Jared Yu

Module 11 Discussion

1. In a logistic regression model with only one regressor and its corresponding regression coefficient . What does mean? Suppose that we obtain the maximum likelihood estimator of . Discuss how to construct a confidence interval of the odds ratio.

Ans:

The question seems slightly unclear, so for the sake of the discussion I will be assuming that it is referring to

To understand the parameter (which in this case is ), we can look at the input into the sigmoid function,

which is the fitted value of the linear predictor at a particular value of (i.e., ). We can look also at the fitted value for , which is as follows,

Then taking the difference between these two we get,

The functions are also known as the *log-odds*, so we can rewrite the difference as follows,

Taking the antilog of the above difference yields the *odds ratio*,

In the textbook, it is said to be, “the estimated increase in the probability of success associated with a one-unit change in the value of the predictor variable.” This can be generalized to a -unit change as a change of . So, for example, if for example is associated with height in centimeters, then a 1-centimeter increase in height is associated with a change in probability of success for the response .

To construct a confidence interval for the odds ratio, we first need to look at the test for the individual model coefficient. The hypothesis test is as follows,

The test statistic is as follows,

which follows a standard normal distribution. The standard error in the denominator can be found by finding the covariance matrix of ,

where is the Hessian matrix found from the log-likelihood function. From here, it can be seen that the confidence interval for would be as follows,

In this confidence interval, we have that is the critical value at the confidence level for the standard normal distribution. We have from before that , so the above confidence interval can be changed to the following,

The above gives a confidence interval of the odds ratio.

1. Consider a logistic regression model with a linear predictor that includes an interaction term, say . Does the odds ratio associated with the regressor have the same interpretation as in the case where the linear predictor does not have the interaction term? Provide mathematical development to support your answer.

Ans:

We have now that the (fitted) logistic regression model is as follows,

Using a similar log-odds function as before from problem 1, we have the following for the odds ratio of ,

Then again taking the antilog, we obtain

It can be seen quite evidently that when excluding the interaction term, we’d only be left with , which is the same as before with only in the model. Therefore, the interpretation is different, and we’d have to consider the value of .