ROC CURVES´ THRESHOLD ESTIMATION BASED ON COST FUNCTION

Konstantina Skaltsa, Lluís Jover, Josep L. Carrasco

Public Health Department, Faculty of Medicine, University of Barcelona, Spain

Medical diagnostic tests are used to classify subjects as healthy or diseased. The classification rule is usually set by determining classifying subjects with result above or below it as diseased. Various ways of estimating the separating threshold have been proposed in the literature. We worked on continuous measurements and approached the estimation based on the already proposed cost function i.e. by assigning costs to every possible classification result. It was thought as a more general way to handle the misclassifications by providing more or less weight to one or the other type of error and then opt for minimise the total cost of the test. We arrive at a general equation which consists of the density ratio of the diseased versus healthy subjects at the optimum threshold and depends on the costs assigned and the disease prevalence. Therefore, by substituting the densities (either known analytically or empirically estimated), an optimum threshold formula can be derived. The threshold estimator and its standard error are derived for the specific case where both diseased and healthy populations are normally distributed with unequal variances. The asymptotic normality and the confidence interval coverage are evaluated by means of a simulation study. In terms of utility, there are some potential benefits behind the confidence interval construction: 1) it allows taking into account the reliability of the estimated threshold when classifying 2) the limits of the confidence interval can be used as complementary thresholds to improve the predictive ability of the test. A case example to illustrate the procedure is provided.