

7b. Quattro aspetti del teorema di Bayes

Quattro aspetti del teorema di Bayes si rivelano della massima importanza per la sua comprensione:

- il collegamento che esiste tra l'espressione del teorema di Bayes in termini di probabilità e l'espressione in termini di numero dei casi osservati;
- .- la risposta all'obiezione che in una classificazione dicotomica si perde il senso del risultato numerico;
- le strategie per la scelta della soglia che discrimina tra sani e malati;
- gli effetti della specificità e delle prevalenza della malattia sul valore predittivo del test.

Per illustrare il primo aspetto, ovvero il collegamento che esiste tra l'espressione del teorema di Bayes in termini di probabilità e l'espressione in termini di numero dei casi osservati, è necessario riprendere la tabella di classificazione (6.1)

	T+	T-
M+	$P(T+M+)$ [sensibilità]	$P(T-M+)$ [1 – sensibilità]
M-	$P(T+M-)$ [1 – specificità]	$P(T-M-)$ [specificità]

e ricordare che $P(M+) = \text{prevalenza [della malattia]}$ e $P(M-) = 1 - \text{prevalenza}$.

Si passa dalle probabilità al numero di casi osservati moltiplicando le grandezze della prima riga per $P(M+)$ e le grandezze della seconda riga per $P(M-)$, ottenendo la seguente tabella

	T+	T-
M+	$P(T+M+) \cdot P(M+)$ [sensibilità · prevalenza] [VP]	$P(T-M+) \cdot P(M+)$ [(1 – sensibilità) · prevalenza] [FN]
M-	$P(T+M-) \cdot P(M-)$ [(1 – specificità) · (1 – prevalenza)] [FP]	$P(T-M-) \cdot P(M-)$ [(specificità) · (1 – prevalenza)] [VN]

(7.1)

che mette in evidenza le corrispondenze che ci permettono di ricalcolare tutte le grandezze in termini di numero dei casi osservati, che sono riassunte nella tabella (7.2) che segue:

Grandezza	Calcolata come...
<i>sensibilità</i> (positività del test nei malati)	VP / (VP+FN)
<i>specificità</i> (negatività del test nei sani)	VN / (VN+FP)
<i>prevalenza</i> (numero dei malati)	(VP+FN) / (VP+FN+FP+VN)
<i>valore predittivo del test positivo</i> (probabilità di essere malato per un soggetto con il test positivo)	(VP / (VP+FP))
<i>valore predittivo del test negativo</i> (probabilità di essere sano per un soggetto con il test negativo)	VN / (VN+FN)

(7.2)

Il seguente esempio tratto da Galen e Gambino [1] si riferisce ai risultati della determinazione dell'alfa-fetoproteina nel cancro del fegato (malati, M+) e in altri disordini (sani, M-)

	T+	T-
M+	90	17
M-	39	2079

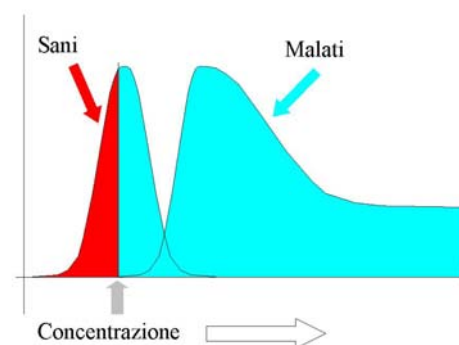
dai quali è possibile calcolare mediante le definizioni fornite nella (7.2):

- sensibilità = $VP / (VP+FN) = 90 / (90+17) = 0,841$
- specificità = $VN / (VN+FP) = 2079 / (2079+39) = 0,982$
- prevalenza = $(VP+FN) / (VP+FN+FP+VN) = (90+17) / (90+17+39+2079) = 0,048$
- valore predittivo del test positivo = $VP / (VP+FP) = 90 / (90+39) = 0,698$
- valore predittivo del test negativo = $VN / (VN+FN) = 2079 / (2079+17) = 0,992$

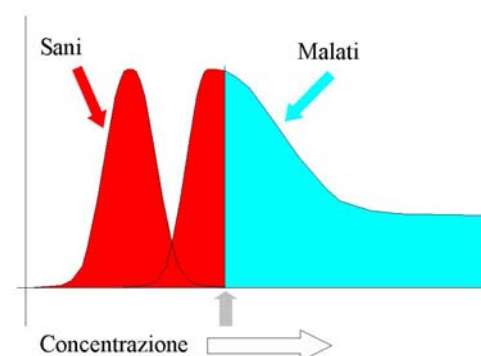
L'obiezione che in una classificazione dicotomica si perde il senso del risultato numerico assume più o meno questa forma: “Perfetto, ma avendo fissato un valore soglia di 40 U/L per la ALT, per la diagnosi di epatite sapere che, a parità di valore predittivo del test positivo, il mio paziente ha 80 U/L o 800 U/L di ALT è molto diverso!” Questa obiezione è valida solo apparentemente: infatti è possibile in ogni momento cambiare la soglia scegliendo quella più opportuna in relazione all'obiettivo clinico.

