World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:13, No:01, 2019

Comparison of Receiver Operating Characteristic Curve Smoothing Methods

Authors : D. Sigirli

Abstract : The Receiver Operating Characteristic (ROC) curve is a commonly used statistical tool for evaluating the diagnostic performance of screening and diagnostic test with continuous or ordinal scale results which aims to predict the presence or absence probability of a condition, usually a disease. When the test results were measured as numeric values, sensitivity and specificity can be computed across all possible threshold values which discriminate the subjects as diseased and non-diseased. There are infinite numbers of possible decision thresholds along the continuum of the test results. The ROC curve presents the trade-off between sensitivity and the 1-specificity as the threshold changes. The empirical ROC curve which is a non-parametric estimator of the ROC curve is robust and it represents data accurately. However, especially for small sample sizes, it has a problem of variability and as it is a step function there can be different false positive rates for a true positive rate value and vice versa. Besides, the estimated ROC curve being in a jagged form, since the true ROC curve is a smooth curve, it underestimates the true ROC curve. Since the true ROC curve is assumed to be smooth, several smoothing methods have been explored to smooth a ROC curve. These include using kernel estimates, using log-concave densities, to fit parameters for the specified density function to the data with the maximum-likelihood fitting of univariate distributions or to create a probability distribution by fitting the specified distribution to the data nd using smooth versions of the empirical distribution functions. In the present paper, we aimed to propose a smooth ROC curve estimation based on the boundary corrected kernel function and to compare the performances of ROC curve smoothing methods for the diagnostic test results coming from different distributions in different sample sizes. We performed simulation study to compare the performances of different methods for different scenarios with 1000 repetitions. It is seen that the performance of the proposed method was typically better than that of the empirical ROC curve and only slightly worse compared to the binormal model when in fact the underlying samples were generated from the normal distribution.

Keywords : empirical estimator, kernel function, smoothing, receiver operating characteristic curve
Conference Title : ICCMSA 2019 : International Conference on Computational Modeling, Simulation and Analysis
Conference Location : Paris, France
Conference Dates : January 24-25, 2019