Weighted ROC analysis

Toby Dylan Hocking

October 9, 2024

1 Introduction

In binary classification, we are given n observations. For each observation $i \in \{1, \ldots, n\}$ we have an input/feature $x_i \in \mathcal{X}$ and output/label $y_i \in \{-1, 1\}$. For example, say that $\mathcal{X}$ is the space of all photographs, and we want to find a binary classifier that predicts whether a particular photograph x_i contains a car ($y_i = 1$) or does not contain a car ($y_i = -1$).

In weighted binary classification we also have observation-specific weights $w_i \in \mathbb{R}_+$ which are the cost of making an error in predicting that observation. Thus the goal is to find a classifier $c : \mathcal{X} \rightarrow \{-1, 1\}$ that minimizes the weighted zero-one loss on a set of test data

$$\underset{c}{\text{minimize}} \sum_{i \in \text{test}} I\left[c(x_i) \neq y_i\right] w_i, \tag{1}$$

where I is the indicator function that is 0 for a correct prediction, and 1 otherwise.

Instead of directly learning a classification function c, binary classifiers often instead learn a score function $f : \mathcal{X} \rightarrow \mathbb{R}$. Large values are more likely to be positive $y_i = 1$ and small values are more likely to be negative. One way of evaluating such a model is by using the weighted Receiver Operating Characteristic (ROC) curve, as explained in the next section.

2 Weighted ROC curve

Let $\hat{y}_i = f(x_i) \in \mathbb{R}$ be the predicted score for each observation $i \in \{1, \ldots, n\}$, let $\mathcal{I}_1 = \{i : y_i = 1\}$ be the set of positive examples and let $\mathcal{I}_{-1} = \{i : y_i = -1\}$ be the set of negative examples. Then the total positive weight is $W_1 = \sum_{i \in \mathcal{I}_1} w_i$ and the total negative weight is $W_{-1} = \sum_{i \in \mathcal{I}_{-1}} w_i$.

For any threshold $\tau \in \mathbb{R}$, define the thresholding function $t_\tau : \mathbb{R} \rightarrow \{-1, 1\}$ as

$$t_\tau(\hat{y}) = \begin{cases} 1 & \text{if } \hat{y} \geq \tau \\ -1 & \text{if } \hat{y} < \tau. \end{cases} \tag{2}$$

We define the weighted false positive count as

$$\text{FP}(\tau) = \sum_{i \in \mathcal{I}_{-1}} I\left[t_\tau(\hat{y}_i) \neq -1\right] w_i \tag{3}$$

and the weighted false negative count as

$$\text{FN}(\tau) = \sum_{i \in \mathcal{I}_1} I\left[t_\tau(\hat{y}_i) \neq 1\right] w_i. \tag{4}$$

