Abstract

This research aimed to firstly study relationship and influence of two-way interaction between respond proportions, reliability in explained variable and misclassification rate in dependent variable with the number of explained variables, size of samples, and true coefficient of determination which was appointed to be five adjusted coefficients of determination$\left(R_{adj}^2\right)$. The studied R_{adj}^2 was the R_{adj}^2 introduced by Mittlbock and Schemper$\left(R_{O,adj,MS}^2, R_{l,adj,MS}^2\right)$, Shtatland ($R_{l,adj,SAS_{AIC}}^2$), and Liao and McGee $\left(R_{O,adj,LM}^2, R_{l,adj,LM}^2\right)$ for latent binary logistic regression model. Secondly, the research also aimed to study a bias of adjusted coefficient of determination. In terms of the analysis, Monte Carlo simulation technique which varied the explained variables (1, 5, and 10) and size of samples (50, 250, 500 and 1000) was used in the experiment. According to 1000 times repeated experiment, this research revealed the following results.

First, the response proportion, the reliability in explained variable and misclassification rate in dependent variable were associated to R_{adj}^2 . Second, the two-way interaction between those three mentioned factors and the number of explained variables, sample sizes, and true coefficient of determination effected to R_{adj}^2 as well. The $R_{O,adj,LM}^2$ was the best statistic in terms of giving the lowest relationship between the response proportion with the number of explained variables, size of samples and true coefficient of determination. Conversely, it was shown the highest relationship between the reliability rate in explained variable with the number of explained variables and sample sizes. Moreover, the $R_{O,adj,LM}^2$ value showed a negative relationship between the misclassification rates in dependent variable with the number of explained variables and sample sizes. According to these results, it could be considered as a good characteristic of R_{adj}^2 value.

According to study about the bias of R^2_{adj}, this research found that the $R^2_{O,adj,LM}$ provided the best result in all conditions of investigation. Moreover, it was shown that the bias of R^2_{adj} was reduced when the reliability in explained variable were increased. On the other hand, according to misclassification rate in dependent variable, the propensity of R^2_{adj} was increased when there was increase in misclassification rate in dependent variable.