

CURVA ROC

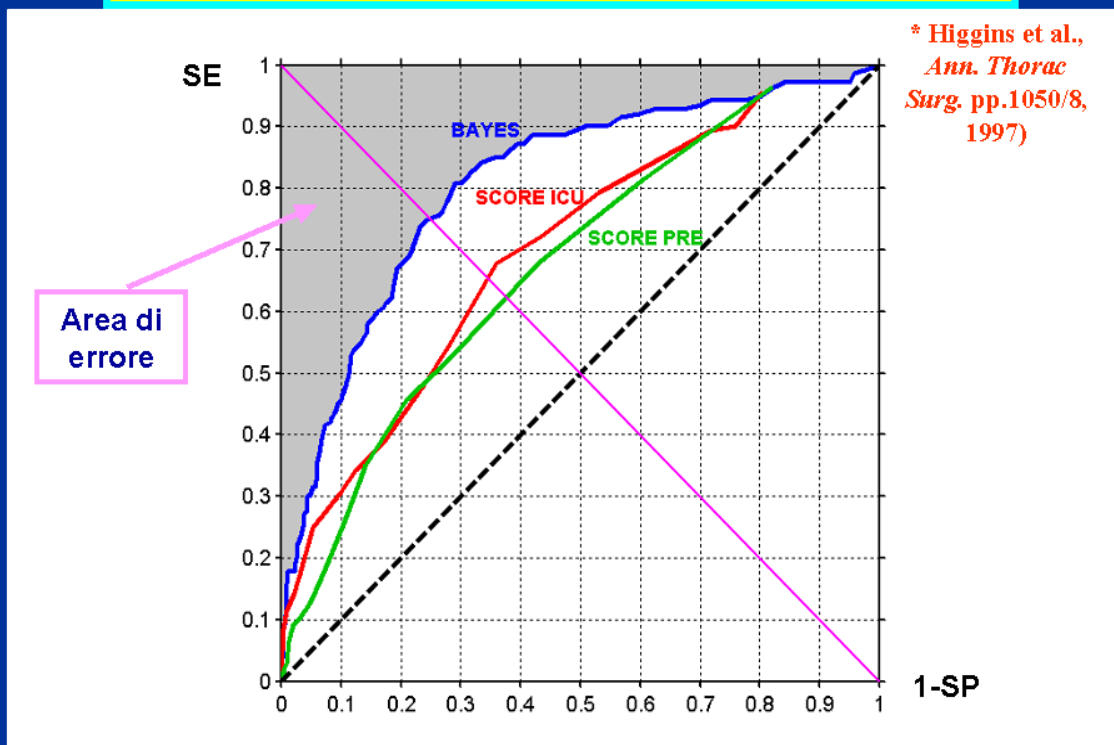
RECEIVER OPERATING CHARACTERISTIC

Ad ogni soglia di decisione corrisponde una coppia (SE, SP)

Riportando in ascisse 1-SP che rappresenta la frazione di FP e in ordinate SE che rappresenta la frazione di VP, si può tracciare una curva i cui punti rappresentano le prestazioni del metodo di decisione (test diagnostico) per ogni possibile scelta della soglia di decisione clinica.

Nel grafico seguente sono riportate tre curve ROC di due punteggi diagnostici (linee verdi e rosse) e di un metodo bayesiano, utilizzati in terapia intensiva post-cardiochirurgica per predire il rischio di complicanza postoperatoria.

Curve ROC classificatore bayesiano e score clinici*



Il grafico si trova in un diagramma quadrato di area unitaria.

La linea diagonale nera tratteggiata a 45 gradi rappresenta il puro caso: è come decidere lanciando una moneta.

La linea diagonale color magenta a 45 gradi rappresenta uguali SE e SP.

Le curve ROC partono sempre dal punto (SE=0, SP=1) che equivale ad affermare che tutti i pazienti non andranno incontro a complicanze cioè assenza di FP e tutti le complicanze stimate come FN: nell'esempio precedente la soglia è completamente a destra delle due popolazioni. La curva termina nel punto (SE=1, SP=0) che equivale ad affermare che si complicheranno tutti i pazienti (FN=0, tutti i normali come FP e soglia dell'esempio completamente a sinistra)

Il punto di ottimo è quello in alto a sinistra (SE=1, SP=1) per il quale la regola di decisione (test diagnostico) è infallibile (nessun FP e FN).

Il vertice del quadrato in basso a destra è un assurdo in quanto rappresenta il punto della "perfetta contraddizione" a una decisione certa, cioè è come se sapendo esattamente la diagnosi decidessimo in modo contrario. Tutto il triangolo a destra della diagonale tratteggiata non può contenere la curva ROC, essendo in pratica più improbabile, e quindi contraddittorio, del lancio della moneta.

