) = \Lambda(A|B) \cdot O(A) \qquad (8.7)$$

La grandezza $\Lambda(A|B)$ si chiama rapporto di verosimiglianza[3] e, dall'inglese likelihood-ratio, viene indicato sovente come LR, e pertanto la (8.7) può essere letta in analogia con la (8.3) come

$$\textit{odds(della causa|dato l'effetto)} = \textit{rapporto di verosimiglianza} \cdot \textit{odds(della causa)} \qquad (8.8)$$

ovvero in analogia con la (8.4) come

$$\textit{odds(a posteriori)} = \textit{rapporto di verosimiglianza} \cdot \textit{odds(a priori)} \qquad (8.9)$$

ove il rapporto di verosimiglianza è la misura della verosimiglianza che collega la probabilità a priori alla probabilità a posteriori.

Nel caso di un test di laboratorio, il rapporto di verosimiglianza è per definizione il rapporto tra la probabilità che si verifichi lo specifico risultato in un individuo che ha la malattia e la probabilità che si verifichi lo specifico risultato in una persona che non ha la malattia:
- un test con rapporto di verosimiglianza di 1,0 non fornisce alcuna informazione, e i sui risultati non influiscono sulla probabilità post-test di malattia (che rimane identica alla probabilità pre-test della malattia);
- un rapporto di verosimiglianza > 1,0 incrementa la probabilità post-test di malattia; più grande è il rapporto di verosimiglianza, maggiore è l'informazione fornita da un risultato positivo del test (maggiore è la probabilità di essere malato dopo avere eseguito il test);
- un rapporto di verosimiglianza < 1,0 diminuisce la probabilità post-test di malattia; più piccolo è il rapporto di verosimiglianza, maggiore è l'informazione fornita da un un risultato negativo del test (minore è la probabilità di essere malato dopo avere eseguito il test).

In base alla definizioni date di odds e di rapporto di verosimiglianza (semplicemente sostituendo ad A la malattia M e a B il test T), possiamo formulare dalla (8.7) il valore predittivo di un test positivo ovvero la probabilità di essere malato per un soggetto con un test positivo

$$O(M+|T+) = \Lambda(M+|T+) \cdot O(M+) \qquad (8.10)$$

nella quale il rapporto di verosimiglianza per un test positivo è espresso come

$$LR+ = \Lambda(M+|T+) = \frac{P(T+|M+)}{P(T+|M-)} \qquad (8.11)$$

[2] probabilità = 4 / (1 + 4) = 0,8

[3] altrimenti detto *fattore di Bayes*

PROBLEMA 14
Il rapporto di verosimiglianza LR+ (LR per un test positivo) è uguale al rapporto tra il risultato del test nei malati e nei sani.
Esprimere LR+ in termini di sensibilità e di specificità[4].

Possiamo formulare dalla (8.7) anche il valore predittivo di un test negativo ovvero la probabilità di essere malato per un soggetto con un test negativo

$$O(M+|T-) = \Lambda(M+|T-) \cdot O(M+) \qquad (8.12)$$

nella quale il rapporto di verosimiglianza per un negativo è espresso come

$$LR- = \Lambda(M+|T-) = \frac{P(T-|M+)}{P(T-|M-)} \qquad (8.13)$$

PROBLEMA 15
Il rapporto di verosimiglianza LR- (LR per un test negativo) è uguale al rapporto tra il risultato del test nei malati e nei sani.
Esprimere LR- in termini di sensibilità e di specificità[5].

Combinando la (8.10) con la (8.11) abbiamo che in definitiva in termini di odds il valore predittivo di un test positivo è

$$O(M+|T+) = \frac{P(T+|M+)}{P(T+|M-)} \cdot O(M+) \qquad (8.14)$$

e combinando la (8.12) con la (8.13) abbiamo che in termini di odds il valore predittivo di un test negativo è

$$O(M+|T-) = \frac{P(T-|M+)}{P(T-|M-)} \cdot O(M+) \qquad (8.15)$$

PROBLEMA 16
Un test del sangue occulto nelle feci per la diagnosi del cancro colorettale ha le seguenti caratteristiche: sensibilità = 50% (0,50), specificità = 97% (0,97). La prevalenza della malattia è uguale allo 0,3% (0,003).
Calcolare valore predittivo del test positivo utilizzando la (6.4) e confrontarlo con quello ottenuto applicando la (8.14). Calcolare valore predittivo del test negativo utilizzando la (6.6) e confrontarlo con quello ottenuto applicando la (8.15).

