Stat637 Homework 3 [1]

Arthur Lui

4 February 2015

1 Titanic

The odds ratio is the ratio between odds, and not probabilities. The correct interpretation should be, "the *odds* of survival for females is 11.4 times that for males".

$$\begin{aligned}
\text{Odds} = \quad & \frac{p_f/(1-p_f)}{p_m/(1-p_m)} = && 11.4 \\
\implies \quad & \frac{2.9}{p_m/(1-p_m)} = && 11.4 \\
\implies \quad & p_m = && \frac{2.9/11.4}{1+2.9/11.4} \\
\implies \quad & p_m = && \mathbf{.2027972}
\end{aligned}$$

$$\begin{aligned}
\text{Odds for females} = \quad & \frac{p_f}{1-p_f} = && 2.9 \\
\implies \quad & p_f = && \frac{2.9}{1+2.9} \\
\implies \quad & p_f = && \mathbf{.7435897}
\end{aligned}$$

Therefore, the proportion of males that survived = **.2027972**, and the proportion of females that survived = **.7435897**.

2 Crime

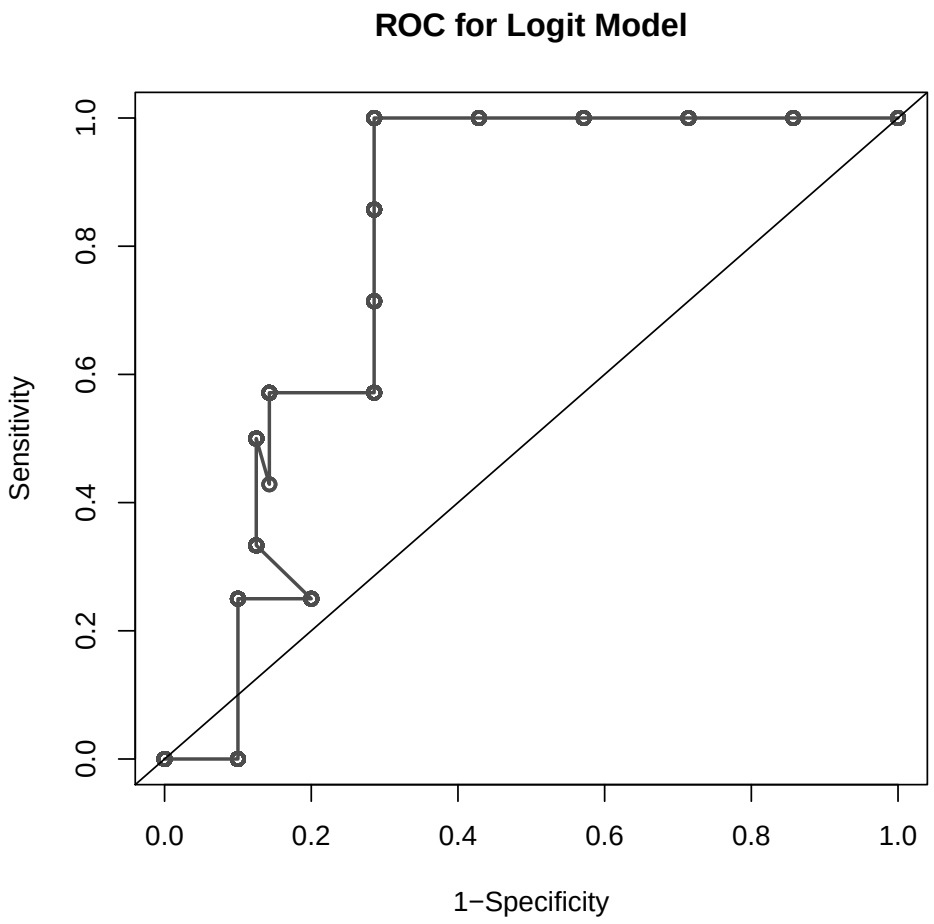
The conditional distribution these statistics refer to is X|Y. Assuming the set of all races only includes "Black" and "White",

$$P[Y = W|X = W] = \frac{P(X = W|Y = W)P(Y = W)}{P(X = W|Y = W)P(Y = W) + P(X = W|Y = B)P(Y = B)},$$

and the only statistics unknown in the above formula is P[Y=B] (or P[Y=W] = 1-P[Y=B]). So, the only other statistic needed is the proportion of victims that are white.

[1] https://github.com/luiarthur/Fall2014/blob/master/Stat637/3/sleep.R

3 Sensitivity & Specificity



On average, the model predicts better than guessing. This means that the logit model was a good model choice.