Logistic Regression

How to calculate, manually, odds and probability of a new observation

The file bank.csv includes data on a sample of 20 banks. The “Financial Condition” column records the judgment of an expert on the financial condition of each bank. The outcome variable (response or target) takes one of the two possible values-weak or strong.

The following is a sample of our dataset.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Obs | Financial Condition | TotCap/Assets | TotExp/Assets | TotLns&Lses/Assets |
| 1 | 0 | 12.5 | 0.09 | 0.3 |
| 2 | 0 | 9.6 | 0.09 | 0.73 |
| 3 | 1 | 4.3 | 0.11 | 0.69 |
| 4 | 0 | 18.3 | 0.08 | 0.49 |
| 5 | 1 | 9.1 | 0.13 | 0.74 |
| 6 | 0 | 20.5 | 0.12 | 0.8 |

We applied logistic regression on only “*TotExp/Assets”* and “*TotLns&Lses/Assets”*. The summary of coefficients of Logistic Regression model is below. We want to use this model to manually classify a new bank as financially weak or strong.

| | Estimate. | Std. Error | z value | Pr(>&#124;z&#124;) |

|:---------------------- |--------------:|--------------:|-------------:|---------------------------:|

|(Intercept) | -14.72101| 6.674899 | -2.205428| 0.0274241 |

|TotExp.Assets | 89.83392| 47.780807| 1.880125| 0.0600910 |

|TotLns.Lses.Assets| 8.37132 | 5.778804 | 1.448625 | 0.1474423 |

The Logit equation is:

Where ln(Odds) is the natural log (*loge*) of Odds.

Now let’s assume the following observation is about a new bank and we want to classify it as weak or strong based on the given predictors’ value

|  |  |  |  |
| --- | --- | --- | --- |
| Financial Condition | TotCap/Assets | TotExp/Assets | TotLns&Lses/Assets |
| ? | 19.1 | 0.10 | 0.60 |

Remember that we ignored TotCap/Assets

Step 1: Compute the logit.

Replacing selected predictors with their values

Logit =

Logit = -0.714826

Then

Using XL exp() function it automatically raise a number to the power of e (2.71828182845904)

And

So the probability of this new financial company is strong is 33% setting the cutoff as 50% then it is weak!