Vediamo cosa accade per il valore predittivo del test positivo. Applicando la (6.4) avremo il calcolo espresso in dettaglio mediante il teorema di Bayes:

[4] Dalle definizioni della tabella (6.1) si ricava immediatamente che LR+ = *sensibilità / (1 – specificità).*
[5] Dalle definizioni della tabella (6.1) si ricava immediatamente che LR- = *(1 – sensibilità) / specificità.*

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

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

Applicando la (8.14) avremo:

$$O(M+) = P(M+) / (1 - (P(M+)) = 0,003 / (1 - 0,003) = 0,003009027$$

$$LR+ = sensibilità / (1 - specificità) = 0,50 / (1 - 0,97) = 16,6667$$

quindi

$$O(M+|T+) = 0,003009027 \cdot 16,6667 = 0,05015$$

e riconvertendo gli odds in probabilità avremo

$$P = odds / (1 + odds) = 0,05015 / (1 + 0,05015) = 0,048 = \text{valore predittivo del test positivo}$$

Quindi le due soluzioni sono, come atteso, numericamente identiche.

Vediamo cosa accade per il valore predittivo del test negativo. Applicando la (6.6) avremo il calcolo espresso in dettaglio mediate il teorema di Bayes:

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

$$P(M-|T-) = \frac{0,97 \cdot 0,997}{(0,97 \cdot 0,997) + (0,50 \cdot 0,003)} = 0,998 = \text{valore predittivo del test negativo}$$

Applicando la (8.15) avremo:

$$O(M+) = P(M+) / (1 - P(M+)) = 0,003 / (1 - 0,003) = 0,003009027$$

$$LR- = (1 - sensibilità) / specificità = (1 - 0,50) / 0,97 = 0,515464$$

quindi

$$O(M+|T-) = 0,003009027 \cdot 0,515464 = 0,001551$$

e riconvertendo gli odds in probabilità avremo

$$P = odds / (1 + odds) = 0{,}001551 / (1 + 0{,}001551) = 0{,}002 = \text{valore predittivo del test negativo}$$

Quindi le due soluzioni sono, come atteso, identiche tranne che per il fatto che, come probabilmente sarà stato notato, con la (6.6) il valore predittivo del test negativo è la probabilità di essere sano per un soggetto con un test negativo, mentre con gli odds (8.15) il valore predittivo del test negativo sono gli odds di essere malato per un soggetto con un test negativo (ovvero, ritrasformando gli odds in probabilità, la probabilità di essere malato per un soggetto con un test negativo). Le differenze nell'espressione dei risultati del test sono riassunte nella seguente tabella:

Formulazione del teorema	***Valore predittivo del test positivo***	***Valore predittivo del test negativo***
In termini di probabilità	P(M+|T+) probabilità di essere malato per un soggetto con un test positivo	P(M-|T-) probabilità di essere sano per un soggetto con un test negativo
In termini di odds e likelihood ratio	O(M+|T+) odds di essere malato per un soggetto con un test positivo	O(M+|T-) odds di essere malato per un soggetto con un test negativo

Anche se l'approccio in termini di odds, con l'utilizzo del rapporto di verosimiglianza, é basato su una espressione della probabilità da bookmaker, e piuttosto lontana dalla mentalità di chi non è abituato alle scommesse, va notato che:
- gli odds sono immediatamente riconvertibili in probabilità, e questo consente di superare facilmente il problema (come vedremo appena avanti con il diagramma di Fagan);
- l'approccio ha il vantaggio di esprimere i risultati del test positivo e del test negativo sotto forma della stessa grandezza (odds di essere malato), semplificando almeno da questo punto di vista la comprensione delle conclusioni;
- gli odds forniscono un modo di utilizzo dell'approccio bayesiano alternativo allo sviluppo analitico delle formule del teorema (uno sviluppo peraltro oggigiorno reso banale dalla possibilità di implementarne le formule anche sul tabellone elettronico di un dispositivo palmare).

Per chi opera in condizioni tali per cui è impensabile utilizzare uno strumento di calcolo elettronico, il nomogramma di Fagan semplifica i calcoli: è sufficiente collegare con una retta il valore della probabilità a priori, sulla sinistra, con il valore di LR+ (o il valore di LR-), al centro, per ottenere, nel punto di intersezione con la retta di destra, la probabilità a posteriori, ovvero la probabilità a priori aggiornata sulla base della verosimiglianza che caratterizza il test. Da notare che, proprio per superare il problema della espressione in odds, nel nomogramma di Fagan le grandezze sia in ingresso (pre-test) sia in uscita (post-test) sono espresse in termini di probabilità.