1]Ian Langmore[1]

ROC Curves

[

January 2, 2013

Contents

Abstract

1 Notation

Following notation in [1], we define Y, N as the sets of outcomes that are predicted positive/negative, and p, n as the outcomes that are positive/negative.

2 ROC for $-x$

Consider a logistic regression model that predicts the probability of p, as a function of x as $\text{logit}(x)$.

For every threshold level, the models with x and $-x$ give opposite predictions.

$$\begin{aligned}\mathrm{P}_x[Y \mid p] &= 1 - \mathrm{P}_x[N \mid p] = 1 - \mathrm{P}_{-x}[Y \mid p] \\ \mathrm{P}_x[Y \mid n] &= 1 - \mathrm{P}_x[N \mid n] = 1 - \mathrm{P}_{-x}[Y \mid n].\end{aligned}$$

The ROC curve of the x model is the graph $(\mathrm{P}_x[Y \mid p], \mathrm{P}_x[Y \mid n]) = (w, f(w))$. The ROC curve of the $-x$ model is the graph $(z, g(z))$. The above equations show that

$$w = 1 - z, \qquad f(w) = 1 - g(z) = 1 - g(1 - w),$$

therefore

$$\begin{aligned}AUC_x &= \int_0^1 f(w)\,\mathrm{d}w = \int_0^1 (1 - g(1-w))\,\mathrm{d}w = 1 - \int_0^1 g(1-w)\,\mathrm{d}w = 1 - \int_0^1 g(z)\,\mathrm{d}z \\ &= 1 - AUC_{-x}.\end{aligned}$$

*ianlangmore@gmail.com

References

[1] T. Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27:861–874, 2006.