

# Using the Receiver Operating Characteristic (ROC) curve to analyze a classification model

## Background

Before explaining what a ROC curve is, we need to recall the definitions of sensitivity and specificity.

Suppose that we are testing people through blood samples to know whether they have a specific disease or not. In this case the sensitivity will be the proportion of true positives given that the condition is present  $P(+|D)$ . The specificity is the proportion of true negatives,  $P(-|\sim D)$ , the proportion of people who tested negative among those who do not have the disease.

$$\text{Sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

$$\text{Specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}$$

## Description

In a Receiver Operating Characteristic (ROC) curve the true positive rate (Sensitivity) is plotted in function of the false positive rate (1-Specificity) for different cut-off points. Each point on the ROC plot represents a sensitivity/specificity pair corresponding to a particular decision threshold. A test with perfect discrimination (no overlap in the two distributions) has a ROC plot that passes through the upper left corner (100% sensitivity, 100% specificity). Therefore the closer the ROC plot is to the upper left corner, the higher the overall accuracy of the test (Zweig & Campbell, 1993).

## Construction of a ROC curve<sup>1</sup>

Suppose that in a study of 125 people we use different cutoff values to determine if they have a certain condition:

Cutoff=0.5	D	~D
+	18	1
-	14	92
<b>Total</b>	<b>32</b>	<b>93</b>

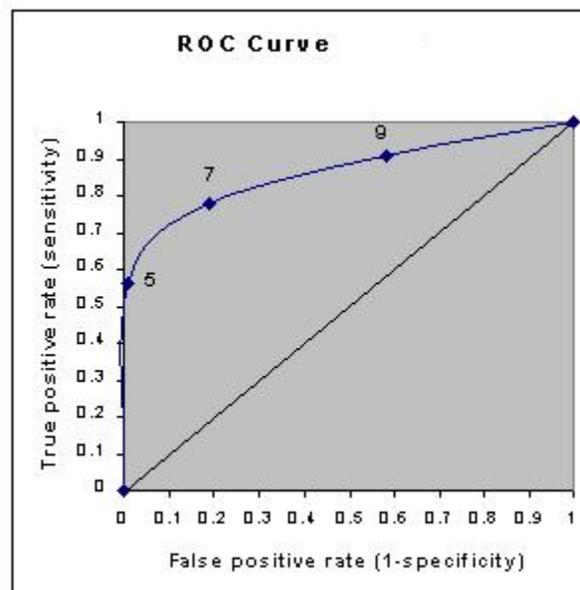
We can verify that the sensitivity is  $18/32=.56$  and the specificity is  $92/93=.99$

For a similar table but different cutoff values we could get

Cutoff=0.7	D	~D
+	25	18
-	7	75
<b>Total</b>	<b>32</b>	<b>93</b>

Here the sensitivity is  $25/32=0.78$  and the specificity is  $75/93=0.81$ .

We repeat this process over and over again, and we replace the specificity by the 1- sensitivity = proportion of false positives. The generated pairs (x,y) are the ones that appear on the ROC curve.



## Properties

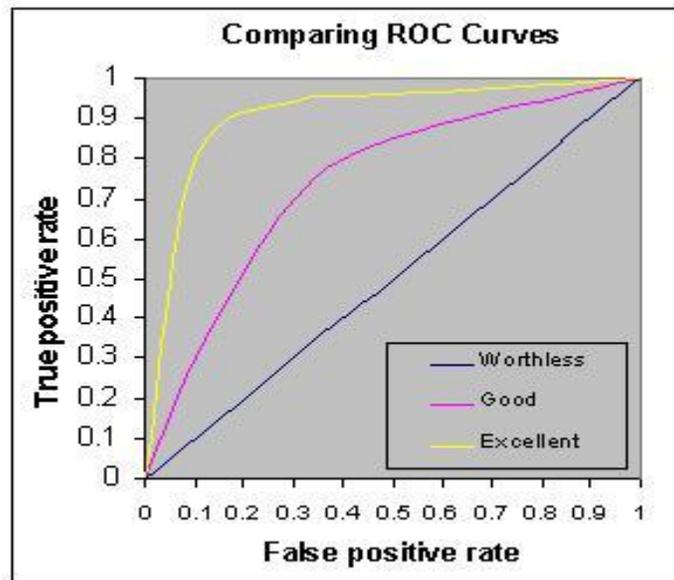
An ROC curve demonstrates several things:

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 upper-left 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 slope of the tangent line at a cutpoint gives the likelihood ratio (LR) for that value of the test.
5. The area under the curve is a measure of accuracy.

### Judging a ROC curve

The graph below shows three ROC curves representing excellent, good, and worthless tests plotted on the same graph. 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)



### A final note of historical interest

You may be wondering where the name "Receiver Operating Characteristic" came from. ROC analysis is part of a field called "Signal Detection Theory" 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.