ROC Curve Analysis

In almost all fields of human activity, there is often a need to discriminate between good and bad, presence and absence. Various tests have been designed to meet this objective. The ROC curve technique has been designed to attain two objectives in this regard.

ROC - Receiver Operating Characteristic

Various tests have been designed to meet this objective. The ROC curve technique has been designed to attain two objectives in this regard.

First, it can be used to calibrate (in some sense) a test so that it is able to perform the discrimination activity well. Second, it can be used to choose between tests and specify best among them.

What is ROC?

If one applies to a bank for credit, it is most likely that the bank will calculate a credit score out of the applicant's background. A higher score could indicate a good customer with minimal chance of default. The banker could refuse credit if the score is low. Often a credit score cut-off is used below which the application is rejected.

It is not difficult to see that there is always an element of risk here – risk of committing two types of errors. A good prospective customer (one who would not default) could be refused credit by the bank and a bad one could be approved credit. Clearly the banker would like the cut-off be fixed in a manner that chances of both the errors [1] are minimized if not entirely eliminated.

The ROC Curve was invented during the WW2 to help radars detect weak signals from aircrafts

While complete elimination is impossible, the ROC curve analysis [2] is a technique which contributes to this endeavour. A related problem is the question of choosing between methods of identifying good/bad customers should there be a choice. The ROC curve analysis technique can be of use even here.

The Plot

In order to draw the ROC curve, the concepts of ‘Sensitivity’ and ‘Specificity’ are used – the curve actually is the plot of sensitivity (in the y axis) against 1- specificity (in the x axis) for different values of the cut-off.

To understand these concepts, assume that we select a sample of z customers of the bank by retrospective sampling method. Further suppose that m and n of these are good and bad (defaulting) customers respectively (m+n=z). Next, we use the credit scale on these customers and calculate their credit scores. Then we use the cut-off and label customers good or bad according to whether the credit score is above or below the cut-off. Out of m good customers, the test classified x of them as good while the remaining m-x were classified as bad.

In the parlance of ROC curve, x is termed as TP (for true positive meaning that the credit scale was able to identify these customers as good correctly) while m-x in termed as FN (for false negative). Further suppose that out of n bad customers, the test classified y of them as bad while the remaining n-y were classified as good. In ROC parlance, y is termed as TN (for true negative) while n-y in termed as FP (for false positive).

Actual vs Credit Scale

		Actual status of costumers	
		Good	**Bad**
As predicted by credit scale	**Good**	x	n-y
	Bad	m-x	y
Total		m	n

Sensitivity and Specificity

The probability that among the good customers the test will identify a customer as good is known as ‘Sensitivity’ of the test for that cut-off (given by x/m). On the other hand, among the bad customers, the probability that the test will identify a customer as bad is known as ‘Specificity’ (given by y/n) of the test again for the same cut-off.

Shape of a Good Curve

To draw the curve, the sensitivity and specificity are determined for a range of cut-offs. Then sensitivity and 1-specificity are plotted. A good test is one in which the curve is closer to the upper left corner.

So far so good but how does one determine the optimal cut-off? The banker would like to determine that cut-off for which sensitivity is high and 1-specificity is low – ideally 100% sensitivity with 100% specificity. That is easier said than done as the best of the curves is not a vertical line but one which rises steeply initially and then slowly. The highest point on the curve has 100% sensitivity and 0% specificity. In other words as one of sensitivity or specificity increases, the other decreases and vice versa. The problem of determining the ideal cut-off is to choose one depending upon the extent of sensitivity and specificity that the decision maker is comfortable with. Having thus fixed a cut-off, the banker can then use it for evaluating fresh credit applications.

The fact that the best of the tests has a curve which rises steeply initially is used to choose between tests. A test can be called the best if its corresponding ROC curve is higher than others.

Related pages:

Type I and Type II Errors [1]

Experimental Error [3]

Links:
[1] https://explorable.com/type-I-error, [2] http://gim.unmc.edu/dxtests/ROC1.htm, [3] https://explorable.com/experimental-error, [4] https://explorable.com/, [5] https://explorable.com/roc-curve-analysis