Optimal cutoff point from ROC curve

Suppose we are devising a screening test for some disease $D \in \{0, 1\}$. Continuous measurements from patients Y's are obtained, and patients are deemed "positive" (or at "high risk") if $Y \geq t$ for some threshold t.

Let $T(t)$ and $F(t)$ be the true positive rate (sensitivity) and false positive rate (1 - specificity) at threshold level t, respectively. The area under curve (AUC) of the receiver operating characteristic (ROC) curve is defined as

$$\mathrm{AUC} = \int_0^1 T(F)\mathrm{d}F.$$

We can write T as a function of F. This is possible because $F(t)$ is monotone increasing with respect to t. Furthermore, assume that $F(t) = \mathbb{P}(Y \geq t|D = 0)$ has derivative $f(t) \geq 0$. Then the AUC can be computed as

$$\begin{aligned}
\mathrm{AUC} &= \int_0^1 T(F)\mathrm{d}F \\
&= \int_0^1 \mathbb{P}(Y \geq t|D = 1)\mathrm{d}\mathbb{P}(Y \geq t|D = 0) \\
&= \int_{\mathcal{T}} T(t)|f(t)|\mathrm{d}t \\
&= \mathbb{P}(Y_D > Y_{\bar{D}}|\mathcal{D}).
\end{aligned}$$

Given the true disease status $\mathcal{D}$, AUC is the probability that the test result Y_D from a randomly selected case is larger than the test result $Y_{\bar{D}}$ from a randomly selected control.

This is a nice interpretation, huh! It makes sense because if Y_D and $Y_{\bar{D}}$ can be easily told apart, the screening test will likely be quite effective.

Now what if we'd like to choose a cutpoint from all the possible values? Should we always choose the point closest to the upper left corner?

The short answer is no. It depends on the disease prevalence π and the cost for false positives c_+ and false negatives c_-. Naturally, for a disease status d and action $a \in \{0, 1\}$, we have a loss function $L(d, a)$ defined as

$$\begin{aligned}
&L(0, 0) = 1, \quad && L(0, 1) = c_+, \\
&L(1, 0) = c_-, \quad && L(1, 1) = 0.
\end{aligned}$$

As a decision rule with threshold value t, we define $m_t(Y) := 1\{Y \geq t\}$. The risks for different disease status are

$$
\begin{aligned}
R(0, m_t) &= \mathbb{E}_{D=0} L(0, m_t(Y)) = \mathbb{P}(Y \geq t | D = 0) c_+ = F(t) c_+, \\
R(1, m_t) &= \mathbb{E}_{D=1} L(1, m_t(Y)) = \mathbb{P}(Y < t | D = 1) c_+ = (1 - T(t)) c_-.
\end{aligned}
$$

With prior distribution $\mathbb{P}(D = d) = \pi^d (1 - \pi)^{1-d}$, the Bayes risk for m_t is

$$
r_\pi(m_t) = \pi c_- (1 - T(t)) + (1 - \pi) c_+ F(t).
$$

To find a point on the ROC curve that minimises the expression, we set the derivative with respect to F to zero, hence

$$
\frac{\partial T}{\partial F} = \frac{1 - \pi}{\pi} \frac{c_+}{c_-}.
$$

That is, the optimal point is such whose tangent line has the slope above.