We define the weighted false positive rate as

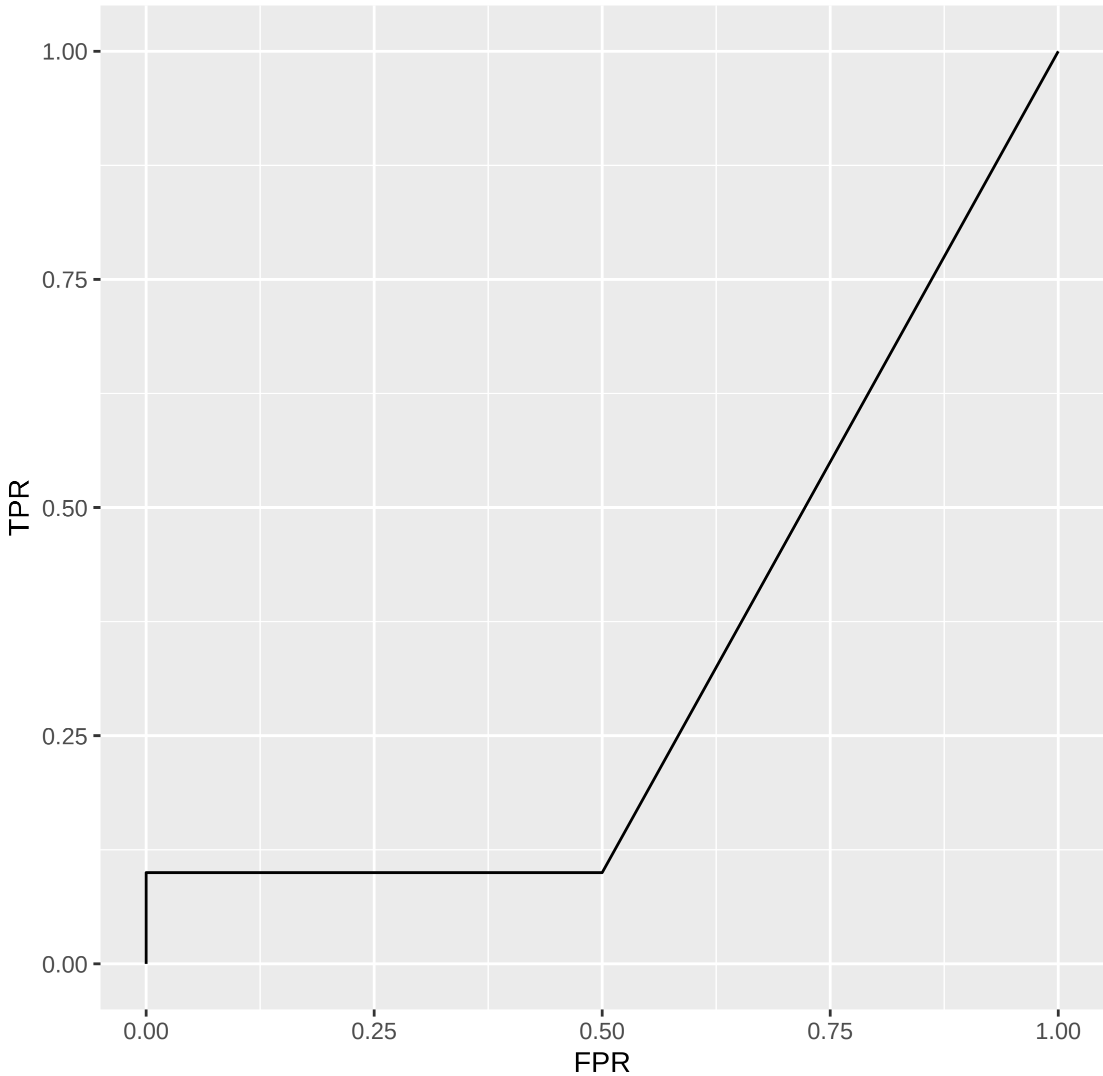
$$\text{FPR}(\tau) = \frac{1}{W_{-1}} \sum_{i \in \mathcal{I}_{-1}} I\left[t_\tau(\hat{y}_i) \neq -1\right] w_i \tag{5}$$

and the weighted true positive rate as

$$\text{TPR}(\tau) = \frac{1}{W_1} \sum_{i \in \mathcal{I}_1} I\left[t_\tau(\hat{y}_i) = 1\right] w_i. \tag{6}$$

A weighted ROC curve is drawn by plotting $\text{FPR}(\tau)$ and $\text{TPR}(\tau)$ for all thresholds $\tau \in \mathbb{R}$. It can be computed and plotted using the R code

```
> y <- c(-1, -1, 1, 1, 1)
> w <- c(1, 1, 1, 4, 5)
> y.hat <- c(1, 2, 3, 1, 1)
> library(WeightedROC)
> tp.fp <- WeightedROC(y.hat, y, w)
> library(ggplot2)
> ggplot()+
+   geom_path(aes(FPR, TPR), data=tp.fp)+
+   coord_equal()
```



3 Weighted AUC

The Area Under the Curve (AUC) may be computed using the R code

```
> WeightedAUC(tp.fp)
[1] 0.325
```