Le strategie per la scelta della soglia che discrimina tra sani e malati hanno una notevole rilevanza, e sono discusse ampiamente da Galen e Gambino [2].

La massima sensibilità (idealmente del 100%) è richiesta quando (i) la malattia è grave e non si vuole correre il rischio di non diagnosticarla, (ii) la malattia è curabile, e (iii) i falsi-positivi non comportano gravi danni psicologici e/o economici.



Un esempio è rappresentato dal feocromocitoma. La malattia può essere fatale se misconosciuta. Se diagnosticata è curabile al 100%. In caso di positività è possibile ripetere il test e/o confermarla con test di secondo livello. Ma lo stesso vale per l'epatite C, per la quale vale il principio per cui non si vuole correre il rischio di non diagnosticarla, la ricerca degli anticorpi anti-epatite C può dare dei falsi positivi, ma esistono test di secondo livello (ricerca del RNA del virus dell'epatite C nel sangue con tecniche di biologia molecolare).

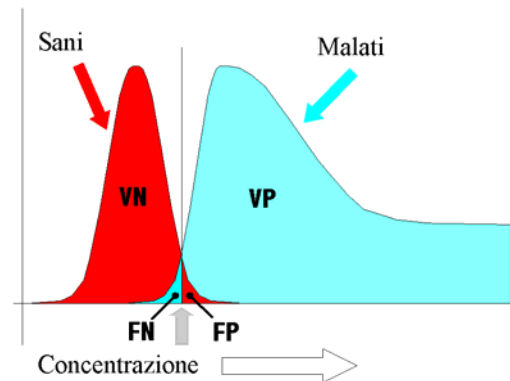


La massima specificità (idealmente del 100%) è richiesta quando (i) la malattia è grave ma non è curabile, (ii) il sapere che la malattia è esclusa ha valore psicologico e/o per la salute pubblica, e (iii) i falsi-positivi possono comportare gravi danni psicologici e/o economici.

Un esempio è rappresentato dalla corea di Huntington.

La malattia non è curabile. Se un caso non viene diagnosticato, l'evoluzione della malattia porterà il paziente di nuovo all'attenzione del medico. D'altra parte un falso positivo porterebbe a danni rilevanti al paziente.

La massima efficienza è richiesta quando (i) la malattia è grave ma curabile, inoltre i falsi-positivi e i falsi-negativi sono parimenti gravi e/o dannosi. Un esempio è rappresentato dall'infarto del miocardio. Se un caso viene trascurato, ne può derivare un danno. Un danno grave può derivare anche da una diagnosi di infarto in assenza di questo. La massima efficienza rappresenta la scelta più oculata.



Infine il massimo valore predittivo è essenziale quando il trattamento di un falso positivo potrebbe avere conseguenze oltremodo gravi. Un esempio è rappresentato dal cancro del polmone. Un test per questa malattia dovrebbe avere un valore predittivo (del test positivo) dal 100%, in quanto i solo trattamenti conosciuti sono la lobectomia e la radioterapia. Ma se tali terapie fossero eseguite su un paziente senza il cancro, le conseguenze sarebbero oltremodo gravi.

Gli effetti della specificità e delle prevalenza della malattia sul valore predittivo del test possono essere illustrati con il seguente esempio.

PROBLEMA 10

E' stato messo a punto un nuovo test per la diagnosi prenatale, su liquido amniotico, di una malattia genetica, per la quale è nota una prevalenza di 3 su mille neonati. Il test ha una sensibilità del 100%. Per la specificità del test, che è pari al 99,5%, sono riportati anche gli intervalli di confidenza al 95%, che rappresentano i limiti all'interno dei quali si colloca l'incertezza della stima: essi vanno da 99,0 a 99,9.

Quale è il grado di incertezza che può essere attribuito al valore predittivo del test positivo?