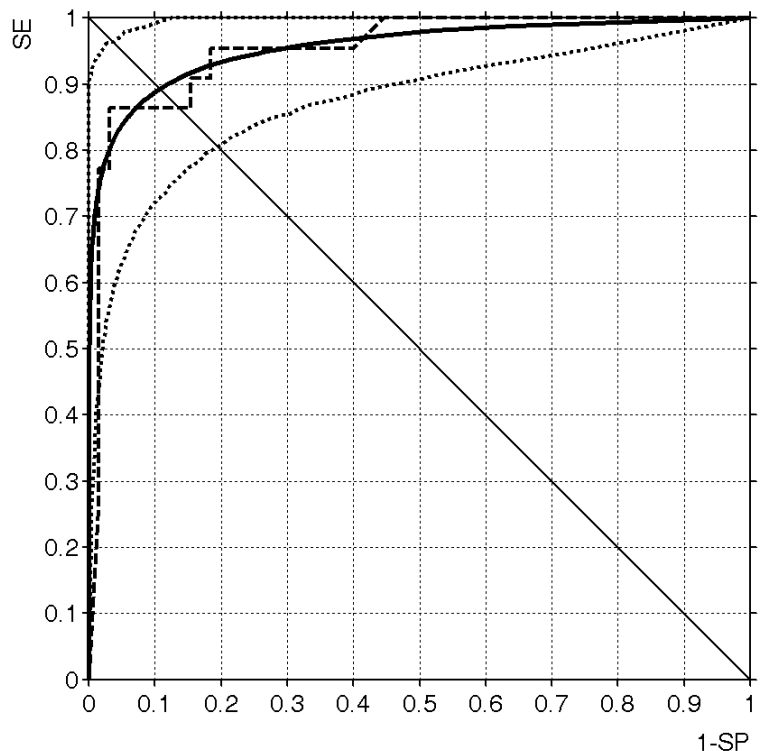
Pertanto, più la curva ROC è arcuata verso quel punto, migliore è il test decisionale. L'area grigia sopra la curva ROC rappresenta pertanto l'errore connesso con l'uso del test stesso.

Attraverso la curva ROC si può dunque valutare l'accuratezza globale di un metodo diagnostico con più di due risultati, tenendo conto anche della soglia di decisione. Essa è rappresentata dall'area (bianca nel grafico) che si trova sotto la curva ROC. Tale area è compresa tra 0,5 e 1. E' consuetudine ritenere sufficientemente accurato un test che ha un'area sotto la curva ROC superiore a 0,8.

Procedendo con rigore, secondo i dettami dell'inferenza statistica, siccome la curva ROC è stata ricavata utilizzando un campione rappresentativo di una popolazione più ampia, occorre definire un intervallo di confidenza entro il quale la curva ROC relativa alla popolazione giace con un prefissato livello di confidenza, p. es. al 95%. Il grafico sotto illustra tali intervalli di confidenza in un esempio tratto dalla diagnosi di melanoma ottenuta con un approccio bayesiano incorporante informazioni

campionarie relative a variabili morfometriche e colorimetriche di nei e melanomi.

Nel grafico, la curva ROC ottenuta dal campione è rappresentata dalla linea tratteggiata; poiché il campione ha bassa numerosità (23 melanomi e 47 nei benigni) tale curva ha una conformazione a scalini, che corrispondono a soglie discrete. La linea continua è ottenuta approssimando la linea tratteggiata con una tecnica modellistica che permette di ottenere una curva continua. Le linee tratteggiate sopra e sotto la curva, definiscono l'intervallo di confidenza al 95%.



La notevole ampiezza di tale intervallo è determinata principalmente dalla scarsa numerosità campionaria.

Se, ad esempio, si suppone di usare il metodo diagnostico bayesiano fissando la soglia in corrispondenza del punto della curva ROC con ugual SE e SP (incrocio tra la linea diagonale e la curva ROC), sulla linea continua campionaria notiamo che tali valori sono all'incirca entrambi uguali al 90%, cioè si avrebbero solo il 10% di FP e FN. In realtà però, l'ampio intervallo di confidenza ci obbliga a riferirci alla curva tratteggiata sottostante, cioè quella ottenibile nelle peggiori condizioni, che indica valori $SE=SP$ circa uguali solo all'80%.

In definitiva, possiamo affermare che con un 95% di probabilità, usando la soglia di decisione corrispondente al punto $SE=SP$, il test diagnostico è in grado di discriminare correttamente almeno l'80% di formazioni cutanee sospette di essere melanomi.