logisticRegressionCoefficientsAndEvaluativeMetrics

October 15, 2019

0.1 Coefficients in logistic regression

We saw that logistic regression is one way of modeling the probability data points belong to one of two classes. The model used a data transform—the logistic function. The logistic function transformed the an unbounded linear function $(\beta_0 + \beta_1 X + \cdots)$ onto the interval $[0, 1]$.

The model looked like the following

$$p(y_i = 1|X_i, \beta) = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_1 x_{in}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_1 x_{in}}}$$

By manipulating the above function we found we could make the right hand side above look linear

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_1 x_{in}$$

where $p = p(y_i = 1|X_i, \beta)$

0.2 Goal

Our next goal is to understand and interpret the coefficients β_i.

0.2.1 Taking a similar approach as linear regression

In multiple linear regression we can interpret βs the following way. A change in the variable X by one unit would result in a corresponding change in Y of β units.

We can derive the reason for this interpretation and then apply the same procedure to our logistic regression.

Consider the following multiple linear regression

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \tag{1}$$
$$\epsilon \sim N(0, \sigma^2) \tag{2}$$

To understand where the above "1-unit change" interpretation comes from, we can take the difference between two regression models: The above regression model for x values x_1^* and x_2^* and a second regression model for x values x_1^* and $x_2^* + 1$.

The first regression model is

$$y^1 = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \epsilon \tag{3}$$
$$\epsilon \sim N(0, \sigma^2) \tag{4}$$

and the second model is

$$y^2 = \beta_0 + \beta_1 x_1^* + \beta_2 (x_2^* + 1) + \epsilon \tag{5}$$
$$\epsilon \sim N(0, \sigma^2) \tag{6}$$

The difference between these two models is

$$y^2 - y^1 = \beta_0 + \beta_1 x_1^* + \beta_2 (x_2^* + 1) + \epsilon - (\beta_0 + \beta_1 x_1^* + \beta_2 x_2^* + \epsilon) = \beta_2$$

When we change the x_2 value by one unit, the difference in y values equals β_2.

Lets apply the same reasoning to our logistic regression. Our first logistic regression will take x values x_1^* and x_2^* and out second model will take x values x_1^* and $x_2^* + 1$.

Model 1 is

$$\log\left(\frac{p^1}{1 - p^1}\right) = \beta_0 + \beta_1 x_1^* + \beta_2 x_2^*$$

and the second model is

$$\log\left(\frac{p^2}{1 - p^2}\right) = \beta_0 + \beta_1 x_1^* + \beta_2 (x_2^* + 1)$$

The difference between model 2 and model 1 equals

$$\log\left(\frac{p^2}{1 - p^2}\right) - \log\left(\frac{p^1}{1 - p^1}\right) = \beta_2$$

We can make the expression on the left side of the equals simpler, using properties of the logarithm.

$$\log\left(\frac{\frac{p^2}{1-p^2}}{\frac{p^1}{1-p^1}}\right) = \beta_2$$

The expression on the left is called the **log odds ratio**. Often both sides are exponentiated.

$$\frac{\frac{p^2}{1-p^2}}{\frac{p^1}{1-p^1}} = e^{\beta_2}$$

The expression on the left is called the **odds ratio**.

0.3 β

When β equals 0

$$\frac{\frac{p^2}{1-p^2}}{\frac{p^1}{1-p^1}} = e^0 = 1$$

The ratio on the left can equal one only if both p^2 and p^1 are equal. When β equals 0 the corresponding explanatory variable has no affect on the probability a data point belongs to group 1 versus group 0.

When β is a positive number, for example $\beta = 2$, then for a one unit increase in the x value the odds this data point belongs to group 1 increases by e^2 or roughly 7.3. This is a bit unsatisfying. Odds are more difficult to interpret. Typically, I like to use the following analogy when interpreting coefficients in logistic regression.

Suppose the original probability of belonging to group 1 versus 0 was 50% $(p^1 = 0.50)$. Then

$$\begin{aligned}
\frac{\frac{p^2}{1-p^2}}{\frac{0.50}{1-0.50}} &= e^2 && (7)\\
\frac{p^2}{1-p^2} &= 7.3 && (8)\\
p^2(1+7.3) &= 7.3 && (9)\\
p^2 &= 7.3/(1+7.3) = 88\% && (10)
\end{aligned}$$

The probability increases from 50% to 88%.

```
[ ]:
```