Comparing the Expected Misclassification Cost for Two Classifiers Based on Estimates From the Same Sample

James Troendle[*], National Heart, Lung, and Blood Institute

Kai Yu, Tsingua University

Peter Westfall, Texas Tech University

Gene Pennello, U.S. Food and Drug Administration

Enrique Schisterman, National Institute of Child Health and Human Development

Abstract

Consider the problem of comparing two binary classifiers, evaluated on the same sample. McNemar's test can be used to compare overall predictive accuracy. However, to evaluate the classifiers in a clinically relevant manner, expected misclassification cost should be accounted for. We show that a Wald type test can be constructed for this purpose. We further derive a likelihood ratio test for comparison of two classifiers based on expected misclassification cost. The null distribution of the test statistic is approximated by simulation from strategically chosen parameter values. The properties of the tests are examined through simulation of correlated classification indicators. The Wald type test has approximate type I error control while maintaining a power advantage over the likelihood ratio test and is therefore recommended for most applications. If conservative error control is desired, the likelihood ratio test calibrated from several strategically chosen parameter values is recommended. The methods are illustrated on a nested case-control study on pre-eclampsia and a prospective cohort study on coronary heart disease. An interval of misclassification cost ratios for which the Wald test rejects the null hypothesis of equal expected misclassification cost is reported.

---

[*] Presenting author