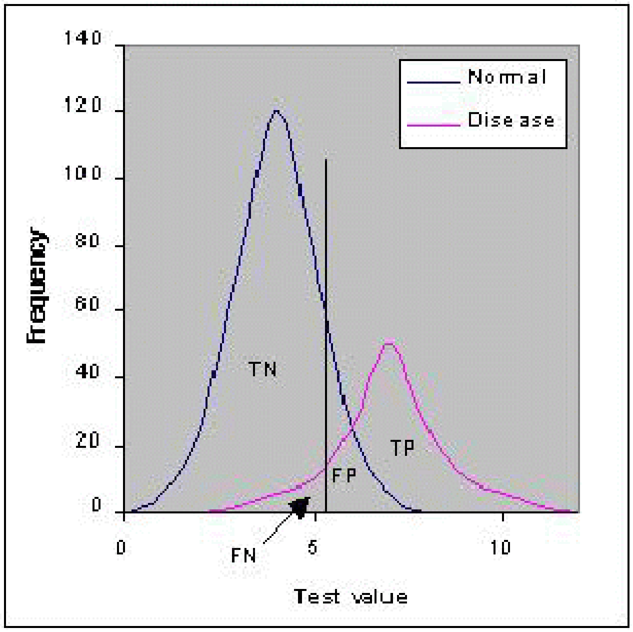
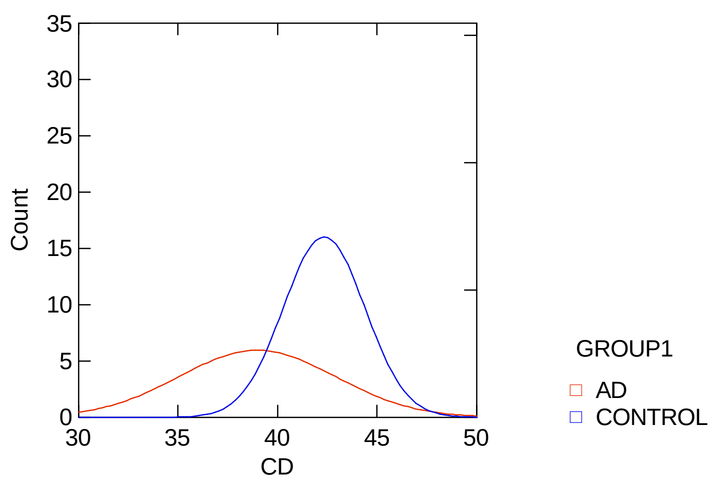
Introduction to ROC Curves



The sensitivity and specificity of a diagnostic test depends on more than just the "quality" of the test--they also depend on the definition of what constitutes an abnormal test. Look at the the idealized graph at right showing the number of patients with and without a disease arranged according to the value of a diagnostic test. This distributions overlap--the test (like most) does not distinguish normal from disease with 100% accuracy. The area of overlap indicates where the test cannot distinguish normal from disease. In practice, we choose a cutpoint (indicated by the vertical black line) above which we consider the test to be abnormal and below which we consider the test to be normal. The position of the cutpoint will determine the number of true positive, true negatives, false positives and false negatives. We may wish to use different cutpoints for different clinical situations if we wish to minimize one of the erroneous types of test results.

Si supponga di avere la distribuzione SPECT del nucleo caudato di 2 popolazioni (AD e controlli). Il grafico delle 2 distribuzioni è il successivo



Le medie dei 2 gruppi sono le seguenti: AD=38.986 e CONTROL=42.346.

Poiché le 2 distribuzioni si sovrappongono qualsiasi criterio si scelga ci saranno falsi positivi e negativi.

Un modo di studiare l'andamento dei falsi positivi e dei falsi negativi a seconda del criterio usato è quello di usare le curve ROC. Receiver Operating Characteristic. It is a plot of the true positive rate against the false positive rate for the different possible cutpoints of a diagnostic test.

An ROC curve demonstrates several things:
1. It shows the tradeoff between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity).
2. The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test.
3. The closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test.

