A ROC guided approach to suggest new cutoff values for CV and NPI

Significance

Let $\boldsymbol{X}(t) = \begin{pmatrix} X_1(t) & \dots & X_d(t) \end{pmatrix}^\top$ be a d-dimensional vector of possibly time-dependent vector of covariates, and let D be a 0-1 binary outcome such as the presence of stroke. We are interested in estimating the risk score function $p(\boldsymbol{x})$, where $p(\boldsymbol{x}) = P(D = 1|\boldsymbol{X}(t) = \boldsymbol{x}), 0 < p(\boldsymbol{x}) < 1$, is the disease probability given $\boldsymbol{X} = \boldsymbol{x}$. One common approach is to consider the logistic regression that links $p(\boldsymbol{x})$ to a linear predictor function of $\boldsymbol{x}$, say $g(\boldsymbol{x})$, where $g : \mathcal{X} \mapsto \mathbb{R}$ is the inverse-logit link function. Then $g(\boldsymbol{x})$ can be regarded as a single marker that summarizes information from $\boldsymbol{X}$. Without loss of generality, we assume that a higher value of $g(\boldsymbol{x})$ is more indicative of $D = 1$. Given a threshold c, we define the true positive rate (TPR) of $g(\boldsymbol{x})$ as $P\{g(\boldsymbol{x}) > c|D = 1\}$ and the false positive rate (FPR) of $g(\boldsymbol{x})$ as $P\{g(\boldsymbol{x}) > c|D = 0\}$. The receiver operating characteristic (ROC) curve of $g(\boldsymbol{x})$ is created by plotting the TPR against the FPR at for all c. The ROC curve is a helpful tool for quantitative assessment of the model, $g(\boldsymbol{x})$. A perfect model that completely separate the disease outcomes would have 100% TPR and the area under the ROC curve (AUC) would be one. On the contrary, a completely ineffective model would result in an ROC curve that closely follows the 45° diagonal line and the AUC would be close to 0.5. Thus AUC is commonly used as a summary measure of the ROC curve, and it indicates the overall performance of the model in terms of prediction accuracy. Motivated by the intuition that the optimal model should have an optimized AUC, we propose an ROC-guided approach to suggest new cutoff points for the clinicians to interpret pupillometry readings.

Approach

We first describe the proposed ROC-guided approach to suggest a new cutoff point for the NPi grade, then describe the extensions. Suppose among the d time-dependent covariate, let $X_1(t) = \text{NPi}(t)$ be the NPi grade at time t. We dichotomy $X_1(t)$ according to a fixed cutoff point τ so that $I\{\text{NPi}(t) < \tau\}$ is included in the model instead of the continuous value NPi(t). The logistic regression model for $p(\boldsymbol{x})$ has the form:

$$g^{-1}\{p(\boldsymbol{x})\} = \alpha + \boldsymbol{X}(t)^\top \boldsymbol{\beta} = \alpha + \beta_1 I\{\text{NPi}(t) < \tau\} + \ldots + \beta_d X_d(t),$$

where α is the intercept and $\boldsymbol{\beta}$ is a d-dimensional vector of regression coefficient. Baseline covaraites such as age, gender, and race will be included as $X_i(t) \equiv X_i(0)$, for $i > 1$. We propose to choose a cutoff point τ that results in an optimized AUC. The following steps outlined the procedure

Step 1. Split the data into training and testing datasets by subject id's.

Step 2. Initialize $m = 1$, fit the logistic regression model under the training set, given a cutoff point τ_m.

Step 3. Calculate the AUC using the testing set and store it as A_m.

Step 4. Repeat Steps 2 and 3 for $m = 2, \ldots, M$, where M is the total number of cutoff points to be tested.

Simple random sampling can be used to split the data in Step 1 so that the training set consists of roughly 80% of the original data. Alternatively, cross-validation based approaches can be implemented for the split. Given the size of END-PANIC data, the proportions of the training and the testing datasets will provide a good balance between modeling building and validating. Depending on the available resources, the cutoff points, $\{\tau_1, \ldots, \tau_M\}$, can be M equally spaced points between 0 and 5 or a subset of the observed NPi values. Given the configurations, the recommended cutoff point is then $\tau_{\arg\max_m A_m}$.

The proposed ROC-guided approach is particularly feasible for our applications as the "true" cutoff point that separates the normal and abnormal is unknown. Moreover, the proposed cutoff point is expected to give favorable predictive results as it is chosen to maximizes the AUC. The AUC

(Pepe et al., 2003)

To evaluating the prediction performance of the proposed method, we consider the following extension of concordance index and Brier score.

The dichotomous covariate, $I\{\text{NPi}_1(t) < \tau\}$, and τ_m can be modified accordingly when NPi is replaced with CV. The recommended cutoff points will compared with the currently used cutoff points of 3 and 0.8 for NPi and CV, respectively.

Since both the NPi and CV are readily available from one pupillometry reading. It is also a clinical interest to recommend treatments considering NPi and CV jointly. However, directly incorporate them into the logistic model could run into the multicollinearity issue due to the correlation between NPi and CV. To avoid this, we propose to construct the indicator $I\{\text{NPi}(t) < \tau_1, \text{CV}(t) < \tau_2\}$ for cutoff pairs (τ_1, τ_2) to replace $X_1(t)$. The cutoff pair provide a more conservative of abnormality as it requires both NPi and CV to exceed the cutoff values. A possible disadvantage of using the AUC to evaluate models is that it obscures information. For example, under limited medical resources models with lower FPR might are of more interest. In this case, a partial AUC can be considered to accommodate a truncated range of clinically relevant values of FPR.

To accommodate both the left and right eye reading, a binvaraite model can be considered. In this case, we propose to use

$$g^{-1}\{p_i(\boldsymbol{x})\} = \alpha_i + \boldsymbol{X}_i(t)^\top \boldsymbol{\beta} = \alpha + \beta_1 X_1(t) + \ldots + \beta_d X_d(t),$$

where $i = 1, 2$ gives the left and right eye reading, and α_i has the similar role of a random effect used to account for the heterogeneity between eyes. In this case, the proposed modification can still be applied.

References

Pepe, M. S. et al. (2003), *The statistical evaluation of medical tests for classification and prediction*, Medicine.