Colloquium February 27th, 2025 | Ben Brewer

Title: Statistical Methods for Cutoff Estimation under Tree and Umbrella Ordering Settings

Abstract: In many biomarker studies, the assays conducted might refer to patients coming from three or more groups. Typically, there exists a "healthy" group, a "mild disease" group, and an "aggressive disease" group. It could further be the case that two or more subgroups lie in one of these categories. For some such biomarkers, the typical stochastic ordering assumption might not be justified for all disease classes under study, and usual ROC methodologies that involve ROC surfaces and hypersurfaces are inadequate. Different types of orderings may be appropriate depending on the setting, and these may involve a number of ambiguously ordered groups that stochastically exhibit larger (or lower) marker scores than the remaining groups. This type of ordering is called tree (or umbrella) ordering. One example is cancer biomarker studies, which include different stages of cancer. Such configurations can also be found in other disease settings, such as tuberculosis.

Recently, there has been scientific interest on ROC methods that can accommodate these so-called 'tree' or 'umbrella' orderings. However, there is limited work discussing the estimation of cutoffs in such settings. Herein, we discuss the estimation and inference around optimized cutoffs when accounting for such configurations that might involve multiple disease categories/stages. We explore different cutoff alternatives and provide parametric, flexible parametric, and non-parametric kernel-based approaches for estimation and inference. We also propose a novel cutoff selection measure, the Max-Min distance, and investigate its theoretical properties. We evaluate our approaches using simulations and illustrate them through a real data set that involves tuberculosis patients.