4. The area under the curve is a measure of text accuracy.

Per ottenere gli Hit Rate e i False Positive possiamo calcolare il numero di pazienti identificati dal test scegliendo vari cutoff. Un modo per ottenere questi dati è quello di usare le caratteristiche della curva normale

Z	M->Z	Z->
0,000	0,000	0,500
0,670	0,249	0,251
1,000	0,341	0,159
1,280	0,400	0,100
1,645	0,450	0,050
2,000	0,477	0,023
2,330	0,490	0,010
3,000	0,499	0,001

Scegliendo vari cut-off (7) corrispondenti a diversi valori z, sono state calcolate le frequenze di pazienti considerati malati (HIT) e la frequenza di controlli considerati malati (FP). Sulla base dei vari ottenuti è stata costruita una tabella come la successiva

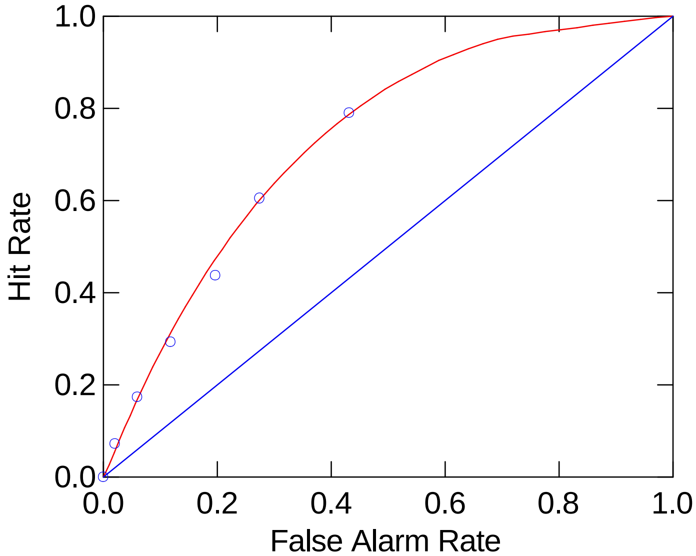
VOI$	FP	HIT_AD	RATING	z
CD	29	35	0	0,00
CD	8	31	1	1,00
CD	4	28	2	1,28
CD	4	24	3	1,64
CD	3	20	4	2,00
CD	2	17	5	2,33
CD	1	12	6	3,00

Attraverso questa tabella è possibile analizzare la bontà di un test, riconducendo l'analisi alla Signal Detection Theory (SDT).

Plotting and Intrepretating an ROC Curve

Per fare questo è neccessario disegnare la funzione degli HIT in funzione dei FP. Quello che si ottiene è una ROC curve del tipo successivo.

Receiver Operating Characteristic (ROC) Plot



The Area Under an ROC Curve

The accuracy of the test depends on how well the test separates the group being tested into those with and without the disease in question. Accuracy is measured by the area under the ROC curve. An area of 1 represents a perfect test; an area of .5 represents a worthless test. A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system:

.90-1 = excellent (A)
.80-.90 = good (B)

.70-.80 = fair (C)
.60-.70 = poor (D)
.50-.60 = fail (F)

(Wigton RS, Connor JL, Centor RM. Transportability of a decision rule for the diagnosis of streptococcal pharyngitis. Arch Intern Med. 1986;146:81-83.)

The area under the curve is the percentage of randomly drawn pairs for which this is true (that is, the test correctly classifies the two patients in the random pair). Two methods are commonly used: a non-parametric method based on constructing trapeziods under the curve as an approximation of area and a parametric method using a maximum likelihood estimator to fit a smooth curve to the data points. Both methods are available as computer programs and give an estimate of area and standard error that can be used to compare different tests or the same test in different patient populations.
For more on quantitative ROC analysis, see Metz CE. Basic principles of ROC analysis. Sem Nuc Med. 1978;8:283-298.

ROC analysis is part of a field called "Signal Dectection Therory" developed during World War II for the analysis of radar images. Radar operators had to decide whether a blip on the screen represented an enemy target, a friendly ship, or just noise. Signal detection theory measures the ability of radar receiver operators to make these important distinctions. Their ability to do so was called the Receiver Operating Characteristics. It was not until the 1970's that signal detection theory was recognized as useful for interpreting medical test results.

Nell’esempio indicato prima i valori forniti da Systat sono i seguenti:

```
VOI$='CD Case frequencies determined by value of variable HIT_AD.
 D-Prime      D Sub-A     Sakitt D     SD-Ratio     ROC Area
  0.802        0.890        0.903        0.789        0.736
```

In particolare il valore dell’area ROC non sembra particolarmente buono, cioè la capacità del test di discriminare tra patologia e non-patologia.