Appendix – Proofs of Theorems 1-3

Proof of Theorem 1

Let $\hat{C} = I(T(X) \geq u)$ and consider the *PPV* of this classifier,

$$PPV(u) = \frac{\pi_1 P(T(X) \geq u \mid C=1)}{\pi_1 P(T(X) \geq u \mid C=1) + \pi_0 P(T(X) \geq u \mid C=0)} \ .$$

We will first show $PPV(u)$ is an increasing function of by equivalently showing that

$r(u) = \dfrac{P(T(X) \geq u \mid C=1)}{P(T(X) \geq u \mid C=0)}$ is an increasing function of u. With $T(X) = \log\big(h_1(X)/h_0(X)\big)$,

$$r(u) = \frac{\int_{T(x)\geq u} h_1(x)dx}{\int_{T(x)\geq u} h_0(x)dx} = \frac{\int_{T(x)\geq u} \big(h_1(x)/h_0(x)\big)h_0(x)dx}{\int_{T(x)\geq u} h_0(x)dx} = E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, T(X) \geq u \right\}.$$

For any $u_1 < u_2$, it follows from the law of total probability that

$$\begin{aligned}
r(u_1) &= E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, T(X) \geq u_1 \right\} \\
&= P\left\{T(X) \geq u_2 \middle| C=0, T(X) \geq u_1\right\} \times E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, T(X) \geq u_2 \right\} \\
&\qquad + P\left\{u_1 \leq T(X) < u_2 \middle| C=0, T(X) \geq u_1\right\} \times E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, u_1 \leq T(X) < u_2 \right\} \\
&= \omega r(u_2) + (1-\omega) r^*
\end{aligned}$$

where $\omega = P\left\{T(X) \geq u_2 \middle| C=0, T(X) \geq u_1\right\}$ and $r^* = E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, u_1 \leq T(X) < u_2 \right\}$. By

definition,

$$r(u_2) = E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, T(X) \geq u_2 \right\} \geq e^{u_2} \geq E\left\{ \frac{h_1(x)}{h_0(x)} \middle| C=0, u_1 < T(X) < u_2 \right\} = r^* \ .$$

Because $r(u_1)$ is a weighted average of $r(u_2)$ and r^*, this implies $r(u_1) \le r(u_2)$ and this shows that $r(u)$ and hence $PPV(u)$ are increasing in u. Similarly it can be shown that $NPV(u)$ is decreasing in u and it therefore follows that the PROC is monotone.

Proof of Theorem 2

If $T(X)$ is the log-likelihood ratio then Theorem 1 implies the PROC curve is increasing, which implies[16] that $h_G(u) \le h_F(u)$ and $\bar{h}_F(u) \le \bar{h}_G(u)$ hold. The theorem then follows from necessary and sufficient conditions given in section 4.

Proof of Theorem 3

Define R_1 to be the set of all (x_1, x_2) at the second stage for which the classifier predicts class 1, and let R_1^c denote the complement of R_1. Then,

$$\begin{aligned}
OER_2(u) &= \pi_0 FPR_2(u) + \pi_1 FNR_2(u) \\
&= \pi_0 P(T_1(X_1) \ge u_1 \,|\, C=0) + \pi_0 P(u_0 < T_1(X_1) < u_1,\, (X_1, X_2) \in R_1 \,|\, C=0) \\
&\qquad + \quad \pi_1 P(T_1(X_1) < u_0 \,|\, C=1) + \pi_1 P(u_0 < T_1(X_1) < u_1,\, (X_1, X_2) \in R_1^c \,|\, C=1) \\
&= \pi_0 P(T_1(X_1) \ge u_1 \,|\, C=0) + \pi_1 P(T_1(X_1) < u_0 \,|\, C=1) + \pi_1 P(u_0 < T_1(X_1) < u_1 \,|\, C=1) \\
&\qquad + \quad \iint_{R_1} I\{u_0 < T_1(x_1) < u_1\}\{\pi_0 h_0(x_1, x_2) - \pi_1 h_1(x_1, x_2)\}\, dx_1 dx_2 ,
\end{aligned}$$

It is clear that to minimize $OER_2(u)$ we should take R_1 to include all points (x_1, x_2) for which $h_1(x_1, x_2) / h_0(x_1, x_2) > \pi_0 / \pi_1$, which is equivalent to the claim in the theorem.