

Sephora-Clotilde Zoro

12 March 2023

Stat 499 Final Write-Up

My DRP was about using statistics to evaluate and choose thresholds for diagnostic tests meant to classify patients as ill or healthy, especially when a “Gold Standard” test is costly or difficult to administer.

Tests may be measured on binary, ordinal, or continuous scales, but we focused on tests measuring on a continuous scale. Thresholds are used to transform the continuous result (e.g., amount of biomarker) into a binary one (positive or negative for a condition). These tests will have some error. The True Positive Rate/Fraction (1-Type II Error) and False Positive Rate/Fraction (Type I Error) will vary based on biomarker threshold chosen. The Receiver Operating Characteristic Curve (ROC Curve) is the set of these TPRs and FPRs for every possible threshold.

The Area under the Curve (AUC) of an ROC curve gives information about the usefulness of the diagnostic test. When the AUC is 0.5, the FPR = TPR for any threshold, which makes the test totally uninformative. When it is 1.0, there exists a threshold for which the test has no error at all. This would be expected of our Gold Standard tests, but most tests have AUCs between these two values, and generally, the further the AUC is from 0.5 the better it is at classification.

Thresholds which optimize test performance may be chosen in a variety of ways— Commonly thresholds are chosen by minimizing error (this can be done by weighing errors differently as well), and they can also be chosen by minimizing the Cost function of a patient. Cost is a function of monetary cost of test administration, cost to patient by misdiagnoses, population prevalence of the illness, and cost of disease treatment in the case of correct positive diagnosis.

When comparing ROC curves for related biomarkers, AUC is a useful metric for evaluating the usefulness of one test against the other. However, especially when AUCs do not have a major difference between them, you must test to see whether the difference between the tests performances is significant. This is most easily done by using the `roc.test(...)` function of the `pROC` r package, and this is sufficient for the purposes of the DRP.