La risposta può essere ricavata dalla seguente tabella, nella quale il valore predittivo del test positivo è stato calcolato anche per i due valori di specificità (99,95 e 99,0%) che rappresentano i limiti superiore e inferiore all'interno dei quali si colloca l'incertezza con cui è stata ottenuta la misura della specificità.

<i>Sensibilità</i>	<i>Specificità</i>	<i>Prevalenza</i>	<i>Valore predittivo del test positivo</i>	<i>Valore predittivo del test negativo</i>
1,000	1,000	0,003	1,000	1,000
1,000	0,999	0,003	0,751	1,000
1,000	0,990	0,003	0,231	1,000
1,000	0,900	0,003	0,029	1,000

Come si vede se, a fronte di una specificità stimata del 99,5%, la specificità "vera" fosse del 99,9% il valore predittivo del test positivo sarebbe del 75,1%, mentre se la specificità "vera" fosse del 99,0% il valore predittivo del test positivo sarebbe del 23,1%. Ora dire che un test di diagnosi prenatale positivo ha un valore predittivo del 75% o che ha un valore predittivo del 23% sono due cose ben diverse.

PROBLEMA 11

E' stato messo a punto un nuovo test per la diagnosi dell'artrite reumatoide. Il test viene provato in via preliminare una clinica universitaria, presso la quale la metà circa dei pazienti è affetta da artrite reumatoide. Per questo motivo vi è la rara opportunità di provarlo su ben 500 pazienti affetti da artrite reumatoide e su 500 pazienti affetti da altre patologie reumatologiche diverse dall'artrite reumatoide. Il test dimostra di avere una sensibilità del 100% e una specificità del 99%.

Il test viene immesso sul mercato includendo nel bugiardino i dati di sensibilità e di specificità, che sono in assoluto i migliori per un test per la diagnosi di artrite reumatoide. Dopo l'immissione in commercio, il test si diffonde rapidamente proprio per queste sue notevoli caratteristiche. Tuttavia l'utilizzo del test da parte dei medici di medicina generale dimostra risultati meno soddisfacenti di quanto atteso.

Esiste un razionale che spieghi perché nella pratica il test risulta meno soddisfacente di quanto prevedibile in base ai dati ottenuti nella fase preliminare?

La risposta può essere ricavata dalla seguente tabella, nella quale il valore predittivo del test positivo è stato calcolato a diversi valori di prevalenza della malattia.

<i>Sensibilità</i>	<i>Specificità</i>	<i>Prevalenza</i>	<i>Valore predittivo del test positivo</i>	<i>Valore predittivo del test negativo</i>
1,000	0,990	0,500	0,990	1,000
1,000	0,990	0,050	0,840	1,000
1,000	0,990	0,005	0,334	1,000
1,000	0,990	0,001	0,048	1,000

Consideriamo il valore predittivo del test positivo. Nel corso delle prove presso la clinica universitaria, il test è stato applicato ad una casistica di 500 malati e di 500 sani: quindi con una prevalenza della malattia del 50%. Il valore predittivo del test positivo corrispondente, quindi la probabilità di essere malato per un soggetto con un test positivo, era del 99%. Una volta entrato nell'uso, il test è stato poi applicato a casistiche non selezionate, con prevalenza della malattia minore. Dalla tabella appare evidente come l'applicazione del test con una prevalenza della malattia più "realistica", con valori della prevalenza che vanno dal 5 all'1 per mille, comporta un valore predittivo del test positivo che va dal 33,4% al 4,8%, o in altre parole comporta il fatto che su 100 soggetti con un test positivo la maggior parte (rispettivamente 67 e 95) sono soggetti sani. Quindi non un falso positivo su 100

Le possibili conseguenze determinate dalla dipendenza del valore predittivo del test dalla prevalenza della malattia sono ben descritte in Gerhardt e Keller [3] e comportano la necessità, nel corso della messa a punto e della applicazione di un test diagnostico, di valutarne tutti gli aspetti.

[1] Galen RS, Gambino RS. *Oltre il concetto di normalità: il valore predittivo e l'efficienza delle diagnosi mediche*. Piccin Editore, 1980, p. 37.

[2] Galen RS, Gambino RS. *Oltre il concetto di normalità: il valore predittivo e l'efficienza delle diagnosi mediche*. Piccin Editore, 1980, pp. 49-80.

[3] Gerhardt W, Keller H. *Evaluation of test data from clinical studies*. Scand J Clin Lab Invest, 46, 1986 (supplement 181), pp. 39-42.

